Red Hat Enterprise Linux 9

Configuring and managing networking

Managing network interfaces and advanced networking features
Red Hat Enterprise Linux 9 Configuring and managing networking

Managing network interfaces and advanced networking features
Abstract

Using the networking capabilities of Red Hat Enterprise Linux (RHEL), you can configure your host to meet your organization’s network and security requirements. For example: You can configure bonds, VLANs, bridges, tunnels and other network types to connect the host to the network. IPSec and WireGuard provide secure VPNs between hosts and networks. RHEL also supports advanced networking features, such as policy-based routing and MultiPath TCP (MPTCP).
# Table of Contents

## MAKING OPEN SOURCE MORE INCLUSIVE

1. PROVIDING FEEDBACK ON RED HAT DOCUMENTATION ......................................................... 9

## CHAPTER 1. IMPLEMENTING CONSISTENT NETWORK INTERFACE NAMING

1. HOW THE UDEV DEVICE MANAGER RENAMES NETWORK INTERFACES ............................. 11
2. NETWORK INTERFACE NAMING POLICIES ........................................................................... 12
3. NETWORK INTERFACE NAMING SCHEMES ....................................................................... 13
4. CUSTOMIZING THE PREFIX FOR ETHERNET INTERFACES DURING INSTALLATION ............ 14
5. CONFIGURING USER-DEFINED NETWORK INTERFACE NAMES BY USING UDEV RULES ........... 16
6. CONFIGURING USER-DEFINED NETWORK INTERFACE NAMES BY USING SYSTEMD LINK FILES ................................................................. 18

## CHAPTER 2. CONFIGURING AN ETHERNET CONNECTION

1. CONFIGURING AN ETHERNET CONNECTION BY USING NMCLI ............................................. 20
2. CONFIGURING AN ETHERNET CONNECTION BY USING THE NMCLI INTERACTIVE EDITOR .......... 23
3. CONFIGURING AN ETHERNET CONNECTION BY USING NMTUI ............................................ 26
4. CONFIGURING AN ETHERNET CONNECTION BY USING CONTROL-CENTER ...................... 29
5. CONFIGURING AN ETHERNET CONNECTION BY USING NM-CONNECTION-EDITOR .......... 32
6. CONFIGURING AN ETHERNET CONNECTION WITH A STATIC IP ADDRESS BY USING NMSTATECTL ................................................................. 34
7. CONFIGURING AN ETHERNET CONNECTION WITH A STATIC IP ADDRESS BY USING THE NETWORK RHEL SYSTEM ROLE WITH AN INTERFACE NAME ................................................................. 37
8. CONFIGURING AN ETHERNET CONNECTION WITH A STATIC IP ADDRESS BY USING THE NETWORK RHEL SYSTEM ROLE WITH A DEVICE PATH ................................................................. 38
9. CONFIGURING AN ETHERNET CONNECTION WITH A DYNAMIC IP ADDRESS BY USING NMSTATECTL ................................................................. 40
10. CONFIGURING AN ETHERNET CONNECTION WITH A DYNAMIC IP ADDRESS BY USING THE NETWORK RHEL SYSTEM ROLE WITH AN INTERFACE NAME ................................................................. 42
11. CONFIGURING AN ETHERNET CONNECTION WITH A DYNAMIC IP ADDRESS BY USING THE NETWORK RHEL SYSTEM ROLE WITH A DEVICE PATH ................................................................. 43
12. CONFIGURING MULTIPLE ETHERNET INTERFACES BY USING A SINGLE CONNECTION PROFILE BY INTERFACE NAME ................................................................. 45
13. CONFIGURING A SINGLE CONNECTION PROFILE FOR MULTIPLE ETHERNET INTERFACES USING PCI IDS ................................................................. 46

## CHAPTER 3. CONFIGURING NETWORK BONDING

1. UNDERSTANDING THE DEFAULT BEHAVIOR OF CONTROLLER AND PORT INTERFACES ......... 48
2. UPSTREAM SWITCH CONFIGURATION DEPENDING ON THE BONDING MODES .................... 48
3. CONFIGURING A NETWORK BOND BY USING NMCLI .......................................................... 49
4. CONFIGURING A NETWORK BOND BY USING THE RHEL WEB CONSOLE .......................... 52
5. CONFIGURING A NETWORK BOND BY USING NMTUI ....................................................... 55
6. CONFIGURING A NETWORK BOND BY USING NM-CONNECTION-EDITOR ....................... 58
7. CONFIGURING A NETWORK BOND BY USING NMSTATECTL .............................................. 60
8. CONFIGURING A NETWORK BOND BY USING THE NETWORK RHEL SYSTEM ROLE .......... 62
9. CREATING A NETWORK BOND TO ENABLE SWITCHING BETWEEN AN ETHERNET AND WIRELESS CONNECTION WITHOUT INTERRUPTING THE VPN ................................................................. 64
10. THE DIFFERENT NETWORK BONDING MODES ................................................................. 67
11. THE XMIT_HASH_POLICY BONDING PARAMETER ......................................................... 68

## CHAPTER 4. CONFIGURING NETWORK TEAMING

1. MIGRATING A NETWORK TEAM CONFIGURATION TO NETWORK BOND ............................. 71
2. UNDERSTANDING THE DEFAULT BEHAVIOR OF CONTROLLER AND PORT INTERFACES .......... 74
3. UNDERSTANDING THE TEAMD SERVICE, RUNNERS, AND LINK-WATCHERS ............................. 74
4. CONFIGURING A NETWORK TEAM BY USING NMCLI .......................................................... 75
4.5. CONFIGURING A NETWORK TEAM BY USING THE RHEL WEB CONSOLE 78
4.6. CONFIGURING A NETWORK TEAM BY USING NM-CONNECTION-EDITOR 82

CHAPTER 5. CONFIGURING VLAN TAGGING .......................................................... 85
5.1. CONFIGURING VLAN TAGGING BY USING NMCLI 85
5.2. CONFIGURING NESTED VLANS BY USING NMCLI 87
5.3. CONFIGURING VLAN TAGGING BY USING THE RHEL WEB CONSOLE 89
5.4. CONFIGURING VLAN TAGGING BY USING NMTUI 91
5.5. CONFIGURING VLAN TAGGING BY USING NM-CONNECT EDITOR 95
5.6. CONFIGURING VLAN TAGGING BY USING NMSTATECTL 97
5.7. CONFIGURING VLAN TAGGING BY USING THE NETWORK RHEL SYSTEM ROLE 99
5.8. ADDITIONAL RESOURCES 101

CHAPTER 6. CONFIGURING A NETWORK BRIDGE .............................................. 102
6.1. CONFIGURING A NETWORK BRIDGE BY USING NMCLI 102
6.2. CONFIGURING A NETWORK BRIDGE BY USING THE RHEL WEB CONSOLE 105
6.3. CONFIGURING A NETWORK BRIDGE BY USING NMTUI 107
6.4. CONFIGURING A NETWORK BRIDGE BY USING NM-CONNECTION-EDITOR 111
6.5. CONFIGURING A NETWORK BRIDGE BY USING NMSTATECTL 113
6.6. CONFIGURING A NETWORK BRIDGE BY USING THE NETWORK RHEL SYSTEM ROLE 115

CHAPTER 7. SETTING UP AN IPSEC VPN ......................................................... 118
7.1. CONFIGURING A VPN CONNECTION WITH CONTROL CENTER 118
7.2. CONFIGURING A VPN CONNECTION USING NM-CONNECTION EDITOR 122
7.3. CONFIGURING AUTOMATIC DETECTION AND USAGE OF ESP HARDWARE OFFLOAD TO ACCELERATE AN IPSEC CONNECTION 125
7.4. CONFIGURING ESP HARDWARE OFFLOAD ON A BOND TO ACCELERATE AN IPSEC CONNECTION 126

CHAPTER 8. SETTING UP A WIREGUARD VPN .................................................. 128
8.1. PROTOCOLS AND PRIMITIVES USED BY WIREGUARD 128
8.2. HOW WIREGUARD USES TUNNEL IP ADDRESSES, PUBLIC KEYS, AND REMOTE ENDPOINTS 129
8.3. USING A WIREGUARD CLIENT BEHIND NAT AND FIREWALLS 129
8.4. CREATING PRIVATE AND PUBLIC KEYS TO BE USED IN WIREGUARD CONNECTIONS 129
8.5. CONFIGURING A WIREGUARD SERVER BY USING NMCLI 130
8.6. CONFIGURING A WIREGUARD SERVER BY USING NMTUI 133
8.7. CONFIGURING A WIREGUARD SERVER BY USING NM-CONNECTION-EDITOR 136
8.8. CONFIGURING A WIREGUARD SERVER BY USING THE WG-QUICK SERVICE 138
8.9. CONFIGURING FIREWALLD ON A WIREGUARD SERVER BY USING THE COMMAND LINE 140
8.10. CONFIGURING FIREWALLD ON A WIREGUARD SERVER BY USING THE GRAPHICAL INTERFACE 141
8.11. CONFIGURING A WIREGUARD CLIENT BY USING NMCLI 141
8.12. CONFIGURING A WIREGUARD CLIENT BY USING NMTUI 144
8.13. CONFIGURING A WIREGUARD CLIENT BY USING NM-CONNECTION-EDITOR 148
8.14. CONFIGURING A WIREGUARD CLIENT BY USING THE WG-QUICK SERVICE 150

CHAPTER 9. CONFIGURING IP TUNNELS ............................................................ 154
9.1. CONFIGURING AN IPIP TUNNEL USING NMCLI TO ENCAPSULATE IPV4 TRAFFIC IN IPV4 PACKETS 154
9.2. CONFIGURING A GRE TUNNEL USING NMCLI TO ENCAPSULATE LAYER-3 TRAFFIC IN IPV4 PACKETS 157
9.3. CONFIGURING A GRETAP TUNNEL TO TRANSFER ETHERNET FRAMES OVER IPV4 159
9.4. ADDITIONAL RESOURCES 162

CHAPTER 10. USING A VXLAN TO CREATE A VIRTUAL LAYER-2 DOMAIN FOR VMS .................................................... 163
10.1. BENEFITS OF VXLANS 163
10.2. CONFIGURING THE ETHERNET INTERFACE ON THE HOSTS 164
10.3. CREATING A NETWORK BRIDGE WITH A VXLAN ATTACHED 165
CHAPTER 22. CONFIGURING THE ORDER OF DNS SERVERS

CHAPTER 21. MANUALLY CONFIGURING THE /ETC/RESOLV.CONF FILE

CHAPTER 20. RUNNING DHCPCLIENT EXIT HOOKS USING NETWORKMANAGER A DISPATCHER SCRIPT

CHAPTER 19. CONFIGURING NETWORKMANAGER DHCP SETTINGS

CHAPTER 18. CHANGING A HOSTNAME

CHAPTER 17. USING NETWORKMANAGER TO DISABLE IPV6 FOR A SPECIFIC CONNECTION

CHAPTER 16. CREATING A DUMMY INTERFACE

CHAPTER 15. CONFIGURING NETWORKMANAGER TO IGNORE CERTAIN DEVICES

CHAPTER 14. GETTING STARTED WITH IPVLAN

CHAPTER 13. USING MACSEC TO ENCRYPT LAYER-2 TRAFFIC IN THE SAME PHYSICAL NETWORK

CHAPTER 12. CONFIGURING RHEL AS A WPA2 OR WPA3 PERSONAL ACCESS POINT

CHAPTER 11. MANAGING WIFI CONNECTIONS

CHAPTER 10. SUPPORTED WIFI SECURITY TYPES

CHAPTER 10. CONFIGURING VIRTUAL MACHINES TO USE VXLAN

Table of Contents
22.1. HOW NETWORKMANAGER ORDERS DNS SERVERS IN /ETC/RESOLV.CONF .......................... 208
   Default values of DNS priority parameters .................................................. 208
   Valid DNS priority values: ............................................................................. 208
22.2. SETTING A NETWORKMANAGER-WIDE DEFAULT DNS SERVER PRIORITY VALUE .... 209
22.3. SETTING THE DNS PRIORITY OF A NETWORKMANAGER CONNECTION .............. 210

CHAPTER 23. USING DIFFERENT DNS SERVERS FOR DIFFERENT DOMAINS .................. 211
23.1. USING DNMASQ IN NETWORKMANAGER TO SEND DNS REQUESTS FOR A SPECIFIC
      DOMAIN TO A SELECTED DNS SERVER ....................................................... 211
23.2. USING SYSTEMD-RESOLVED IN NETWORKMANAGER TO SEND DNS REQUESTS FOR A
      SPECIFIC DOMAIN TO A SELECTED DNS SERVER ..................................... 213

CHAPTER 24. MANAGING THE DEFAULT GATEWAY SETTING ........................................ 216
24.1. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING NMCLI 216
24.2. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING THE NMCLI
      INTERACTIVE MODE ..................................................................................... 217
24.3. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING NM-CONNECTION-
      EDITOR .......................................................................................................... 218
24.4. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING CONTROL-CENTER 220
24.5. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING NMSTATECTL 221
24.6. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING THE NETWORK RHEL
      SYSTEM ROLE .............................................................................................. 222
24.7. HOW NETWORKMANAGER MANAGES MULTIPLE DEFAULT GATEWAYS ............ 223
24.8. CONFIGURING NETWORKMANAGER TO AVOID USING A SPECIFIC PROFILE TO PROVIDE A DEFAULT
      GATEWAY ....................................................................................................... 225
24.9. FIXING UNEXPECTED ROUTING BEHAVIOR DUE TO MULTIPLE DEFAULT GATEWAYS 225

CHAPTER 25. CONFIGURING STATIC ROUTES .............................................................. 228
25.1. EXAMPLE OF A NETWORK THAT REQUIRES STATIC ROUTES ............................... 228
25.2. HOW TO USE THE NMCLI COMMAND TO CONFIGURE A STATIC ROUTE ............ 230
25.3. CONFIGURING A STATIC ROUTE BY USING NMCLI ........................................... 231
25.4. CONFIGURING A STATIC ROUTE BY USING NMTUI .......................................... 232
25.5. CONFIGURING A STATIC ROUTE BY USING CONTROL-CENTER ......................... 234
25.6. CONFIGURING A STATIC ROUTE BY USING NM-CONNECTION-EDITOR ............. 236
25.7. CONFIGURING A STATIC ROUTE BY USING THE NMCLI INTERACTIVE MODE ........ 237
25.8. CONFIGURING A STATIC ROUTE BY USING NMSTATECTL ............................... 239
25.9. CONFIGURING A STATIC ROUTE BY USING THE NETWORK RHEL SYSTEM ROLE .... 240

CHAPTER 26. CONFIGURING POLICY-BASED ROUTING TO DEFINE ALTERNATIVE ROUTES .... 243
26.1. ROUTING TRAFFIC FROM A SPECIFIC SUBNET TO A DIFFERENT DEFAULT GATEWAY BY USING
      NMCLI ........................................................................................................... 243
26.2. ROUTING TRAFFIC FROM A SPECIFIC SUBNET TO A DIFFERENT DEFAULT GATEWAY BY USING THE
      NETWORK RHEL SYSTEM ROLE .................................................................... 247

CHAPTER 27. REUSING THE SAME IP ADDRESS ON DIFFERENT INTERFACES ............... 251
27.1. PERMANENTLY REUSING THE SAME IP ADDRESS ON DIFFERENT INTERFACES .... 251
27.2. TEMPORARILY REUSING THE SAME IP ADDRESS ON DIFFERENT INTERFACES .... 252
27.3. ADDITIONAL RESOURCES ............................................................................ 254

CHAPTER 28. STARTING A SERVICE WITHIN AN ISOLATED VRF NETWORK .................. 255
28.1. CONFIGURING A VRF DEVICE ....................................................................... 255
28.2. STARTING A SERVICE WITHIN AN ISOLATED VRF NETWORK ......................... 256

CHAPTER 29. CONFIGURING ETHTOOL SETTINGS IN NETWORKMANAGER CONNECTION
      PROFILES ........................................................................................................ 259
29.1. CONFIGURING AN ETHTOOL OFFLOAD FEATURE BY USING NMCLI ................. 259
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>29.2. Configuring an Ethtool Offload Feature by Using the Network RHEL System Role</td>
<td>260</td>
</tr>
<tr>
<td>29.3. Configuring an Ethtool Coalesce Settings by Using Nmcli</td>
<td>261</td>
</tr>
<tr>
<td>29.4. Configuring an Ethtool Coalesce Settings by Using the Network RHEL System Role</td>
<td>262</td>
</tr>
<tr>
<td>29.5. Increasing the Ring Buffer Size to Reduce a High Packet Drop Rate by Using Nmcli</td>
<td>264</td>
</tr>
<tr>
<td>29.6. Increasing the Ring Buffer Size to Reduce a High Packet Drop Rate by Using the Network RHEL System Role</td>
<td>265</td>
</tr>
<tr>
<td>Chapter 30. Introduction to NetworkManager Debugging</td>
<td>268</td>
</tr>
<tr>
<td>30.1. Introduction to NetworkManager Reapply Method</td>
<td>268</td>
</tr>
<tr>
<td>30.2. Setting the NetworkManager Log Level</td>
<td>270</td>
</tr>
<tr>
<td>30.3. Temporarily Setting Log Levels at Run Time Using Nmcli</td>
<td>271</td>
</tr>
<tr>
<td>30.4. Viewing NetworkManager Logs</td>
<td>272</td>
</tr>
<tr>
<td>30.5. Debugging Levels and Domains</td>
<td>272</td>
</tr>
<tr>
<td>Chapter 31. Using LLDP to Debug Network Configuration Problems</td>
<td>274</td>
</tr>
<tr>
<td>31.1. Debugging an Incorrect VLAN Configuration Using LLDP Information</td>
<td>274</td>
</tr>
<tr>
<td>Chapter 32. Linux Traffic Control</td>
<td>277</td>
</tr>
<tr>
<td>32.1. Overview of Queuing Disciplines</td>
<td>277</td>
</tr>
<tr>
<td>32.2. Introduction to Connection Tracking</td>
<td>277</td>
</tr>
<tr>
<td>32.3. Inspecting Qdiscs of a Network Interface Using the TC Utility</td>
<td>278</td>
</tr>
<tr>
<td>32.4. Updating the Default Qdisc</td>
<td>279</td>
</tr>
<tr>
<td>32.5. Temporarily Setting the Current Qdisc of a Network Interface Using the TC Utility</td>
<td>279</td>
</tr>
<tr>
<td>32.6. Permanently Setting the Current Qdisc of a Network Interface Using NetworkManager</td>
<td>280</td>
</tr>
<tr>
<td>32.7. Configuring the Rate Limiting of Packets by Using the TC-CTINFO Utility</td>
<td>281</td>
</tr>
<tr>
<td>32.8. Available Qdiscs in RHEL</td>
<td>285</td>
</tr>
<tr>
<td>Chapter 33. Authenticating a RHEL Client to the Network by Using the 802.1x Standard with a Certificate Stored on the File System</td>
<td>287</td>
</tr>
<tr>
<td>33.1. Configuring 802.1x Network Authentication on an Existing Ethernet Connection by Using Nmcli</td>
<td>287</td>
</tr>
<tr>
<td>33.2. Configuring a Static Ethernet Connection with 802.1x Network Authentication by Using NmStatectl</td>
<td>288</td>
</tr>
<tr>
<td>33.3. Configuring a Static Ethernet Connection with 802.1x Network Authentication by Using the Network RHEL System Role</td>
<td>290</td>
</tr>
<tr>
<td>33.4. Configuring a WiFi Connection with 802.1x Network Authentication by Using the Network RHEL System Role</td>
<td>292</td>
</tr>
<tr>
<td>Chapter 34. Setting Up an 802.1x Network Authentication Service for LAN Clients by Using Hostapd with Freeradius Backend</td>
<td>295</td>
</tr>
<tr>
<td>34.1. Prerequisites</td>
<td>295</td>
</tr>
<tr>
<td>34.2. Setting Up the Bridge on the Authenticator</td>
<td>295</td>
</tr>
<tr>
<td>34.3. Certificate Requirements by Freeradius</td>
<td>296</td>
</tr>
<tr>
<td>34.4. Creating a Set of Certificates on a Freeradius Server for Testing Purposes</td>
<td>297</td>
</tr>
<tr>
<td>34.5. Configuring Freeradius to Authenticate Network Clients Securely by Using Eap</td>
<td>299</td>
</tr>
<tr>
<td>34.6. Configuring Hostapd as an Authenticator in a Wired Network</td>
<td>302</td>
</tr>
<tr>
<td>34.7. Testing Eap-TTLS Authentication Against a Freeradius Server or Authenticator</td>
<td>305</td>
</tr>
<tr>
<td>34.8. Testing Eap-Tls Authentication Against a Freeradius Server or Authenticator</td>
<td>306</td>
</tr>
<tr>
<td>34.9. Blocking and Allowing Traffic Based on Hostapd Authentication Events</td>
<td>308</td>
</tr>
<tr>
<td>Chapter 35. Getting Started with Multipath TCP</td>
<td>311</td>
</tr>
<tr>
<td>35.1. Understanding Mptcp</td>
<td>311</td>
</tr>
</tbody>
</table>
35.2. PREPARING RHEL TO ENABLE MPTCP SUPPORT
35.3. USING IPROUTE2 TO TEMPORARILY CONFIGURE AND ENABLE MULTIPLE PATHS FOR MPTCP APPLICATIONS
35.4. PERMANENTLY CONFIGURING MULTIPLE PATHS FOR MPTCP APPLICATIONS
35.5. MONITORING MPTCP SUB-FLOWS
35.6. DISABLING MULTIPATH TCP IN THE KERNEL

CHAPTER 36. MANAGING THE MPTCPD SERVICE
36.1. CONFIGURING MPTCPD
36.2. MANAGING APPLICATIONS WITH MPTCPIZE TOOL
36.3. ENABLING MPTCP SOCKETS FOR A SERVICES USING THE MPTCPIZE UTILITY

CHAPTER 37. NETWORKMANAGER CONNECTION PROFILES IN KEYFILE FORMAT
37.1. THE KEYFILE FORMAT OF NETWORKMANAGER PROFILES
37.2. USING NMCLI TO CREATE KEYFILE CONNECTION PROFILES IN OFFLINE MODE
37.3. MANUALLY CREATING A NETWORKMANAGER PROFILE IN KEYFILE FORMAT
37.4. THE DIFFERENCES IN INTERFACE RENAMING WITH PROFILES IN IFCFG AND KEYFILE FORMAT
37.5. MIGRATING NETWORKMANAGER PROFILES FROM IFCFG TO KEYFILE FORMAT

CHAPTER 38. SYSTEMD NETWORK TARGETS AND SERVICES
38.1. DIFFERENCES BETWEEN THE NETWORK AND NETWORK-ONLINE SYSTEMD TARGET
38.2. OVERVIEW OF NETWORKMANAGER-WAIT-ONLINE
38.3. CONFIGURING A SYSTEMD SERVICE TO START AFTER THE NETWORK HAS BEEN STARTED

CHAPTER 39. INTRODUCTION TO NMSTATE
39.1. USING THE LIBNMSTATE LIBRARY IN A PYTHON APPLICATION
39.2. UPDATING THE CURRENT NETWORK CONFIGURATION USING NMSTATECTL
39.3. THE NMSTATE SYSTEMD SERVICE
39.4. NETWORK STATES FOR THE NETWORK RHEL SYSTEM ROLE
39.5. ADDITIONAL RESOURCES

CHAPTER 40. CAPTURING NETWORK PACKETS
40.1. USING XDPDUMP TO CAPTURE NETWORK PACKETS INCLUDING PACKETS DROPPED BY XDP PROGRAMS
40.2. ADDITIONAL RESOURCES

CHAPTER 41. UNDERSTANDING THE EBPF NETWORKING FEATURES IN RHEL 9
41.1. OVERVIEW OF NETWORKING EBPF FEATURES IN RHEL 9
  XDP
  AF_XDP
  Traffic Control
  Socket filter
  Control Groups
  Stream Parser
  SO_REUSEPORT socket selection
  Flow dissector
  TCP Congestion Control
  Routes with encapsulation
  Socket lookup
41.2. OVERVIEW OF XDP FEATURES IN RHEL 9 BY NETWORK CARDS

CHAPTER 42. NETWORK TRACING USING THE BPF COMPILER COLLECTION
42.1. INSTALLING THE BCC-TOOLS PACKAGE
42.2. DISPLAYING TCP CONNECTIONS ADDED TO THE KERNEL’S ACCEPT QUEUE
42.3. TRACING OUTGOING TCP CONNECTION ATTEMPTS
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>42.4. MEASURING THE LATENCY OF OUTGOING TCP CONNECTIONS</td>
<td>345</td>
</tr>
<tr>
<td>42.5. DISPLAYING DETAILS ABOUT TCP PACKETS AND SEGMENTS THAT WERE DROPPED BY THE KERNEL</td>
<td>345</td>
</tr>
<tr>
<td>42.6. TRACING TCP SESSIONS</td>
<td>346</td>
</tr>
<tr>
<td>42.7. TRACING TCP RETRANSMISSIONS</td>
<td>347</td>
</tr>
<tr>
<td>42.8. DISPLAYING TCP STATE CHANGE INFORMATION</td>
<td>347</td>
</tr>
<tr>
<td>42.9. SUMMARIZING AND AGGREGATING TCP TRAFFIC SENT TO SPECIFIC SUBNETS</td>
<td>348</td>
</tr>
<tr>
<td>42.10. DISPLAYING THE NETWORK THROUGHPUT BY IP ADDRESS AND PORT</td>
<td>349</td>
</tr>
<tr>
<td>42.11. TRACING ESTABLISHED TCP CONNECTIONS</td>
<td>349</td>
</tr>
<tr>
<td>42.12. TRACING IPV4 AND IPV6 LISTEN ATTEMPTS</td>
<td>350</td>
</tr>
<tr>
<td>42.13. SUMMARIZING THE SERVICE TIME OF SOFT INTERRUPTS</td>
<td>350</td>
</tr>
<tr>
<td>42.14. SUMMARIZING PACKETS SIZE AND COUNT ON A NETWORK INTERFACE</td>
<td>351</td>
</tr>
<tr>
<td>42.15. ADDITIONAL RESOURCES</td>
<td>352</td>
</tr>
<tr>
<td><strong>CHAPTER 43. CONFIGURING NETWORK DEVICES TO ACCEPT TRAFFIC FROM ALL MAC ADDRESSES</strong></td>
<td>353</td>
</tr>
<tr>
<td>43.1. TEMPORARILY CONFIGURING A DEVICE TO ACCEPT ALL TRAFFIC</td>
<td>353</td>
</tr>
<tr>
<td>43.2. PERMANENTLY CONFIGURING A NETWORK DEVICE TO ACCEPT ALL TRAFFIC USING NMCLI</td>
<td>354</td>
</tr>
<tr>
<td>43.3. PERMANENTLY CONFIGURING A NETWORK DEVICE TO ACCEPT ALL TRAFFIC USING NMSTATECTL</td>
<td>354</td>
</tr>
<tr>
<td><strong>CHAPTER 44. MIRRORING A NETWORK INTERFACE BY USING NMCLI</strong></td>
<td>356</td>
</tr>
<tr>
<td><strong>CHAPTER 45. USING NMSTATE-AUTOCONF TO AUTOMATICALLY CONFIGURE THE NETWORK STATE USING LLDP</strong></td>
<td>358</td>
</tr>
<tr>
<td>45.1. USING NMSTATE-AUTOCONF TO AUTOMATICALLY CONFIGURE NETWORK INTERFACES</td>
<td>358</td>
</tr>
<tr>
<td><strong>CHAPTER 46. CONFIGURING 802.3 LINK SETTINGS</strong></td>
<td>361</td>
</tr>
<tr>
<td>46.1. CONFIGURING 802.3 LINK SETTINGS USING THE NMCLI UTILITY</td>
<td>361</td>
</tr>
<tr>
<td><strong>CHAPTER 47. GETTING STARTED WITH DPDK</strong></td>
<td>363</td>
</tr>
<tr>
<td>47.1. INSTALLING THE DPDK PACKAGE</td>
<td>363</td>
</tr>
<tr>
<td>47.2. ADDITIONAL RESOURCES</td>
<td>363</td>
</tr>
<tr>
<td><strong>CHAPTER 48. GETTING STARTED WITH TIPC</strong></td>
<td>364</td>
</tr>
<tr>
<td>48.1. THE ARCHITECTURE OF TIPC</td>
<td>364</td>
</tr>
<tr>
<td>48.2. LOADING THE TIPC MODULE WHEN THE SYSTEM BOOTS</td>
<td>364</td>
</tr>
<tr>
<td>48.3. CREATING A TIPC NETWORK</td>
<td>365</td>
</tr>
<tr>
<td>48.4. ADDITIONAL RESOURCES</td>
<td>366</td>
</tr>
<tr>
<td><strong>CHAPTER 49. AUTOMATICALLY CONFIGURING NETWORK INTERFACES IN PUBLIC CLOUDS USING NM-CLOUD-SETUP</strong></td>
<td>367</td>
</tr>
<tr>
<td>49.1. CONFIGURING AND PRE-DEPLOYING NM-CLOUD-SETUP</td>
<td>367</td>
</tr>
<tr>
<td>49.2. UNDERSTANDING THE ROLE OF IMDSV2 AND NM-CLOUD-SETUP IN THE RHEL EC2 INSTANCE</td>
<td>368</td>
</tr>
</tbody>
</table>
Red Hat is committed to replacing problematic language in our code, documentation, and web properties. We are beginning with these four terms: master, slave, blacklist, and whitelist. Because of the enormity of this endeavor, these changes will be implemented gradually over several upcoming releases. For more details, see our CTO Chris Wright’s message.
PROVIDING FEEDBACK ON RED HAT DOCUMENTATION

We appreciate your feedback on our documentation. Let us know how we can improve it.

Submitting feedback through Jira (account required)

1. Log in to the Jira website.
2. Click Create in the top navigation bar
3. Enter a descriptive title in the Summary field.
4. Enter your suggestion for improvement in the Description field. Include links to the relevant parts of the documentation.
5. Click Create at the bottom of the dialogue.
CHAPTER 1. IMPLEMENTING CONSISTENT NETWORK INTERFACE NAMING

The udev device manager implements consistent device naming in Red Hat Enterprise Linux. The device manager supports different naming schemes and, by default, assigns fixed names based on firmware, topology, and location information.

Without consistent device naming, the Linux kernel assigns names to network interfaces by combining a fixed prefix and an index. The index increases as the kernel initializes the network devices. For example, eth0 represents the first Ethernet device being probed on start-up. If you add another network interface controller to the system, the assignment of the kernel device names is no longer fixed because, after a reboot, the devices can initialize in a different order. In that case, the kernel can name the devices differently.

To solve this problem, udev assigns consistent device names. This has the following advantages:

- Device names are stable across reboots.
- Device names stay fixed even if you add or remove hardware.
- Defective hardware can be seamlessly replaced.
- The network naming is stateless and does not require explicit configuration files.

**WARNING**

Generally, Red Hat does not support systems where consistent device naming is disabled. For exceptions, see the [Is it safe to set net.ifnames=0](#) solution.

1.1. HOW THE UDEV DEVICE MANAGER RENAMES NETWORK INTERFACES

To implement a consistent naming scheme for network interfaces, the udev device manager processes the following rule files in the listed order:

   This file exists only if you install the `initscripts-rename-device` package. The `/usr/lib/udev/rules.d/60-net.rules` file defines that the deprecated `/usr/lib/udev/rename_device` helper utility searches for the HWADDR parameter in `/etc/sysconfig/network-scripts/ifcfg-*` files. If the value set in the variable matches the MAC address of an interface, the helper utility renames the interface to the name set in the DEVICE parameter of the ifcfg file.

   If the system uses only NetworkManager connection profiles in keyfile format, udev skips this step.

2. Only on Dell systems: `/usr/lib/udev/rules.d/71-biosdevname.rules`  
   This file exists only if the `biosdevname` package is installed, and the rules file defines that the biosdevname utility renames the interface according to its naming policy, if it was not renamed in the previous step.
NOTE

Install and use biosdevname only on Dell systems.

3. /usr/lib/udev/rules.d/75-net-description.rules
   This file defines how udev examines the network interface and sets the properties in udev-internal variables. These variables are then processed in the next step by the /usr/lib/udev/rules.d/80-net-setup-link.rules file. Some of the properties can be undefined.

   This file calls the net_setup_link builtin of the udev service, and udev renames the interface based on the order of the policies in the NamePolicy parameter in the /usr/lib/systemd/network/99-default.link file. For further details, see Network interface naming policies.

   If none of the policies applies, udev does not rename the interface.

Additional resources

- Why are systemd network interface names different between major RHEL versions

1.2. NETWORK INTERFACE NAMING POLICIES

By default, the udev device manager uses the /usr/lib/systemd/network/99-default.link file to determine which device naming policies to apply when it renames interfaces. The NamePolicy parameter in this file defines which policies udev uses and in which order:

   NamePolicy=keep kernel database onboard slot path

The following table describes the different actions of udev based on which policy matches first as specified by the NamePolicy parameter:

<table>
<thead>
<tr>
<th>Policy</th>
<th>Description</th>
<th>Example name</th>
</tr>
</thead>
<tbody>
<tr>
<td>keep</td>
<td>If the device already has a name that was assigned in the user space, udev does not rename this device. For example, this is the case if the name was assigned during device creation or by a rename operation.</td>
<td></td>
</tr>
<tr>
<td>kernel</td>
<td>If the kernel indicates that a device name is predictable, udev does not rename this device.</td>
<td>lo</td>
</tr>
<tr>
<td>database</td>
<td>This policy assigns names based on mappings in the udev hardware database. For details, see the hwdb(7) man page.</td>
<td>idrac</td>
</tr>
<tr>
<td>onboard</td>
<td>Device names incorporate firmware or BIOS-provided index numbers for onboard devices.</td>
<td>eno1</td>
</tr>
<tr>
<td>slot</td>
<td>Device names incorporate firmware or BIOS-provided PCI Express (PCIe) hot-plug slot-index numbers.</td>
<td>ens1</td>
</tr>
<tr>
<td>Policy</td>
<td>Description</td>
<td>Example name</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>path</td>
<td>Device names incorporate the physical location of the connector of the hardware.</td>
<td>enp1s0</td>
</tr>
<tr>
<td>mac</td>
<td>Device names incorporate the MAC address. By default, Red Hat Enterprise Linux does not use this policy, but administrators can enable it.</td>
<td>enx525400d5e0f0b</td>
</tr>
</tbody>
</table>

Additional resources

- How the udev device manager renames network interfaces
- systemd.link(5) man page

1.3. NETWORK INTERFACE NAMING SCHEMES

The udev device manager uses certain stable interface attributes to generate consistent device names. For details about the naming schemes for different device types and platforms, see the systemd.net-naming-scheme(7) man page.

1.4. CUSTOMIZING THE PREFIX FOR ETHERNET INTERFACES DURING INSTALLATION

If you do not want to use the default device-naming policy for Ethernet interfaces, you can set a custom device prefix during the Red Hat Enterprise Linux (RHEL) installation.

**IMPORTANT**

Red Hat supports systems with customized Ethernet prefixes only if you set the prefix during the RHEL installation. Using the prefixdevname utility on already deployed systems is not supported.

If you set a device prefix during the installation, the udev service uses the `<prefix><index>` format for Ethernet interfaces after the installation. For example, if you set the prefix `net`, the service assigns the names `net0`, `net1`, and so on to the Ethernet interfaces.

The udev service appends the index to the custom prefix, and preserves the index values of known Ethernet interfaces. If you add an interface, udev assigns an index value that is one greater than the previously-assigned index value to the new interface.

**Prerequisites**

- The prefix consists of ASCII characters.
- The prefix is an alphanumeric string.
- The prefix is shorter than 16 characters.
• The prefix does not conflict with any other well-known network interface prefix, such as eth, eno, ens, and em.

Procedure

1. Boot the Red Hat Enterprise Linux installation media.

2. In the boot manager, follow these steps:
   a. Select the Install Red Hat Enterprise Linux <version> entry.
   b. Press Tab to edit the entry.
   c. Append net.ifnames.prefix=<prefix> to the kernel options.
   d. Press Enter to start the installation program.

3. Install Red Hat Enterprise Linux.

Verification

• To verify the interface names, display the network interfaces:

```
# ip link show
...
2: net0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
  mode DEFAULT group default qlen 1000
    link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
  ... 
```

Additional resources

• Performing a standard RHEL 9 installation

1.5. CONFIGURING USER-DEFINED NETWORK INTERFACE NAMES BY USING UDEV RULES

You can use udev rules to implement custom network interface names that reflect your organization’s requirements.

Procedure

1. Identify the network interface that you want to rename:

```
# ip link show
...
enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
  mode DEFAULT group default qlen 1000
    link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
  ... 
```

Record the MAC address of the interface.

2. Display the device type ID of the interface:

...
3. Create the `/etc/udev/rules.d/70-persistent-net.rules` file, and add a rule for each interface that you want to rename:

```bash
SUBSYSTEM=="net", ACTION=="add", ATTR{address}=="<MAC_address>", ATTR{type} =="<device_type_id>", NAME="<new_interface_name>"
```

**IMPORTANT**

Use only `70-persistent-net.rules` as a file name if you require consistent device names during the boot process. The `dracut` utility adds a file with this name to the `initrd` image if you regenerate the RAM disk image.

For example, use the following rule to rename the interface with MAC address `00:00:5e:00:53:1a` to `provider0`:

```bash
SUBSYSTEM=="net", ACTION=="add", ATTR{address}=="00:00:5e:00:53:1a", ATTR{type} =="1", NAME="provider0"
```

4. Optional: Regenerate the `initrd` RAM disk image:

```bash
# dracut -f
```

You require this step only if you need networking capabilities in the RAM disk. For example, this is the case if the root file system is stored on a network device, such as iSCSI.

5. Identify which NetworkManager connection profile uses the interface that you want to rename:

```bash
# nmcli -f device,name connection show
DEVICE    NAME
enp1s0    example_profile
...
```

6. Unset the `connection.interface-name` property in the connection profile:

```bash
# nmcli connection modify example_profile connection.interface-name ""
```

7. Temporarily, configure the connection profile to match both the new and the previous interface name:

```bash
# nmcli connection modify example_profile match.interface-name "provider0 enp1s0"
```

8. Reboot the system:

```bash
# reboot
```

9. Verify that the device with the MAC address that you specified in the link file has been renamed to `provider0`:

```bash
# ip link show
```
provider0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP mode DEFAULT group default qlen 1000
   link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
...

10. Configure the connection profile to match only the new interface name:

```bash
# nmcli connection modify example_profile match.interface-name "provider0"
```

You have now removed the old interface name from the connection profile.

11. Reactivate the connection profile:

```bash
# nmcli connection up example_profile
```

Additional resources

- udev(7) man page

1.6. CONFIGURING USER-DEFINED NETWORK INTERFACE NAMES BY USING SYSTEMD LINK FILES

You can use `systemd` link files to implement custom network interface names that reflect your organization’s requirements.

Prerequisites

- You must meet one of these conditions: NetworkManager does not manage this interface, or the corresponding connection profile uses the keyfile format.

Procedure

1. Identify the network interface that you want to rename:

```bash
# ip link show
...
enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP mode DEFAULT group default qlen 1000
   link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
...
```

Record the MAC address of the interface.

2. If it does not already exist, create the `/etc/systemd/network/` directory:

```bash
# mkdir -p /etc/systemd/network/
```

3. For each interface that you want to rename, create a `70-*.link` file in the `/etc/systemd/network/` directory with the following content:

```ini
[Match]
MACAddress=<MAC_address>
```
**IMPORTANT**

Use a file name with a 70- prefix to keep the file names consistent with the udev rules-based solution.

For example, create the `/etc/systemd/network/70-provider0.link` file with the following content to rename the interface with MAC address `00:00:5e:00:53:1a` to `provider0`:

```
[Match]
MACAddress=00:00:5e:00:53:1a

[Link]
Name=provider0
```

4. Optional: Regenerate the `initrd` RAM disk image:

```
# dracut -f
```

You require this step only if you need networking capabilities in the RAM disk. For example, this is the case if the root file system is stored on a network device, such as iSCSI.

5. Identify which NetworkManager connection profile uses the interface that you want to rename:

```
# nmcli -f device,name connection show
DEVICE   NAME
enp1s0   example_profile
... 
```

6. Unset the `connection.interface-name` property in the connection profile:

```
# nmcli connection modify example_profile connection.interface-name "" 
```

7. Temporarily, configure the connection profile to match both the new and the previous interface name:

```
# nmcli connection modify example_profile match.interface-name "provider0 enp1s0"
```

8. Reboot the system:

```
# reboot
```

9. Verify that the device with the MAC address that you specified in the link file has been renamed to `provider0`:

```
# ip link show
provider0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP mode DEFAULT group default qlen 1000
   link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
... 
```
10. Configure the connection profile to match only the new interface name:

```
# nmcli connection modify example_profile match.interface-name "provider0"
```

You have now removed the old interface name from the connection profile.

11. Reactivate the connection profile.

```
# nmcli connection up example_profile
```

Additional resources

- `systemd.link(5)` man page

1.7. ASSIGNING ALTERNATIVE NAMES TO A NETWORK INTERFACE BY USING SYSTEMD LINK FILES

With alternative interface naming, the kernel can assign additional names to network interfaces. You can use these alternative names in the same way as the normal interface names in commands that require a network interface name.

Prerequisites

- You must use ASCII characters for the alternative name.
- The alternative name must be shorter than 128 characters.

Procedure

1. Display the network interface names and their MAC addresses:

```
# ip link show
...
enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
 mode DEFAULT group default qlen 1000
   link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
...
```

Record the MAC address of the interface to which you want to assign an alternative name.

2. If it does not already exist, create the `/etc/systemd/network/` directory:

```
# mkdir -p /etc/systemd/network/
```

3. For each interface that must have an alternative name, create a `.link` file in the `/etc/systemd/network/` directory with the following content:

```
[Match]
MACAddress=<MAC_address>

[Link]
```
AlternativeName=<alternative_interface_name_1>
AlternativeName=<alternative_interface_name_2>
AlternativeName=<alternative_interface_name_n>

For example, create the `/etc/systemd/network/70-altname.link` file with the following content to assign `provider` as an alternative name to the interface with MAC address `00:00:5e:00:53:1a`:

```
[Match]
MACAddress=00:00:5e:00:53:1a

[Link]
AlternativeName=provider
```

4. Regenerate the `initrd` RAM disk image:

   ```
   # dracut -f
   ```

5. Reboot the system:

   ```
   # reboot
   ```

**Verification**

- Use the alternative interface name. For example, display the IP address settings of the device with the alternative name `provider`:

   ```
   # ip address show provider
   2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
     link/ether 00:00:5e:00:53:1a brd ff:ff:ff:ff:ff:ff
     altname provider
     ... 
   ```

**Additional resources**

- [What is AlternativeNamesPolicy in Interface naming scheme?](#)
CHAPTER 2. CONFIGURING AN ETHERNET CONNECTION

NetworkManager creates a connection profile for each Ethernet adapter that is installed in a host. By default, this profile uses DHCP for both IPv4 and IPv6 connections. Modify this automatically-created profile or add a new one in the following cases:

- The network requires custom settings, such as a static IP address configuration.
- You require multiple profiles because the host roams among different networks.

Red Hat Enterprise Linux provides administrators different options to configure Ethernet connections. For example:

- Use `nmcli` to configure connections on the command line.
- Use `nmtui` to configure connections in a text-based user interface.
- Use the GNOME Settings menu or `nm-connection-editor` application to configure connections in a graphical interface.
- Use `nmstatectl` to configure connections through the Nmstate API.
- Use RHEL system roles to automate the configuration of connections on one or multiple hosts.

**NOTE**

If you want to manually configure Ethernet connections on hosts running in the Microsoft Azure cloud, disable the `cloud-init` service or configure it to ignore the network settings retrieved from the cloud environment. Otherwise, `cloud-init` will override on the next reboot the network settings that you have manually configured.

2.1. CONFIGURING AN ETHERNET CONNECTION BY USING NMCLI

If you connect a host to the network over Ethernet, you can manage the connection’s settings on the command line by using the `nmcli` utility.

**Prerequisites**

- A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.

**Procedure**

1. List the NetworkManager connection profiles:

   ```
   # nmcli connection show
   NAME                UUID                                  TYPE      DEVICE
   Wired connection 1  a5eb6490-cc20-3668-81f8-0314a2713f75  ethernet  enp1s0
   ```

   By default, NetworkManager creates a profile for each NIC in the host. If you plan to connect this NIC only to a specific network, adapt the automatically-created profile. If you plan to connect this NIC to networks with different settings, create individual profiles for each network.

2. If you want to create an additional connection profile, enter:
# nmcli connection add con-name <connection-name> ifname <device-name> type ethernet

Skip this step to modify an existing profile.

3. Optional: Rename the connection profile:

   # nmcli connection modify "Wired connection 1" connection.id "Internal-LAN"

On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

4. Display the current settings of the connection profile:

   # nmcli connection show Internal-LAN
   ...
   connection.interface-name:     enp1s0
   connection.autoconnect:        yes
   ipv4.method:                   auto
   ipv6.method:                   auto
   ...

5. Configure the IPv4 settings:

   - To use DHCP, enter:

     # nmcli connection modify Internal-LAN ipv4.method auto

     Skip this step if ipv4.method is already set to auto (default).

   - To set a static IPv4 address, network mask, default gateway, DNS servers, and search domain, enter:

     # nmcli connection modify Internal-LAN ipv4.method manual ipv4.addresses 192.0.2.1/24 ipv4.gateway 192.0.2.254 ipv4.dns 192.0.2.200 ipv4.dns-search example.com

6. Configure the IPv6 settings:

   - To use stateless address autoconfiguration (SLAAC), enter:

     # nmcli connection modify Internal-LAN ipv6.method auto

     Skip this step if ipv6.method is already set to auto (default).

   - To set a static IPv6 address, network mask, default gateway, DNS servers, and search domain, enter:


7. To customize other settings in the profile, use the following command:
# nmcli connection modify <connection-name> <setting> <value>

Enclose values with spaces or semicolons in quotes.

8. Activate the profile:

# nmcli connection up Internal-LAN

Verification

1. Display the IP settings of the NIC:

```
# ip address show enp1s0
2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
group default qlen 1000
    link/ether 52:54:00:17:b8:b6 brd ff:ff:ff:ff:ff:ff
    inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute enp1s0
        valid_lft forever preferred_lft forever
    inet6 2001:db8:1::fffe/64 scope global noprefixroute
        valid_lft forever preferred_lft forever
```

2. Display the IPv4 default gateway:

```
# ip route show default
default via 192.0.2.254 dev enp1s0 proto static metric 102
```

3. Display the IPv6 default gateway:

```
# ip -6 route show default
default via 2001:db8:1::ffee dev enp1s0 proto static metric 102 pref medium
```

4. Display the DNS settings:

```
# cat /etc/resolv.conf
search example.com
nameserver 192.0.2.200
nameserver 2001:db8:1::ffbb
```

If multiple connection profiles are active at the same time, the order of `nameserver` entries depend on the DNS priority values in these profile and the connection types.

5. Use the `ping` utility to verify that this host can send packets to other hosts:

```
# ping <host-name-or-IP-address>
```

Troubleshooting

- Verify that the network cable is plugged-in to the host and a switch.
- Check whether the link failure exists only on this host or also on other hosts connected to the same switch.
• Verify that the network cable and the network interface are working as expected. Perform hardware diagnosis steps and replace defect cables and network interface cards.

• If the configuration on the disk does not match the configuration on the device, starting or restarting NetworkManager creates an in-memory connection that reflects the configuration of the device. For further details and how to avoid this problem, see the NetworkManager duplicates a connection after restart of NetworkManager service solution.

Additional resources

• nm-settings(5) man page
• Configuring NetworkManager to avoid using a specific profile to provide a default gateway
• Configuring the order of DNS servers

2.2. CONFIGURING AN ETHERNET CONNECTION BY USING THE NMCLI INTERACTIVE EDITOR

If you connect a host to the network over Ethernet, you can manage the connection’s settings on the command line by using the nmcli utility.

Prerequisites

• A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.

Procedure

1. List the NetworkManager connection profiles:

   # nmcli connection show

   NAME                UUID                                  TYPE      DEVICE
   Wired connection 1  a5eb6490-cc20-3668-81f8-0314a27f3f75  ethernet  enp1s0

   By default, NetworkManager creates a profile for each NIC in the host. If you plan to connect this NIC only to a specific network, adapt the automatically-created profile. If you plan to connect this NIC to networks with different settings, create individual profiles for each network.

2. Start nmcli in interactive mode:

   • To create an additional connection profile, enter:

   # nmcli connection edit type ethernet con-name "<connection-name>"

   • To modify an existing connection profile, enter:

   # nmcli connection edit con-name "<connection-name>"

3. Optional: Rename the connection profile:

   nmcli> set connection.id Internal-LAN
On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

Do not use quotes to set an ID that contains spaces to avoid that nmcli makes the quotes part of the name. For example, to set Example Connection as ID, enter `set connection.id Example Connection`.

4. Display the current settings of the connection profile:

```
        nmcli> print
... connection.interface-name: enp1s0
      connection.autoconnect: yes
    ipv4.method: auto
    ipv6.method: auto
...
```

5. If you create a new connection profile, set the network interface:

```
        nmcli> set connection.interface-name enp1s0
```

6. Configure the IPv4 settings:

- To use DHCP, enter:

```
        nmcli> set ipv4.method auto
```

Skip this step if `ipv4.method` is already set to `auto` (default).

- To set a static IPv4 address, network mask, default gateway, DNS servers, and search domain, enter:

```
        nmcli> ipv4.addresses 192.0.2.1/24
Do you also want to set 'ipv4.method' to 'manual'? [yes]: yes
        nmcli> ipv4.gateway 192.0.2.254
        nmcli> ipv4.dns 192.0.2.200
        nmcli> ipv4.dns-search example.com
```

7. Configure the IPv6 settings:

- To use stateless address autoconfiguration (SLAAC), enter:

```
        nmcli> set ipv6.method auto
```

Skip this step if `ipv6.method` is already set to `auto` (default).

- To set a static IPv6 address, network mask, default gateway, DNS servers, and search domain, enter:

```
        nmcli> ipv6.addresses 2001:db8:1::fffe/64
Do you also want to set 'ipv6.method' to 'manual'? [yes]: yes
        nmcli> ipv6.gateway 2001:db8:1::fffe
        nmcli> ipv6.dns 2001:db8:1::ffbb
        nmcli> ipv6.dns-search example.com
```
8. Save and activate the connection:

```
nmcli> save persistent
```

9. Leave the interactive mode:

```
nmcli> quit
```

**Verification**

1. Display the IP settings of the NIC:

```
# ip address show enp1s0
2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
group default qlen 1000
  link/ether 52:54:00:17:b8:b6 brd ff:ff:ff:ff:ff:ff
  inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute enp1s0
    valid_lft forever preferred_lft forever
  inet6 2001:db8:1::fffe/64 scope global noprefixroute
    valid_lft forever preferred_lft forever
```

2. Display the IPv4 default gateway:

```
# ip route show default
default via 192.0.2.254 dev enp1s0 proto static metric 102
```

3. Display the IPv6 default gateway:

```
# ip -6 route show default
default via 2001:db8:1::ffee dev enp1s0 proto static metric 102 pref medium
```

4. Display the DNS settings:

```
# cat /etc/resolv.conf
search example.com
nameserver 192.0.2.200
nameserver 2001:db8:1::ffbb
```

If multiple connection profiles are active at the same time, the order of `nameserver` entries depend on the DNS priority values in these profile and the connection types.

5. Use the `ping` utility to verify that this host can send packets to other hosts:

```
# ping <host-name-or-IP-address>
```

**Troubleshooting**

- Verify that the network cable is plugged-in to the host and a switch.
- Check whether the link failure exists only on this host or also on other hosts connected to the same switch.
Verify that the network cable and the network interface are working as expected. Perform hardware diagnosis steps and replace defect cables and network interface cards.

If the configuration on the disk does not match the configuration on the device, starting or restarting NetworkManager creates an in-memory connection that reflects the configuration of the device. For further details and how to avoid this problem, see the NetworkManager duplicates a connection after restart of NetworkManager service solution

Additional resources

- `nm-settings(5)` man page
- `nmcli(1)` man page
- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- Configuring the order of DNS servers

### 2.3. CONFIGURING AN ETHERNET CONNECTION BY USING NMTUI

If you connect a host to the network over Ethernet, you can manage the connection’s settings in a text-based user interface by using the `nmtui` application. Use `nmtui` to create new profiles and to update existing ones on a host without a graphical interface.

**NOTE**

In **nmtui**:

- Navigate by using the cursor keys.
- Press a button by selecting it and hitting **Enter**.
- Select and deselect checkboxes by using **Space**.

**Prerequisites**

- A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.

**Procedure**

1. If you do not know the network device name you want to use in the connection, display the available devices:

   ```bash
   # nmcli device status
   DEVICE  TYPE    STATE       CONNECTION
   enp1s0  ethernet  unavailable --
   ...
   ```

2. Start **nmtui**:

   ```bash
   # nmtui
   ```

3. Select **Edit a connection** and press **Enter**.
4. Choose whether to add a new connection profile or to modify an existing one:
   - To create a new profile:
     i. Press the Add button.
     ii. Select Ethernet from the list of network types, and press Enter.
   - To modify an existing profile, select the profile from the list, and press Enter.

5. Optional: Update the name of the connection profile.
   On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

6. If you create a new connection profile, enter the network device name into the Device field.

7. Depending on your environment, configure the IP address settings in the IPv4 configuration and IPv6 configuration areas accordingly. For this, press the button next to these areas, and select:
   - **Disabled**, if this connection does not require an IP address.
   - **Automatic**, if a DHCP server dynamically assigns an IP address to this NIC.
   - **Manual**, if the network requires static IP address settings. In this case, you must fill further fields:
     i. Press the Show button next to the protocol you want to configure to display additional fields.
     ii. Press the Add button next to Addresses, and enter the IP address and the subnet mask in Classless Inter-Domain Routing (CIDR) format.
        If you do not specify a subnet mask, NetworkManager sets a /32 subnet mask for IPv4 addresses and /64 for IPv6 addresses.
     iii. Enter the address of the default gateway.
     iv. Press the Add button next to DNS servers, and enter the DNS server address.
     v. Press the Add button next to Search domains, and enter the DNS search domain.
8. Press the **OK** button to create and automatically activate the new connection.

9. Press the **Back** button to return to the main menu.

10. Select **Quit**, and press **Enter** to close the **nmtui** application.

**Verification**

1. Display the IP settings of the NIC:

```
# ip address show enp1s0
```

```
2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
   link/ether 52:54:00:17:b8:b6 brd ff:ff:ff:ff:ff:ff
```

Figure 2.1. Example of an Ethernet connection with static IP address settings
2. Display the IPv4 default gateway:

```
# ip route show default
default via 192.0.2.254 dev enp1s0 proto static metric 102
```

3. Display the IPv6 default gateway:

```
# ip -6 route show default
default via 2001:db8:1::ffee dev enp1s0 proto static metric 102 pref medium
```

4. Display the DNS settings:

```
# cat /etc/resolv.conf
search example.com
nameserver 192.0.2.200
nameserver 2001:db8:1::ffbb
```

If multiple connection profiles are active at the same time, the order of nameserver entries depend on the DNS priority values in these profile and the connection types.

5. Use the `ping` utility to verify that this host can send packets to other hosts:

```
# ping <host-name-or-IP-address>
```

Troubleshooting

- Verify that the network cable is plugged-in to the host and a switch.

- Check whether the link failure exists only on this host or also on other hosts connected to the same switch.

- Verify that the network cable and the network interface are working as expected. Perform hardware diagnosis steps and replace defect cables and network interface cards.

- If the configuration on the disk does not match the configuration on the device, starting or restarting NetworkManager creates an in-memory connection that reflects the configuration of the device. For further details and how to avoid this problem, see the NetworkManager duplicates a connection after restart of NetworkManager service solution.

Additional resources

- Configuring NetworkManager to avoid using a specific profile to provide a default gateway

- Configuring the order of DNS servers

2.4. CONFIGURING AN ETHERNET CONNECTION BY USING CONTROL-CENTER
If you connect a host to the network over Ethernet, you can manage the connection’s settings with a graphical interface by using the GNOME Settings menu.

Note that control-center does not support as many configuration options as the nm-connection-editor application or the nmcli utility.

Prerequisites

- A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.
- GNOME is installed.

Procedure

1. Press the Super key, enter Settings, and press Enter.
2. Select Network in the navigation on the left.
3. Choose whether to add a new connection profile or to modify an existing one:
   - To create a new profile, click the + button next to the Ethernet entry.
   - To modify an existing profile, click the gear icon next to the profile entry.
4. Optional: On the Identity tab, update the name of the connection profile. On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.
5. Depending on your environment, configure the IP address settings on the IPv4 and IPv6 tabs accordingly:
   - To use DHCP or IPv6 stateless address autoconfiguration (SLAAC), select Automatic (DHCP) as method (default).
   - To set a static IP address, network mask, default gateway, DNS servers, and search domain, select Manual as method, and fill the fields on the tabs.
6. Depending on whether you add or modify a connection profile, click the Add or Apply button to save the connection.
   The GNOME control-center automatically activates the connection.

Verification

1. Display the IP settings of the NIC:
2. Display the IPv4 default gateway:

```bash
# ip route show default
default via 192.0.2.254 dev enp1s0 proto static metric 102
```

3. Display the IPv6 default gateway:

```bash
# ip -6 route show default
default via 2001:db8:1::ffee dev enp1s0 proto static metric 102 pref medium
```

4. Display the DNS settings:

```bash
# cat /etc/resolv.conf
search example.com
nameserver 192.0.2.200
nameserver 2001:db8:1::ffbb
```

If multiple connection profiles are active at the same time, the order of nameserver entries depend on the DNS priority values in these profile and the connection types.

5. Use the ping utility to verify that this host can send packets to other hosts:

```bash
# ping <host-name-or-IP-address>
```

Troubleshooting steps

- Verify that the network cable is plugged-in to the host and a switch.
- Check whether the link failure exists only on this host or also on other hosts connected to the same switch.
- Verify that the network cable and the network interface are working as expected. Perform hardware diagnosis steps and replace defect cables and network interface cards.
- If the configuration on the disk does not match the configuration on the device, starting or restarting NetworkManager creates an in-memory connection that reflects the configuration of the device. For further details and how to avoid this problem, see the NetworkManager duplicates a connection after restart of NetworkManager service solution.

Additional Resources

- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- Configuring the order of DNS servers
2.5. CONFIGURING AN ETHERNET CONNECTION BY USING NM-CONNECTION-EDITOR

If you connect a host to the network over Ethernet, you can manage the connection’s settings with a graphical interface by using the `nm-connection-editor` application.

Prerequisites

- A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.
- GNOME is installed.

Procedure

1. Open a terminal, and enter:

   
   $ nm-connection-editor

2. Choose whether to add a new connection profile or to modify an existing one:

   - To create a new profile:
     i. Click the + button
     ii. Select Ethernet as connection type, and click Create.
   - To modify an existing profile, double-click the profile entry.

3. Optional: Update the name of the profile in the Connection Name field.
   On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

4. If you create a new profile, select the device on the Ethernet tab:

5. Depending on your environment, configure the IP address settings on the IPv4 Settings and IPv6 Settings tabs accordingly:

   - To use DHCP or IPv6 stateless address autoconfiguration (SLAAC), select Automatic (DHCP) as method (default).
   - To set a static IP address, network mask, default gateway, DNS servers, and search domain, select Manual as method, and fill the fields on the tabs:
6. Click **Save**.

7. Close **nm-connection-editor**.

**Verification**

1. Display the IP settings of the NIC:

   ```
   # ip address show enp1s0
   2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
   group default qlen 1000
   link/ether 52:54:00:17:b8:b6 brd ff:ff:ff:ff:ff:ff
   inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute enp1s0
   valid_lft forever preferred_lft forever
   inet6 2001:db8:1::fffe/64 scope global noprefixroute
   valid_lft forever preferred_lft forever
   ```

2. Display the IPv4 default gateway:

   ```
   # ip route show default
   default via 192.0.2.254 dev enp1s0 proto static metric 102
   ```

3. Display the IPv6 default gateway:

   ```
   # ip -6 route show default
   default via 2001:db8:1::ffe6/64 dev enp1s0 proto static metric 102 pref medium
   ```

4. Display the DNS settings:

   ```
   # cat /etc/resolv.conf
   search example.com
   nameserver 192.0.2.200
   nameserver 2001:db8:1::ffbb
   ```

   If multiple connection profiles are active at the same time, the order of **nameserver** entries depend on the DNS priority values in these profile and the connection types.

5. Use the **ping** utility to verify that this host can send packets to other hosts:

   ```
   # ping <host-name-or-IP-address>
   ```

**Troubleshooting steps**

- Verify that the network cable is plugged-in to the host and a switch.
• Check whether the link failure exists only on this host or also on other hosts connected to the same switch.

• Verify that the network cable and the network interface are working as expected. Perform hardware diagnosis steps and replace defect cables and network interface cards.

• If the configuration on the disk does not match the configuration on the device, starting or restarting NetworkManager creates an in-memory connection that reflects the configuration of the device. For further details and how to avoid this problem, see the NetworkManager duplicates a connection after restart of NetworkManager service solution.

Additional Resources

• Configuring NetworkManager to avoid using a specific profile to provide a default gateway

• Configuring the order of DNS servers

2.6. CONFIGURING AN ETHERNET CONNECTION WITH A STATIC IP ADDRESS BY USING NMSTATECTL

Use the nmstatectl utility to configure an Ethernet connection through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, nmstatectl automatically rolls back the changes to avoid leaving the system in an incorrect state.

Prerequisites

• A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.

• The nmstate package is installed.

Procedure

1. Create a YAML file, for example ~/create-ethernet-profile.yml, with the following content:

```yaml
---
interfaces:
- name: enp1s0
type: ethernet
state: up
ipv4:
  enabled: true
  address:
    - ip: 192.0.2.1
      prefix-length: 24
dhcp: false
ipv6:
  enabled: true
  address:
    - ip: 2001:db8:1::1
      prefix-length: 64
      autoconf: false
dhcp: false
routes:
  config:
```

Red Hat Enterprise Linux 9 Configuring and managing networking

34
These settings define an Ethernet connection profile for the `enp1s0` device with the following settings:

- A static IPv4 address - 192.0.2.1 with the /24 subnet mask
- A static IPv6 address - 2001:db8:1::1 with the /64 subnet mask
- An IPv4 default gateway - 192.0.2.254
- An IPv6 default gateway - 2001:db8:1::ffe
- An IPv4 DNS server - 192.0.2.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com

**NOTE**

You can define `identifier: mac-address` and `mac-address: <ACTUAL-NIC-MACADDRESS>` properties directly to the `interfaces` property to identify the network interface card by its MAC address instead of its name (such as `enp1s0`).

For example:

```yaml
---
interfaces:
  - name: <profile-name>
    type: ethernet
    identifier: mac-address
    mac-address: <ACTUAL-NIC-MACADDRESS>
    state: up
    ipv4:
      enabled: true
      auto-dns: true
      auto-gateway: true
      auto-routes: true
      dhcp: true
---
```

2. Apply the settings to the system:
# nmstatectl apply ~/create-ethernet-profile.yml

**Verification**

1. Display the current state in YAML format:
   
   ```
   # nmstatectl show enp1s0
   ```

2. Display the IP settings of the NIC:
   
   ```
   # ip address show enp1s0
   2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
group default qlen 1000
   link/ether 52:54:00:17:b8:b6 brd ff:ff:ff:ff:ff:ff
   inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute enp1s0
       valid_lft forever preferred_lft forever
   inet6 2001:db8:1::fffe/64 scope global noprefixroute
       valid_lft forever preferred_lft forever
   ```

3. Display the IPv4 default gateway:
   
   ```
   # ip route show default
   default via 192.0.2.254 dev enp1s0 proto static metric 102
   ```

4. Display the IPv6 default gateway:
   
   ```
   # ip -6 route show default
   default via 2001:db8:1::ffee dev enp1s0 proto static metric 102 pref medium
   ```

5. Display the DNS settings:
   
   ```
   # cat /etc/resolv.conf
   search example.com
   nameserver 192.0.2.200
   nameserver 2001:db8:1::ffbb
   ```

   If multiple connection profiles are active at the same time, the order of `nameserver` entries depend on the DNS priority values in these profile and the connection types.

6. Use the `ping` utility to verify that this host can send packets to other hosts:
   
   ```
   # ping <host-name-or-IP-address>
   ```

**Additional resources**

- `nmstatectl(8)` man page
- `/usr/share/doc/nmstate/examples/` directory
2.7. CONFIGURING AN ETHERNET CONNECTION WITH A STATIC IP ADDRESS BY USING THE NETWORK RHEL SYSTEM ROLE WITH AN INTERFACE NAME

You can remotely configure an Ethernet connection by using the `network` RHEL system role.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
- A physical or virtual Ethernet device exists in the server’s configuration.
- The managed nodes use NetworkManager to configure the network.

Procedure

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure an Ethernet connection with static IP
      ansible.builtin.include_role:
        name: rhel-system-roles.network
        vars:
          network_connections:
            - name: enp1s0
              interface_name: enp1s0
              type: ethernet
              autoconnect: yes
              ip:
                address:
                  - 192.0.2.1/24
                  - 2001:db8:1::1/64
              gateway4: 192.0.2.254
              gateway6: 2001:db8:1::ffe
              dns:
                - 192.0.2.200
                - 2001:db8:1::ffbb
              dns_search:
                - example.com
              state: up
```

These settings define an Ethernet connection profile for the `enp1s0` device with the following settings:

- A static IPv4 address - `192.0.2.1` with a `/24` subnet mask
- A static IPv6 address - `2001:db8:1::1` with a `/64` subnet mask
• An IPv4 default gateway - 192.0.2.254
• An IPv6 default gateway - 2001:db8:1::fffe
• An IPv4 DNS server - 192.0.2.200
• An IPv6 DNS server - 2001:db8:1::ffbb
• A DNS search domain - example.com

2. Validate the playbook syntax:

   $ ansible-playbook --syntax-check ~/playbook.yml

   Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

   $ ansible-playbook ~/playbook.yml

Additional resources

• /usr/share/ansible/roles/rhel-system-roles.network/README.md file
• /usr/share/doc/rhel-system-roles/network/ directory

2.8. CONFIGURING AN ETHERNET CONNECTION WITH A STATIC IP ADDRESS BY USING THE NETWORK RHEL SYSTEM ROLE WITH A DEVICE PATH

You can remotely configure an Ethernet connection using the network RHEL system role.

You can identify the device path with the following command:

   # udevadm info /sys/class/net/<device_name> | grep ID_PATH=

Prerequisites

• You have prepared the control node and the managed nodes
• You are logged in to the control node as a user who can run playbooks on the managed nodes.
• The account you use to connect to the managed nodes has sudo permissions on them.
• A physical or virtual Ethernet device exists in the server’s configuration.
• The managed nodes use NetworkManager to configure the network.

Procedure

1. Create a playbook file, for example ~/playbook.yml, with the following content:
These settings define an Ethernet connection profile with the following settings:

- A static IPv4 address - 192.0.2.1 with a /24 subnet mask
- A static IPv6 address - 2001:db8:1::1 with a /64 subnet mask
- An IPv4 default gateway - 192.0.2.254
- An IPv6 default gateway - 2001:db8:1::ffe
- An IPv4 DNS server - 192.0.2.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com

The *match* parameter in this example defines that Ansible applies the play to devices that match PCI ID 0000:00:0[1-3].0, but not 0000:00:02.0. For further details about special modifiers and wild cards you can use, see the *match* parameter description in the /usr/share/ansible/roles/rhel-system-roles.network/README.md file.

2. Validate the playbook syntax:

```
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.
3. Run the playbook:

```
$ ansible-playbook ~/playbook.yml
```

Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory

### 2.9. CONFIGURING AN ETHERNET CONNECTION WITH A DYNAMIC IP ADDRESS BY USING NMSTATECTL

Use the `nmstatectl` utility to configure an Ethernet connection through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.

**Prerequisites**

- A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server’s configuration.
- A DHCP server is available in the network.
- The `nmstate` package is installed.

**Procedure**

1. Create a YAML file, for example `~/create-ethernet-profile.yml`, with the following content:

```yaml
---
interfaces:
- name: enp1s0
type: ethernet
state: up
ipv4:
  enabled: true
  auto-dns: true
  auto-gateway: true
  auto-routes: true
  dhcp: true
ipv6:
  enabled: true
  auto-dns: true
  auto-gateway: true
  auto-routes: true
  autoconf: true
  dhcp: true
```

These settings define an Ethernet connection profile for the `enp1s0` device. The connection retrieves IPv4 addresses, IPv6 addresses, default gateway, routes, DNS servers, and search domains from a DHCP server and IPv6 stateless address autoconfiguration (SLAAC).
NOTE

You can define `identifier: mac-address` and `mac-address: <ACTUAL-NIC-MACADDRESS>` properties directly to the `interfaces` property to identify the network interface card by its MAC address instead of its name (such as `enp1s0`).

For example:

```yaml
---
interfaces:
  - name: <profile-name>
    type: ethernet
    identifier: mac-address
    mac-address: <ACTUAL-NIC-MACADDRESS>
    state: up
    ipv4:
      enabled: true
      auto-dns: true
      auto-gateway: true
      auto-routes: true
      dhcp: true
    ... 
```

2. Apply the settings to the system:

```
# nmstatectl apply ~/create-ethernet-profile.yml
```

Verification

1. Display the current state in YAML format:

```
# nmstatectl show enp1s0
```

2. Display the IP settings of the NIC:

```
# ip address show enp1s0
2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
    group default qlen 1000
    link/ether 52:54:00:17:b8:b6 brd ff:ff:ff:ff:ff:ff
    inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute enp1s0
      valid_lft forever preferred_lft forever
    inet6 2001:db8:1::fffe/64 scope global noprefixroute
      valid_lft forever preferred_lft forever
```

3. Display the IPv4 default gateway:

```
# ip route show default
default via 192.0.2.254 dev enp1s0 proto static metric 102
```

4. Display the IPv6 default gateway:

```
# ip -6 route show default
default via 2001:db8:1::fffe dev enp1s0 proto static metric 102 pref medium
```
5. Display the DNS settings:

```bash
# cat /etc/resolv.conf
search example.com
nameserver 192.0.2.200
nameserver 2001:db8:1::ffbb
```

If multiple connection profiles are active at the same time, the order of `nameserver` entries depend on the DNS priority values in these profile and the connection types.

6. Use the `ping` utility to verify that this host can send packets to other hosts:

```bash
# ping <host-name-or-IP-address>
```

### Additional resources

- `nmstatectl(8)` man page
- `/usr/share/doc/nmstate/examples/` directory

## 2.10. Configuring an Ethernet Connection with a Dynamic IP Address by Using the Network RHEL System Role with an Interface Name

You can remotely configure an Ethernet connection using the `network` RHEL system role. For connections with dynamic IP address settings, NetworkManager requests the IP settings for the connection from a DHCP server.

### Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
- A physical or virtual Ethernet device exists in the server’s configuration.
- A DHCP server is available in the network
- The managed nodes use NetworkManager to configure the network.

### Procedure

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure an Ethernet connection with dynamic IP
      ansible.builtin.include_role:
        name: rhel-system-roles.network
      vars:
```
These settings define an Ethernet connection profile for the `enp1s0` device. The connection retrieves IPv4 addresses, IPv6 addresses, default gateway, routes, DNS servers, and search domains from a DHCP server and IPv6 stateless address autoconfiguration (SLAAC).

2. Validate the playbook syntax:

```bash
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```bash
$ ansible-playbook ~/playbook.yml
```

Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory

### 2.11. Configuring an Ethernet Connection with a Dynamic IP Address by Using the Network RHEL System Role with a Device Path

You can remotely configure an Ethernet connection using the `network` RHEL system role. For connections with dynamic IP address settings, NetworkManager requests the IP settings for the connection from a DHCP server.

You can identify the device path with the following command:

```bash
# udevadm info /sys/class/net/<device_name> | grep ID_PATH=
```

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
- A physical or virtual Ethernet device exists in the server’s configuration.
A DHCP server is available in the network.

The managed hosts use NetworkManager to configure the network.

**Procedure**

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

   ```yaml
   ---
   - name: Configure the network
     hosts: managed-node-01.example.com
     tasks:
       - name: Configure an Ethernet connection with dynamic IP
         ansible.builtin.include_role:
           name: rhel-system-roles.network
         vars:
           network_connections:
             - name: example
               match:
                 path:
                   - pci-0000:00:0[1-3].0
                   - &lpci-0000:00:02.0
               type: ethernet
               autoconnect: yes
               ip:
                 dhcp4: yes
                 auto6: yes
               state: up
   ```

   These settings define an Ethernet connection profile. The connection retrieves IPv4 addresses, IPv6 addresses, default gateway, routes, DNS servers, and search domains from a DHCP server and IPv6 stateless address autoconfiguration (SLAAC).

   The `match` parameter defines that Ansible applies the play to devices that match PCI ID `0000:00:0[1-3].0`, but not `0000:00:02.0`.

2. Validate the playbook syntax:

   ```bash
   $ ansible-playbook --syntax-check ~/playbook.yml
   ```

   Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

   ```bash
   $ ansible-playbook ~/playbook.yml
   ```

**Additional resources**

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory
2.12. CONFIGURING MULTIPLE ETHERNET INTERFACES BY USING A SINGLE CONNECTION PROFILE BY INTERFACE NAME

In most cases, one connection profile contains the settings of one network device. However, NetworkManager also supports wildcards when you set the interface name in connection profiles. If a host roams between Ethernet networks with dynamic IP address assignment, you can use this feature to create a single connection profile that you can use for multiple Ethernet interfaces.

Prerequisites

- Multiple physical or virtual Ethernet devices exist in the server’s configuration.
- A DHCP server is available in the network.
- No connection profile exists on the host.

Procedure

1. Add a connection profile that applies to all interface names starting with `enp`:

```
# nmcli connection add con-name Example connection.multi-connect multiple
    match.interface-name enp* type ethernet
```

Verification

1. Display all settings of the single connection profile:

```
# nmcli connection show Example
connection.id:                      Example
  connection.multi-connect:         3 (multiple)
  match.interface-name:            enp*
```

3 indicates the number of interfaces active on the connection profile at the same time, and not the number of network interfaces in the connection profile. The connection profile uses all devices that match the pattern in the `match.interface-name` parameter and, therefore, the connection profiles have the same Universally Unique Identifier (UUID).

2. Display the status of the connections:

```
# nmcli connection show
NAME                     UUID                    TYPE     DEVICE
  Example 6f22402e-c0cc-49cf-b702-eaf0cd5ea7d1  ethernet  enp7s0
  Example 6f22402e-c0cc-49cf-b702-eaf0cd5ea7d1  ethernet  enp8s0
  Example 6f22402e-c0cc-49cf-b702-eaf0cd5ea7d1  ethernet  enp9s0
```

Additional resources

- `nmcli(1)` man page
- `nm-settings(5)` man page
2.13. CONFIGURING A SINGLE CONNECTION PROFILE FOR MULTIPLE ETHERNET INTERFACES USING PCI IDs

The PCI ID is a unique identifier of the devices connected to the system. The connection profile adds multiple devices by matching interfaces based on a list of PCI IDs. You can use this procedure to connect multiple device PCI IDs to the single connection profile.

Prerequisites

- Multiple physical or virtual Ethernet devices exist in the server’s configuration.
- A DHCP server is available in the network.
- No connection profile exists on the host.

Procedure

1. Identify the device path. For example, to display the device paths of all interfaces starting with `enp`, enter:

   ```bash
   # udevadm info /sys/class/net/enp* | grep ID_PATH=
   ...
   E: ID_PATH=pci-0000:07:00.0
   E: ID_PATH=pci-0000:08:00.0
   ```

2. Add a connection profile that applies to all PCI IDs matching the `0000:00:0[7-8].0` expression:

   ```bash
   # nmcli connection add type ethernet connection.multi-connect multiple match.path "pci-0000:07:00.0 pci-0000:08:00.0" con-name Example
   ```

Verification

1. Display the status of the connection:

   ```bash
   # nmcli connection show
   NAME   UUID     TYPE        DEVICE
   Example      9cee0958-512f-4203-9d3d-b57af1d88466  ethernet  enp7s0
   Example      9cee0958-512f-4203-9d3d-b57af1d88466  ethernet  enp8s0
   ...
   ```

2. To display all settings of the connection profile:

   ```bash
   # nmcli connection show Example
   connection.id:               Example
   ... connection.multi-connect: 3 (multiple)
   match.path:                  pci-0000:07:00.0,pci-0000:08:00.0
   ...                            
   ```

   This connection profile uses all devices with a PCI ID which match the pattern in the `match.path` parameter and, therefore, the connection profiles have the same Universally Unique Identifier (UUID).

Additional resources
Additional resources

- `nmcli(1)` man page
- `nm-settings(5)` man page
CHAPTER 3. CONFIGURING NETWORK BONDING

A network bond is a method to combine or aggregate physical and virtual network interfaces to provide a logical interface with higher throughput or redundancy. In a bond, the kernel handles all operations exclusively. You can create bonds on different types of devices, such as Ethernet devices or VLANs.

Red Hat Enterprise Linux provides administrators different options to configure team devices. For example:

- Use `nmcli` to configure bond connections using the command line.
- Use the RHEL web console to configure bond connections using a web browser.
- Use `nmtui` to configure bond connections in a text-based user interface.
- Use the `nm-connection-editor` application to configure bond connections in a graphical interface.
- Use `nmstatectl` to configure bond connections through the Nmstate API.
- Use RHEL system roles to automate the bond configuration on one or multiple hosts.

3.1. UNDERSTANDING THE DEFAULT BEHAVIOR OF CONTROLLER AND PORT INTERFACES

Consider the following default behavior when managing or troubleshooting team or bond port interfaces using the NetworkManager service:

- Starting the controller interface does not automatically start the port interfaces.
- Starting a port interface always starts the controller interface.
- Stopping the controller interface also stops the port interface.
- A controller without ports can start static IP connections.
- A controller without ports waits for ports when starting DHCP connections.
- A controller with a DHCP connection waiting for ports completes when you add a port with a carrier.
- A controller with a DHCP connection waiting for ports continues waiting when you add a port without carrier.

3.2. UPSTREAM SWITCH CONFIGURATION DEPENDING ON THE BONDING MODES

Depending on the bonding mode you want to use, you must configure the ports on the switch:

<table>
<thead>
<tr>
<th>Bonding mode</th>
<th>Configuration on the switch</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - balance-rr</td>
<td>Requires static EtherChannel enabled, not Link Aggregation Control Protocol (LACP)-negotiated.</td>
</tr>
</tbody>
</table>
Bonding mode | Configuration on the switch
---|---
1 - active-backup | No configuration required on the switch.
2 - balance-xor | Requires static EtherChannel enabled, not LACP-negotiated.
3 - broadcast | Requires static EtherChannel enabled, not LACP-negotiated.
4 - 802.3ad | Requires LACP-negotiated EtherChannel enabled.
5 - balance-tlb | No configuration required on the switch.
6 - balance-alb | No configuration required on the switch.

For details how to configure your switch, see the documentation of the switch.

**IMPORTANT**

Certain network bonding features, such as the fail-over mechanism, do not support direct cable connections without a network switch. For further details, see the Is bonding supported with direct connection using crossover cables? KCS solution.

### 3.3. CONFIGURING A NETWORK BOND BY USING NMCLI

To configure a network bond on the command line, use the `nmcli` utility.

**Prerequisites**

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bond, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bridge, or VLAN devices as ports of the bond, you can either create these devices while you create the bond or you can create them in advance as described in:
  - Configuring a network team by using nmcli
  - Configuring a network bridge by using nmcli
  - Configuring VLAN tagging by using nmcli

**Procedure**

1. Create a bond interface:

   ```bash
   # nmcli connection add type bond con-name bond0 ifname bond0 bond.options "mode=active-backup"
   ```

   This command creates a bond named `bond0` that uses the `active-backup` mode.
To additionally set a Media Independent Interface (MII) monitoring interval, add the `miimon=interval` option to the `bond.options` property, for example:

```bash
# nmcli connection add type bond con-name bond0 ifname bond0 bond.options "mode=active-backup,miimon=1000"
```

2. Display the network interfaces, and note names of interfaces you plan to add to the bond:

```bash
# nmcli device status

<table>
<thead>
<tr>
<th>DEVICE</th>
<th>TYPE</th>
<th>STATE</th>
<th>CONNECTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>enp7s0</td>
<td>ethernet</td>
<td>disconnected</td>
<td>--</td>
</tr>
<tr>
<td>enp8s0</td>
<td>ethernet</td>
<td>disconnected</td>
<td>--</td>
</tr>
<tr>
<td>bridge0</td>
<td>bridge</td>
<td>connected</td>
<td>bridge0</td>
</tr>
<tr>
<td>bridge1</td>
<td>bridge</td>
<td>connected</td>
<td>bridge1</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

In this example:

- `enp7s0` and `enp8s0` are not configured. To use these devices as ports, add connection profiles in the next step.
- `bridge0` and `bridge1` have existing connection profiles. To use these devices as ports, modify their profiles in the next step.

3. Assign interfaces to the bond:

   a. If the interfaces you want to assign to the bond are not configured, create new connection profiles for them:

   ```bash
   # nmcli connection add type ethernet slave-type bond con-name bond0-port1 ifname enp7s0 master bond0
   # nmcli connection add type ethernet slave-type bond con-name bond0-port2 ifname enp8s0 master bond0
   
   These commands create profiles for `enp7s0` and `enp8s0`, and add them to the `bond0` connection.
   
   b. To assign an existing connection profile to the bond:

      i. Set the `master` parameter of these connections to `bond0`:

      ```bash
      # nmcli connection modify bridge0 master bond0
      # nmcli connection modify bridge1 master bond0
      
      These commands assign the existing connection profiles named `bridge0` and `bridge1` to the `bond0` connection.
      
      ii. Reactivate the connections:

      ```bash
      # nmcli connection up bridge0
      # nmcli connection up bridge1
      
      4. Configure the IPv4 settings:

      - To use this bond device as a port of other devices, enter:
# nmcli connection modify bond0 ipv4.method disabled

- To use DHCP, no action is required.
- To set a static IPv4 address, network mask, default gateway, and DNS server to the bond0 connection, enter:

  ```
  # nmcli connection modify bond0 ipv4.addresses '192.0.2.1/24' ipv4.gateway '192.0.2.254' ipv4.dns '192.0.2.253' ipv4.dns-search 'example.com' ipv4.method manual
  ```

5. Configure the IPv6 settings:

- To use this bond device as a port of other devices, enter:

  ```
  # nmcli connection modify bond0 ipv6.method disabled
  ```

- To use stateless address autoconfiguration (SLAAC), no action is required.
- To set a static IPv6 address, network mask, default gateway, and DNS server to the bond0 connection, enter:

  ```
  # nmcli connection modify bond0 ipv6.addresses '2001:db8:1::1/64' ipv6.gateway '2001:db8:1::fffe' ipv6.dns '2001:db8:1::fffd' ipv6.dns-search 'example.com' ipv6.method manual
  ```

6. Optional: If you want to set any parameters on the bond ports, use the following command:

  ```
  # nmcli connection modify bond0-port1 bond-port.<parameter> <value>
  ```

7. Activate the connection:

  ```
  # nmcli connection up bond0
  ```

8. Verify that the ports are connected, and the CONNECTION column displays the port’s connection name:

  ```
  # nmcli device
  DEVICE   TYPE      STATE      CONNECTION
  ...   enp7s0   ethernet  connected  bond0-port1
        enp8s0   ethernet  connected  bond0-port2
  ```

When you activate any port of the connection, NetworkManager also activates the bond, but not the other ports of it. You can configure that Red Hat Enterprise Linux enables all ports automatically when the bond is enabled:

a. Enable the connection.autoconnect-slaves parameter of the bond’s connection:

  ```
  # nmcli connection modify bond0 connection.autoconnect-slaves 1
  ```

b. Reactivate the bridge:
# nmcli connection up bond0

Verification

1. Temporarily remove the network cable from the host.
   Note that there is no method to properly test link failure events using software utilities. Tools that deactivate connections, such as `nmcli`, show only the bonding driver’s ability to handle port configuration changes and not actual link failure events.

2. Display the status of the bond:

   # cat /proc/net/bonding/bond0

Additional resources

- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- Network bonding documentation

3.4. CONFIGURING A NETWORK BOND BY USING THE RHEL WEB CONSOLE

Use the RHEL web console to configure a network bond if you prefer to manage network settings using a web browser-based interface.

Prerequisites

- You are logged in to the RHEL web console.
- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as members of the bond, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bridge, or VLAN devices as members of the bond, create them in advance as described in:
  - Configuring a network team by using the RHEL web console
  - Configuring a network bridge by using the RHEL web console
  - Configuring VLAN tagging by using the RHEL web console

Procedure

1. Select the Networking tab in the navigation on the left side of the screen.
2. Click Add bond in the Interfaces section.
3. Enter the name of the bond device you want to create.
4. Select the interfaces that should be members of the bond.
5. Select the mode of the bond.
If you select **Active backup**, the web console shows the additional field **Primary** in which you can select the preferred active device.

6. Set the link monitoring mode. For example, when you use the **Adaptive load balancing** mode, set it to **ARP**.

7. Optional: Adjust the monitoring interval, link up delay, and link down delay settings. Typically, you only change the defaults for troubleshooting purposes.

8. Click **Apply**.

9. By default, the bond uses a dynamic IP address. If you want to set a static IP address:
   a. Click the name of the bond in the **Interfaces** section.
b. Click **Edit** next to the protocol you want to configure.

c. Select **Manual** next to **Addresses**, and enter the IP address, prefix, and default gateway.

d. In the **DNS** section, click the + button, and enter the IP address of the DNS server. Repeat this step to set multiple DNS servers.

e. In the **DNS search domains** section, click the + button, and enter the search domain.

f. If the interface requires static routes, configure them in the **Routes** section.

g. Click **Apply**

**Verification**

1. Select the **Networking** tab in the navigation on the left side of the screen, and check if there is incoming and outgoing traffic on the interface:

   ![IPv4 settings](image)

   **Interfaces**

<table>
<thead>
<tr>
<th>Name</th>
<th>IP address</th>
<th>Sending</th>
<th>Receiving</th>
</tr>
</thead>
<tbody>
<tr>
<td>bond0</td>
<td>192.0.2/24</td>
<td>1.11 Mbps</td>
<td>61.2 Mbps</td>
</tr>
</tbody>
</table>

2. Temporarily remove the network cable from the host.
Note that there is no method to properly test link failure events using software utilities. Tools that deactivate connections, such as the web console, show only the bonding driver’s ability to handle member configuration changes and not actual link failure events.

3. Display the status of the bond:

```
# cat /proc/net/bonding/bond0
```

### 3.5. CONFIGURING A NETWORK BOND BY USING NMTUI

The `nmtui` application provides a text-based user interface for NetworkManager. You can use `nmtui` to configure a network bond on a host without a graphical interface.

**NOTE**

In `nmtui`:

- Navigate by using the cursor keys.
- Press a button by selecting it and hitting **Enter**.
- Select and deselect checkboxes by using **Space**.

**Prerequisites**

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bond, the physical or virtual Ethernet devices must be installed on the server.

**Procedure**

1. If you do not know the network device names on which you want configure a network bond, display the available devices:

```
# nmcli device status
```

<table>
<thead>
<tr>
<th>DEVICE</th>
<th>TYPE</th>
<th>STATE</th>
<th>CONNECTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>enp7s0</td>
<td>ethernet</td>
<td>unavailable</td>
<td>--</td>
</tr>
<tr>
<td>enp8s0</td>
<td>ethernet</td>
<td>unavailable</td>
<td>--</td>
</tr>
</tbody>
</table>

```
```

2. Start `nmtui`:

```
# nmtui
```

3. Select **Edit a connection**, and press **Enter**.

4. Press the **Add** button.

5. Select **Bond** from the list of network types, and press **Enter**.

6. Optional: Enter a name for the NetworkManager profile to be created. On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.
7. Enter the bond device name to be created into the **Device** field.

8. Add ports to the bond to be created:
   
   a. Press the **Add** button next to the **Slaves** list.
   
   b. Select the type of the interface you want to add as port to the bond, for example, **Ethernet**.
   
   c. Optional: Enter a name for the NetworkManager profile to be created for this bond port.
   
   d. Enter the port’s device name into the **Device** field.
   
   e. Press the **OK** button to return to the window with the bond settings.

   **Figure 3.1. Adding an Ethernet device as port to a bond**

   ![Adding an Ethernet device as port to a bond](image)

   f. Repeat these steps to add more ports to the bond.

9. Set the bond mode. Depending on the value you set, **nmtui** displays additional fields for settings that are related to the selected mode.

10. Depending on your environment, configure the IP address settings in the **IPv4 configuration** and **IPv6 configuration** areas accordingly. For this, press the button next to these areas, and select:

   - **Disabled**, if the bond does not require an IP address.
   
   - **Automatic**, if a DHCP server or stateless address autoconfiguration (SLAAC) dynamically assigns an IP address to the bond.
   
   - **Manual**, if the network requires static IP address settings. In this case, you must fill further fields:
     
     i. Press the **Show** button next to the protocol you want to configure to display additional fields.
     
     ii. Press the **Add** button next to **Addresses**, and enter the IP address and the subnet mask in Classless Inter-Domain Routing (CIDR) format.

     If you do not specify a subnet mask, NetworkManager sets a /32 subnet mask for IPv4 addresses and /64 for IPv6 addresses.
     
     iii. Enter the address of the default gateway.
     
     iv. Press the **Add** button next to **DNS servers**, and enter the DNS server address.
v. Press the **Add** button next to **Search domains**, and enter the DNS search domain.

Figure 3.2. Example of a bond connection with static IP address settings
11. Press the **OK** button to create and automatically activate the new connection.

12. Press the **Back** button to return to the main menu.

13. Select **Quit**, and press **Enter** to close the **nmtui** application.

**Verification**

1. Temporarily remove the network cable from the host.
   
   Note that there is no method to properly test link failure events using software utilities. Tools that deactivate connections, such as **nmcli**, show only the bonding driver’s ability to handle port configuration changes and not actual link failure events.

2. Display the status of the bond:

   ```bash
   # cat /proc/net/bonding/bond0
   ```

**3.6. CONFIGURING A NETWORK BOND BY USING NM-CONNECTION-EDITOR**

If you use Red Hat Enterprise Linux with a graphical interface, you can configure network bonds using the **nm-connection-editor** application.

Note that **nm-connection-editor** can add only new ports to a bond. To use an existing connection profile as a port, create the bond by using the **nmcli** utility as described in **Configuring a network bond by using nmcli**.

**Prerequisites**

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bond, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bond, or VLAN devices as ports of the bond, ensure that these devices are not already configured.

**Procedure**

1. Open a terminal, and enter **nm-connection-editor**:

   ```bash
   $ nm-connection-editor
   ```

2. Click the **+** button to add a new connection.

3. Select the **Bond** connection type, and click **Create**.

4. On the **Bond** tab:

   a. Optional: Set the name of the bond interface in the **Interface name** field.

   b. Click the **Add** button to add a network interface as a port to the bond.

      i. Select the connection type of the interface. For example, select **Ethernet** for a wired connection.
ii. Optional: Set a connection name for the port.

iii. If you create a connection profile for an Ethernet device, open the Ethernet tab, and select in the Device field the network interface you want to add as a port to the bond. If you selected a different device type, configure it accordingly. Note that you can only use Ethernet interfaces in a bond that are not configured.

iv. Click Save.

c. Repeat the previous step for each interface you want to add to the bond:

![Editing Bond connection 1](image)

Bonded connections:
- bond0-port1
- bond0-port2

Optional: Set other options, such as the Media Independent Interface (MII) monitoring interval.

5. Configure the IP address settings on both the IPv4 Settings and IPv6 Settings tabs:
   - To use this bridge device as a port of other devices, set the Method field to Disabled.
   - To use DHCP, leave the Method field at its default, Automatic (DHCP).
   - To use static IP settings, set the Method field to Manual and fill the fields accordingly:

6. Click Save.


Verification

1. Temporarily remove the network cable from the host.
Note that there is no method to properly test link failure events using software utilities. Tools that deactivate connections, such as `nmcli`, show only the bonding driver’s ability to handle port configuration changes and not actual link failure events.

2. Display the status of the bond:

```
# cat /proc/net/bonding/bond0
```

Additional resources

- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- Configuring a network team by using `nm-connection-editor`
- Configuring a network bridge by using `nm-connection-editor`
- Configuring VLAN tagging by using `nm-connection-editor`

### 3.7. CONFIGURING A NETWORK BOND BY USING NMSTATECTL

Use the `nmstatectl` utility to configure a network bond through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.

Depending on your environment, adjust the YAML file accordingly. For example, to use different devices than Ethernet adapters in the bond, adapt the `base-face` attribute and `type` attributes of the ports you use in the bond.

#### Prerequisites

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports in the bond, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bridge, or VLAN devices as ports in the bond, set the interface name in the `port` list, and define the corresponding interfaces.
- The `nmstate` package is installed.

#### Procedure

1. Create a YAML file, for example `~/create-bond.yml`, with the following content:

```yaml
---
interfaces:
- name: bond0
  type: bond
  state: up
  ipv4:
    enabled: true
    address:
      - ip: 192.0.2.1
        prefix-length: 24
        dhcp: false
```
These settings define a network bond with the following settings:

- Network interfaces in the bond: **enp1s0** and **enp7s0**
- Mode: **active-backup**
- Static IPv4 address: **192.0.2.1** with a /24 subnet mask
- Static IPv6 address: **2001:db8:1::1** with a /64 subnet mask
- IPv4 default gateway: **192.0.2.254**
- IPv6 default gateway: **2001:db8:1::fffe**
- IPv4 DNS server: **192.0.2.200**
- IPv6 DNS server: **2001:db8:1::ffbb**
- DNS search domain: **example.com**

2. Apply the settings to the system:
# nmstatectl apply ~/create-bond.yml

Verification

1. Display the status of the devices and connections:

   ```
   # nmcli device status
   DEVICE      TYPE      STATE      CONNECTION
   bond0       bond      connected  bond0
   ```

2. Display all settings of the connection profile:

   ```
   # nmcli connection show bond0
   connection.id: bond0
   connection.uuid: 79cbc3bd-302e-4b1f-ad89-f12533b818ee
   connection.stable-id: --
   connection.type: bond
   connection.interface-name: bond0
   ...
   ```

3. Display the connection settings in YAML format:

   ```
   # nmstatectl show bond0
   ```

Additional resources

- `nmstatectl(8)` man page
- `/usr/share/doc/nmstate/examples/` directory

3.8. CONFIGURING A NETWORK BOND BY USING THE NETWORK RHEL SYSTEM ROLE

You can remotely configure a network bond by using the `network` RHEL system role.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
- Two or more physical or virtual network devices are installed on the server.

Procedure

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

   ```yaml
   ---
   - name: Configure the network
     hosts: managed-node-01.example.com
   ```
These settings define a network bond with the following settings:

- A static IPv4 address - 192.0.2.1 with a /24 subnet mask
- A static IPv6 address - 2001:db8:1::1 with a /64 subnet mask
- An IPv4 default gateway - 192.0.2.254
- An IPv6 default gateway - 2001:db8:1::ffffe
- An IPv4 DNS server - 192.0.2.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com
- Ports of the bond - enp7s0 and enp8s0
3.9. CREATING A NETWORK BOND TO ENABLE SWITCHING BETWEEN AN ETHERNET AND WIRELESS CONNECTION WITHOUT INTERRUPTING THE VPN

RHEL users who connect their workstation to their company’s network typically use a VPN to access remote resources. However, if the workstation switches between an Ethernet and Wi-Fi connection, for example, if you release a laptop from a docking station with an Ethernet connection, the VPN connection is interrupted. To avoid this problem, you can create a network bond that uses the Ethernet and Wi-Fi connection in active-backup mode.

Prerequisites

- The host contains an Ethernet and a Wi-Fi device.
- An Ethernet and Wi-Fi NetworkManager connection profile has been created and both connections work independently.

This procedure uses the following connection profiles to create a network bond named bond0:

- Docking_station associated with the enp11s0u1 Ethernet device
- Wi-Fi associated with the wlp1s0 Wi-Fi device

Procedure

1. Create a bond interface in active-backup mode:

```bash
# nmcli connection add type bond con-name bond0 ifname bond0 bond.options
"mode=active-backup"
```

This command names both the interface and connection profile bond0.
2. Configure the IPv4 settings of the bond:
   - If a DHCP server in your network assigns IPv4 addresses to hosts, no action is required.
   - If your local network requires static IPv4 addresses, set the address, network mask, default gateway, DNS server, and DNS search domain to the bond0 connection:

   ```
   # nmcli connection modify bond0 ipv4.addresses '192.0.2.1/24'
   # nmcli connection modify bond0 ipv4.gateway '192.0.2.254'
   # nmcli connection modify bond0 ipv4.dns '192.0.2.253'
   # nmcli connection modify bond0 ipv4.dns-search 'example.com'
   # nmcli connection modify bond0 ipv4.method manual
   ```

3. Configure the IPv6 settings of the bond:
   - If your router or a DHCP server in your network assigns IPv6 addresses to hosts, no action is required.
   - If your local network requires static IPv6 addresses, set the address, network mask, default gateway, DNS server, and DNS search domain to the bond0 connection:

   ```
   # nmcli connection modify bond0 ipv6.addresses '2001:db8:1::1/64'
   # nmcli connection modify bond0 ipv6.gateway '2001:db8:1::fffe'
   # nmcli connection modify bond0 ipv6.dns '2001:db8:1::fffd'
   # nmcli connection modify bond0 ipv6.dns-search 'example.com'
   # nmcli connection modify bond0 ipv6.method manual
   ```

4. Display the connection profiles:

   ```
   # nmcli connection show
   NAME             UUID                                  TYPE      DEVICE
   Docking_station  256dd073-fecc-339d-91ae-9834a00407f9  ethernet  enp11s0u1
   Wi-Fi            1f1531c7-8737-4c60-91af-2d21164417e8  wifi      wlp1s0
   ...
   ```

   You require the names of the connection profiles and the Ethernet device name in the next steps.

5. Assign the connection profile of the Ethernet connection to the bond:

   ```
   # nmcli connection modify Docking_station master bond0
   ```

6. Assign the connection profile of the Wi-Fi connection to the bond:

   ```
   # nmcli connection modify Wi-Fi master bond0
   ```

7. If your Wi-Fi network uses MAC filtering to allow only MAC addresses on a allow list to access the network, configure that NetworkManager dynamically assigns the MAC address of the active port to the bond:

   ```
   # nmcli connection modify bond0 +bond.options fail_over_mac=1
   ```

   With this setting, you must set only the MAC address of the Wi-Fi device to the allow list instead of the MAC address of both the Ethernet and Wi-Fi device.
8. Set the device associated with the Ethernet connection as primary device of the bond:

```bash
# nmcli con modify bond0 +bond.options "primary=enp11s0u1"
```

With this setting, the bond always uses the Ethernet connection if it is available.

9. Configure that NetworkManager automatically activates ports when the bond0 device is activated:

```bash
# nmcli connection modify bond0 connection.autoconnect-slaves 1
```

10. Activate the bond0 connection:

```bash
# nmcli connection up bond0
```

**Verification**

- Display the currently active device, the status of the bond and its ports:

```bash
# cat /proc/net/bonding/bond0
```

```
Ethernet Channel Bonding Driver: v3.7.1 (April 27, 2011)

Bonding Mode: fault-tolerance (active-backup) (fail_over_mac active)
Primary Slave: enp11s0u1 (primary_reselect always)
**Currently Active Slave: enp11s0u1**
**MII Status:** up
MII Polling Interval (ms): 1
Up Delay (ms): 0
Down Delay (ms): 0
Peer Notification Delay (ms): 0

Slave Interface: enp11s0u1
**MII Status:** up
Speed: 1000 Mbps
Duplex: full
Link Failure Count: 0
Permanent HW addr: 00:53:00:59:da:b7
Slave queue ID: 0

Slave Interface: wlp1s0
**MII Status:** up
Speed: Unknown
Duplex: Unknown
Link Failure Count: 2
Permanent HW addr: 00:53:00:b3:22:ba
Slave queue ID: 0
```

**Additional resources**

- [Configuring an Ethernet connection](#)
- [Managing Wi-Fi connections](#)
- [Configuring network bonding](#)
3.10. THE DIFFERENT NETWORK BONDING MODES

The Linux bonding driver provides link aggregation. Bonding is the process of aggregating multiple network interfaces in parallel to provide a single logical bonded interface. The actions of a bonded interface depend on the bonding policy that is also known as mode. The different modes provide either load-balancing or hot standby services.

The following modes exist:

Balance-rr (Mode 0)

Balance-rr uses the round-robin algorithm that sequentially transmits packets from the first available port to the last one. This mode provides load balancing and fault tolerance. This mode requires switch configuration of a port aggregation group, also called EtherChannel or similar port grouping. An EtherChannel is a port link aggregation technology to group multiple physical Ethernet links to one logical Ethernet link.

The drawback of this mode is that it is not suitable for heavy workloads and if TCP throughput or ordered packet delivery is essential.

Active-backup (Mode 1)

Active-backup uses the policy that determines that only one port is active in the bond. This mode provides fault tolerance and does not require any switch configuration. If the active port fails, an alternate port becomes active. The bond sends a gratuitous address resolution protocol (ARP) response to the network. The gratuitous ARP forces the receiver of the ARP frame to update their forwarding table. The Active-backup mode transmits a gratuitous ARP to announce the new path to maintain connectivity for the host.

The primary option defines the preferred port of the bonding interface.

Balance-xor (Mode 2)

Balance-xor uses the selected transmit hash policy to send the packets. This mode provides load balancing, fault tolerance, and requires switch configuration to set up an Etherchannel or similar port grouping. To alter packet transmission and balance transmit, this mode uses the xmit_hash_policy option. Depending on the source or destination of traffic on the interface, the interface requires an additional load-balancing configuration. See description xmit_hash_policy bonding parameter.

Broadcast (Mode 3)

Broadcast uses a policy that transmits every packet on all interfaces. This mode provides fault tolerance and requires a switch configuration to set up an EtherChannel or similar port grouping. The drawback of this mode is that it is not suitable for heavy workloads and if TCP throughput or ordered packet delivery is essential.

802.3ad (Mode 4)

802.3ad uses the same-named IEEE standard dynamic link aggregation policy. This mode provides fault tolerance. This mode requires switch configuration to set up a Link Aggregation Control Protocol (LACP) port grouping. This mode creates aggregation groups that share the same speed and duplex settings and utilizes all ports in the active aggregator. Depending on the source or destination of traffic on the interface, this mode requires an additional load-balancing configuration.
By default, the port selection for outgoing traffic depends on the transmit hash policy. Use the \texttt{xmit\_hash\_policy} option of the transmit hash policy to change the port selection and balance transmit.

The difference between the \texttt{802.3ad} and the \textbf{Balance-xor} is compliance. The \texttt{802.3ad} policy negotiates LACP between the port aggregation groups. See description \texttt{xmit\_hash\_policy} bonding parameter

\textbf{Balance-\texttt{tlb}} (Mode 5)

\texttt{Balance-\texttt{tlb}} uses the transmit load balancing policy. This mode provides fault tolerance, load balancing, and establishes channel bonding that does not require any switch support.

The active port receives the incoming traffic. In case of failure of the active port, another one takes over the MAC address of the failed port. To decide which interface processes the outgoing traffic, use one of the following modes:

- \texttt{Value 0}: Uses the hash distribution policy to distribute traffic without load balancing
- \texttt{Value 1}: Distributes traffic to each port by using load balancing

With the bonding option \texttt{tlb\_dynamic\_lb=0}, this bonding mode uses the \texttt{xmit\_hash\_policy} bonding option to balance transmit. The \texttt{primary} option defines the preferred port of the bonding interface.

See description \texttt{xmit\_hash\_policy} bonding parameter.

\textbf{Balance-\texttt{alb}} (Mode 6)

\texttt{Balance-\texttt{alb}} uses an adaptive load balancing policy. This mode provides fault tolerance, load balancing, and does not require any special switch support.

This mode includes balance-transmit load balancing (\texttt{balance-\texttt{tlb}}) and receive-load balancing for IPv4 and IPv6 traffic. The bonding intercepts ARP replies sent by the local system and overwrites the source hardware address of one of the ports in the bond. ARP negotiation manages the receive-load balancing. Therefore, different ports use different hardware addresses for the server.

The \texttt{primary} option defines the preferred port of the bonding interface. With the bonding option \texttt{tlb\_dynamic\_lb=0}, this bonding mode uses the \texttt{xmit\_hash\_policy} bonding option to balance transmit. See description \texttt{xmit\_hash\_policy} bonding parameter.

\textbf{Additional resources}

- \texttt{/usr/share/doc/kernel-doc-<version>/Documentation/networking/bonding.rst} provided by the \texttt{kernel-doc} package
- \texttt{/usr/share/doc/kernel-doc-<version>/Documentation/networking/bonding.txt} provided by the \texttt{kernel-doc} package
- Which bonding modes work when used with a bridge that virtual machine guests or containers connect to?
- How are the values for different policies in "\texttt{xmit\_hash\_policy}" bonding parameter calculated?

\textbf{3.11. THE XMIT\_HASH\_POLICY BONDING PARAMETER}

The \texttt{xmit\_hash\_policy} load balancing parameter selects the transmit hash policy for a node selection in the \texttt{balance-xor}, \texttt{802.3ad}, \texttt{balance-alb}, and \texttt{balance-\texttt{tlb}} modes. It is only applicable to mode 5 and 6 if the \texttt{tlb\_dynamic\_lb} parameter is \texttt{0}. The possible values of this parameter are \texttt{layer2}, \texttt{layer2+3},
layer3+4, encap2+3, encap3+4, and vlan+srcmac.

Refer the table for details:

<table>
<thead>
<tr>
<th>Policy or Network layers</th>
<th>Layer2</th>
<th>Layer2+3</th>
<th>Layer3+4</th>
<th>encap2+3</th>
<th>encap3+4</th>
<th>VLAN+srcmac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses</td>
<td>XOR of source and destination MAC addresses and Ethernet protocol type</td>
<td>XOR of source and destination MAC addresses and IP addresses</td>
<td>XOR of source and destination ports and IP addresses</td>
<td>XOR of source and destination MAC addresses and IP addresses inside a supported tunnel, for example, VXLAN. This mode relies on skb_flow_dissect() function to obtain the header fields</td>
<td>XOR of source and destination ports and IP addresses inside a supported tunnel, for example, VXLAN. This mode relies on skb_flow_dissect() function to obtain the header fields</td>
<td>XOR of VLAN ID and source MAC vendor and source MAC device</td>
</tr>
<tr>
<td>Placement of traffic</td>
<td>All traffic to a particular network peer on the same underlying network interface</td>
<td>All traffic to a particular IP address on the same underlying network interface</td>
<td>All traffic to a particular IP address and port on the same underlying network interface</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Primary choice</strong></td>
<td>If network traffic is between this system and multiple other systems in the same broadcast domain</td>
<td>If network traffic between this system and multiple other systems goes through a default gateway</td>
<td>If network traffic between this system and another system uses the same IP addresses but goes through multiple ports</td>
<td>The encapsulated traffic is between the source system and multiple other systems using multiple IP addresses</td>
<td>If the bond carries network traffic, from multiple containers or virtual machines (VM), that expose their MAC address directly to the external network such as the bridge network, and you can not configure a switch for Mode 2 or Mode 4</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>Secondary choice</strong></td>
<td>If network traffic is mostly between this system and multiple other systems behind a default gateway</td>
<td>If network traffic is mostly between this system and another system</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Compliant</strong></td>
<td>802.3ad</td>
<td>802.3ad</td>
<td>Not 802.3ad</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Default policy</strong></td>
<td>This is the default policy if no configuration is provided</td>
<td>For non-IP traffic, the formula is the same as for the layer2 transmit policy</td>
<td>For non-IP traffic, the formula is the same as for the layer2 transmit policy</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 4. CONFIGURING NETWORK TEAMING

A network team is a method to combine or aggregate physical and virtual network interfaces to provide a logical interface with higher throughput or redundancy. Network teaming uses a small kernel module to implement fast handling of packet flows and a user-space service for other tasks. This way, network teaming is an easily extensible and scalable solution for load-balancing and redundancy requirements.

Red Hat Enterprise Linux provides administrators different options to configure team devices. For example:

- Use `nmcli` to configure teams connections using the command line.
- Use the RHEL web console to configure team connections using a web browser.
- Use the `nm-connection-editor` application to configure team connections in a graphical interface.

**IMPORTANT**

Network teaming is deprecated in Red Hat Enterprise Linux 9. Consider using the network bonding driver as an alternative. For details, see Configuring network bonding.

4.1. MIGRATING A NETWORK TEAM CONFIGURATION TO NETWORK BOND

Network teaming is deprecated in Red Hat Enterprise Linux 9. If you already have a working network team configured, for example because you upgraded from an earlier RHEL version, you can migrate the configuration to a network bond that is managed by NetworkManager.

**IMPORTANT**

The `team2bond` utility only converts the network team configuration to a bond. Afterwards, you must manually configure further settings of the bond, such as IP addresses and DNS configuration.

**Prerequisites**

- The `team-team0` NetworkManager connection profile is configured and manages the `team0` device.
- The `teamd` package is installed.

**Procedure**

1. Optional: Display the IP configuration of the `team-team0` NetworkManager connection:

```bash
# nmcli connection show team-team0 | egrep "^ip"
... ipv4.method: manual
ipv4.dns: 192.0.2.253
ipv4.dns-search: example.com
ipv4.addresses: 192.0.2.1/24
ipv4.gateway: 192.0.2.254
... 
```
ipv6.method: manual
ipv6.dns: 2001:db8:1::fffd
ipv6.dns-search: example.com
ipv6.addresses: 2001:db8:1::/64
ipv6.gateway: 2001:db8:1::fffe

2. Export the configuration of the team0 device to a JSON file:

```bash
# teamdctl team0 config dump actual > /tmp/team0.json
```

3. Remove the network team. For example, if you configured the team in NetworkManager, remove the team-team0 connection profile and the profiles of associated ports:

```bash
# nmcli connection delete team-team0
# nmcli connection delete team-team0-port1
# nmcli connection delete team-team0-port2
```

4. Run the team2bond utility in dry-run mode to display nmcli commands that set up a network bond with similar settings as the team device:

```bash
# team2bond --config=/tmp/team0.json --rename=bond0
nmcli con add type bond ifname bond0 bond.options "mode=active-backup,num_grat_arp=1,num_unsol_na=1,miimon=100,miimon=100"
nmcli con add type ethernet ifname enp7s0 master bond0
nmcli con add type ethernet ifname enp8s0 master bond0
```

The first command contains two miimon options because the team configuration file contained two link_watch entries. Note that this does not affect the creation of the bond.

If you bound services to the device name of the team and want to avoid updating or breaking these services, omit the --rename=bond0 option. In this case, team2bond uses the same interface name for the bond as for the team.

5. Verify that the options for the bond the team2bond utility suggested are correct.

6. Create the bond. You can execute the suggested nmcli commands or re-run the team2bond command with the --exec-cmd option:

```bash
# team2bond --config=/tmp/team0.json --rename=bond0 --exec-cmd
```

Connection 'bond-bond0' (0241a531-0c72-4202-80df-73eadfc126b5) successfully added.
Connection 'bond-slave-enp7s0' (38489729-b624-4606-a784-1ccf01e2f6d6) successfully added.
Connection 'bond-slave-enp8s0' (de97ec06-7daa-4298-9a71-9d4c7909d3a1) successfully added.

You require the name of the bond connection profile (bond-bond0) in the next steps.

7. Set the IPv4 settings that were previously configured on team-team0 to the bond-bond0 connection:

```bash
# nmcli connection modify bond-bond0 ipv4.addresses '192.0.2.1/24'
# nmcli connection modify bond-bond0 ipv4.gateway '192.0.2.254'
# nmcli connection modify bond-bond0 ipv4.dns '192.0.2.253'
```
8. Set the IPv6 settings that were previously configured on team-team0 to the bond-bond0 connection:

```
# nmcli connection modify bond-bond0 ipv6.addresses '2001:db8:1::1/64'
# nmcli connection modify bond-bond0 ipv6.gateway '2001:db8:1::fffe'
# nmcli connection modify bond-bond0 ipv6.dns '2001:db8:1::fffd'
# nmcli connection modify bond-bond0 ipv6.dns-search 'example.com'
# nmcli connection modify bond-bond0 ipv6.method manual
```

9. Activate the connection:

```
# nmcli connection up bond-bond0
```

Verification

1. Display the IP configuration of the bond-bond0 NetworkManager connection:

```
# nmcli connection show bond-bond0 | egrep "^ip"
...
ipv4.method: manual
ipv4.dns: 192.0.2.253
ipv4.dns-search: example.com
ipv4.addresses: 192.0.2.1/24
ipv4.gateway: 192.0.2.254
...
ipv6.method: manual
ipv6.dns: 2001:db8:1::fffd
ipv6.dns-search: example.com
ipv6.addresses: 2001:db8:1::1/64
ipv6.gateway: 2001:db8:1::fffe
...
```

2. Display the status of the bond:

```
# cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v5.13.0-0.rc7.51.el9.x86_64
Bonding Mode: fault-tolerance (active-backup)
Primary Slave: None
Currently Active Slave: enp7s0
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 0
Down Delay (ms): 0
Peer Notification Delay (ms): 0
Slave Interface: enp7s0
MII Status: up
Speed: Unknown
Duplex: Unknown
Link Failure Count: 0
```
Permanent HW addr: 52:54:00:bf:b1:a9
Slave queue ID: 0
Slave Interface: enp8s0
MII Status: up
Speed: Unknown
Duplex: Unknown
Link Failure Count: 0
Permanent HW addr: 52:54:00:04:36:0f
Slave queue ID: 0

In this example, both ports are up.

3. To verify that bonding failover works:
   a. Temporarily remove the network cable from the host. Note that there is no method to properly test link failure events using the command line.
   b. Display the status of the bond:

   ```
   # cat /proc/net/bonding/bond0
   ```

4.2. UNDERSTANDING THE DEFAULT BEHAVIOR OF CONTROLLER AND PORT INTERFACES

Consider the following default behavior when managing or troubleshooting team or bond port interfaces using the NetworkManager service:

- Starting the controller interface does not automatically start the port interfaces.
- Starting a port interface always starts the controller interface.
- Stopping the controller interface also stops the port interface.
- A controller without ports can start static IP connections.
- A controller without ports waits for ports when starting DHCP connections.
- A controller with a DHCP connection waiting for ports completes when you add a port with a carrier.
- A controller with a DHCP connection waiting for ports continues waiting when you add a port without carrier.

4.3. UNDERSTANDING THE TEAMD SERVICE, RUNNERS, AND LINK-WATCHERS

The team service, teamd, controls one instance of the team driver. This instance of the driver adds instances of a hardware device driver to form a team of network interfaces. The team driver presents a network interface, for example team0, to the kernel.

The teamd service implements the common logic to all methods of teaming. Those functions are unique to the different load sharing and backup methods, such as round-robin, and implemented by separate units of code referred to as runners. Administrators specify runners in JavaScript Object Notation.
(JSON) format, and the JSON code is compiled into an instance of teamd when the instance is created. Alternatively, when using NetworkManager, you can set the runner in the team.runner parameter, and NetworkManager auto-creates the corresponding JSON code.

The following runners are available:

- **broadcast**: Transmits data over all ports.
- **roundrobin**: Transmits data over all ports in turn.
- **activebackup**: Transmits data over one port while the others are kept as a backup.
- **loadbalance**: Transmits data over all ports with active Tx load balancing and Berkeley Packet Filter (BPF)-based Tx port selectors.
- **random**: Transmits data on a randomly selected port.
- **lacp**: Implements the 802.3ad Link Aggregation Control Protocol (LACP).

The teamd services uses a link watcher to monitor the state of subordinate devices. The following link-watchers are available:

- **ethtool**: The libteam library uses the ethtool utility to watch for link state changes. This is the default link-watcher.
- **arp_ping**: The libteam library uses the arp_ping utility to monitor the presence of a far-end hardware address using Address Resolution Protocol (ARP).
- **nsna_ping**: On IPv6 connections, the libteam library uses the Neighbor Advertisement and Neighbor Solicitation features from the IPv6 Neighbor Discovery protocol to monitor the presence of a neighbor’s interface.

Each runner can use any link watcher, with the exception of lacp. This runner can only use the ethtool link watcher.

### 4.4. CONFIGURING A NETWORK TEAM BY USING NMCLI

To configure a network team on the command line, use the nmcli utility.

**IMPORTANT**

Network teaming is deprecated in Red Hat Enterprise Linux 9. Consider using the network bonding driver as an alternative. For details, see Configuring network bonding.

**Prerequisites**

- The teamd and NetworkManager-team packages are installed.
- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the team, the physical or virtual Ethernet devices must be installed on the server and connected to a switch.
- To use bond, bridge, or VLAN devices as ports of the team, you can either create these devices while you create the team or you can create them in advance as described in:
Configuring a network bond by using nmcli

Configuring a network bridge by using nmcli

Configuring VLAN tagging by using nmcli

Procedure

1. Create a team interface:

```
# nmcli connection add type team con-name team0 ifname team0 team.runner activebackup
```

This command creates a network team named `team0` that uses the `activebackup` runner.

2. Optionally, set a link watcher. For example, to set the `ethtool` link watcher in the `team0` connection profile:

```
# nmcli connection modify team0 team.link-watchers "name=ethtool"
```

Link watchers support different parameters. To set parameters for a link watcher, specify them space-separated in the `name` property. Note that the name property must be surrounded by quotation marks. For example, to use the `ethtool` link watcher and set its `delay-up` parameter to 2500 milliseconds (2.5 seconds):

```
# nmcli connection modify team0 team.link-watchers "name=ethtool delay-up=2500"
```

To set multiple link watchers and each of them with specific parameters, the link watchers must be separated by a comma. The following example sets the `ethtool` link watcher with the `delay-up` parameter and the `arp_ping` link watcher with the `source-host` and `target-host` parameter:

```
# nmcli connection modify team0 team.link-watchers "name=ethtool delay-up=2, name=arp_ping source-host=192.0.2.1 target-host=192.0.2.2"
```

3. Display the network interfaces, and note the names of the interfaces you want to add to the team:

```
# nmcli device status
DEVICE  TYPE      STATE         CONNECTION
enp7s0  ethernet  disconnected  --
enp8s0  ethernet  disconnected  --
bond0  bond      connected  bond0
bond1  bond      connected  bond1
...
```

In this example:

- `enp7s0` and `enp8s0` are not configured. To use these devices as ports, add connection profiles in the next step. Note that you can only use Ethernet interfaces in a team that are not assigned to any connection.

- `bond0` and `bond1` have existing connection profiles. To use these devices as ports, modify their profiles in the next step.
4. Assign the port interfaces to the team:
   a. If the interfaces you want to assign to the team are not configured, create new connection profiles for them:

```bash
# nmcli connection add type ethernet slave-type team con-name team0-port1
    ifname enp7s0 master team0
# nmcli connection add type ethernet slave-type team con-name team0-port2
    ifname enp8s0 master team0
```

These commands create profiles for `enp7s0` and `enp8s0`, and add them to the `team0` connection.

b. To assign an existing connection profile to the team:
   i. Set the `master` parameter of these connections to `team0`:

```bash
# nmcli connection modify bond0 master team0
# nmcli connection modify bond1 master team0
```

These commands assign the existing connection profiles named `bond0` and `bond1` to the `team0` connection.

   ii. Reactivate the connections:

```bash
# nmcli connection up bond0
# nmcli connection up bond1
```

5. Configure the IPv4 settings:
   - To use this team device as a port of other devices, enter:

```bash
# nmcli connection modify team0 ipv4.method disabled
```

   - To use DHCP, no action is required.

   - To set a static IPv4 address, network mask, default gateway, and DNS server to the `team0` connection, enter:

```bash
# nmcli connection modify team0 ipv4.addresses '192.0.2.1/24' ipv4.gateway
    '192.0.2.254' ipv4.dns '192.0.2.253' ipv4.dns-search 'example.com' ipv4.method manual
```

6. Configure the IPv6 settings:
   - To use this team device as a port of other devices, enter:

```bash
# nmcli connection modify team0 ipv6.method disabled
```

   - To use stateless address autoconfiguration (SLAAC), no action is required.

   - To set a static IPv6 address, network mask, default gateway, and DNS server to the `team0` connection, enter:
# nmcli connection modify team0 ipv6.addresses '2001:db8:1::1/64' ipv6.gateway '2001:db8:1::fffe' ipv6.dns '2001:db8:1::fffd' ipv6.dns-search 'example.com' ipv6.method manual

7. Activate the connection:

```bash
# nmcli connection up team0
```

Verification

- Display the status of the team:

```bash
# teamdctl team0 state
setup:  
  runner: activebackup  
ports:  
enp7s0  
  link watches:   
    link summary: up  
    instance[link_watch_0]:  
      name: ethtool  
      link: up  
      down count: 0  
enp8s0  
  link watches:   
    link summary: up  
    instance[link_watch_0]:  
      name: ethtool  
      link: up  
      down count: 0  
runner:  
active port: enp7s0
```

In this example, both ports are up.

Additional resources

- [Configuring NetworkManager to avoid using a specific profile to provide a default gateway](#)
- [Understanding the teamd service, runners, and link-watchers](#)
- [nm-settings(5) man page](#)
- [teamd.conf(5) man page](#)

### 4.5. CONFIGURING A NETWORK TEAM BY USING THE RHEL WEB CONSOLE

Use the RHEL web console to configure a network team if you prefer to manage network settings using a web browser-based interface.
Network teaming is deprecated in Red Hat Enterprise Linux 9. Consider using the network bonding driver as an alternative. For details, see Configuring network bonding.

**Prerequisites**

- The `teamd` and `NetworkManager-team` packages are installed.
- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the team, the physical or virtual Ethernet devices must be installed on the server and connected to a switch.
- To use bond, bridge, or VLAN devices as ports of the team, create them in advance as described in:
  - Configuring a network bond by using the RHEL web console
  - Configuring a network bridge by using the RHEL web console
  - Configuring VLAN tagging by using the RHEL web console

**Procedure**

1. Select the **Networking** tab in the navigation on the left side of the screen.
2. Click **Add team** in the **Interfaces** section.
3. Enter the name of the team device you want to create.
4. Select the interfaces that should be ports of the team.
5. Select the runner of the team.
   - If you select **Load balancing** or **802.3ad LACP**, the web console shows the additional field **Balancer**.
6. Set the link watcher:
   - If you select **Ethtool**, additionally, set a link up and link down delay.
   - If you set **ARP ping** or **NSNA ping**, additionally, set a ping interval and ping target.
7. Click **Apply**.

8. By default, the team uses a dynamic IP address. If you want to set a static IP address:
   
   a. Click the name of the team in the **Interfaces** section.
   
   b. Click **Edit** next to the protocol you want to configure.
   
   c. Select **Manual** next to **Addresses**, and enter the IP address, prefix, and default gateway.
   
   d. In the **DNS** section, click the + button, and enter the IP address of the DNS server. Repeat this step to set multiple DNS servers.
   
   e. In the **DNS search domains** section, click the + button, and enter the search domain.
   
   f. If the interface requires static routes, configure them in the **Routes** section.
g. Click **Apply**

### Verification

1. Select the **Networking** tab in the navigation on the left side of the screen, and check if there is incoming and outgoing traffic on the interface.

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Add bond</th>
<th>Add team</th>
<th>Add bridge</th>
<th>Add VLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Name</strong></td>
<td><strong>IP address</strong></td>
<td><strong>Sending</strong></td>
<td><strong>Receiving</strong></td>
<td></td>
</tr>
<tr>
<td>team0</td>
<td>192.0.2.1/24</td>
<td>1.11 Mbps</td>
<td>612 Mbps</td>
<td></td>
</tr>
</tbody>
</table>

2. Display the status of the team:

```bash
# teamdctl team0 state
setup:
  runner: activebackup
ports:
enp7s0
  link watches:
    link summary: up
    instance[link_watch_0]:
      name: ethtool
      link: up
```
In this example, both ports are up.

Additional resources

- Network team runners

### 4.6. CONFIGURING A NETWORK TEAM BY USING NM-CONNECTION-EDITOR

If you use Red Hat Enterprise Linux with a graphical interface, you can configure network teams using the `nm-connection-editor` application.

Note that `nm-connection-editor` can add only new ports to a team. To use an existing connection profile as a port, create the team using the `nmcli` utility as described in Configuring a network team by using `nmcli`.

**IMPORTANT**

Network teaming is deprecated in Red Hat Enterprise Linux 9. Consider using the network bonding driver as an alternative. For details, see Configuring network bonding.

**Prerequisites**

- The `teamd` and `NetworkManager-team` packages are installed.
- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the team, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bond, or VLAN devices as ports of the team, ensure that these devices are not already configured.

**Procedure**

1. Open a terminal, and enter `nm-connection-editor`:

   ```bash
   $ nm-connection-editor
   ```

2. Click the + button to add a new connection.

3. Select the **Team** connection type, and click **Create**.
4. On the **Team** tab:
   a. Optional: Set the name of the team interface in the **Interface name** field.
   b. Click the **Add** button to add a new connection profile for a network interface and adding the profile as a port to the team.
      i. Select the connection type of the interface. For example, select **Ethernet** for a wired connection.
      ii. Optional: Set a connection name for the port.
      iii. If you create a connection profile for an Ethernet device, open the **Ethernet** tab, and select in the **Device** field the network interface you want to add as a port to the team. If you selected a different device type, configure it accordingly. Note that you can only use Ethernet interfaces in a team that are not assigned to any connection.
      iv. Click **Save**.
   c. Repeat the previous step for each interface you want to add to the team.

   ![Editing Team connection 1](image)

<table>
<thead>
<tr>
<th>Connection name</th>
<th>Team connection 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td><strong>Team</strong></td>
</tr>
<tr>
<td>Interface name</td>
<td>team0</td>
</tr>
<tr>
<td>MTU</td>
<td>automatic</td>
</tr>
<tr>
<td>Teamed connections</td>
<td>team0-port1</td>
</tr>
<tr>
<td></td>
<td>team0-port2</td>
</tr>
<tr>
<td>Add</td>
<td></td>
</tr>
<tr>
<td>Edit</td>
<td></td>
</tr>
</tbody>
</table>

d. Click the **Advanced** button to set advanced options to the team connection.
   i. On the **Runner** tab, select the runner.
   ii. On the **Link Watcher** tab, set the link watcher and its optional settings.
   iii. Click **OK**.

5. Configure the IP address settings on both the **IPv4 Settings** and **IPv6 Settings** tabs:
   - To use this bridge device as a port of other devices, set the **Method** field to **Disabled**.
   - To use DHCP, leave the **Method** field at its default, **Automatic (DHCP)**.
   - To use static IP settings, set the **Method** field to **Manual** and fill the fields accordingly:
6. Click **Save**.

7. Close **nm-connection-editor**.

**Verification**

- Display the status of the team:

```
# teamdctl team0 state
setup:
  runner: activebackup
ports:
enp7s0
  link watches:
    link summary: up
    instance[link_watch_0]:
      name: ethtool
      link: up
      down count: 0
enp8s0
  link watches:
    link summary: up
    instance[link_watch_0]:
      name: ethtool
      link: up
      down count: 0
runner:
  active port: enp7s0
```

**Additional resources**

- Configuring a network bond by using nm-connection-editor
- Configuring a network team by using nm-connection-editor
- Configuring VLAN tagging by using nm-connection-editor
- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- Understanding the teamd service, runners, and link-watchers
- NetworkManager duplicates a connection after restart of NetworkManager service
CHAPTER 5. CONFIGURING VLAN TAGGING

A Virtual Local Area Network (VLAN) is a logical network within a physical network. The VLAN interface tags packets with the VLAN ID as they pass through the interface, and removes tags of returning packets. You create VLAN interfaces on top of another interface, such as Ethernet, bond, team, or bridge devices. These interfaces are called the **parent interface**.

Red Hat Enterprise Linux provides administrators different options to configure VLAN devices. For example:

- Use `nmcli` to configure VLAN tagging using the command line.
- Use the RHEL web console to configure VLAN tagging using a web browser.
- Use `nmtui` to configure VLAN tagging in a text-based user interface.
- Use the `nm-connection-editor` application to configure connections in a graphical interface.
- Use `nmstatetctl` to configure connections through the Nmstate API.
- Use RHEL system roles to automate the VLAN configuration on one or multiple hosts.

### 5.1. CONFIGURING VLAN TAGGING BY USING NMCLI

You can configure Virtual Local Area Network (VLAN) tagging on the command line using the `nmcli` utility.

**Prerequisites**

- The interface you plan to use as a parent to the virtual VLAN interface supports VLAN tags.
- If you configure the VLAN on top of a bond interface:
  - The ports of the bond are up.
  - The bond is not configured with the `fail_over_mac=follow` option. A VLAN virtual device cannot change its MAC address to match the parent’s new MAC address. In such a case, the traffic would still be sent with the incorrect source MAC address.
  - The bond is usually not expected to get IP addresses from a DHCP server or IPv6 auto-configuration. Ensure it by setting the `ipv4.method=disable` and `ipv6.method=ignore` options while creating the bond. Otherwise, if DHCP or IPv6 auto-configuration fails after some time, the interface might be brought down.
  - The switch, the host is connected to, is configured to support VLAN tags. For details, see the documentation of your switch.

**Procedure**

1. Display the network interfaces:

   ```bash
   # nmcli device status
   DEVICE   TYPE      STATE         CONNECTION
   enp1s0    ethernet  disconnected  enp1s0
   ```
2. Create the VLAN interface. For example, to create a VLAN interface named `vlan10` that uses `enp1s0` as its parent interface and that tags packets with VLAN ID `10`, enter:

```
# nmcli connection add type vlan con-name vlan10 ifname vlan10 vlan.parent enp1s0 vlan.id 10
```

Note that the VLAN must be within the range from `0` to `4094`.

3. By default, the VLAN connection inherits the maximum transmission unit (MTU) from the parent interface. Optionally, set a different MTU value:

```
# nmcli connection modify vlan10 ethernet.mtu 2000
```

4. Configure the IPv4 settings:
   - To use this VLAN device as a port of other devices, enter:
     
     ```
     # nmcli connection modify vlan10 ipv4.method disabled
     ```
   - To use DHCP, no action is required.
   - To set a static IPv4 address, network mask, default gateway, and DNS server to the `vlan10` connection, enter:
     
     ```
     # nmcli connection modify vlan10 ipv4.addresses '192.0.2.1/24' ipv4.gateway '192.0.2.254' ipv4.dns '192.0.2.253' ipv4.method manual
     ```

5. Configure the IPv6 settings:
   - To use this VLAN device as a port of other devices, enter:
     
     ```
     # nmcli connection modify vlan10 ipv6.method disabled
     ```
   - To use stateless address autoconfiguration (SLAAC), no action is required.
   - To set a static IPv6 address, network mask, default gateway, and DNS server to the `vlan10` connection, enter:
     
     ```
     ```

6. Activate the connection:

```
# nmcli connection up vlan10
```

**Verification**

- Verify the settings:
# ip -d addr show vlan10
4: vlan10@enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue
   state UP group default qlen 1000
   link/ether 52:54:00:72:2f:6e brd ff:ff:ff:ff:ff:ff promiscuity 0
   vlan protocol 802.1Q id 10 <REORDER_HDR> numtxqueues 1 numrxqueues 1
   gso_max_size 65536 gso_max_segs 65535
   inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute vlan10
      valid_lft forever preferred_lft forever
   inet6 2001:db8:1::1/32 scope global noprefixroute
      valid_lft forever preferred_lft forever
   inet6 fe80::8dd7:9030:6f8e:89e6/64 scope link noprefixroute
      valid_lft forever preferred_lft forever

Additional resources
- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- nm-settings(5) man page

5.2. CONFIGURING NESTED VLANS BY USING NMCLI

802.1ad is a protocol used for Virtual Local Area Network (VLAN) tagging. It is also known as Q-in-Q tagging. You can use this technology to create multiple VLAN tags within a single Ethernet frame to achieve the following benefits:

- Increased network scalability by creating multiple isolated network segments within a VLAN. This enables you to segment and organize large networks into smaller, manageable units.
- Improved traffic management by isolating and controlling different types of network traffic. This can improve the network performance and reduce network congestion.
- Efficient resource utilization by enabling the creation of smaller, more targeted network segments.
- Enhanced security by isolating network traffic and reducing the risk of unauthorized access to sensitive data.

Prerequisites
- The interface you plan to use as a parent to the virtual VLAN interface supports VLAN tags.
- If you configure the VLAN on top of a bond interface:
  - The ports of the bond are up.
  - The bond is not configured with the fail_over_mac=follow option. A VLAN virtual device cannot change its MAC address to match the parent’s new MAC address. In such a case, the traffic would still be sent with the incorrect source MAC address.
  - The bond is usually not expected to get IP addresses from a DHCP server or IPv6 auto-configuration. Ensure it by setting the ipv4.method/disable and ipv6.method=ignore options while creating the bond. Otherwise, if DHCP or IPv6 auto-configuration fails after some time, the interface might be brought down.
The switch, the host is connected to, is configured to support VLAN tags. For details, see the documentation of your switch.

Procedure

1. Display the physical network devices:

   ```
   # nmcli device status
   DEVICE   TYPE      STATE         CONNECTION
   enp1s0  ethernet  connected      enp1s0
   ...
   ```

2. Create the base VLAN interface. For example, to create a base VLAN interface named `vlan10` that uses `enp1s0` as its parent interface and that tags packets with VLAN ID `10`, enter:

   ```
   # nmcli connection add type vlan con-name vlan10 dev enp1s0 vlan.id 10
   ```

   Note that the VLAN must be within the range from 0 to 4094.

3. By default, the VLAN connection inherits the maximum transmission unit (MTU) from the parent interface. Optionally, set a different MTU value:

   ```
   # nmcli connection modify vlan10 ethernet.mtu 2000
   ```

4. Create the nested VLAN interface on top of the base VLAN interface:

   ```
   # nmcli connection add type vlan con-name vlan10.20 dev enp1s0.10 id 20 vlan.protocol 802.1ad
   ```

   This command creates a new VLAN connection with a name of `vlan10.20` and a VLAN ID of `20` on the parent VLAN connection `vlan10`. The `dev` option specifies the parent network device. In this case it is `enp1s0.10`. The `vlan.protocol` option specifies the VLAN encapsulation protocol. In this case it is `802.1ad` (Q-in-Q).

5. Configure the IPv4 settings of the nested VLAN interface:

   - To use DHCP, no action is required.
   - To set a static IPv4 address, network mask, default gateway, and DNS server to the `vlan10.20` connection, enter:

     ```
     # nmcli connection modify vlan10.20 ipv4.method manual ipv4.addresses 192.0.2.1/24 ipv4.gateway 192.0.2.254 ipv4.dns 192.0.2.200
     ```

6. Configure the IPv6 settings of the nested VLAN interface:

   - To use stateless address autoconfiguration (SLAAC), no action is required.
   - To set a static IPv4 address, network mask, default gateway, and DNS server to the `vlan10` connection, enter:

     ```
     # nmcli connection modify vlan10 ipv4.addresses '192.0.2.1/24' ipv4.gateway '192.0.2.254' ipv4.dns '192.0.2.253' ipv4.method manual
     ```
7. Activate the profile:

```
# nmcli connection up vlan10.20
```

Verification

1. Verify the configuration of the nested VLAN interface:

```
# ip -d addr show enp1s0.10.20
10: enp1s0.10.20@enp1s0.10: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default qlen 1000
   link/ether 52:54:00:d2:74:3e brd ff:ff:ff:ff:ff:ff promiscuity 0 minmtu 0 maxmtu 65535
   vlan protocol 802.1ad id 20 <REORDER_HDR> numtxqueues 1 numrxqueues 1
   gso_max_size 65536 gso_max_segs 65535 tso_max_size 65536 tso_max_segs 65535
gro_max_size 65536
   inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute enp1s0.10.20
      valid_lft forever preferred_lft forever
   inet6 2001:db8:1::1/32 scope global noprefixroute
      valid_lft forever preferred_lft forever
   inet6 fe80::ce3b:84c5:9ef8:d0e6/64 scope link noprefixroute
      valid_lft forever preferred_lft forever
```

Additional resources

- `nm-settings(5)` man page

5.3. CONFIGURING VLAN TAGGING BY USING THE RHEL WEB CONSOLE

Use the RHEL web console to configure VLAN tagging if you prefer to manage network settings using a web browser-based interface.

Prerequisites

- The interface you plan to use as a parent to the virtual VLAN interface supports VLAN tags.
- If you configure the VLAN on top of a bond interface:
  - The ports of the bond are up.
  - The bond is not configured with the `fail_over_mac=follow` option. A VLAN virtual device cannot change its MAC address to match the parent’s new MAC address. In such a case, the traffic would still be sent with the incorrect source MAC address.
  - The bond is usually not expected to get IP addresses from a DHCP server or IPv6 auto-configuration. Ensure it by disabling the IPv4 and IPv6 protocol creating the bond. Otherwise, if DHCP or IPv6 auto-configuration fails after some time, the interface might be brought down.
- The switch, the host is connected to, is configured to support VLAN tags. For details, see the documentation of your switch.

Procedure
1. Select the **Networking** tab in the navigation on the left side of the screen.

2. Click **Add VLAN** in the **Interfaces** section.

3. Select the parent device.

4. Enter the VLAN ID.

5. Enter the name of the VLAN device or keep the automatically-generated name.

6. Click **Apply**.

7. By default, the VLAN device uses a dynamic IP address. If you want to set a static IP address:
   a. Click the name of the VLAN device in the **Interfaces** section.
   b. Click **Edit** next to the protocol you want to configure.
   c. Select **Manual** next to **Addresses**, and enter the IP address, prefix, and default gateway.
   d. In the **DNS** section, click the + button, and enter the IP address of the DNS server. Repeat this step to set multiple DNS servers.
   e. In the **DNS search domains** section, click the + button, and enter the search domain.
   f. If the interface requires static routes, configure them in the **Routes** section.
g. Click **Apply**

**Verification**

- Select the **Networking** tab in the navigation on the left side of the screen, and check if there is incoming and outgoing traffic on the interface:

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Add bond</th>
<th>Add team</th>
<th>Add bridge</th>
<th>Add VLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>IP address</td>
<td>Sending</td>
<td>Receiving</td>
<td></td>
</tr>
<tr>
<td>enp1s0.10</td>
<td>192.0.2/24</td>
<td>1.11 Mbps</td>
<td>612 Mbps</td>
<td></td>
</tr>
</tbody>
</table>

**5.4. CONFIGURING VLAN TAGGING BY USING NMTUI**

The **nmtui** application provides a text-based user interface for NetworkManager. You can use **nmtui** to configure VLAN tagging on a host without a graphical interface.
NOTE

In nmtui:

- Navigate by using the cursor keys.
- Press a button by selecting it and hitting Enter.
- Select and deselect checkboxes by using Space.

Prerequisites

- The interface you plan to use as a parent to the virtual VLAN interface supports VLAN tags.
- If you configure the VLAN on top of a bond interface:
  - The ports of the bond are up.
  - The bond is not configured with the fail_over_mac=follow option. A VLAN virtual device cannot change its MAC address to match the parent’s new MAC address. In such a case, the traffic would still be sent with the then incorrect source MAC address.
  - The bond is usually not expected to get IP addresses from a DHCP server or IPv6 auto-configuration. Ensure it by setting the ipv4.method=disable and ipv6.method=ignore options while creating the bond. Otherwise, if DHCP or IPv6 auto-configuration fails after some time, the interface might be brought down.
- The switch the host is connected to is configured to support VLAN tags. For details, see the documentation of your switch.

Procedure

1. If you do not know the network device name on which you want configure VLAN tagging, display the available devices:

   ```
   # nmcli device status
   DEVICE   TYPE   STATE     CONNECTION
   enp1s0   ethernet unavailable   --
   ...  
   ```

2. Start nmtui:

   ```
   # nmtui
   ```

3. Select Edit a connection, and press Enter.

4. Press the Add button.

5. Select VLAN from the list of network types, and press Enter.

6. Optional: Enter a name for the NetworkManager profile to be created. On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

7. Enter the VLAN device name to be created into the Device field.
8. Enter the name of the device on which you want to configure VLAN tagging into the Parent field.

9. Enter the VLAN ID. The ID must be within the range from 0 to 4094.

10. Depending on your environment, configure the IP address settings in the IPv4 configuration and IPv6 configuration areas accordingly. For this, press the button next to these areas, and select:

- **Disabled**, if this VLAN device does not require an IP address or you want to use it as a port of other devices.

- **Automatic**, if a DHCP server or stateless address autoconfiguration (SLAAC) dynamically assigns an IP address to the VLAN device.

- **Manual**, if the network requires static IP address settings. In this case, you must fill further fields:
  
  i. Press the Show button next to the protocol you want to configure to display additional fields.

  ii. Press the Add button next to Addresses, and enter the IP address and the subnet mask in Classless Inter-Domain Routing (CIDR) format.

  If you do not specify a subnet mask, NetworkManager sets a /32 subnet mask for IPv4 addresses and /64 for IPv6 addresses.

  iii. Enter the address of the default gateway.

  iv. Press the Add button next to DNS servers, and enter the DNS server address.

  v. Press the Add button next to Search domains, and enter the DNS search domain.

Figure 5.1. Example of a VLAN connection with static IP address settings
11. Press the **OK** button to create and automatically activate the new connection.

12. Press the **Back** button to return to the main menu.

13. Select **Quit**, and press **Enter** to close the **nmtui** application.

**Verification**

- Verify the settings:
5.5. CONFIGURING VLAN TAGGING BY USING NM-CONNECTION-EDITOR

You can configure Virtual Local Area Network (VLAN) tagging in a graphical interface using the `nm-connection-editor` application.

Prerequisites

- The interface you plan to use as a parent to the virtual VLAN interface supports VLAN tags.
- If you configure the VLAN on top of a bond interface:
  - The ports of the bond are up.
  - The bond is not configured with the `fail_over_mac=follow` option. A VLAN virtual device cannot change its MAC address to match the parent’s new MAC address. In such a case, the traffic would still be sent with the incorrect source MAC address.
- The switch, the host is connected to, is configured to support VLAN tags. For details, see the documentation of your switch.

Procedure

1. Open a terminal, and enter `nm-connection-editor`:

   ```
   $ nm-connection-editor
   ```

2. Click the + button to add a new connection.

3. Select the VLAN connection type, and click Create.

4. On the VLAN tab:
   a. Select the parent interface.
   b. Select the VLAN id. Note that the VLAN must be within the range from 0 to 4094.
   c. By default, the VLAN connection inherits the maximum transmission unit (MTU) from the parent interface. Optionally, set a different MTU value.
   d. Optionally, set the name of the VLAN interface and further VLAN-specific options.
5. Configure the IP address settings on both the IPv4 Settings and IPv6 Settings tabs:
   - To use this bridge device as a port of other devices, set the Method field to Disabled.
   - To use DHCP, leave the Method field at its default, Automatic (DHCP).
   - To use static IP settings, set the Method field to Manual and fill the fields accordingly:

6. Click Save.


Verification

1. Verify the settings:

```
# ip -d addr show vlan10
4: vlan10@enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue state UP group default qlen 1000
    link/ether 52:54:00:72:2F:6E brd ff:ff:ff:ff:ff:ff promiscuity 0
    vlan protocol 802.1Q id 10 <REORDER_HDR> numtxqueues 1 numrxqueues 1
gso_max_size 65536 gso_max_segs 65535
    inet 192.0.2.21/24 brd 192.0.2.254 scope global noprefixroute vlan10
        valid_lft forever preferred_lft forever
    inet 2001:db8:1::1/32 scope global noprefixroute
        valid_lft forever preferred_lft forever
    inet6 fe80::8dd7:9030:6f8e:89e6/64 scope link noprefixroute
        valid_lft forever preferred_lft forever
```

Red Hat Enterprise Linux 9 Configuring and managing networking
5.6. CONFIGURING VLAN TAGGING BY USING NMSTATECTL

Use the `nmstatectl` utility to configure Virtual Local Area Network VLAN through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.

Depending on your environment, adjust the YAML file accordingly. For example, to use different devices than Ethernet adapters in the VLAN, adapt the `base-iface` attribute and `type` attributes of the ports you use in the VLAN.

Prerequisites

- To use Ethernet devices as ports in the VLAN, the physical or virtual Ethernet devices must be installed on the server.
- The `nmstate` package is installed.

Procedure

1. Create a YAML file, for example `~/create-vlan.yml`, with the following content:

```yaml
---
interfaces:
- name: vlan10
type: vlan
state: up
ipv4:
  enabled: true
  address:
    - ip: 192.0.2.1
      prefix-length: 24
dhcp: false
ipv6:
  enabled: true
  address:
    - ip: 2001:db8:1::1
      prefix-length: 64
autoconf: false
dhcp: false
vlan:
  base-iface: enp1s0
  id: 10
- name: enp1s0
type: ethernet
state: up
routes:
  config:
    - destination: 0.0.0.0/0
```
These settings define a VLAN with ID 10 that uses the **enp1s0** device. As the child device, the VLAN connection has the following settings:

- A static IPv4 address - **192.0.2.1** with the **/24** subnet mask
- A static IPv6 address - **2001:db8:1::1** with the **/64** subnet mask
- An IPv4 default gateway - **192.0.2.254**
- An IPv6 default gateway - **2001:db8:1::fffe**
- An IPv4 DNS server - **192.0.2.200**
- An IPv6 DNS server - **2001:db8:1::ffbb**
- A DNS search domain - **example.com**

2. Apply the settings to the system:

   ```
   # nmstatectl apply ~/create-vlan.yml
   ```

**Verification**

1. Display the status of the devices and connections:

   ```
   # nmcli device status
   DEVICE   TYPE      STATE      CONNECTION
   vlan10   vlan      connected  vlan10
   ```

2. Display all settings of the connection profile:

   ```
   # nmcli connection show vlan10
   connection.id:   vlan10
   connection.uuid: 1722970f-788e-4f81-bd7d-a86bf21c9df5
   connection.stable-id: --
   connection.type:   vlan
   connection.interface-name: vlan10
   ```

3. Display the connection settings in YAML format:
# nmstatectl show vlan0

Additional resources

- nmstatectl(8) man page
- /usr/share/doc/nmstate/examples/ directory

5.7. CONFIGURING VLAN TAGGING BY USING THE NETWORK RHEL SYSTEM ROLE

You can use the network RHEL system role to configure VLAN tagging. This example adds an Ethernet connection and a VLAN with ID 10 on top of this Ethernet connection. As the child device, the VLAN connection contains the IP, default gateway, and DNS configurations.

Depending on your environment, adjust the play accordingly. For example:

- To use the VLAN as a port in other connections, such as a bond, omit the ip attribute, and set the IP configuration in the child configuration.
- To use team, bridge, or bond devices in the VLAN, adapt the interface_name and type attributes of the ports you use in the VLAN.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has sudo permissions on them.

Procedure

1. Create a playbook file, for example ~/playbook.yml, with the following content:

```yaml
---
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure a VLAN that uses an Ethernet connection
      ansible.builtin.include_role:
        name: rhel-system-roles.network
      vars:
        network_connections:
          # Add an Ethernet profile for the underlying device of the VLAN
          - name: enp1s0
            type: ethernet
            interface_name: enp1s0
            autoconnect: yes
            state: up
            ip:
              dhcp4: no
              auto6: no
```

```
These settings define a VLAN to operate on top of the `enp1s0` device. The VLAN interface has the following settings:

- A static IPv4 address - 192.0.2.1 with a /24 subnet mask
- A static IPv6 address - 2001:db8:1::1 with a /64 subnet mask
- An IPv4 default gateway - 192.0.2.254
- An IPv6 default gateway - 2001:db8:1::ffe
- An IPv4 DNS server - 192.0.2.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com
- VLAN ID - 10
  The `parent` attribute in the VLAN profile configures the VLAN to operate on top of the `enp1s0` device. As the child device, the VLAN connection contains the IP, default gateway, and DNS configurations.

2. Validate the playbook syntax:

```bash
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```bash
$ ansible-playbook ~/playbook.yml
```

Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory
5.8. ADDITIONAL RESOURCES

- VLANs for sysadmins: The basics
CHAPTER 6. CONFIGURING A NETWORK BRIDGE

A network bridge is a link-layer device which forwards traffic between networks based on a table of MAC addresses. The bridge builds the MAC addresses table by listening to network traffic and thereby learning what hosts are connected to each network. For example, you can use a software bridge on a Red Hat Enterprise Linux host to emulate a hardware bridge or in virtualization environments, to integrate virtual machines (VM) to the same network as the host.

A bridge requires a network device in each network the bridge should connect. When you configure a bridge, the bridge is called controller and the devices it uses ports.

You can create bridges on different types of devices, such as:

- Physical and virtual Ethernet devices
- Network bonds
- Network teams
- VLAN devices

Due to the IEEE 802.11 standard which specifies the use of 3-address frames in Wi-Fi for the efficient use of airtime, you cannot configure a bridge over Wi-Fi networks operating in Ad-Hoc or Infrastructure modes.

6.1. CONFIGURING A NETWORK BRIDGE BY USING NMCLI

To configure a network bridge on the command line, use the nmcli utility.

Prerequisites

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bridge, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bond, or VLAN devices as ports of the bridge, you can either create these devices while you create the bridge or you can create them in advance as described in:
  - Configuring a network team by using nmcli
  - Configuring a network bond by using nmcli
  - Configuring VLAN tagging by using nmcli

Procedure

1. Create a bridge interface:

   ```
   # nmcli connection add type bridge con-name bridge0 ifname bridge0
   ```

   This command creates a bridge named bridge0, enter:

2. Display the network interfaces, and note the names of the interfaces you want to add to the bridge:

   ```
   #
   ```
# nmcli device status

<table>
<thead>
<tr>
<th>DEVICE</th>
<th>TYPE</th>
<th>STATE</th>
<th>CONNECTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>enp7s0</td>
<td>ethernet</td>
<td>disconnected</td>
<td>--</td>
</tr>
<tr>
<td>enp8s0</td>
<td>ethernet</td>
<td>disconnected</td>
<td>--</td>
</tr>
<tr>
<td>bond0</td>
<td>bond</td>
<td>connected</td>
<td>bond0</td>
</tr>
<tr>
<td>bond1</td>
<td>bond</td>
<td>connected</td>
<td>bond1</td>
</tr>
</tbody>
</table>

In this example:

- *enp7s0* and *enp8s0* are not configured. To use these devices as ports, add connection profiles in the next step.

- *bond0* and *bond1* have existing connection profiles. To use these devices as ports, modify their profiles in the next step.

3. Assign the interfaces to the bridge.

   a. If the interfaces you want to assign to the bridge are not configured, create new connection profiles for them:

   ```
   # nmcli connection add type ethernet slave-type bridge con-name bridge0-port1
   ifname enp7s0 master bridge0
   # nmcli connection add type ethernet slave-type bridge con-name bridge0-port2
   ifname enp8s0 master bridge0
   ```

   These commands create profiles for *enp7s0* and *enp8s0*, and add them to the *bridge0* connection.

   b. If you want to assign an existing connection profile to the bridge:

      i. Set the *master* parameter of these connections to *bridge0*:

      ```
      # nmcli connection modify bond0 master bridge0
      # nmcli connection modify bond1 master bridge0
      ```

      These commands assign the existing connection profiles named *bond0* and *bond1* to the *bridge0* connection.

      ii. Reactivate the connections:

      ```
      # nmcli connection up bond0
      # nmcli connection up bond1
      ```

4. Configure the IPv4 settings:

   - To use this bridge device as a port of other devices, enter:

     ```
     # nmcli connection modify bridge0 ipv4.method disabled
     ```

   - To use DHCP, no action is required.

   - To set a static IPv4 address, network mask, default gateway, and DNS server to the *bridge0* connection, enter:
5. Configure the IPv6 settings:
   - To use this bridge device as a port of other devices, enter:
     ```bash
     # nmcli connection modify bridge0 ipv6.method disabled
     ``
   - To use stateless address autoconfiguration (SLAAC), no action is required.
   - To set a static IPv6 address, network mask, default gateway, and DNS server to the bridge0 connection, enter:
     ```bash
     # nmcli connection modify bridge0 ipv6.addresses '2001:db8:1::1/64' ipv6.gateway '2001:db8:1::fffe' ipv6.dns '2001:db8:1::fffd' ipv6.dns-search 'example.com'
     ipv6.method manual
     ```

6. Optional: Configure further properties of the bridge. For example, to set the Spanning Tree Protocol (STP) priority of bridge0 to 16384, enter:
   ```bash
   # nmcli connection modify bridge0 bridge.priority '16384'
   ``
   By default, STP is enabled.

7. Activate the connection:
   ```bash
   # nmcli connection up bridge0
   ```

8. Verify that the ports are connected, and the CONNECTION column displays the port’s connection name:
   ```bash
   # nmcli device
   DEVICE  TYPE      STATE      CONNECTION
   ...      enp7s0  ethernet  connected  bridge0-port1
   ...      enp8s0  ethernet  connected  bridge0-port2
   ``
   When you activate any port of the connection, NetworkManager also activates the bridge, but not the other ports of it. You can configure that Red Hat Enterprise Linux enables all ports automatically when the bridge is enabled:
   a. Enable the connection.autoconnect-slaves parameter of the bridge connection:
      ```bash
      # nmcli connection modify bridge0 connection.autoconnect-slaves 1
      ```
   b. Reactivate the bridge:
      ```bash
      # nmcli connection up bridge0
      ```
Use the `ip` utility to display the link status of Ethernet devices that are ports of a specific bridge:

```
# ip link show master bridge0
3: enp7s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master
    bridge0 state UP mode DEFAULT group default qlen 1000
    link/ether 52:54:00:62:61:0e brd ff:ff:ff:ff:ff:ff
4: enp8s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master
    bridge0 state UP mode DEFAULT group default qlen 1000
    link/ether 52:54:00:9e:f1:ce brd ff:ff:ff:ff:ff:ff
```

Use the `bridge` utility to display the status of Ethernet devices that are ports of any bridge device:

```
# bridge link show
3: enp7s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge0 state
    forwarding priority 32 cost 100
4: enp8s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge0 state
    listening priority 32 cost 100
5: enp9s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge1 state
    forwarding priority 32 cost 100
6: enp11s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge1 state
    blocking priority 32 cost 100
...```

To display the status for a specific Ethernet device, use the `bridge link show dev ethernet_device_name` command.

Additional resources

- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- `nm-settings(5)` man page
- `bridge(8)` man page
- NetworkManager duplicates a connection after restart of NetworkManager service
- How to configure a bridge with VLAN information?

### 6.2. Configuring a Network Bridge by Using the RHEL Web Console

Use the RHEL web console to configure a network bridge if you prefer to manage network settings using a web browser-based interface.

**Prerequisites**

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bridge, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bond, or VLAN devices as ports of the bridge, you can either create these devices while you create the bridge or you can create them in advance as described in:
Configuring a network team using the RHEL web console

Configuring a network bond by using the RHEL web console

Configuring VLAN tagging by using the RHEL web console

Procedure

1. Select the Networking tab in the navigation on the left side of the screen.

2. Click Add bridge in the Interfaces section.

3. Enter the name of the bridge device you want to create.

4. Select the interfaces that should be ports of the bridge.

5. Optional: Enable the Spanning tree protocol (STP) feature to avoid bridge loops and broadcast radiation.

6. Click Apply.

7. By default, the bridge uses a dynamic IP address. If you want to set a static IP address:
   a. Click the name of the bridge in the Interfaces section.
   b. Click Edit next to the protocol you want to configure.
   c. Select Manual next to Addresses, and enter the IP address, prefix, and default gateway.
   d. In the DNS section, click the + button, and enter the IP address of the DNS server. Repeat this step to set multiple DNS servers.
   e. In the DNS search domains section, click the + button, and enter the search domain.
   f. If the interface requires static routes, configure them in the Routes section.
6.3. CONFIGURING A NETWORK BRIDGE BY USING NMTUI

The `nmtui` application provides a text-based user interface for NetworkManager. You can use `nmtui` to configure a network bridge on a host without a graphical interface.

**Verification**

1. Select the Networking tab in the navigation on the left side of the screen, and check if there is incoming and outgoing traffic on the interface:

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Add bond</th>
<th>Add team</th>
<th>Add bridge</th>
<th>Add VLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>IP address</td>
<td>Sending</td>
<td>Receiving</td>
<td></td>
</tr>
<tr>
<td>bridge0</td>
<td>192.0.2.1/24</td>
<td>1.11 Mbps</td>
<td>61.2 Mbps</td>
<td></td>
</tr>
</tbody>
</table>
NOTE

In `nmtui`:

- Navigate by using the cursor keys.
- Press a button by selecting it and hitting `Enter`.
- Select and deselect checkboxes by using `Space`.

**Prerequisites**

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bridge, the physical or virtual Ethernet devices must be installed on the server.

**Procedure**

1. If you do not know the network device names on which you want configure a network bridge, display the available devices:

   ```bash
   # nmcli device status
   DEVICE    TYPE      STATE                   CONNECTION
   enp7s0    ethernet  unavailable             --
   enp8s0    ethernet  unavailable             --
   ...
   #
   ``

2. Start `nmtui`:

   ```bash
   # nmtui
   #
   ``

3. Select **Edit a connection**, and press **Enter**.

4. Press the **Add** button.

5. Select **Bridge** from the list of network types, and press **Enter**.

6. Optional: Enter a name for the NetworkManager profile to be created. On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

7. Enter the bridge device name to be created into the **Device** field.

8. Add ports to the bridge to be created:
   
   a. Press the **Add** button next to the **Slaves** list.
   
   b. Select the type of the interface you want to add as port to the bridge, for example, **Ethernet**.
   
   c. Optional: Enter a name for the NetworkManager profile to be created for this bridge port.
   
   d. Enter the port’s device name into the **Device** field.
   
   e. Press the **OK** button to return to the window with the bridge settings.
f. Repeat these steps to add more ports to the bridge.

9. Depending on your environment, configure the IP address settings in the **IPv4 configuration** and **IPv6 configuration** areas accordingly. For this, press the button next to these areas, and select:

- **Disabled**, if the bridge does not require an IP address.
- **Automatic**, if a DHCP server or stateless address autoconfiguration (SLAAC) dynamically assigns an IP address to the bridge.
- **Manual**, if the network requires static IP address settings. In this case, you must fill further fields:
  
  i. Press the **Show** button next to the protocol you want to configure to display additional fields.

  ii. Press the **Add** button next to **Addresses**, and enter the IP address and the subnet mask in Classless Inter-Domain Routing (CIDR) format.
   - If you do not specify a subnet mask, NetworkManager sets a /32 subnet mask for IPv4 addresses and /64 for IPv6 addresses.

  iii. Enter the address of the default gateway.

  iv. Press the **Add** button next to **DNS servers**, and enter the DNS server address.

  v. Press the **Add** button next to **Search domains**, and enter the DNS search domain.
10. Press the **OK** button to create and automatically activate the new connection.

11. Press the **Back** button to return to the main menu.

12. Select **Quit**, and press **Enter** to close the **nmtui** application.

**Verification**

1. Use the **ip** utility to display the link status of Ethernet devices that are ports of a specific bridge:

   ```bash
   # ip link show master bridge0
   3: enp7s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master bridge0 state UP mode DEFAULT group default qlen 1000
   link/ether 52:54:00:62:61:0e brd ff:ff:ff:ff:ff:ff
   4: enp8s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master bridge0 state UP mode DEFAULT group default qlen 1000
   link/ether 52:54:00:9e:f1:ce brd ff:ff:ff:ff:ff:ff
   ```

2. Use the **bridge** utility to display the status of Ethernet devices that are ports of any bridge device:

   ```bash
   # bridge link show
   3: enp7s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge0 state forwarding priority 32 cost 100
   ```
To display the status for a specific Ethernet device, use the `bridge link show dev ethernet_device_name` command.

### 6.4. Configuring a Network Bridge by Using NM-Connection-Editor

If you use Red Hat Enterprise Linux with a graphical interface, you can configure network bridges using the `nm-connection-editor` application.

Note that `nm-connection-editor` can add only new ports to a bridge. To use an existing connection profile as a port, create the bridge using the `nmcli` utility as described in Configuring a network bridge by using `nmcli`.

**Prerequisites**

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports of the bridge, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bond, or VLAN devices as ports of the bridge, ensure that these devices are not already configured.

**Procedure**

1. Open a terminal, and enter `nm-connection-editor`:

   ```
   $ nm-connection-editor
   ```

2. Click the `+` button to add a new connection.

3. Select the **Bridge** connection type, and click **Create**.

4. On the **Bridge** tab:
   a. Optional: Set the name of the bridge interface in the **Interface name** field.
   b. Click the **Add** button to create a new connection profile for a network interface and adding the profile as a port to the bridge.
      i. Select the connection type of the interface. For example, select **Ethernet** for a wired connection.
      ii. Optionally, set a connection name for the port device.
      iii. If you create a connection profile for an Ethernet device, open the **Ethernet** tab, and select in the **Device** field the network interface you want to add as a port to the bridge. If you selected a different device type, configure it accordingly.
      iv. Click **Save**.
c. Repeat the previous step for each interface you want to add to the bridge.

![Editing Bridge connection 1](image)

5. Optional: Configure further bridge settings, such as Spanning Tree Protocol (STP) options.

6. Configure the IP address settings on both the IPv4 Settings and IPv6 Settings tabs:
   - To use this bridge device as a port of other devices, set the Method field to Disabled.
   - To use DHCP, leave the Method field at its default, Automatic (DHCP).
   - To use static IP settings, set the Method field to Manual and fill the fields accordingly:

![Editing Bridge connection 1](image)

7. Click Save.


**Verification**

- Use the `ip` utility to display the link status of Ethernet devices that are ports of a specific bridge.

```
# ip link show master bridge0
3: enp7s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master bridge0 state UP mode DEFAULT group default qlen 1000
    link/ether 52:54:00:62:61:0e brd ff:ff:ff:ff:ff:ff
4: enp8s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master bridge0 state UP mode DEFAULT group default qlen 1000
    link/ether 52:54:00:9e:f1:ce brd ff:ff:ff:ff:ff:ff
```
Use the `bridge` utility to display the status of Ethernet devices that are ports in any bridge device:

```
# bridge link show
3: enp7s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge0 state forwarding priority 32 cost 100
4: enp8s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge0 state listening priority 32 cost 100
5: enp9s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge1 state forwarding priority 32 cost 100
6: enp11s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 master bridge1 state blocking priority 32 cost 100
...```

To display the status for a specific Ethernet device, use the `bridge link show dev ethernet_device_name` command.

Additional resources

- Configuring a network bond by using `nm-connection-editor`
- Configuring a network team by using `nm-connection-editor`
- Configuring VLAN tagging by using `nm-connection-editor`
- Configuring NetworkManager to avoid using a specific profile to provide a default gateway
- How to configure a bridge with VLAN information?

6.5. CONFIGURING A NETWORK BRIDGE BY USING NMSTATECTL

Use the `nmstatectl` utility to configure a network bridge through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.

Depending on your environment, adjust the YAML file accordingly. For example, to use different devices than Ethernet adapters in the bridge, adapt the `base-iface` attribute and `type` attributes of the ports you use in the bridge.

Prerequisites

- Two or more physical or virtual network devices are installed on the server.
- To use Ethernet devices as ports in the bridge, the physical or virtual Ethernet devices must be installed on the server.
- To use team, bond, or VLAN devices as ports in the bridge, set the interface name in the `port` list, and define the corresponding interfaces.
- The `nmstate` package is installed.

Procedure

1. Create a YAML file, for example `~/create-bridge.yml`, with the following content:
These settings define a network bridge with the following settings:

- Network interfaces in the bridge: `enp1s0` and `enp7s0`
- Spanning Tree Protocol (STP): Enabled
- Static IPv4 address: `192.0.2.1` with the `/24` subnet mask
• Static IPv6 address: 2001:db8:1::1 with the /64 subnet mask
• IPv4 default gateway: 192.0.2.254
• IPv6 default gateway: 2001:db8:1::ffe
• IPv4 DNS server: 192.0.2.200
• IPv6 DNS server: 2001:db8:1::ffbb
• DNS search domain: example.com

2. Apply the settings to the system:

```bash
# nmstatectl apply ~/create-bridge.yml
```

Verification

1. Display the status of the devices and connections:

```bash
# nmcli device status
DEVICE TYPE STATE CONNECTION
bridge0 bridge connected bridge0
```

2. Display all settings of the connection profile:

```bash
# nmcli connection show bridge0
connection.id: bridge0
connection.uuid: e2cc9206-75a2-4622-89cf-1252926060a9
connection.stable-id: --
connection.type: bridge
connection.interface-name: bridge0
...
```

3. Display the connection settings in YAML format:

```bash
# nmstatectl show bridge0
```

Additional resources

- `nmstatectl(8)` man page
- `/usr/share/doc/nmstate/examples/` directory
- How to configure a bridge with VLAN information?

## 6.6. CONFIGURING A NETWORK BRIDGE BY USING THE NETWORK RHEL SYSTEM ROLE

You can remotely configure a network bridge by using the `network` RHEL system role.

### Prerequisites
You have prepared the control node and the managed nodes

You are logged in to the control node as a user who can run playbooks on the managed nodes.

The account you use to connect to the managed nodes has **sudo** permissions on them.

Two or more physical or virtual network devices are installed on the server.

**Procedure**

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure a network bridge that uses two Ethernet ports
      ansible.builtin.include_role:
        name: rhel-system-roles.network
      vars:
        network_connections:
          # Define the bridge profile
          - name: bridge0
            type: bridge
            interface_name: bridge0
            ip:
              address:
                - "192.0.2.1/24"
                - "2001:db8:1::1/64"
              gateway4: 192.0.2.254
              gateway6: 2001:db8:1::ffe
            dns:
              - 192.0.2.200
              - 2001:db8:1::ffbb
            dns_search:
              - example.com
            state: up

          # Add an Ethernet profile to the bridge
          - name: bridge0-port1
            interface_name: enp7s0
            type: ethernet
            controller: bridge0
            port_type: bridge
            state: up

          # Add a second Ethernet profile to the bridge
          - name: bridge0-port2
            interface_name: enp8s0
            type: ethernet
            controller: bridge0
            port_type: bridge
            state: up
```

These settings define a network bridge with the following settings:
• A static IPv4 address - 192.0.2.1 with a /24 subnet mask
• A static IPv6 address - 2001:db8:1::1 with a /64 subnet mask
• An IPv4 default gateway - 192.0.2.254
• An IPv6 default gateway - 2001:db8:1::fffe
• An IPv4 DNS server - 192.0.2.200
• An IPv6 DNS server - 2001:db8:1::ffbb
• A DNS search domain - example.com
• Ports of the bridge - enp7s0 and enp8s0

NOTE
Set the IP configuration on the bridge and not on the ports of the Linux bridge.

2. Validate the playbook syntax:

   $ ansible-playbook --syntax-check ~/playbook.yml

   Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

   $ ansible-playbook ~/playbook.yml

Additional resources

• /usr/share/ansible/roles/rhel-system-roles.network/README.md file
• /usr/share/doc/rhel-system-roles/network/ directory
CHAPTER 7. SETTING UP AN IPSEC VPN

A virtual private network (VPN) is a way of connecting to a local network over the internet. IPsec provided by Libreswan is the preferred method for creating a VPN. Libreswan is a user-space IPsec implementation for VPN. A VPN enables the communication between your LAN, and another, remote LAN by setting up a tunnel across an intermediate network such as the internet. For security reasons, a VPN tunnel always uses authentication and encryption. For cryptographic operations, Libreswan uses the NSS library.

7.1. CONFIGURING A VPN CONNECTION WITH CONTROL-CENTER

If you use Red Hat Enterprise Linux with a graphical interface, you can configure a VPN connection in the GNOME control-center.

Prerequisites

- The NetworkManager-libreswan-gnome package is installed.

Procedure

1. Press the Super key, type Settings, and press Enter to open the control-center application.
2. Select the Network entry on the left.
3. Click the + icon.
4. Select VPN.
5. Select the Identity menu entry to see the basic configuration options:
   - General
     - Gateway – The name or IP address of the remote VPN gateway.
   - Authentication
   - Type
     - IKEv2 (Certificate) – client is authenticated by certificate. It is more secure (default).
     - IKEv1 (XAUTH) – client is authenticated by user name and password, or a pre-shared key (PSK).

The following configuration settings are available under the Advanced section:
### Figure 7.1. Advanced options of a VPN connection

<table>
<thead>
<tr>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Security</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase1 Algorithms:</td>
</tr>
<tr>
<td>Phase2 Algorithms:</td>
</tr>
<tr>
<td>Disable PFS:</td>
</tr>
<tr>
<td>Phase1 Lifetime:</td>
</tr>
<tr>
<td>Phase2 Lifetime:</td>
</tr>
<tr>
<td>Disable rekeying:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Remote Network:</td>
</tr>
<tr>
<td>narrowing:</td>
</tr>
<tr>
<td>Enable fragmentation: yes</td>
</tr>
<tr>
<td>Enable MOBIKE: no</td>
</tr>
</tbody>
</table>

Apply
WARNING
When configuring an IPsec-based VPN connection using the `gnome-control-center` application, the Advanced dialog displays the configuration, but it does not allow any changes. As a consequence, users cannot change any advanced IPsec options. Use the `nm-connection-editor` or `nmcli` tools instead to perform configuration of the advanced properties.

Identification
- **Domain** – If required, enter the Domain Name.

Security
- **Phase1 Algorithms** – corresponds to the `ike` Libreswan parameter – enter the algorithms to be used to authenticate and set up an encrypted channel.
- **Phase2 Algorithms** – corresponds to the `esp` Libreswan parameter – enter the algorithms to be used for the IPsec negotiations.
  Check the **Disable PFS** field to turn off Perfect Forward Secrecy (PFS) to ensure compatibility with old servers that do not support PFS.
- **Phase1 Lifetime** – corresponds to the `ikelifetime` Libreswan parameter – how long the key used to encrypt the traffic will be valid.
- **Phase2 Lifetime** – corresponds to the `salifetime` Libreswan parameter – how long a particular instance of a connection should last before expiring.
  Note that the encryption key should be changed from time to time for security reasons.
- **Remote network** – corresponds to the `rightsubnet` Libreswan parameter – the destination private remote network that should be reached through the VPN.
  Check the **narrowing** field to enable narrowing. Note that it is only effective in IKEv2 negotiation.
- **Enable fragmentation** – corresponds to the `fragmentation` Libreswan parameter – whether or not to allow IKE fragmentation. Valid values are `yes` (default) or `no`.
- **Enable Mobike** – corresponds to the `mobike` Libreswan parameter – whether to allow Mobility and Multihoming Protocol (MOBIKE, RFC 4555) to enable a connection to migrate its endpoint without needing to restart the connection from scratch. This is used on mobile devices that switch between wired, wireless, or mobile data connections. The values are `no` (default) or `yes`.

6. Select the IPv4 menu entry:
   **IPv4 Method**
   - **Automatic (DHCP)** – Choose this option if the network you are connecting to uses a DHCP server to assign dynamic IP addresses.
   - **Link-Local Only** – Choose this option if the network you are connecting to does not have a DHCP server and you do not want to assign IP addresses manually. Random addresses will be assigned as per RFC 3927 with prefix 169.254/16.
• **Manual** – Choose this option if you want to assign IP addresses manually.

• **Disable** – IPv4 is disabled for this connection.

**DNS**

In the DNS section, when **Automatic** is ON, switch it to **OFF** to enter the IP address of a DNS server you want to use separating the IPs by comma.

**Routes**

Note that in the Routes section, when **Automatic** is ON, routes from DHCP are used, but you can also add additional static routes. When **OFF**, only static routes are used.

• **Address** – Enter the IP address of a remote network or host.

• **Netmask** – The netmask or prefix length of the IP address entered above.

• **Gateway** – The IP address of the gateway leading to the remote network or host entered above.

• **Metric** – A network cost, a preference value to give to this route. Lower values will be preferred over higher values.

**Use this connection only for resources on its network**

Select this check box to prevent the connection from becoming the default route. Selecting this option means that only traffic specifically destined for routes learned automatically over the connection or entered here manually is routed over the connection.

7. To configure IPv6 settings in a VPN connection, select the IPv6 menu entry:

**IPv6 Method**

• **Automatic** – Choose this option to use IPv6 Stateless Address AutoConfiguration (SLAAC) to create an automatic, stateless configuration based on the hardware address and Router Advertisements (RA).

• **Automatic, DHCP only** – Choose this option to not use RA, but request information from DHCPv6 directly to create a stateful configuration.

• **Link-Local Only** – Choose this option if the network you are connecting to does not have a DHCP server and you do not want to assign IP addresses manually. Random addresses will be assigned as per RFC 4862 with prefix FE80::0.

• **Manual** – Choose this option if you want to assign IP addresses manually.

• **Disable** – IPv6 is disabled for this connection.

Note that DNS, Routes, **Use this connection only for resources on its network** are common to IPv4 settings.

8. Once you have finished editing the VPN connection, click the **Add** button to customize the configuration or the **Apply** button to save it for the existing one.

9. Switch the profile to **ON** to active the VPN connection.

**Additional resources**

• nm-settings-libreswan(5)
7.2. CONFIGURING A VPN CONNECTION USING NM-CONNECTION-EDITOR

If you use Red Hat Enterprise Linux with a graphical interface, you can configure a VPN connection in the nm-connection-editor application.

Prerequisites

- The NetworkManager-libreswan-gnome package is installed.
- If you configure an Internet Key Exchange version 2 (IKEv2) connection:
  - The certificate is imported into the IPsec network security services (NSS) database.
  - The nickname of the certificate in the NSS database is known.

Procedure

1. Open a terminal, and enter:
   
   `$ nm-connection-editor`

2. Click the + button to add a new connection.

3. Select the IPsec based VPN connection type, and click Create.

4. On the VPN tab:

   a. Enter the host name or IP address of the VPN gateway into the Gateway field, and select an authentication type. Based on the authentication type, you must enter different additional information:

      - **IKEv2 (Certificate)** authenticates the client by using a certificate, which is more secure. This setting requires the nickname of the certificate in the IPsec NSS database
      
      - **IKEv1 (XAUTH)** authenticates the user by using a user name and password (pre-shared key). This setting requires that you enter the following values:
        - User name
        - Password
        - Group name
        - Secret

   b. If the remote server specifies a local identifier for the IKE exchange, enter the exact string in the Remote ID field. In the remote server runs Libreswan, this value is set in the server’s leftid parameter.
c. Optionally, configure additional settings by clicking the **Advanced** button. You can configure the following settings:

- **Identification**
  - **Domain** – If required, enter the domain name.

- **Security**
  - **Phase1 Algorithms** corresponds to the `ike` Libreswan parameter. Enter the algorithms to be used to authenticate and set up an encrypted channel.
  
  - **Phase2 Algorithms** corresponds to the `esp` Libreswan parameter. Enter the algorithms to be used for the IPsec negotiations. Check the **Disable PFS** field to turn off Perfect Forward Secrecy (PFS) to ensure compatibility with old servers that do not support PFS.

  - **Phase1 Lifetime** corresponds to the `ikelifetime` Libreswan parameter. This parameter defines how long the key used to encrypt the traffic is valid.

  - **Phase2 Lifetime** corresponds to the `salifetime` Libreswan parameter. This parameter defines how long a security association is valid.

- **Connectivity**
Remote network corresponds to the rightsubnet Libreswan parameter and defines the destination private remote network that should be reached through the VPN. Check the narrowing field to enable narrowing. Note that it is only effective in the IKEv2 negotiation.

Enable fragmentation corresponds to the fragmentation Libreswan parameter and defines whether or not to allow IKE fragmentation. Valid values are yes (default) or no.

Enable Mobike corresponds to the mobike Libreswan parameter. The parameter defines whether to allow Mobility and Multihoming Protocol (MOBIKE) (RFC 4555) to enable a connection to migrate its endpoint without needing to restart the connection from scratch. This is used on mobile devices that switch between wired, wireless or mobile data connections. The values are no (default) or yes.

5. On the IPv4 Settings tab, select the IP assignment method and, optionally, set additional static addresses, DNS servers, search domains, and routes.

6. Save the connection.


NOTE

When you add a new connection by clicking the + button, NetworkManager creates a new configuration file for that connection and then opens the same dialog that is used for editing an existing connection. The difference between these dialogs is that an existing connection profile has a Details menu entry.

Additional resources

- nm-settings-libreswan(5) man page
7.3. CONFIGURING AUTOMATIC DETECTION AND USAGE OF ESP HARDWARE OFFLOAD TO ACCELERATE AN IPSEC CONNECTION

Offloading Encapsulating Security Payload (ESP) to the hardware accelerates IPsec connections over Ethernet. By default, Libreswan detects if hardware supports this feature and, as a result, enables ESP hardware offload. In case that the feature was disabled or explicitly enabled, you can switch back to automatic detection.

Prerequisites

- The network card supports ESP hardware offload.
- The network driver supports ESP hardware offload.
- The IPsec connection is configured and works.

Procedure

1. Edit the Libreswan configuration file in the `/etc/ipsec.d/` directory of the connection that should use automatic detection of ESP hardware offload support.

2. Ensure the `nic-offload` parameter is not set in the connection’s settings.

3. If you removed `nic-offload`, restart the `ipsec` service:

   ```bash
   # systemctl restart ipsec
   ```

Verification

If the network card supports ESP hardware offload support, following these steps to verify the result:

1. Display the `tx_ipsec` and `rx_ipsec` counters of the Ethernet device the IPsec connection uses:

   ```bash
   # ethtool -S enp1s0 | egrep "_ipsec"
   tx_ipsec: 10
   rx_ipsec: 10
   ```

2. Send traffic through the IPsec tunnel. For example, ping a remote IP address:

   ```bash
   # ping -c 5 remote_ip_address
   ```

3. Display the `tx_ipsec` and `rx_ipsec` counters of the Ethernet device again:

   ```bash
   # ethtool -S enp1s0 | egrep "_ipsec"
   tx_ipsec: 15
   rx_ipsec: 15
   ```

   If the counter values have increased, ESP hardware offload works.

Additional resources

- Configuring a VPN with IPsec
7.4. CONFIGURING ESP HARDWARE OFFLOAD ON A BOND TO ACCELERATE AN IPSEC CONNECTION

Offloading Encapsulating Security Payload (ESP) to the hardware accelerates IPsec connections. If you use a network bond for fail-over reasons, the requirements and the procedure to configure ESP hardware offload are different from those using a regular Ethernet device. For example, in this scenario, you enable the offload support on the bond, and the kernel applies the settings to the ports of the bond.

Prerequisites

- All network cards in the bond support ESP hardware offload.
- The network driver supports ESP hardware offload on a bond device. In RHEL, only the ixgbe driver supports this feature.
- The bond is configured and works.
- The bond uses the active-backup mode. The bonding driver does not support any other modes for this feature.
- The IPsec connection is configured and works.

Procedure

1. Enable ESP hardware offload support on the network bond:
   
   ```
   # nmcli connection modify bond0 ethtool.feature-esp-hw-offload on
   ```
   
   This command enables ESP hardware offload support on the `bond0` connection.

2. Reactivate the `bond0` connection:
   
   ```
   # nmcli connection up bond0
   ```

3. Edit the Libreswan configuration file in the `/etc/ipsec.d/` directory of the connection that should use ESP hardware offload, and append the `nic-offload=yes` statement to the connection entry:
   
   ```
   conn example
   ...
   nic-offload=yes
   ```

4. Restart the `ipsec` service:
   
   ```
   # systemctl restart ipsec
   ```

Verification

1. Display the active port of the bond:
   
   ```
   # grep "Currently Active Slave" /proc/net/bonding/bond0
   Currently Active Slave: enp1s0
   ```

2. Display the `tx_ipsec` and `rx_ipsec` counters of the active port:
   
   ```
   tx_ipsec
   ```
3. Send traffic through the IPsec tunnel. For example, ping a remote IP address:

```bash
# ping -c 5 remote_ip_address
```

4. Display the `tx_ipsec` and `rx_ipsec` counters of the active port again:

```bash
# ethtool -S enp1s0 | egrep "_ipsec"
  tx_ipsec: 15
  rx_ipsec: 15
```

If the counter values have increased, ESP hardware offload works.

Additional resources

- Configuring network bonding
- Configuring a VPN with IPsec section in the Securing networks document
CHAPTER 8. SETTING UP A WIREGUARD VPN

WireGuard is a high-performance VPN solution that runs in the Linux kernel. It uses modern cryptography and is easier to configure than many other VPN solutions. Additionally, WireGuard’s small codebase reduces the surface for attacks and, therefore, improves security. For authentication and encryption, WireGuard uses keys similar to SSH.

IMPORTANT

WireGuard is provided as a Technology Preview only. Technology Preview features are not supported with Red Hat production Service Level Agreements (SLAs), might not be functionally complete, and Red Hat does not recommend using them for production. These previews provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.

See Technology Preview Features Support Scope on the Red Hat Customer Portal for information about the support scope for Technology Preview features.

Note that all hosts that participate in a WireGuard VPN are peers. This documentation uses the terms client to describe hosts that establish a connection and server to describe the host with the fixed hostname or IP address that the clients connect to and optionally route all traffic through this server.

To set up a WireGuard VPN, you must complete the following steps. You can perform most steps by using different options:

1. Create public and private keys for every host in the VPN.
2. Configure the WireGuard server by using nmcli, nmtui, nm-connection-editor, or the wg-quick service.
3. Configure firewalld on the WireGuard server by using the command line or graphical interface.
4. Configure the WireGuard client by using nmcli, nm-connection-editor, or the wg-quick service.

WireGuard operates on the network layer (layer 3). Therefore, you cannot use DHCP and must assign static IP addresses or IPv6 link-local addresses to the tunnel devices on both the server and clients.

IMPORTANT

You can use WireGuard only if the Federal Information Processing Standard (FIPS) mode in RHEL is disabled.

8.1. PROTOCOLS AND PRIMITIVES USED BY WIREGUARD

WireGuard uses the following protocols and primitives:

- ChaCha20 for symmetric encryption, authenticated with Poly1305, using Authenticated Encryption with Associated Data (AEAD) construction as described in RFC7539
- Curve25519 for Elliptic-curve Diffie–Hellman (ECDH) key exchange
- BLAKE2s for hashing and keyed hashing, as described in RFC7693
- SipHash24 for hash table keys
• HKDF for key derivation, as described in RFC5869

8.2. HOW WIREGUARD USES TUNNEL IP ADDRESSES, PUBLIC KEYS, AND REMOTE ENDPOINTS

When WireGuard sends a network packet to a peer:

1. WireGuard reads the destination IP from the packet and compares it to the list of allowed IP addresses in the local configuration. If the peer is not found, WireGuard drops the packet.

2. If the peer is valid, WireGuard encrypts the packet using the peer’s public key.

3. The sending host looks up the most recent Internet IP address of the host and sends the encrypted packet to it.

When WireGuard receives a packet:

1. WireGuard decrypts the packet using the private key of the remote host.

2. WireGuard reads the internal source address from the packet and looks up whether the IP is configured in the list of allowed IP addresses in the settings for the peer on the local host. If the source IP is on the allowlist, WireGuard accepts the packet. If the IP address is not on the list, WireGuard drops the packet.

The association of public keys and allowed IP addresses is called **Cryptokey Routing Table**. This means that the list of IP addresses behaves similar to a routing table when sending packets, and as a kind of access control list when receiving packets.

8.3. USING A WIREGUARD CLIENT BEHIND NAT AND FIREWALLS

WireGuard uses the UDP protocol and transmits data only when a peer sends packets. Stateful firewalls and network address translation (NAT) on routers track connections to enable a peer behind NAT or a firewall to receive packets.

To keep the connection active, WireGuard supports **persistent keepalives**. This means you can set an interval at which WireGuard sends keepalive packets. By default, the persistent keep-alive feature is disabled to reduce network traffic. Enable this feature on the client if you use the client in a network with NAT or if a firewall closes the connection after some time of inactivity.

8.4. CREATING PRIVATE AND PUBLIC KEYS TO BE USED IN WIREGUARD CONNECTIONS

WireGuard uses base64-encoded private and public keys to authenticate hosts to each other. Therefore, you must create the keys on each host that participates in the WireGuard VPN.

**IMPORTANT**

For secure connections, create different keys for each host, and ensure that you only share the public key with the remote WireGuard host. Do not use the example keys used in this documentation.

Procedure

1. Install the **wireguard-tools** package:
# dnf install wireguard-tools

2. Create a private key and a corresponding public key for the host:

```bash
# wg genkey | tee /etc/wireguard/$HOSTNAME.private.key | wg pubkey > /etc/wireguard/$HOSTNAME.public.key
```

You will need the content of the key files, but not the files themselves. However, Red Hat recommends keeping the files in case that you need to remember the keys in future.

3. Set secure permissions on the key files:

```bash
# chmod 600 /etc/wireguard/$HOSTNAME.private.key /etc/wireguard/$HOSTNAME.public.key
```

4. Display the private key:

```bash
# cat /etc/wireguard/$HOSTNAME.private.key
YFAnE0psglLiAF7XR4abxiwVRnlMfelta10s/c4JXg=
```

You will need the private key to configure the WireGuard connection on the local host. Do not share the private key.

5. Display the public key:

```bash
# cat /etc/wireguard/$HOSTNAME.public.key
UtjqCJ57DeAsgYKRfp7cFGiQqdONRn69u249Fa406BE=
```

You will need the public key to configure the WireGuard connection on the remote host.

Additional resources

- The `wg(8)` man page

**8.5. CONFIGURING A WIREGUARD SERVER BY USING NMCLI**

You can configure the WireGuard server by creating a connection profile in NetworkManager. Use this method to let NetworkManager manage the WireGuard connection.

This procedure assumes the following settings:

- Server:
  - Private key: `YFAnE0psglLiAF7XR4abxiwVRnlMfelta10s/c4JXg=`
  - Tunnel IPv4 address: `192.0.2.1/24`
  - Tunnel IPv6 address: `2001:db8:1::1/32`

- Client:
  - Public key: `bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=`
  - Tunnel IPv4 address: `192.0.2.2/24`
Tunnel IPv6 address: 2001:db8:1::2/32

Prerequisites

- You have generated the public and private key for both the server and client.
- You know the following information:
  - The private key of the server
  - The static tunnel IP addresses and subnet masks of the client
  - The public key of the client
  - The static tunnel IP addresses and subnet masks of the server

Procedure

1. Add a NetworkManager WireGuard connection profile:

   ```
   # nmcli connection add type wireguard con-name server-wg0 ifname wg0 autoconnect no
   ```

   This command creates a profile named `server-wg0` and assigns the virtual interface `wg0` to it. To prevent the connection from starting automatically after you add it without finalizing the configuration, disable the `autoconnect` parameter.

2. Set the tunnel IPv4 address and subnet mask of the server:

   ```
   # nmcli connection modify server-wg0 ipv4.method manual ipv4.addresses 192.0.2.1/24
   ```

3. Set the tunnel IPv6 address and subnet mask of the server:

   ```
   # nmcli connection modify server-wg0 ipv6.method manual ipv6.addresses 2001:db8:1::1/32
   ```

4. Add the server’s private key to the connection profile:

   ```
   # nmcli connection modify server-wg0 wireguard.private-key "YFAnE0psgIdiAF7XR4abxiwVRnlMfeItxu10s/c4JXg="
   ```

5. Set the port for incoming WireGuard connections:

   ```
   # nmcli connection modify server-wg0 wireguard.listen-port 51820
   ```

   Always set a fixed port number on hosts that receive incoming WireGuard connections. If you do not set a port, WireGuard uses a random free port each time you activate the `wg0` interface.

6. Add peer configurations for each client that you want to allow to communicate with this server. You must add these settings manually, because the `nmcli` utility does not support setting the corresponding connection properties.

   a. Edit the `/etc/NetworkManager/system-connections/server-wg0.nmconnection` file, and append:
 allowed-ips=192.0.2.2;2001:db8:1::2;

- The [wireguard-peer.<public_key_of_the_client>] entry defines the peer section of the client, and the section name contains the public key of the client.

- The allowed-ips parameter sets the tunnel IP addresses of the client that are allowed to send data to this server.
  Add a section for each client.

b. Reload the server-wg0 connection profile:

```bash
# nmcli connection load /etc/NetworkManager/system-connections/server-wg0.nmconnection
```

7. Optional: Configure the connection to start automatically, enter:

```bash
# nmcli connection modify server-wg0 autoconnect yes
```

8. Reactivate the server-wg0 connection:

```bash
# nmcli connection up server-wg0
```

Next steps

- Configure the firewalld service on the WireGuard server.

Verification

1. Display the interface configuration of the wg0 device:

```bash
# wg show wg0
interface: wg0
  public key: UtjqCJ57DeAscYKRlfp7cFGiQqdONRn69u249Fa4O6BE=
  private key: (hidden)
  listening port: 51820

  peer: bwnfQcC8/g2i4vvE qcRUM2e6Hi3Nskk6G9t4r26nFVM=
    allowed ips: 192.0.2.2/32, 2001:db8:1::2/128

To display the private key in the output, use the WG_HIDE_KEYS=never wg show wg0 command.
```

2. Display the IP configuration of the wg0 device:

```bash
# ip address show wg0
20: wg0: <POINTOPOINT,NOARP,UP,LOWER_UP> mtu 1420 qdisc noqueue state unknown
   group default qlen 1000
   link/none
   inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute wg0
     valid_lft forever preferred_lft forever
     inet6 2001:db8:1::1/32 scope global noprefixroute
```

Red Hat Enterprise Linux 9 Configuring and managing networking
Valid_LFT forever preferred_LFT forever
inet6 fe80::3ef:8863:1ce2:844/64 scope link noprefixroute
Valid_LFT forever preferred_LFT forever

Additional resources

- The `wg(8)` man page
- The **WireGuard setting** section in the `nm-settings(5)` man page

8.6. CONFIGURING A WIREGUARD SERVER BY USING NMTUI

You can configure the WireGuard server by creating a connection profile in NetworkManager. Use this method to let NetworkManager manage the WireGuard connection.

This procedure assumes the following settings:

- **Server:**
  - Private key: YFAnE0psgliAF7XR4abxiwVRnlMfeltxu10s/c4JXg=
  - Tunnel IPv4 address: 192.0.2.1/24
  - Tunnel IPv6 address: 2001:db8:1::1/32

- **Client:**
  - Public key: bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
  - Tunnel IPv4 address: 192.0.2.2/24
  - Tunnel IPv6 address: 2001:db8:1::2/32

**Prerequisites**

- You have generated the public and private key for both the server and client.
- You know the following information:
  - The private key of the server
  - The static tunnel IP addresses and subnet masks of the client
  - The public key of the client
  - The static tunnel IP addresses and subnet masks of the server
- You installed the **NetworkManager-tui** package.

**Procedure**

1. Start the **nmtui** application:

   ```
   # nmtui
   ```

2. Select **Edit a connection**, and press **Enter**.
3. Select **Add**, and press **Enter**.

4. Select the **WireGuard** connection type in the list, and press **Enter**.

5. In the **Edit connection** window:
   a. Enter the name of the connection and the virtual interface, such as **wg0**, that NetworkManager should assign to the connection.
   
   b. Enter the private key of the server.
   
   c. Set the listen port number, such as **51820**, for incoming WireGuard connections. Always set a fixed port number on hosts that receive incoming WireGuard connections. If you do not set a port, WireGuard uses a random free port each time you activate the interface.
   
   d. Click **Add** next to the **Peers** pane:
      
      i. Enter the public key of the client.
      
      ii. Set the **Allowed IPs** field to the tunnel IP addresses of the client that are allowed to send data to this server.
      
      iii. Select **OK**, and press **Enter**.
   
   e. Select **Show** next to **IPv4 Configuration**, and press **Enter**.
      
      i. Select the IPv4 configuration method **Manual**.
      
      ii. Enter the tunnel IPv4 address and the subnet mask. Leave the **Gateway** field empty.
   
   f. Select **Show** next to **IPv6 Configuration**, and press **Enter**.
      
      i. Select the IPv6 configuration method **Manual**.
      
      ii. Enter the tunnel IPv6 address and the subnet mask. Leave the **Gateway** field empty.
g. Select OK, and press Enter.

6. In the window with the list of connections, select Back, and press Enter.

7. In the NetworkManager TUI main window, select Quit, and press Enter.

Next steps

- Configure the firewalld service on the WireGuard server.

Verification

1. Display the interface configuration of the wg0 device:

```
# wg show wg0
interface: wg0
  public key: UtjqCJ57DeAscYKRfp7cFGiQqdONRn69u249Fa4O6BE=
  private key: (hidden)
  listening port: 51820
  peer: bnwfQcC8/g24vvEgcRUM2e6Hl3Nsk6G9t4r26nFVM=
  allowed ips: 192.0.2.2/32, 2001:db8:1::2/128
```

To display the private key in the output, use the `WG_HIDE_KEYS=never wg show wg0` command.

2. Display the IP configuration of the wg0 device:
Additional resources

- The `wg(8)` man page

### 8.7. CONFIGURING A WIREGUARD SERVER BY USING NM-CONNECTION-EDITOR

You can configure the WireGuard server by creating a connection profile in NetworkManager. Use this method to let NetworkManager manage the WireGuard connection.

**Prerequisites**

- You have generated the public and private key for both the server and client.
- You know the following information:
  - The private key of the server
  - The static tunnel IP addresses and subnet masks of the client
  - The public key of the client
  - The static tunnel IP addresses and subnet masks of the server

**Procedure**

1. Open a terminal, and enter:

   ```bash
   # nm-connection-editor
   ```

2. Add a new connection by clicking the + button.

3. Select the **WireGuard** connection type, and click **Create**.

4. Optional: Update the connection name.

5. On the **General** tab, select **Connect automatically with priority**. Optionally, set a priority value.

6. On the **WireGuard** tab:
   
   a. Enter the name of the virtual interface, such as `wg0`, that NetworkManager should assign to the connection.
   
   b. Enter the private key of the server.
c. Set the listen port number, such as 51820, for incoming WireGuard connections. Always set a fixed port number on hosts that receive incoming WireGuard connections. If you do not set a port, WireGuard uses a random free port each time you activate the interface.

d. Click Add to add peers:
   i. Enter the public key of the client.
   ii. Set the Allowed IPs field to the tunnel IP addresses of the client that are allowed to send data to this server.
   iii. Click Apply.

7. On the IPv4 Settings tab:
   b. Click Add to enter the tunnel IPv4 address and the subnet mask. Leave the Gateway field empty.

8. On the IPv6 Settings tab:
   b. Click Add to enter the tunnel IPv6 address and the subnet mask. Leave the Gateway field empty.

9. Click Save to store the connection profile.

Next steps
   • Configure the firewalld service on the WireGuard server.

Verification

1. Display the interface configuration of the wg0 device:

   # wg show wg0
   interface: wg0
   public key: UtjqCJ57DeAsYKRfp7cF6qiQqONRn69u249Fa4O6BE=
   private key: (hidden)
   listening port: 51820
   peer: bnwfQcC8/g24vEqcRUM2e6H13Ns+k6G9t4r26nFVM=
   allowed ips: 192.0.2.2/32, 2001:db8:1::2/128

   To display the private key in the output, use the WG_HIDE_KEYS=never wg show wg0 command.

2. Display the IP configuration of the wg0 device:

   # ip address show wg0
   20: wg0: <POINTOPOINT,NOARP,UP,LOWER_UP> mtu 1420 qdisc noqueue state UNKNOWN group default qlen 1000
   link/none
inet 192.0.2.1/24 brd 192.0.2.255 scope global noprefixroute wg0
  valid_lft forever preferred_lft forever
inet6 2001:db8:1:1/32 scope global noprefixroute
  valid_lft forever preferred_lft forever
inet6 fe80::3ef:8863:1ce2:844/64 scope link noprefixroute
  valid_lft forever preferred_lft forever

Additional resources
  • The wg(8) man page

8.8. CONFIGURING A WIREGUARD SERVER BY USING THE WG-QUICK SERVICE

You can configure the WireGuard server by creating a configuration file in the /etc/wireguard/ directory. Use this method to configure the service independently from NetworkManager.

This procedure assumes the following settings:

  • Server:
    ▪ Private key: YFAnE0psgIdiAF7XR4abxiwVRnlMfetxu10s/c4JXg=
    ▪ Tunnel IPv4 address: 192.0.2.1/24
    ▪ Tunnel IPv6 address: 2001:db8:1::1/32
  • Client:
    ▪ Public key: bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
    ▪ Tunnel IPv4 address: 192.0.2.2/24
    ▪ Tunnel IPv6 address: 2001:db8:1::2/32

Prerequisites
  • You have generated the public and private key for both the server and client.
  • You know the following information:
    ▪ The private key of the server
    ▪ The static tunnel IP addresses and subnet masks of the client
    ▪ The public key of the client
    ▪ The static tunnel IP addresses and subnet masks of the server

Procedure
  1. Install the wireguard-tools package:

     # dnf install wireguard-tools
2. Create the `/etc/wireguard/wg0.conf` file with the following content:

```plaintext
[Interface]
Address = 192.0.2.1/24, 2001:db8:1::1/32
ListenPort = 51820
PrivateKey = YFAnE0psgldiAF7XR4abxiwVRnlMfeJtxu10s/c4JXg=

[Peer]
PublicKey = bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
AllowedIPs = 192.0.2.2, 2001:db8:1::2
```

- The `[Interface]` section describes the WireGuard settings of the interface on the server:
  - **Address**: A comma-separated list of the server’s tunnel IP addresses.
  - **PrivateKey**: The private key of the server.
  - **ListenPort**: The port on which WireGuard listens for incoming UDP connections.
    - Always set a fixed port number on hosts that receive incoming WireGuard connections.
    - If you do not set a port, WireGuard uses a random free port each time you activate the `wg0` interface.

- Each `[Peer]` section describes the settings of one client:
  - **PublicKey**: The public key of the client.
  - **AllowedIPs**: The tunnel IP addresses of the client that are allowed to send data to this server.

3. Enable and start the WireGuard connection:

```plaintext
# systemctl enable --now wg-quick@wg0
```

The systemd instance name must match the name of the configuration file in the `/etc/wireguard/` directory without the `.conf` suffix. The service also uses this name for the virtual network interface.

Next steps

- Configure the firewalld service on the WireGuard server.

Verification

1. Display the interface configuration of the `wg0` device:

```plaintext
# wg show wg0
interface: wg0
  public key: UtjqCJ57DeAscYKRIfp7cFGiQqdONRn69u249Fa4O6BE=
  private key: (hidden)
  listening port: 51820

peer: bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
  allowed ips: 192.0.2.2/32, 2001:db8:1::2/128
```

To display the private key in the output, use the `WG_HIDE_KEYS=never wg show wg0` command.
2. Display the IP configuration of the **wg0** device:

```
# ip address show wg0
20: wg0: <POINTOPOINT,NOARP,UP,LOWER_UP> mtu 1420 qdisc noqueue state UNKNOWN group default qlen 1000
    link/none
    inet 192.0.2.1/24 scope global wg0
        valid_lft forever preferred_lft forever
    inet6 2001:db8:1::1/32 scope global
        valid_lft forever preferred_lft forever
```

Additional resources

- The **wg(8)** man page
- The **wg-quick(8)** man page

### 8.9. Configuring Firewalld on a WireGuard Server by Using the Command Line

You must configure the **firewalld** service on the WireGuard server to allow incoming connections from clients. Additionally, if clients should be able to use the WireGuard server as the default gateway and route all traffic through the tunnel, you must enable masquerading.

**Procedure**

1. Open the WireGuard port for incoming connections in the **firewalld** service:

   ```
   # firewall-cmd --permanent --add-port=51820/udp --zone=public
   ```

2. If clients should route all traffic through the tunnel and use the WireGuard server as the default gateway, enable masquerading for the **public** zone:

   ```
   # firewall-cmd --permanent --zone=public --add-masquerade
   ```

3. Reload the **firewalld** rules.

   ```
   # firewall-cmd --reload
   ```

**Verification**

- Display the configuration of the **public** zone:

  ```
  # firewall-cmd --list-all
  public (active)
  ... ports: 51820/udp masquerade: yes
  ... 
  ```

Additional resources
8.10. CONFIGURING FIREWALLD ON A WIREGUARD SERVER BY USING THE GRAPHICAL INTERFACE

You must configure the firewalld service on the WireGuard server to allow incoming connections from clients. Additionally, if clients should be able to use the WireGuard server as the default gateway and route all traffic through the tunnel, you must enable masquerading.

Procedure

1. Press the Super key, enter firewall, and select the Firewall application from the results.
2. Select Permanent in the Configuration list.
3. Select the public zone.
4. Allow incoming connections to the WireGuard port:
   a. On the Ports tab, click Add.
   b. Enter the port number you set for incoming WireGuard connections:
   c. Select udp from the Protocol list.
   d. Click OK.
5. If clients should route all traffic through the tunnel and use the WireGuard server as the default gateway:
   a. Navigate to the Masquerading tab of the public zone.
   b. Select Masquerade zone.
6. Select Options → Reload Firewalld.

Verification

- Display the configuration of the public zone:

```
# firewall-cmd --list-all
public (active)
... ports: 51820/udp masquerade: yes ...
```

8.11. CONFIGURING A WIREGUARD CLIENT BY USING NMCLI

You can configure a WireGuard client by creating a connection profile in NetworkManager. Use this method to let NetworkManager manage the WireGuard connection.

This procedure assumes the following settings:

- Client:
Prerequisites

- You have generated the public and private key for both the server and client.
- You know the following information:
  - The private key of the client
  - The static tunnel IP addresses and subnet masks of the client
  - The public key of the server
  - The static tunnel IP addresses and subnet masks of the server

Procedure

1. Add a NetworkManager WireGuard connection profile:

   ```
   # nmcli connection add type wireguard con-name client-wg0 ifname wg0 autoconnect no
   ```

   This command creates a profile named `client-wg0` and assigns the virtual interface `wg0` to it. To prevent the connection from starting automatically after you add it without finalizing the configuration, disable the `autoconnect` parameter.

2. Optional: Configure NetworkManager so that it does not automatically start the `client-wg` connection:

   ```
   # nmcli connection modify client-wg0 autoconnect no
   ```

3. Set the tunnel IPv4 address and subnet mask of the client:

   ```
   # nmcli connection modify client-wg0 ipv4.method manual ipv4.addresses 192.0.2.2/24
   ```

4. Set the tunnel IPv6 address and subnet mask of the client:

   ```
   # nmcli connection modify client-wg0 ipv6.method manual ipv6.addresses 2001:db8:1::2/32
   ```

5. If you want to route all traffic through the tunnel, set the tunnel IP addresses of the server as the default gateway:
# nmcli connection modify client-wg0 ipv4.gateway 192.0.2.1 ipv6.gateway 2001:db8:1::1

Routing all traffic through the tunnel requires that you set, in a later step, the **allowed-ips** on the this client to **0.0.0.0/0;::/0**.

Note that routing all traffic through the tunnel can impact the connectivity to other hosts based on the server routing and firewall configuration.

6. Add the client’s private key to the connection profile:

```
# nmcli connection modify client-wg0 wireguard.private-key "aPUcp5vHz8yMLrzk8SsDyYnV33IhE/k20e52iKJFV0A=
```

7. Add peer configurations for each server that you want to allow to communicate with this client. You must add these settings manually, because the `nmcli` utility does not support setting the corresponding connection properties.

   a. Edit the `/etc/NetworkManager/system-connections/client-wg0.nmconnection` file, and append:

   ```
   [wireguard-peer.UtjqCJ57DeAscYKRfp7cFGiQqdONRn69u249Fa406BE=]
   endpoint=server.example.com:51820
   allowed-ips=192.0.2.1;2001:db8:1::1;
   persistent-keepalive=20
   
   # The [wireguard-peer.<public_key_of_the_server>] entry defines the peer section of the server, and the section name has the public key of the server.
   
   # The endpoint parameter sets the hostname or IP address and the port of the server. The client uses this information to establish the connection.
   
   # The allowed-ips parameter sets a list of IP addresses that can send data to this client. For example, set the parameter to:
   
   # The tunnel IP addresses of the server to allow only the server to communicate with this client. The value in the example above configures this scenario.
   
   # 0.0.0.0/0;::/0; to allow any remote IPv4 and IPv6 address to communicate with this client. Use this setting to route all traffic through the tunnel and use the WireGuard server as default gateway.
   
   # The optional persistent-keepalive parameter defines an interval in seconds in which WireGuard sends a keep alive packet to the server. Set this parameter if you use the client in a network with network address translation (NAT) or if a firewall closes the UDP connection after some time of inactivity.
   
   b. Reload the `client-wg0` connection profile:

```
# nmcli connection load /etc/NetworkManager/system-connections/client-wg0.nmconnection
```

8. Reactivate the `client-wg0` connection:

```
# nmcli connection up client-wg0
```
Verification

1. Ping the IP addresses of the server:

   ```
   # ping 192.0.2.1
   # ping6 2001:db8:1::1
   ```

2. Display the interface configuration of the `wg0` device:

   ```
   # wg show wg0
   interface: wg0
   public key: bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
   private key: (hidden)
   listening port: 51820
   peer: UtlqCJ57DeAscYKRFp7cFGiQqdONRn69u249Fa4O6BE=
   endpoint: server.example.com:51820
   allowed ips: 192.0.2.1/32, 2001:db8:1::1/128
   latest handshake: 1 minute, 41 seconds ago
   transfer: 824 B received, 1.01 KiB sent
   persistent keepalive: every 20 seconds
   ```

   To display the private key in the output, use the `WG_HIDE_KEYS=never wg show wg0` command.

   Note that the output has only the latest handshake and transfer entries if you have already sent traffic through the VPN tunnel.

3. Display the IP configuration of the `wg0` device:

   ```
   # ip address show wg0
   10: wg0: <POINTOPOINT,NOARP,UP,LOWER_UP> mtu 1420 qdisc noqueue state UNKNOWN group default qlen 1000
       link/none
       inet 192.0.2.2/24 brd 192.0.2.255 scope global noprefixroute wg0 valid_lft forever preferred_lft forever
           inet6 2001:db8:1::/64 scope link noprefixroute valid_lft forever preferred_lft forever
           inet6 fe80::73d9:6f51:ea6f:863e/64 scope link noprefixroute valid_lft forever preferred_lft forever
   ```

Additional resources

- The `wg(8)` man page
- The WireGuard setting section in the `nm-settings(5)` man page

8.12. CONFIGURING A WIREGUARD CLIENT BY USING NMTUI

You can configure a WireGuard client by creating a connection profile in NetworkManager. Use this method to let NetworkManager manage the WireGuard connection.

This procedure assumes the following settings:
• Private key: aPUcp5vHz8yMLrzk8SsDyYnV33IhE/k20e52iKJFV0A=
• Tunnel IPv4 address: 192.0.2.2/24
• Tunnel IPv6 address: 2001:db8:1::2/32

● Server:
  • Public key: UtjqCJ57DeAscYKRfp7cFGiQqdONRn69u249Fa4O6BE=
  • Tunnel IPv4 address: 192.0.2.1/24
  • Tunnel IPv6 address: 2001:db8:1::1/32

Prerequisites

● You have generated the public and private key for both the server and client.

● You know the following information:
  • The private key of the client
  • The static tunnel IP addresses and subnet masks of the client
  • The public key of the server
  • The static tunnel IP addresses and subnet masks of the server

● You installed the NetworkManager-tui package

Procedure

1. Start the nmtui application:
   # nmtui

2. Select Edit a connection, and press Enter.

3. Select Add, and press Enter.

4. Select the WireGuard connection type in the list, and press Enter.

5. In the Edit connection window:
   a. Enter the name of the connection and the virtual interface, such as wg0, that NetworkManager should assign to the connection.
   b. Enter the private key of the client.
c. Click **Add** next to the **Peers** pane:
   
i. Enter the public key of the server.

   ii. Set the **Allowed IPs** field. For example, set it to:
       
       - The tunnel IP addresses of the server to allow only the server to communicate with this client.
       
       - 0.0.0.0/0,::/0 to allow any remote IPv4 and IPv6 address to communicate with this client. Use this setting to route all traffic through the tunnel and use the WireGuard server as default gateway.

   iii. Enter the host name or IP address and port of the WireGuard server into the **Endpoint** field. Use the following format: `hostname_or_IP:port_number`

   iv. Optional: If you use the client in a network with network address translation (NAT) or if a firewall closes the UDP connection after some time of inactivity, set a persistent keep alive interval in seconds. In this interval, the client sends a keepalive packet to the server.

   v. Select **OK**, and press **Enter**.

d. Select **Show** next to **IPv4 Configuration**, and press **Enter**.

   i. Select the IPv4 configuration method **Manual**.

   ii. Enter the tunnel IPv4 address and the subnet mask. Leave the **Gateway** field empty.

e. Select **Show** next to **IPv6 Configuration**, and press **Enter**.

   i. Select the IPv6 configuration method **Manual**.

   ii. Enter the tunnel IPv6 address and the subnet mask. Leave the **Gateway** field empty.

f. Optional: Select **Automatically connect**.

g. Select **OK**, and press **Enter**
6. In the window with the list of connections, select **Back**, and press **Enter**.

7. In the **NetworkManager TUI** main window, select **Quit**, and press **Enter**.

**Verification**

1. Ping the IP addresses of the server:
   
   ```
   # ping 192.0.2.1
   # ping6 2001:db8:1::1
   ```

2. Display the interface configuration of the **wg0** device:

   ```
   # wg show wg0
   interface: wg0
   public key: bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
   private key: (hidden)
   listening port: 51820
   peer: UtjqCJ57DeAscYKRfp7cFGiQqdONRn69u249Fa4O6BE=
   endpoint: server.example.com:51820
   allowed ips: 192.0.2.1/32, 2001:db8:1::1/128
   latest handshake: 1 minute, 41 seconds ago
   transfer: 824 B received, 1.01 KiB sent
   persistent keepalive: every 20 seconds
   ```
To display the private key in the output, use the \texttt{WG\_HIDE\_KEYS=never} \texttt{wg show wg0} command.

Note that the output contains only the \textit{latest handshake} and \textit{transfer} entries if you have already sent traffic through the VPN tunnel.

3. Display the IP configuration of the \texttt{wg0} device:

\begin{verbatim}
# ip address show wg0
10: wg0: <POINTOPOINT,NOARP,UP,LOWER_UP> mtu 1420 qdisc noqueue state UNKNOWN group default qlen 1000
   link/none
   inet 192.0.2.2/24 brd 192.0.2.255 scope global noprefixroute wg0
       valid_lft forever preferred_lft forever
   inet6 2001:db8:1::2/32 scope global noprefixroute
       valid_lft forever preferred_lft forever
   inet6 fe80::73d9:6f51:ea6f:863e/64 scope link noprefixroute
       valid_lft forever preferred_lft forever
\end{verbatim}

\textbf{Additional resources}

- The \texttt{wg(8)} man page

\textbf{8.13. CONFIGURING A WIREGUARD CLIENT BY USING NM-CONNECTION-EDITOR}

You can configure a WireGuard client by creating a connection profile in NetworkManager. Use this method to let NetworkManager manage the WireGuard connection.

\textbf{Prerequisites}

- You have generated the public and private key for both the server and client.
- You know the following information:
  - The private key of the client
  - The static tunnel IP addresses and subnet masks of the client
  - The public key of the server
  - The static tunnel IP addresses and subnet masks of the server

\textbf{Procedure}

1. Open a terminal, and enter:

\begin{verbatim}
# nm-connection-editor
\end{verbatim}

2. Add a new connection by clicking the \texttt{+} button.

3. Select the \texttt{WireGuard} connection type, and click \texttt{Create}.

4. Optional: Update the connection name.
5. Optional: On the General tab, select Connect automatically with priority.

6. On the WireGuard tab:
   a. Enter the name of the virtual interface, such as wg0, that NetworkManager should assign to the connection.
   b. Enter client’s private key.
   c. Click Add to add peers:
      i. Enter the public key of the server.
      ii. Set the Allowed IPs field. For example, set it to:
          • The tunnel IP addresses of the server to allow only the server to communicate with this client.
          • 0.0.0.0/0;::/0; to allow any remote IPv4 and IPv6 address to communicate with this client. Use this setting to route all traffic through the tunnel and use the WireGuard server as default gateway. Note that routing all traffic through the tunnel can impact the connectivity to other hosts based on the server routing and firewall configuration.
      iii. Enter the hostname or IP address and port of the WireGuard server into the Endpoint field. Use the following format: hostname_or_IP:port_number
      iv. Optional: If you use the client in a network with network address translation (NAT) or if a firewall closes the UDP connection after some time of inactivity, set a persistent keep alive interval in seconds. In this interval, the client sends a keep alive packet to the server.
   v. Click Apply.

7. On the IPv4 Settings tab:
   b. Click Add to enter the tunnel IPv4 address and the subnet mask.
   c. If you want to route all traffic through the tunnel, set the tunnel IPv4 address of the server in the Gateway field. Otherwise, leave the field empty. Routing all IPv4 traffic through the tunnel requires that you included 0.0.0.0/0 in the Allowed IPs field on this client.

8. On the IPv6 Settings tab:
   b. Click Add to enter the tunnel IPv6 address and the subnet mask.
   c. If you want to route all traffic through the tunnel, set the tunnel IPv6 address of the server in the Gateway field. Otherwise, leave the field empty. Routing all IPv4 traffic through the tunnel requires that you included ::/0 in the Allowed IPs field on this client.

9. Click Save to store the connection profile.
Verification

1. Ping the IP addresses of the server:

```bash
# ping 192.0.2.1
# ping6 2001:db8:1::1
```

2. Display the interface configuration of the `wg0` device:

```bash
# wg show wg0
interface: wg0
   public key: bnwfQcC8/g2i4vvEqcRUM2e6Hi3Nskk6G9t4r26nFVM=
   private key: (hidden)
   listening port: 51820
   peer: UtjqCJ57DeAscYKRlp7cFGiQqdONRn69u249Fa4O6BE=
   endpoint: server.example.com:51820
   allowed ips: 192.0.2.1/32, 2001:db8:1::1/128
   latest handshake: 1 minute, 41 seconds ago
   transfer: 824 B received, 1.01 KiB sent
   persistent keepalive: every 20 seconds
```

To display the private key in the output, use the `WG_HIDE_KEYS=never wg show wg0` command.

Note that the output only has the latest handshake and transfer entries if you have already sent traffic through the VPN tunnel.

3. Display the IP configuration of the `wg0` device:

```bash
# ip address show wg0
10: wg0: <POINTOPOINT,NOARP,UP,LOWER_UP> mtu 1420 qdisc noqueue state UNKNOWN group default qlen 1000
   link/none
      inet 192.0.2.2/24 brd 192.0.2.255 scope global noprefixroute wg0
      valid_lft forever preferred_lft forever
      inet6 2001:db8:1::2/32 scope global noprefixroute
      valid_lft forever preferred_lft forever
      inet6 fe80::73d9:6f51:ea6f:863e/64 scope link noprefixroute
      valid_lft forever preferred_lft forever
```

Additional resources

- The `wg(8)` man page

8.14. CONFIGURING A WIREGUARD CLIENT BY USING THE WG-QUICK SERVICE

You can configure a WireGuard client by creating a configuration file in the `/etc/wireguard/` directory. Use this method to configure the service independently from NetworkManager.

This procedure assumes the following settings:

- Client:
- Private key: aPUcp5vHz8yMLrzk8SsDyYnV33lhE/k20e52iKJFV0A=
- Tunnel IPv4 address: 192.0.2.2/24
- Tunnel IPv6 address: 2001:db8:1::2/32

- Server:
  - Public key: UtjqCJ57DeAscYKRfp7cFGiQqdONRn69u249Fa4O6BE=
  - Tunnel IPv4 address: 192.0.2.1/24
  - Tunnel IPv6 address: 2001:db8:1::1/32

Prerequisites
- You have generated the public and private key for both the server and client.
- You know the following information:
  - The private key of the client
  - The static tunnel IP addresses and subnet masks of the client
  - The public key of the server
  - The static tunnel IP addresses and subnet masks of the server

Procedure
1. Install the wireguard-tools package:
   ```
   # dnf install wireguard-tools
   ```
2. Create the /etc/wireguard/wg0.conf file with the following content:
   ```
   [Interface]
   Address = 192.0.2.2/24, 2001:db8:1::2/32
   PrivateKey = aPUcp5vHz8yMLrzk8SsDyYnV33lhE/k20e52iKJFV0A=

   [Peer]
   PublicKey = UtjqCJ57DeAscYKRfp7cFGiQqdONRn69u249Fa4O6BE=
   AllowedIPs = 192.0.2.1, 2001:db8:1::1
   Endpoint = server.example.com:51820
   PersistentKeepalive = 20
   ```

- The [Interface] section describes the WireGuard settings of the interface on the client:
  - **Address**: A comma-separated list of the client’s tunnel IP addresses.
  - **PrivateKey**: The private key of the client.

- The [Peer] section describes the settings of the server:
  - **PublicKey**: The public key of the server.
- **AllowedIPs**: The IP addresses that are allowed to send data to this client. For example, set the parameter to:
  - The tunnel IP addresses of the server to allow only the server to communicate with this client. The value in the example above configures this scenario.
  - `0.0.0.0/0, ::/0` to allow any remote IPv4 and IPv6 address to communicate with this client. Use this setting to route all traffic through the tunnel and use the WireGuard server as default gateway.

- **Endpoint**: Sets the hostname or IP address and the port of the server. The client uses this information to establish the connection.

- The optional `persistent-keepalive` parameter defines an interval in seconds in which WireGuard sends a keepalive packet to the server. Set this parameter if you use the client in a network with network address translation (NAT) or if a firewall closes the UDP connection after some time of inactivity.

3. Enable and start the WireGuard connection:

```bash
# systemctl enable --now wg-quick@wg0
```

The systemd instance name must match the name of the configuration file in the `/etc/wireguard/` directory without the `.conf` suffix. The service also uses this name for the virtual network interface.

**Verification**

1. Ping the IP addresses of the server:

```bash
# ping 192.0.2.1
# ping6 2001:db8:1::1
```

2. Display the interface configuration of the `wg0` device:

```bash
# wg show wg0
interface: wg0	public key: bnfwQcC8/g2i4vvEqcRUM2e6Hi3Nske6G9t4r26nFVM=
private key: (hidden)
listening port: 51820

peer: UtjqCJ57DcAsyKRFp7cFGdQdONRn69u249Fa4O6BE=
endpoint: server.example.com:51820
allowed ips: 192.0.2.1/32, 2001:db8:1::1/128
latest handshake: 1 minute, 41 seconds ago
transfer: 824 B received, 1.01 KiB sent
persistent keepalive: every 20 seconds
```

To display the private key in the output, use the `WG_HIDE_KEYS=never wg show wg0` command.

Note that the output contains only the `latest handshake` and `transfer` entries if you have already sent traffic through the VPN tunnel.

3. Display the IP configuration of the `wg0` device:
### ip address show wg0

- **10. wg0**: `<POINTOPOINT,NOARP,UP,LOWER_UP>` mtu 1420 qdisc noqueue state UNKNOWN group default qlen 1000
  - link/none
  - inet 192.0.2.2/24 scope global wg0
    - valid_lft forever preferred_lft forever
  - inet6 2001:db8:1::2/32 scope global
    - valid_lft forever preferred_lft forever

#### Additional resources

- The **wg(8)** man page
- The **wg-quick(8)** man page
CHAPTER 9. CONFIGURING IP TUNNELS

Similar to a VPN, an IP tunnel directly connects two networks over a third network, such as the internet. However, not all tunnel protocols support encryption.

The routers in both networks that establish the tunnel requires at least two interfaces:

- One interface that is connected to the local network
- One interface that is connected to the network through which the tunnel is established.

To establish the tunnel, you create a virtual interface on both routers with an IP address from the remote subnet.

NetworkManager supports the following IP tunnels:

- Generic Routing Encapsulation (GRE)
- Generic Routing Encapsulation over IPv6 (IP6GRE)
- Generic Routing Encapsulation Terminal Access Point (GRETAP)
- Generic Routing Encapsulation Terminal Access Point over IPv6 (IP6GRETAP)
- IPv4 over IPv4 (IPIP)
- IPv4 over IPv6 (IPIP6)
- IPv6 over IPv6 (IP6IP6)
- Simple Internet Transition (SIT)

Depending on the type, these tunnels act either on layer 2 or 3 of the Open Systems Interconnection (OSI) model.

9.1. CONFIGURING AN IPIP TUNNEL USING NMCLI TO ENCAPSULATE IPV4 TRAFFIC IN IPV4 PACKETS

An IP over IP (IPIP) tunnel operates on OSI layer 3 and encapsulates IPv4 traffic in IPv4 packets as described in RFC 2003.

IMPORTANT

Data sent through an IPIP tunnel is not encrypted. For security reasons, use the tunnel only for data that is already encrypted, for example, by other protocols, such as HTTPS.

Note that IPIP tunnels support only unicast packets. If you require an IPv4 tunnel that supports multicast, see Configuring a GRE tunnel using nmcli to encapsulate layer-3 traffic in IPv4 packets.

For example, you can create an IPIP tunnel between two RHEL routers to connect two internal subnets over the internet as shown in the following diagram:
Prerequisites

- Each RHEL router has a network interface that is connected to its local subnet.
- Each RHEL router has a network interface that is connected to the internet.
- The traffic you want to send through the tunnel is IPv4 unicast.

Procedure

1. On the RHEL router in network A:
   a. Create an IPIP tunnel interface named `tun0`:

   ```
   # nmcli connection add type ip-tunnel ip-tunnel.mode ipip con-name tun0 ifname tun0 remote 198.51.100.5 local 203.0.113.10
   ```

   The `remote` and `local` parameters set the public IP addresses of the remote and the local routers.
   
   b. Set the IPv4 address to the `tun0` device:

   ```
   # nmcli connection modify tun0 ipv4.addresses '10.0.1.1/30'
   ```

   Note that a `/30` subnet with two usable IP addresses is sufficient for the tunnel.
   
   c. Configure the `tun0` connection to use a manual IPv4 configuration:

   ```
   # nmcli connection modify tun0 ipv4.method manual
   ```
   
   d. Add a static route that routes traffic to the `172.16.0.0/24` network to the tunnel IP on router B:

   ```
   # nmcli connection modify tun0 +ipv4.routes "172.16.0.0/24 10.0.1.2"
   ```
   
   e. Enable the `tun0` connection.
2. On the RHEL router in network B:
   a. Create an IPIP tunnel interface named `tun0`:
      ```
      # nmcli connection add type ip-tunnel ip-tunnel.mode ipip con-name tun0 ifname tun0 remote 203.0.113.10 local 198.51.100.5
      ```
      The `remote` and `local` parameters set the public IP addresses of the remote and local routers.
   b. Set the IPv4 address to the `tun0` device:
      ```
      # nmcli connection modify tun0 ipv4.addresses '10.0.1.2/30'
      ```
   c. Configure the `tun0` connection to use a manual IPv4 configuration:
      ```
      # nmcli connection modify tun0 ipv4.method manual
      ```
   d. Add a static route that routes traffic to the `192.0.2.0/24` network to the tunnel IP on router A:
      ```
      # nmcli connection modify tun0 +ipv4.routes "192.0.2.0/24 10.0.1.1"
      ```
   e. Enable the `tun0` connection.
      ```
      # nmcli connection up tun0
      ```
   f. Enable packet forwarding:
      ```
      # echo "net.ipv4.ip_forward=1" > /etc/sysctl.d/95-IPv4-forwarding.conf
      # sysctl -p /etc/sysctl.d/95-IPv4-forwarding.conf
      ```

Verification

- From each RHEL router, ping the IP address of the internal interface of the other router:
  a. On Router A, ping `172.16.0.1`:
     ```
     # ping 172.16.0.1
     ```
  b. On Router B, ping `192.0.2.1`:
     ```
     # ping 192.0.2.1
     ```

Additional resources
9.2. CONFIGURING A GRE TUNNEL USING NMCLI TO ENCAPSULATE LAYER-3 TRAFFIC IN IPV4 PACKETS

A Generic Routing Encapsulation (GRE) tunnel encapsulates layer-3 traffic in IPv4 packets as described in RFC 2784. A GRE tunnel can encapsulate any layer 3 protocol with a valid Ethernet type.

**IMPORTANT**

Data sent through a GRE tunnel is not encrypted. For security reasons, use the tunnel only for data that is already encrypted, for example, by other protocols, such as HTTPS.

For example, you can create a GRE tunnel between two RHEL routers to connect two internal subnets over the internet as shown in the following diagram:

**NOTE**

The `gre0` device name is reserved. Use `gre1` or a different name for the device.

**Prerequisites**

- Each RHEL router has a network interface that is connected to its local subnet.
- Each RHEL router has a network interface that is connected to the internet.

**Procedure**

1. On the RHEL router in network A:
   a. Create a GRE tunnel interface named `gre1`:

      ```bash
      # nmcli connection add type ip-tunnel ip-tunnel.mode gre con-name gre1 ifname gre1 remote 198.51.100.5 local 203.0.113.10
      ```
The remote and local parameters set the public IP addresses of the remote and the local routers.

b. Set the IPv4 address to the gre1 device:

```bash
# nmcli connection modify gre1 ipv4.addresses '10.0.1.1/30'
```

Note that a /30 subnet with two usable IP addresses is sufficient for the tunnel.

c. Configure the gre1 connection to use a manual IPv4 configuration:

```bash
# nmcli connection modify gre1 ipv4.method manual
```

d. Add a static route that routes traffic to the 172.16.0.0/24 network to the tunnel IP on router B:

```bash
# nmcli connection modify gre1 +ipv4.routes "172.16.0.0/24 10.0.1.2"
```

e. Enable the gre1 connection.

```bash
# nmcli connection up gre1
```

f. Enable packet forwarding:

```bash
# echo "net.ipv4.ip_forward=1" > /etc/sysctl.d/95-IPv4-forwarding.conf
# sysctl -p /etc/sysctl.d/95-IPv4-forwarding.conf
```

2. On the RHEL router in network B:

a. Create a GRE tunnel interface named gre1:

```bash
# nmcli connection add type ip-tunnel ip-tunnel.mode gre con-name gre1 ifname gre1 remote 203.0.113.10 local 198.51.100.5
```

The remote and local parameters set the public IP addresses of the remote and the local routers.

b. Set the IPv4 address to the gre1 device:

```bash
# nmcli connection modify gre1 ipv4.addresses '10.0.1.2/30'
```

c. Configure the gre1 connection to use a manual IPv4 configuration:

```bash
# nmcli connection modify gre1 ipv4.method manual
```

d. Add a static route that routes traffic to the 192.0.2.0/24 network to the tunnel IP on router A:

```bash
# nmcli connection modify gre1 +ipv4.routes "192.0.2.0/24 10.0.1.1"
```

e. Enable the gre1 connection.

```bash
# nmcli connection up gre1
```
Enable packet forwarding:

```
# echo "net.ipv4.ip_forward=1" > /etc/sysctl.d/95-IPv4-forwarding.conf
# sysctl -p /etc/sysctl.d/95-IPv4-forwarding.conf
```

Verification

1. From each RHEL router, ping the IP address of the internal interface of the other router:
   a. On Router A, ping 172.16.0.1:
      
      ```
      # ping 172.16.0.1
      ```
   b. On Router B, ping 192.0.2.1:
      
      ```
      # ping 192.0.2.1
      ```

Additional resources

- `nmcli(1)` man page
- `nm-settings(5)` man page

9.3. CONFIGURING A GRETAP TUNNEL TO TRANSFER ETHERNET FRAMES OVER IPV4

A Generic Routing Encapsulation Terminal Access Point (GRETAP) tunnel operates on OSI level 2 and encapsulates Ethernet traffic in IPv4 packets as described in RFC 2784.

**IMPORTANT**

Data sent through a GRETAP tunnel is not encrypted. For security reasons, establish the tunnel over a VPN or a different encrypted connection.

For example, you can create a GRETAP tunnel between two RHEL routers to connect two networks using a bridge as shown in the following diagram:
NOTE

The `gretap0` device name is reserved. Use `gretap1` or a different name for the device.

Prerequisites

- Each RHEL router has a network interface that is connected to its local network, and the interface has no IP configuration assigned.
- Each RHEL router has a network interface that is connected to the internet.

Procedure

1. On the RHEL router in network A:
   a. Create a bridge interface named `bridge0`:
      
      ```
      # nmcli connection add type bridge con-name bridge0 ifname bridge0
      ```
   b. Configure the IP settings of the bridge:
      
      ```
      # nmcli connection modify bridge0 ipv4.addresses '192.0.2.1/24'
      # nmcli connection modify bridge0 ipv4.method manual
      ```
   c. Add a new connection profile for the interface that is connected to local network to the bridge:
      
      ```
      # nmcli connection add type ethernet slave-type bridge con-name bridge0-port1 ifname enp1s0 master bridge0
      ```
   d. Add a new connection profile for the GRETAP tunnel interface to the bridge:
# nmcli connection add type ip-tunnel ip-tunnel.mode gretap slave-type bridge con-name bridge0-port2 ifname gretap1 remote 198.51.100.5 local 203.0.113.10 master bridge0

The `remote` and `local` parameters set the public IP addresses of the remote and the local routers.

e. Optional: Disable the Spanning Tree Protocol (STP) if you do not need it:

```bash
# nmcli connection modify bridge0 bridge.stp no
```

By default, STP is enabled and causes a delay before you can use the connection.

f. Configure that activating the `bridge0` connection automatically activates the ports of the bridge:

```bash
# nmcli connection modify bridge0 connection.autoconnect-slaves 1
```

g. Active the `bridge0` connection:

```bash
# nmcli connection up bridge0
```

2. On the RHEL router in network B:

   a. Create a bridge interface named `bridge0`:

```bash
# nmcli connection add type bridge con-name bridge0 ifname bridge0
```

   b. Configure the IP settings of the bridge:

```bash
# nmcli connection modify bridge0 ipv4.addresses '192.0.2.2/24'
# nmcli connection modify bridge0 ipv4.method manual
```

   c. Add a new connection profile for the interface that is connected to local network to the bridge:

```bash
# nmcli connection add type ethernet slave-type bridge con-name bridge0-port1 ifname enp1s0 master bridge0
```

   d. Add a new connection profile for the GRETAP tunnel interface to the bridge:

```bash
# nmcli connection add type ip-tunnel ip-tunnel.mode gretap slave-type bridge con-name bridge0-port2 ifname gretap1 remote 203.0.113.10 local 198.51.100.5 master bridge0
```

The `remote` and `local` parameters set the public IP addresses of the remote and the local routers.

e. Optional: Disable the Spanning Tree Protocol (STP) if you do not need it:

```bash
# nmcli connection modify bridge0 bridge.stp no
```
f. Configure that activating the `bridge0` connection automatically activates the ports of the bridge:

```
# nmcli connection modify bridge0 connection.autoconnect-slaves 1
```

g. Active the `bridge0` connection:

```
# nmcli connection up bridge0
```

**Verification**

1. On both routers, verify that the `enp1s0` and `gretap1` connections are connected and that the `CONNECTION` column displays the connection name of the port:

```
# nmcli device

DEVICE  TYPE      STATE      CONNECTION

...  
bridge0  bridge   connected  bridge0
enp1s0   ethernet   connected  bridge0-port1
gretap1  iptunnel   connected  bridge0-port2
```

2. From each RHEL router, ping the IP address of the internal interface of the other router:

   a. On Router A, ping `192.0.2.2`:

      ```
      # ping 192.0.2.2
      ```

   b. On Router B, ping `192.0.2.1`:

      ```
      # ping 192.0.2.1
      ```

**Additional resources**

- `nmcli(1)` man page
- `nm-settings(5)` man page

**9.4. ADDITIONAL RESOURCES**

- `ip-link(8)` man page
CHAPTER 10. USING A VXLAN TO CREATE A VIRTUAL LAYER-2 DOMAIN FOR VMS

A virtual extensible LAN (VXLAN) is a networking protocol that tunnels layer-2 traffic over an IP network using the UDP protocol. For example, certain virtual machines (VMs), that are running on different hosts can communicate over a VXLAN tunnel. The hosts can be in different subnets or even in different data centers around the world. From the perspective of the VMs, other VMs in the same VXLAN are within the same layer-2 domain:

In this example, RHEL-host-A and RHEL-host-B use a bridge, br0, to connect the virtual network of a VM on each host with a VXLAN named vxlan10. Due to this configuration, the VXLAN is invisible to the VMs, and the VMs do not require any special configuration. If you later connect more VMs to the same virtual network, the VMs are automatically members of the same virtual layer-2 domain.

**IMPORTANT**

Just as normal layer-2 traffic, data in a VXLAN is not encrypted. For security reasons, use a VXLAN over a VPN or other types of encrypted connections.

10.1. BENEFITS OF VXLANS

A virtual extensible LAN (VXLAN) provides the following major benefits:

- VXLANs use a 24-bit ID. Therefore, you can create up to 16,777,216 isolated networks. For example, a virtual LAN (VLAN), supports only 4,096 isolated networks.

- VXLANs use the IP protocol. This enables you to route the traffic and virtually run systems in different networks and locations within the same layer-2 domain.

- Unlike most tunnel protocols, a VXLAN is not only a point-to-point network. A VXLAN can learn the IP addresses of the other endpoints either dynamically or use statically-configured forwarding entries.
• Certain network cards support UDP tunnel-related offload features.

Additional resources
• /usr/share/doc/kernel-doc-<kernel_version>/Documentation/networking/vxlan.rst provided by the kernel-doc package

10.2. CONFIGURING THE ETHERNET INTERFACE ON THE HOSTS

To connect a RHEL VM host to the Ethernet, create a network connection profile, configure the IP settings, and activate the profile.

Run this procedure on both RHEL hosts, and adjust the IP address configuration accordingly.

Prerequisites
• The host is connected to the Ethernet.

Procedure
1. Add a new Ethernet connection profile to NetworkManager:

   ```
   # nmcli connection add con-name Example ifname enp1s0 type ethernet
   ```

2. Configure the IPv4 settings:

   ```
   # nmcli connection modify Example ipv4.addresses 198.51.100.2/24 ipv4.method manual ipv4.gateway 198.51.100.254 ipv4.dns 198.51.100.200 ipv4.dns-search example.com
   ```

   Skip this step if the network uses DHCP.

3. Activate the Example connection:

   ```
   # nmcli connection up Example
   ```

Verification
1. Display the status of the devices and connections:

   ```
   # nmcli device status
   DEVICE TYPE STATE CONNECTION
   enp1s0 ethernet connected Example
   ```

2. Ping a host in a remote network to verify the IP settings:

   ```
   # ping RHEL-host-B.example.com
   ```

   Note that you cannot ping the other VM host before you have configured the network on that host as well.

Additional resources
10.3. CREATING A NETWORK BRIDGE WITH A VXLAN ATTACHED

To make a virtual extensible LAN (VXLAN) invisible to virtual machines (VMs), create a bridge on a host, and attach the VXLAN to the bridge. Use NetworkManager to create both the bridge and the VXLAN. You do not add any traffic access point (TAP) devices of the VMs, typically named `vnet*` on the host, to the bridge. The `libvirtd` service adds them dynamically when the VMs start.

Run this procedure on both RHEL hosts, and adjust the IP addresses accordingly.

Procedure

1. Create the bridge `br0`:

   ```
   # nmcli connection add type bridge con-name br0 ifname br0 ipv4.method disabled ipv6.method disabled
   ```

   This command sets no IPv4 and IPv6 addresses on the bridge device, because this bridge works on layer 2.

2. Create the VXLAN interface and attach it to `br0`:

   ```
   # nmcli connection add type vxlan slave-type bridge con-name br0-vxlan10 ifname vxlan10 id 10 local 198.51.100.2 remote 203.0.113.1 master br0
   ```

   This command uses the following settings:
   - `id 10`: Sets the VXLAN identifier.
   - `local 198.51.100.2`: Sets the source IP address of outgoing packets.
   - `remote 203.0.113.1`: Sets the unicast or multicast IP address to use in outgoing packets when the destination link layer address is not known in the VXLAN device forwarding database.
   - `master br0`: Sets this VXLAN connection to be created as a port in the `br0` connection.

   By default, NetworkManager uses 8472 as the destination port. If the destination port is different, additionally, pass the `destination-port <port_number>` option to the command.

3. Activate the `br0` connection profile:

   ```
   # nmcli connection up br0
   ```

4. Open port 8472 for incoming UDP connections in the local firewall:

   ```
   # firewall-cmd --permanent --add-port=8472/udp
   # firewall-cmd --reload
   ```

Verification
Display the forwarding table:

```bash
# bridge fdb show dev vxlan10
2a:53:bd:d5:b3:0a master br0 permanent
00:00:00:00:00:00 dst 203.0.113.1 self permanent...
```

Additional resources

- `nm-settings(5)` man page

### 10.4. CREATING A VIRTUAL NETWORK IN LIBVIRT WITH AN EXISTING BRIDGE

To enable virtual machines (VM) to use the `br0` bridge with the attached virtual extensible LAN (VXLAN), first add a virtual network to the `libvirt` service that uses this bridge.

**Prerequisites**

- You installed the `libvirt` package.
- You started and enabled the `libvirtd` service.
- You configured the `br0` device with the VXLAN on RHEL.

**Procedure**

1. Create the `~/vxlan10-bridge.xml` file with the following content:

   ```xml
   <network>
   <name>vxlan10-bridge</name>
   <forward mode="bridge" />
   <bridge name="br0" />
   </network>
   ```

2. Use the `~/vxlan10-bridge.xml` file to create a new virtual network in `libvirt`:

   ```bash
   # virsh net-define ~/vxlan10-bridge.xml
   ```

3. Remove the `~/vxlan10-bridge.xml` file:

   ```bash
   # rm ~/vxlan10-bridge.xml
   ```

4. Start the `vxlan10-bridge` virtual network:

   ```bash
   # virsh net-start vxlan10-bridge
   ```

5. Configure the `vxlan10-bridge` virtual network to start automatically when the `libvirtd` service starts:

   ```bash
   # virsh net-autostart vxlan10-bridge
   ```
Verification

- Display the list of virtual networks:

```
# virsh net-list
Name              State    Autostart   Persistent
----------------------------------------------------
vxlan10-bridge    active   yes         yes
...`
2. Display the interfaces attached to the `vxlan10-bridge` bridge:

```bash
# ip link show master vxlan10-bridge
18: vxlan10: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue master
    br0 state UNKNOWN mode DEFAULT group default qlen 1000
    link/ether 2a:53:bd:d5:b3:0a brd ff:ff:ff:ff:ff:ff
19: vnet1: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc noqueue master
    br0 state UNKNOWN mode DEFAULT group default qlen 1000
    link/ether 52:54:00:c5:98:1c brd ff:ff:ff:ff:ff:ff
```

Note that the `libvirtd` service dynamically updates the bridge’s configuration. When you start a VM which uses the `vxlan10-bridge` network, the corresponding `vnet*` device on the host appears as a port of the bridge.

3. Use address resolution protocol (ARP) requests to verify whether VMs are in the same VXLAN:
   
   a. Start two or more VMs in the same VXLAN.
   
   b. Send an ARP request from one VM to the other one:

   ```bash
   # arping -c 1 192.0.2.2
   ARPING 192.0.2.2 from 192.0.2.1 enp1s0
   Unicast reply from 192.0.2.2 [52:54:00:c5:98:1c] 1.450ms
   Sent 1 probe(s) (0 broadcast(s))
   Received 1 response(s) (0 request(s), 0 broadcast(s))
   ```

   If the command shows a reply, the VM is in the same layer-2 domain and, in this case in the same VXLAN.

   Install the `iputils` package to use the `arping` utility.

Additional resources

- `virt-install(1)` man page
- `virt-xml(1)` man page
- `virsh(1)` man page
- `arping(8)` man page
CHAPTER 11. MANAGING WIFI CONNECTIONS

RHEL provides multiple utilities and applications to configure and connect to wifi networks, for example:

- Use the `nmcli` utility to configure connections by using the command line.
- Use the `nmtui` application to configure connections in a text-based user interface.
- Use the GNOME system menu to quickly connect to wifi networks that do not require any configuration.
- Use the `GNOME Settings` application to configure connections by using the GNOME application.
- Use the `nm-connection-editor` application to configure connections in a graphical user interface.
- Use the `network` RHEL system role to automate the configuration of connections on one or multiple hosts.

### 11.1. SUPPORTED WIFI SECURITY TYPES

Depending on the security type a wifi network supports, you can transmitted data more or less securely.

**WARNING**

Do not connect to wifi networks that do not use encryption or which support only the insecure WEP or WPA standards.

Red Hat Enterprise Linux 9 supports the following wifi security types:

- **None**: Encryption is disabled, and data is transferred in plain text over the network.
- **Enhanced Open**: With opportunistic wireless encryption (OWE), devices negotiate unique pairwise master keys (PMK) to encrypt connections in wireless networks without authentication.
- **LEAP**: The Lightweight Extensible Authentication Protocol, which was developed by Cisco, is a proprietary version of the extensible authentication protocol (EAP).
- **WPA & WPA2 Personal**: In personal mode, the Wi-Fi Protected Access (WPA) and Wi-Fi Protected Access 2 (WPA2) authentication methods use a pre-shared key.
- **WPA & WPA2 Enterprise**: In enterprise mode, WPA and WPA2 use the EAP framework and authenticate users to a remote authentication dial-in user service (RADIUS) server.
- **WPA3 Personal**: Wi-Fi Protected Access 3 (WPA3) Personal uses simultaneous authentication of equals (SAE) instead of pre-shared keys (PSK) to prevent dictionary attacks. WPA3 uses perfect forward secrecy (PFS).

### 11.2. CONNECTING TO A WIFI NETWORK BY USING NMCLI
You can use the `nmcli` utility to connect to a wifi network. When you attempt to connect to a network for the first time, the utility automatically creates a NetworkManager connection profile for it. If the network requires additional settings, such as static IP addresses, you can then modify the profile after it has been automatically created.

**Prerequisites**

- A wifi device is installed on the host.
- The wifi device is enabled, if it has a hardware switch.

**Procedure**

1. If the wifi radio has been disabled in NetworkManager, enable this feature:

   ```shell
   # nmcli radio wifi on
   ```

2. Optional: Display the available wifi networks:

   ```shell
   # nmcli device wifi list
   IN-USE   BSSID              SSID          MODE   CHAN  RATE        SIGNAL  BARS  SECURITY
   00:53:00:2F:3B:08  Office        Infra  44    270 Mbit/s  57▂▄▆_  WPA2 WPA3
   00:53:00:15:03:BF  --            Infra  1     130 Mbit/s  48▂▄__  WPA2 WPA3
   ```

   The service set identifier (SSID) column contains the names of the networks. If the column shows --, the access point of this network does not broadcast an SSID.

3. Connect to the wifi network:

   ```shell
   # nmcli device wifi connect Office --ask
   Password: wifi-password
   ```

   If you prefer to set the password in the command instead of entering it interactively, use the `password wifi-password` option in the command instead of `--ask`:

   ```shell
   # nmcli device wifi connect Office wifi-password
   ```

   Note that, if the network requires static IP addresses, NetworkManager fails to activate the connection at this point. You can configure the IP addresses in later steps.

4. If the network requires static IP addresses:

   a. Configure the IPv4 address settings, for example:

   ```shell
   # nmcli connection modify Office ipv4.method manual ipv4.addresses 192.0.2.1/24
   ipv4.gateway 192.0.2.254 ipv4.dns 192.0.2.200 ipv4.dns-search example.com
   ```

   b. Configure the IPv6 address settings, for example:

   ```shell
   # nmcli connection modify Office ipv6.method manual ipv6.addresses
   2001:db8:1::1/64 ipv6.gateway 2001:db8:1::ffe ipv6.dns 2001:db8:1::ffbb ipv6.dns-search example.com
   ```

5. Re-activate the connection:
# nmcli connection up Office

Verification

1. Display the active connections:

   ```
   # nmcli connection show --active
   NAME    ID                                    TYPE  DEVICE
   Office  2501eb7e-7b16-4dc6-97ef-7cc460139a58 wifi wlp0s20f3
   ```

   If the output lists the wifi connection you have created, the connection is active.

2. Ping a hostname or IP address:

   ```
   # ping -c 3 example.com
   ```

Additional resources

- [nm-settings-nmcli(5) man page](#)

11.3. CONNECTING TO A WIFI NETWORK BY USING THE GNOME SYSTEM MENU

You can use the GNOME system menu to connect to a wifi network. When you connect to a network for the first time, GNOME creates a NetworkManager connection profile for it. If you configure the connection profile to not automatically connect, you can also use the GNOME system menu to manually connect to a wifi network with an existing NetworkManager connection profile.

**NOTE**

Using the GNOME system menu to establish a connection to a wifi network for the first time has certain limitations. For example, you can not configure IP address settings. In this case first configure the connections:

- In the [GNOME settings](#) application
- In the [nm-connection-editor](#) application
- Using [nmcli](#) commands

Prerequisites

- A wifi device is installed on the host.
- The wifi device is enabled, if it has a hardware switch.

Procedure

1. Open the system menu on the right side of the top bar.
2. Expand the **Wi-Fi Not Connected** entry.
3. Click **Select Network**:
4. Select the wifi network you want to connect to.

5. Click Connect.

6. If this is the first time you connect to this network, enter the password for the network, and click Connect.

Verification

1. Open the system menu on the right side of the top bar, and verify that the wifi network is connected:

   ![Wi-Fi Network Connected](image)

   If the network appears in the list, it is connected.

2. Ping a hostname or IP address:

   ```
   # ping -c 3 example.com
   ```

11.4. CONNECTING TO A WIFI NETWORK BY USING THE GNOME SETTINGS APPLICATION

You can use the GNOME settings application, also named gnome-control-center, to connect to a wifi network and configure the connection. When you connect to the network for the first time, GNOME creates a NetworkManager connection profile for it.

In GNOME settings, you can configure wifi connections for all wifi network security types that RHEL supports.
Prerequisites

- A wifi device is installed on the host.
- The wifi device is enabled, if it has a hardware switch.

Procedure

1. Press the Super key, type Wi-Fi, and press Enter.
2. Click on the name of the wifi network you want to connect to.
3. Enter the password for the network, and click Connect.
4. If the network requires additional settings, such as static IP addresses or a security type other than WPA2 Personal:
   a. Click the gear icon next to the network’s name.
   b. Optional: Configure the network profile on the Details tab to not automatically connect. If you deactivate this feature, you must always manually connect to the network, for example, by using GNOME settings or the GNOME system menu.
   d. On the Security tab, select the authentication of the network, such as WPA3 Personal, and enter the password. Depending on the selected security, the application shows additional fields. Fill them accordingly. For details, ask the administrator of the wifi network.
   e. Click Apply.

Verification

1. Open the system menu on the right side of the top bar, and verify that the wifi network is connected:

   ![Image showing the system menu with a connected wifi network]

   If the network appears in the list, it is connected.
2. Ping a hostname or IP address:

   # ping -c 3 example.com

11.5. CONFIGURING A WIFI CONNECTION BY USING NMTUI
The **nmtui** application provides a text-based user interface for NetworkManager. You can use **nmtui** to connect to a wifi network.

**NOTE**

In **nmtui**:

- Navigate by using the cursor keys.
- Press a button by selecting it and hitting **Enter**.
- Select and deselect checkboxes by using **Space**.

**Procedure**

1. If you do not know the network device name you want to use in the connection, display the available devices:

   ```
   # nmcli device status
   DEVICE     TYPE      STATE                   CONNECTION
   wlp2s0     wifi      unavailable             --
   ...
   ```

2. Start **nmtui**:

   ```
   # nmtui
   ```

3. Select **Edit a connection**, and press **Enter**.

4. Press the **Add** button.

5. Select **Wi-Fi** from the list of network types, and press **Enter**.

6. Optional: Enter a name for the NetworkManager profile to be created. On hosts with multiple profiles, a meaningful name makes it easier to identify the purpose of a profile.

7. Enter the network device name into the **Device** field.

8. Enter the name of the Wi-Fi network, the Service Set Identifier (SSID), into the **SSID** field.

9. Leave the **Mode** field set to its default, **Client**.

10. Select the **Security** field, press **Enter**, and set the authentication type of the network from the list. Depending on the authentication type you have selected, **nmtui** displays different fields.

11. Fill the authentication type-related fields.

12. If the Wi-Fi network requires static IP addresses:

   a. Press the **Automatic** button next to the protocol, and select **Manual** from the displayed list.

   b. Press the **Show** button next to the protocol you want to configure to display additional fields, and fill them.
13. Press the OK button to create and automatically activate the new connection.

14. Press the Back button to return to the main menu.

15. Select Quit, and press Enter to close the nmtui application.

Verification

1. Display the active connections:

   # nmcli connection show --active

   NAME    ID                                    TYPE  DEVICE
   Office  2501eb7e-7b16-4dc6-97ef-7cc460139a58 wifi wlp0s20f3

   If the output lists the wifi connection you have created, the connection is active.

2. Ping a hostname or IP address:

   # ping -c 3 example.com

11.6. CONFIGURING A WIFI CONNECTION BY USING NM-CONNECTION-EDITOR

You can use the nm-connection-editor application to create a connection profile for a wireless network. In this application you can configure all wifi network authentication types that RHEL supports.

By default, NetworkManager enables the auto-connect feature for connection profiles and automatically connects to a saved network if it is available.

Prerequisites
- A wifi device is installed on the host.
- The wifi device is enabled, if it has a hardware switch.

**Procedure**

1. Open a terminal and enter:
   
   ```bash
   # nm-connection-editor
   ```

2. Click the + button to add a new connection.

3. Select the **Wi-Fi** connection type, and click **Create**.

4. Optional: Set a name for the connection profile.

5. Optional: Configure the network profile on the **General** tab to not automatically connect. If you deactivate this feature, you must always manually connect to the network, for example, by using **GNOME settings** or the GNOME system menu.

6. On the **Wi-Fi** tab, enter the service set identifier (SSID) in the **SSID** field.

7. On the **Wi-Fi Security** tab, select the authentication type for the network, such as **WPA3 Personal**, and enter the password. Depending on the selected security, the application shows additional fields. Fill them accordingly. For details, ask the administrator of the wifi network.

8. Configure IPv4 settings on the **IPv4** tab, and IPv6 settings on the **IPv6** tab.

9. Click **Save**.

10. Close the **Network Connections** window.

**Verification**

1. Open the system menu on the right side of the top bar, and verify that the wifi network is connected:

   ![Network Connections](image)

   If the network appears in the list, it is connected.

2. Ping a hostname or IP address:

   ```bash
   # ping -c 3 example.com
   ```
11.7. CONFIGURING A WIFI CONNECTION WITH 802.1X NETWORK AUTHENTICATION BY USING THE NETWORK RHEL SYSTEM ROLE

Using RHEL system role, you can automate the creation of a wifi connection. For example, you can remotely add a wireless connection profile for the `wlp1s0` interface using an Ansible Playbook. The created profile uses the 802.1X standard to authenticate the client to a wifi network. The playbook configures the connection profile to use DHCP. To configure static IP settings, adapt the parameters in the `ip` dictionary accordingly.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
- The network supports 802.1X network authentication.
- You installed the `wpa_supplicant` package on the managed node.
- DHCP is available in the network of the managed node.
- The following files required for TLS authentication exist on the control node:
  - The client key is stored in the `/srv/data/client.key` file.
  - The client certificate is stored in the `/srv/data/client.crt` file.
  - The CA certificate is stored in the `/srv/data/ca.crt` file.

Procedure

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure a wifi connection with 802.1X authentication
  hosts: managed-node-01.example.com
  tasks:
    - name: Copy client key for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/client.key'
        dest: '/etc/pki/tls/private/client.key'
        mode: 0400

    - name: Copy client certificate for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/client.crt'
        dest: '/etc/pki/tls/certs/client.crt'

    - name: Copy CA certificate for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/ca.crt'
        dest: '/etc/pki/ca-trust/source/anchors/ca.crt'

    - block:
```

Prerequisites

You have prepared the control node and the managed nodes
You are logged in to the control node as a user who can run playbooks on the managed nodes.
The account you use to connect to the managed nodes has `sudo` permissions on them.
The network supports 802.1X network authentication.
You installed the `wpa_supplicant` package on the managed node.
DHCP is available in the network of the managed node.
The following files required for TLS authentication exist on the control node:
  - The client key is stored in the `/srv/data/client.key` file.
  - The client certificate is stored in the `/srv/data/client.crt` file.
  - The CA certificate is stored in the `/srv/data/ca.crt` file.

Procedure

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure a wifi connection with 802.1X authentication
  hosts: managed-node-01.example.com
  tasks:
    - name: Copy client key for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/client.key'
        dest: '/etc/pki/tls/private/client.key'
        mode: 0400

    - name: Copy client certificate for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/client.crt'
        dest: '/etc/pki/tls/certs/client.crt'

    - name: Copy CA certificate for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/ca.crt'
        dest: '/etc/pki/ca-trust/source/anchors/ca.crt'

    - block:
```

Prerequisites

You have prepared the control node and the managed nodes
You are logged in to the control node as a user who can run playbooks on the managed nodes.
The account you use to connect to the managed nodes has `sudo` permissions on them.
The network supports 802.1X network authentication.
You installed the `wpa_supplicant` package on the managed node.
DHCP is available in the network of the managed node.
The following files required for TLS authentication exist on the control node:
  - The client key is stored in the `/srv/data/client.key` file.
  - The client certificate is stored in the `/srv/data/client.crt` file.
  - The CA certificate is stored in the `/srv/data/ca.crt` file.

Procedure

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure a wifi connection with 802.1X authentication
  hosts: managed-node-01.example.com
  tasks:
    - name: Copy client key for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/client.key'
        dest: '/etc/pki/tls/private/client.key'
        mode: 0400

    - name: Copy client certificate for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/client.crt'
        dest: '/etc/pki/tls/certs/client.crt'

    - name: Copy CA certificate for 802.1X authentication
      ansible.builtin.copy:
        src: '/srv/data/ca.crt'
        dest: '/etc/pki/ca-trust/source/anchors/ca.crt'

    - block:
```
These settings define a wifi connection profile for the **wlp1s0** interface. The profile uses 802.1X standard to authenticate the client to the wifi network. The connection retrieves IPv4 addresses, IPv6 addresses, default gateway, routes, DNS servers, and search domains from a DHCP server and IPv6 stateless address autoconfiguration (SLAAC).

2. Validate the playbook syntax:

```bash
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```bash
$ ansible-playbook ~/playbook.yml
```

Additional resources

- /usr/share/ansible/roles/rhel-system-roles.network/README.md file
- /usr/share/doc/rhel-system-roles/network/ directory

### 11.8. Configuring a Wifi Connection with 802.1X Network Authentication in an Existing Profile by Using NMCLI

Using the **nmcli** utility, you can configure the client to authenticate itself to the network. For example, you can configure Protected Extensible Authentication Protocol (PEAP) authentication with the Microsoft Challenge-Handshake Authentication Protocol version 2 (MSCHAPv2) in an existing NetworkManager wifi connection profile named **wlp1s0**.
Prerequisites

- The network must have 802.1X network authentication.
- The wifi connection profile exists in NetworkManager and has a valid IP configuration.
- If the client is required to verify the certificate of the authenticator, the Certificate Authority (CA) certificate must be stored in the `/etc/pki/ca-trust/source/anchors/` directory.
- The `wpa_supplicant` package is installed.

Procedure

1. Set the wifi security mode to `wpa-eap`, the Extensible Authentication Protocol (EAP) to `peap`, the inner authentication protocol to `mschapv2`, and the user name:

   ```shell
   # nmcli connection modify wlp1s0 wireless-security.key-mgmt wpa-eap 802-1x.eap peap 802-1x.phase2-auth mschapv2 802-1x.identity user_name
   ```

   Note that you must set the `wireless-security.key-mgmt`, `802-1x.eap`, `802-1x.phase2-auth`, and `802-1x.identity` parameters in a single command.

2. Optionally, store the password in the configuration:

   ```shell
   # nmcli connection modify wlp1s0 802-1x.password password
   ```

   **IMPORTANT**

   By default, NetworkManager stores the password in plain text in the `/etc/sysconfig/network-scripts/keys-connection_name` file, which is readable only by the `root` user. However, plain text passwords in a configuration file can be a security risk.

   To increase the security, set the `802-1x.password-flags` parameter to `0x1`. With this setting, on servers with the GNOME desktop environment or the `nm-applet` running, NetworkManager retrieves the password from these services. In other cases, NetworkManager prompts for the password.

3. If the client needs to verify the certificate of the authenticator, set the `802-1x.ca-cert` parameter in the connection profile to the path of the CA certificate:

   ```shell
   # nmcli connection modify wlp1s0 802-1x.ca-cert /etc/pki/ca-trust/source/anchors/ca.crt
   ```

   **NOTE**

   For security reasons, Red Hat recommends the certificate of the authenticator to enable clients to validate the identity of the authenticator.

4. Activate the connection profile:

   ```shell
   # nmcli connection up wlp1s0
   ```
11.9. MANUALLY SETTING THE WIRELESS REGULATORY DOMAIN

On RHEL, a udev rule executes the setregdomain utility to set the wireless regulatory domain. The utility then provides this information to the kernel.

By default, setregdomain attempts to determine the country code automatically. If this fails, the wireless regulatory domain setting might be wrong. To work around this problem, you can manually set the country code.

IMPORTANT

Manually setting the regulatory domain disables the automatic detection. Therefore, if you later use the computer in a different country, the previously configured setting might no longer be correct. In this case, remove the /etc/sysconfig/regdomain file to switch back to automatic detection or use this procedure to manually update the regulatory domain setting again.

Procedure

1. Optional: Display the current regulatory domain settings:

   # iw reg get
   global
country US: DFS-FCC
   ...

2. Create the /etc/sysconfig/regdomain file with the following content:

   COUNTRY=<country_code>

   Set the COUNTRY variable to an ISO 3166-1 alpha2 country code, such as DE for Germany or US for the United States of America.

3. Set the regulatory domain:

   # setregdomain

Verification

- Display the regulatory domain settings:

  # iw reg get
Additional resources

- `setregdomain(1)` man page
- `iw(8)` man page
- `regulatory.bin(5)` man page
- `ISO 3166 Country Codes`
CHAPTER 12. CONFIGURING RHEL AS A WPA2 OR WPA3 PERSONAL ACCESS POINT

On a host with a wifi device, you can use NetworkManager to configure this host as an access point. Wi-Fi Protected Access 2 (WPA2) and Wi-Fi Protected Access 3 (WPA3) Personal provide secure authentication methods, and wireless clients can use a pre-shared key (PSK) to connect to the access point and use services on the RHEL host and in the network.

When you configure an access point, NetworkManager automatically:

- Configures the `dnsmasq` service to provide DHCP and DNS services for clients
- Enables IP forwarding
- Adds `nftables` firewall rules to masquerade traffic from the wifi device and configures IP forwarding

**Prerequisites**

- The wifi device supports running in access point mode.
- The wifi device is not in use.
- The host has internet access.

**Procedure**

1. List the wifi devices to identify the one that should provide the access point:

   ```sh
   # nmcli device status | grep wifi
   wlp0s20f3   wifi   disconnected   --
   ```

2. Verify that the device supports the access point mode:

   ```sh
   # nmcli -f WIFI-PROPERTIES.AP device show wlp0s20f3
   WIFI-PROPERTIES.AP: yes
   ```

   To use a wifi device as an access point, the device must support this feature.

3. Install the `dnsmasq` and `NetworkManager-wifi` packages:

   ```sh
   # dnf install dnsmasq NetworkManager-wifi
   ```

   NetworkManager uses the `dnsmasq` service to provide DHCP and DNS services to clients of the access point.

4. Create the initial access point configuration:

   ```sh
   # nmcli device wifi hotspot ifname wlp0s20f3 con-name Example-Hotspot ssid Example-Hotspot password "password"
   ```

   This command creates a connection profile for an access point on the `wlp0s20f3` device that provides WPA2 and WPA3 Personal authentication. The name of the wireless network, the Service Set Identifier (SSID), is `Example-Hotspot` and uses the pre-shared key `password`. 
5. Optional: Configure the access point to support only WPA3:

```
# nmcli connection modify Example-Hotspot 802-11-wireless-security.key-mgmt sae
```

6. By default, NetworkManager uses the IP address `10.42.0.1` for the wifi device and assigns IP addresses from the remaining `10.42.0.0/24` subnet to clients. To configure a different subnet and IP address, enter:

```
# nmcli connection modify Example-Hotspot ipv4.addresses 192.0.2.254/24
```

The IP address you set, in this case `192.0.2.254`, is the one that NetworkManager assigns to the wifi device. Clients will use this IP address as default gateway and DNS server.

7. Activate the connection profile:

```
# nmcli connection up Example-Hotspot
```

**Verification**

1. On the server:
   a. Verify that NetworkManager started the `dnsmasq` service and that the service listens on port 67 (DHCP) and 53 (DNS):

   ```
   # ss -tulpn | grep ":53|:67"
   udp  UNCONN 0  0 10.42.0.1:53  0.0.0.0:*  users:("dnsmasq",pid=55905,fd=6)
   udp  UNCONN 0  0 0.0.0.0:67  0.0.0.0:*  users:("dnsmasq",pid=55905,fd=4)
   tcp   LISTEN 0  32 10.42.0.1:53  0.0.0.0:*  users:("dnsmasq",pid=55905,fd=7)
   ```

   b. Display the `nftables` rule set to ensure that NetworkManager enabled forwarding and masquerading for traffic from the `10.42.0.0/24` subnet:

   ```
   # nft list ruleset
table ip nm-shared-wlp0s20f3 {
  chain nat_postrouting {
    type nat hook postrouting priority srcnat; policy accept;
    ip saddr 10.42.0.0/24 ip daddr !== 10.42.0.0/24 masquerade
  }
  chain filter_forward {
    type filter hook forward priority filter; policy accept;
    ip daddr 10.42.0.0/24 ofname "wlp0s20f3" ct state { established, related } accept
    ip saddr 10.42.0.0/24 ofname "wlp0s20f3" accept
    ofname "wlp0s20f3" ofname "wlp0s20f3" accept
    ofname "wlp0s20f3" reject
    ofname "wlp0s20f3" reject
  }
}
```

2. On a client with a wifi adapter:
   a. Display the list of available networks:

   ```
   # nmcli device wifi
   ```
<table>
<thead>
<tr>
<th>IN-USE BSSID</th>
<th>SSID</th>
<th>MODE</th>
<th>CHAN</th>
<th>RATE</th>
<th>SIGNAL</th>
<th>BARS</th>
<th>SECURITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>00:53:00:88:29:04</td>
<td>Example-Hotspot</td>
<td>Infra</td>
<td>11</td>
<td>130 Mbit/s</td>
<td>62</td>
<td>▂ ▄ ▆ _</td>
<td>WPA3</td>
</tr>
</tbody>
</table>

b. Connect to the Example-Hotspot wireless network. See Managing Wi-Fi connections.

c. Ping a host on the remote network or the internet to verify that the connection works:

```bash
# ping -c 3 www.redhat.com
```

Additional resources

- `nm-settings(5)` man page
CHAPTER 13. USING MACSEC TO ENCRYPT LAYER-2 TRAFFIC IN THE SAME PHYSICAL NETWORK

You can use MACsec to secure the communication between two devices (point-to-point). For example, your branch office is connected over a Metro-Ethernet connection with the central office, you can configure MACsec on the two hosts that connect the offices to increase the security.

Media Access Control security (MACsec) is a layer 2 protocol that secures different traffic types over the Ethernet links including:

- dynamic host configuration protocol (DHCP)
- address resolution protocol (ARP)
- Internet Protocol version 4 / 6 (IPv4 / IPv6) and
- any traffic over IP such as TCP or UDP

MACsec encrypts and authenticates all traffic in LANs, by default with the GCM-AES-128 algorithm, and uses a pre-shared key to establish the connection between the participant hosts. If you want to change the pre-shared key, you need to update the NM configuration on all hosts in the network that uses MACsec.

A MACsec connection uses an Ethernet device, such as an Ethernet network card, VLAN, or tunnel device, as parent. You can either set an IP configuration only on the MACsec device to communicate with other hosts only using the encrypted connection, or you can also set an IP configuration on the parent device. In the latter case, you can use the parent device to communicate with other hosts using an unencrypted connection and the MACsec device for encrypted connections.

MACsec does not require any special hardware. For example, you can use any switch, except if you want to encrypt traffic only between a host and a switch. In this scenario, the switch must also support MACsec.

In other words, there are 2 common methods to configure MACsec;

- host to host and
- host to switch then switch to other host(s)

**IMPORTANT**

You can use MACsec only between hosts that are in the same (physical or virtual) LAN.

13.1. CONFIGURING A MACSEC CONNECTION USING NMCLI

You can configure Ethernet interfaces to use MACsec using the `nmcli` utility. For example, you can create a MACsec connection between two hosts that are connected over Ethernet.

**Procedure**

1. On the first host on which you configure MACsec:
   - Create the connectivity association key (CAK) and connectivity-association key name (CKN) for the pre-shared key:
a. Create a 16-byte hexadecimal CAK:

```
# dd if=/dev/urandom count=16 bs=1 2> /dev/null | hexdump -e '1/2 "%04x"'
50b71a8ef0bd5751ea76de6d6c98c03a
```

b. Create a 32-byte hexadecimal CKN:

```
# dd if=/dev/urandom count=32 bs=1 2> /dev/null | hexdump -e '1/2 "%04x"'
f2b4297d39da7330910a74abc0449feb45b5c0b9fc23df1430e1898fcf1c4550
```

2. On both hosts you want to connect over a MACsec connection:

3. Create the MACsec connection:

```
# nmcli connection add type macsec con-name macsec0 ifname macsec0
    connection.autoconnect yes macsec.parent enp1s0 macsec.mode psk
    macsec.mka-cak 50b71a8ef0bd5751ea76de6d6c98c03a
    macsec.mka-ckn f2b4297d39da7330910a74abc0449feb45b5c0b9fc23df1430e1898fcf1c4550
```

Use the CAK and CKN generated in the previous step in the `macsec.mka-cak` and `macsec.mka-ckn` parameters. The values must be the same on every host in the MACsec-protected network.

4. Configure the IP settings on the MACsec connection.

   a. Configure the IPv4 settings. For example, to set a static IPv4 address, network mask, default gateway, and DNS server to the `macsec0` connection, enter:

   ```
   # nmcli connection modify macsec0 ipv4.method manual
   ipv4.addresses '192.0.2.1/24' ipv4.gateway '192.0.2.254' ipv4.dns '192.0.2.253'
   ```

   b. Configure the IPv6 settings. For example, to set a static IPv6 address, network mask, default gateway, and DNS server to the `macsec0` connection, enter:

   ```
   # nmcli connection modify macsec0 ipv6.method manual
   ```

5. Activate the connection:

   ```
   # nmcli connection up macsec0
   ```

Verification

1. Verify that the traffic is encrypted:

   ```
   # tcpdump -nn -i enp1s0
   ```

2. Optional: Display the unencrypted traffic:

   ```
   # tcpdump -nn -i macsec0
   ```

3. Display MACsec statistics:
4. Display individual counters for each type of protection: integrity-only (encrypt off) and encryption (encrypt on)

```
# ip macsec show
```

13.2. CONFIGURING A MACSEC CONNECTION USING NMSTATECTL

You can configure Ethernet interfaces to use MACsec through the `nmstatectl` utility in a declarative way. For example, in a YAML file, you describe the desired state of your network, which is supposed to have a MACsec connection between two hosts connected over Ethernet. The `nmstatectl` utility interprets the YAML file and deploys persistent and consistent network configuration across the hosts.

Using the MACsec security standard for securing communication at the link layer, also known as layer 2 of the Open Systems Interconnection (OSI) model provides the following notable benefits:

- Encryption at layer 2 eliminates the need for encrypting individual services at layer 7. This reduces the overhead associated with managing a large number of certificates for each endpoint on each host.
- Point-to-point security between directly connected network devices such as routers and switches.
- No changes needed for applications and higher-layer protocols.

Prerequisites

- A physical or virtual Ethernet Network Interface Controller (NIC) exists in the server configuration.
- The `nmstate` package is installed.

Procedure

1. On the first host on which you configure MACsec, create the connectivity association key (CAK) and connectivity-association key name (CKN) for the pre-shared key:
   a. Create a 16-byte hexadecimal CAK:
      ```
      # dd if=/dev/urandom count=16 bs=1 2> /dev/null | hexdump -e '1/2 "%04x"'
      50b71a8ef0bd5751ea76de6d6c98c03a
      ```
   b. Create a 32-byte hexadecimal CKN:
      ```
      # dd if=/dev/urandom count=32 bs=1 2> /dev/null | hexdump -e '1/2 "%04x"'
      f2b4297d39da7330910a74abc0449febeb45b5c0b9fc23df1430e1898f1c4550
      ```

2. On both hosts that you want to connect over a MACsec connection, complete the following steps:
   a. Create a YAML file, for example `create-macsec-connection.yml`, with the following settings:
---
routes:
  config:
  - destination: 0.0.0.0/0
    next-hop-interface: macsec0
    next-hop-address: 192.0.2.2
    table-id: 254
  - destination: 192.0.2.2/32
    next-hop-interface: macsec0
    next-hop-address: 0.0.0.0
    table-id: 254
dns-resolver:
  config:
    search:
    - example.com
    server:
    - 192.0.2.200
    - 2001:db8:1::ffbb
interfaces:
  - name: macsec0
    type: macsec
    state: up
    ipv4:
      enabled: true
      address:
      - ip: 192.0.2.1
        prefix-length: 32
    ipv6:
      enabled: true
      address:
      - ip: 2001:db8:1::1
        prefix-length: 64
    macsec:
      encrypt: true
      base-iface: enp0s1
      mka-cak: 50b71a8ef0bd5751ea76de6d6c98c03a
      mka-ckn: f2b4297d39da7330910a74abc0449feb45b5c0b9fc23df1430e1898fcf1c4550
      port: 0
      validation: strict
      send-sci: true

b. Use the CAK and CKN generated in the previous step in the `mka-cak` and `mka-ckn` parameters. The values must be the same on every host in the MACsec-protected network.

c. Optional: In the same YAML configuration file, you can also configure the following settings:
   
   - A static IPv4 address - 192.0.2.1 with the /32 subnet mask
   - A static IPv6 address - 2001:db8:1::1 with the /64 subnet mask
   - An IPv4 default gateway - 192.0.2.2
   - An IPv4 DNS server - 192.0.2.200
   - An IPv6 DNS server - 2001:db8:1::ffbb
3. Apply the settings to the system:

```
# nmstatectl apply create-macsec-connection.yml
```

**Verification**

1. Display the current state in YAML format:

```
# **nmstatectl show macsec0
```

2. Verify that the traffic is encrypted:

```
# tcpdump -nn -i enp0s1
```

3. Optional: Display the unencrypted traffic:

```
# tcpdump -nn -i macsec0
```

4. Display MACsec statistics:

```
# ip macsec show
```

5. Display individual counters for each type of protection: integrity-only (encrypt off) and encryption (encrypt on)

```
# ip -s macsec show
```

**Additional resources**

- MACsec: a different solution to encrypt network traffic

**13.3. ADDITIONAL RESOURCES**

- MACsec: a different solution to encrypt network traffic blog.
CHAPTER 14. GETTING STARTED WITH IPVLAN

IPVLAN is a driver for a virtual network device that can be used in container environment to access the host network. IPVLAN exposes a single MAC address to the external network regardless the number of IPVLAN device created inside the host network. This means that a user can have multiple IPVLAN devices in multiple containers and the corresponding switch reads a single MAC address. IPVLAN driver is useful when the local switch imposes constraints on the total number of MAC addresses that it can manage.

14.1. IPVLAN MODES

The following modes are available for IPVLAN:

- **L2 mode**
  In IPVLAN L2 mode, virtual devices receive and respond to address resolution protocol (ARP) requests. The netfilter framework runs only inside the container that owns the virtual device. No netfilter chains are executed in the default namespace on the containerized traffic. Using L2 mode provides good performance, but less control on the network traffic.

- **L3 mode**
  In L3 mode, virtual devices process only L3 traffic and above. Virtual devices do not respond to ARP request and users must configure the neighbour entries for the IPVLAN IP addresses on the relevant peers manually. The egress traffic of a relevant container is landed on the netfilter POSTROUTING and OUTPUT chains in the default namespace while the ingress traffic is threaded in the same way as L2 mode. Using L3 mode provides good control but decreases the network traffic performance.

- **L3S mode**
  In L3S mode, virtual devices process the same way as in L3 mode, except that both egress and ingress traffics of a relevant container are landed on netfilter chain in the default namespace. L3S mode behaves in a similar way to L3 mode but provides greater control of the network.

**NOTE**

The IPVLAN virtual device does not receive broadcast and multicast traffic in case of L3 and L3S modes.

14.2. COMPARISON OF IPVLAN AND MACVLAN

The following table shows the major differences between MACVLAN and IPVLAN:

<table>
<thead>
<tr>
<th>MACVLAN</th>
<th>IPVLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uses MAC address for each MACVLAN device.</td>
<td>Uses single MAC address which does not limit the number of IPVLAN devices.</td>
</tr>
<tr>
<td>Note that, if a switch reaches the maximum number of MAC addresses it can store in its MAC table, connectivity can be lost.</td>
<td></td>
</tr>
<tr>
<td>Netfilter rules for a global namespace cannot affect traffic to or from a MACVLAN device in a child namespace.</td>
<td>It is possible to control traffic to or from a IPVLAN device in L3 mode and L3S mode.</td>
</tr>
</tbody>
</table>
Both IPVLAN and MACVLAN do not require any level of encapsulation.

14.3. CREATING AND CONFIGURING THE IPVLAN DEVICE USING IPROUTE2

This procedure shows how to set up the IPVLAN device using `iproute2`.

Procedure

1. To create an IPVLAN device, enter the following command:

   ```
   # ip link add link real_NIC_device name IPVLAN_device type ipvlan mode l2
   ```

   Note that network interface controller (NIC) is a hardware component which connects a computer to a network.

   **Example 14.1. Creating an IPVLAN device**

   ```
   # ip link add link enp0s31f6 name my_ipvlan type ipvlan mode l2
   # ip link
   47: my_ipvlan@enp0s31f6: <BROADCAST,MULTICAST> mtu 1500 qdisc noop state DOWN mode DEFAULT group default qlen 1000 link/ether e8:6a:6e:8a:a2:44 brd ff:ff:ff:ff:ff:ff
   ```

2. To assign an IPv4 or IPv6 address to the interface, enter the following command:

   ```
   # ip addr add dev IPVLAN_device IP_address/subnet_mask_prefix
   ```

3. In case of configuring an IPVLAN device in L3 mode or L3S mode, make the following setups:
   a. Configure the neighbor setup for the remote peer on the remote host:

   ```
   # ip neigh add dev peer_device IPVLAN_device_IP_address lladdr MAC_address
   ```

   where MAC_address is the MAC address of the real NIC on which an IPVLAN device is based on.

   b. Configure an IPVLAN device for L3 mode with the following command:

   ```
   # ip route add dev <real_NIC_device> <peer_IP_address/32>
   ```

   For L3S mode:

   ```
   # ip route add dev real_NIC_device peer_IP_address/32
   ```

   where IP-address represents the address of the remote peer.

4. To set an IPVLAN device active, enter the following command:

   ```
   # ip link set dev IPVLAN_device up
   ```
5. To check if the IPVLAN device is active, execute the following command on the remote host:

```
# ping IP_address
```

where the \textit{IP\_address} uses the IP address of the IPVLAN device.
CHAPTER 15. CONFIGURING NETWORKMANAGER TO IGNORE CERTAIN DEVICES

By default, NetworkManager manages all devices. To ignore certain devices, you can configure NetworkManager by setting as `unmanaged`.

15.1. CONFIGURING THE LOOPBACK INTERFACE BY USING NMCLI

By default, NetworkManager does not manage the loopback (`lo`) interface. After creating a connection profile for the `lo` interface, you can configure this device by using NetworkManager. Some of the examples are as follows:

- Assign additional IP addresses to the `lo` interface
- Define DNS addresses
- Change the Maximum Transmission Unit (MTU) size of the `lo` interface

Procedure

1. Create a new connection of type `loopback`:

   ```
   # nmcli connection add con-name example-loopback type loopback
   ```

2. Configure custom connection settings, for example:

   a. To assign an additional IP address to the interface, enter:

   ```
   # nmcli connection modify example-loopback +ipv4.addresses 192.0.2.1/24
   ```

   NOTE

   NetworkManager manages the `lo` interface by always assigning the IP addresses `127.0.0.1` and `::1` that are persistent across the reboots. You cannot override `127.0.0.1` and `::1`. However, you can assign additional IP addresses to the interface.

   b. To set a custom Maximum Transmission Unit (MTU), enter:

   ```
   # nmcli con mod example-loopback loopback.mtu 16384
   ```

   c. To set an IP address to your DNS server, enter:

   ```
   # nmcli connection modify example-loopback ipv4.dns 192.0.2.0
   ```

   If you set a DNS server in the loopback connection profile, this entry is always available in the `/etc/resolv.conf` file. The DNS server entry remains independent of whether or not the host roams between different networks.

3. Activate the connection:

   ```
   # nmcli connection up example-loopback
   ```
Verification

1. Display the settings of the lo interface:

```
# ip address show lo
```

```
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 16384 qdisc noqueue state UNKNOWN group default qlen 1000
link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00 inet 127.0.0.1/8 scope host lo valid_lft forever preferred_lft forever inet 192.0.2.1/24 brd 192.0.2.255 scope global lo valid_lft forever preferred_lft forever inet6 ::1/128 scope host valid_lft forever preferred_lft forever
```

2. Verify the DNS address:

```
# cat /etc/resolv.conf
```

... nameserver 192.0.2.0 ...

15.2. PERMANENTLY CONFIGURING A DEVICE AS UNMANAGED IN NETWORKMANAGER

You can permanently configure devices as unmanaged based on several criteria, such as the interface name, MAC address, or device type.

To temporarily configure network devices as unmanaged, see Temporarily configuring a device as unmanaged in NetworkManager.

Procedure

1. Optional: Display the list of devices to identify the device or MAC address you want to set as unmanaged:

```
# ip link show
```

```
2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP mode DEFAULT group default qlen 1000
link/ether 52:54:00:74:79:56 brd ff:ff:ff:ff:ff:ff
```

2. Create the /etc/NetworkManager/conf.d/99-unmanaged-devices.conf file with the following content:

- To configure a specific interface as unmanaged, add:

```
[keyfile]
unmanaged-devices=interface-name:enp1s0
```

- To configure a device with a specific MAC address as unmanaged, add:
To configure all devices of a specific type as unmanaged, add:

```
unmanaged-devices=type:ethernet
```

To set multiple devices as unmanaged, separate the entries in the `unmanaged-devices` parameter with a semicolon, for example:

```
unmanaged-devices=interface-name:enp1s0;interface-name:enp7s0
```

3. Reload the `NetworkManager` service:

```
# systemctl reload NetworkManager
```

Verification

- Display the list of devices:

```
# nmcli device status
DEVICE  TYPE      STATE      CONNECTION
enp1s0  ethernet  unmanaged  --
...  
```

The `unmanaged` state next to the `enp1s0` device indicates that NetworkManager does not manage this device.

Troubleshooting

- If the device is not shown as `unmanaged`, display the NetworkManager configuration:

```
# NetworkManager --print-config
```

```
[unmanaged-devices=interface-name:enp1s0]
```

If the output does not match the settings that you configured, ensure that no configuration file with a higher priority overrides your settings. For details about how NetworkManager merges multiple configuration files, see the `NetworkManager.conf(5)` man page.

### 15.3. TEMPORARILY CONFIGURING A DEVICE AS UNMANAGED IN NETWORKMANAGER

You can temporarily configure devices as `unmanaged`.

Use this method, for example, for testing purposes. To permanently configure network devices as `unmanaged`, see Permanently configuring a device as unmanaged in NetworkManager.
Procedure

1. Optional: Display the list of devices to identify the device you want to set as unmanaged:

   ```bash
   # nmcli device status
   DEVICE  TYPE      STATE         CONNECTION
   enp1s0  ethernet  disconnected  --
   ...  
   ```

2. Set the enp1s0 device to the unmanaged state:

   ```bash
   # nmcli device set enp1s0 managed no
   ```

Verification

- Display the list of devices:

  ```bash
  # nmcli device status
  DEVICE  TYPE      STATE      CONNECTION
  enp1s0  ethernet  unmanaged  --
  ...  
  ```

  The unmanaged state next to the enp1s0 device indicates that NetworkManager does not manage this device.

Additional resources

- NetworkManager.conf(5) man page
CHAPTER 16. CREATING A DUMMY INTERFACE

As a Red Hat Enterprise Linux user, you can create and use dummy network interfaces for debugging and testing purposes. A dummy interface provides a device to route packets without actually transmitting them. It enables you to create additional loopback-like devices managed by NetworkManager and makes an inactive SLIP (Serial Line Internet Protocol) address look like a real address for local programs.

16.1. CREATING A DUMMY INTERFACE WITH BOTH AN IPV4 AND IPV6 ADDRESS USING NMCLI

You can create a dummy interface with various settings, such as IPv4 and IPv6 addresses. After creating the interface, NetworkManager automatically assigns it to the default public firewall zone.

Procedure

- Create a dummy interface named dummy0 with static IPv4 and IPv6 addresses:

```bash
# nmcli connection add type dummy ifname dummy0 ipv4.method manual ipv4.addresses 192.0.2.1/24 ipv6.method manual ipv6.addresses 2001:db8:2::1/64
```

**NOTE**

To configure a dummy interface without IPv4 and IPv6 addresses, set both the ipv4.method and ipv6.method parameters to disabled. Otherwise, IP auto-configuration fails, and NetworkManager deactivates the connection and removes the device.

Verification

- List the connection profiles:

```bash
# nmcli connection show
```

<table>
<thead>
<tr>
<th>NAME</th>
<th>UUID</th>
<th>TYPE</th>
<th>DEVICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>dummy-dummy0</td>
<td>aaf6eb56-73e5-4746-9037-ee42ca8a65</td>
<td>dummy</td>
<td>dummy0</td>
</tr>
</tbody>
</table>

Additional resources

- `nm-settings(5)` man page
CHAPTER 17. USING NETWORKMANAGER TO DISABLE IPV6 FOR A SPECIFIC CONNECTION

On a system that uses NetworkManager to manage network interfaces, you can disable the IPv6 protocol if the network only uses IPv4. If you disable IPv6, NetworkManager automatically sets the corresponding sysctl values in the Kernel.

NOTE

If disabling IPv6 using kernel tunables or kernel boot parameters, additional consideration must be given to system configuration. For more information, see the How do I disable or enable the IPv6 protocol in RHEL? article.

17.1. DISABLING IPV6 ON A CONNECTION USING NMCLI

You can use the nmcli utility to disable the IPv6 protocol on the command line.

Prerequisites

- The system uses NetworkManager to manage network interfaces.

Procedure

1. Optionally, display the list of network connections:

```
# nmcli connection show
NAME   UUID                         TYPE     DEVICE
Example 7a7e0151-9c18-4e6f-89ee-65bb2d64d365  ethernet  enp1s0 ...
```

2. Set the ipv6.method parameter of the connection to disabled:

```
# nmcli connection modify Example ipv6.method "disabled"
```

3. Restart the network connection:

```
# nmcli connection up Example
```

Verification

1. Display the IP settings of the device:

```
# ip address show enp1s0
2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP
       group default qlen 1000
       link/ether 52:54:00:6b:74:be brd ff:ff:ff:ff:ff:ff
       inet 192.0.2.1/24 brd 192.10.2.255 scope global noprefixroute enp1s0
       valid_lft forever preferred_lft forever
```

   If no inet6 entry is displayed, IPv6 is disabled on the device.

2. Verify that the /proc/sys/net/ipv6/conf/enp1s0/disable_ipv6 file now contains the value 1:
The value 1 means that IPv6 is disabled for the device.
CHAPTER 18. CHANGING A HOSTNAME

The hostname of a system is the name on the system itself. You can set the name when you install RHEL, and you can change it afterwards.

18.1. CHANGING A HOSTNAME USING NMCLI

You can use the `nmcli` utility to update the system hostname. Note that other utilities might use a different term, such as static or persistent hostname.

**Procedure**

1. Optional: Display the current hostname setting:

   ```
   # nmcli general hostname
   old-hostname.example.com
   ```

2. Set the new hostname:

   ```
   # nmcli general hostname new-hostname.example.com
   ```

3. NetworkManager automatically restarts the `systemd-hostnamed` to activate the new name. For the changes to take effect, reboot the host:

   ```
   # reboot
   ```

   Alternatively, if you know which services use the hostname:

   a. Restart all services that only read the hostname when the service starts:

   ```
   # systemctl restart <service_name>
   ```

   b. Active shell users must re-login for the changes to take effect.

**Verification**

- Display the hostname:

  ```
  # nmcli general hostname
  new-hostname.example.com
  ```

18.2. CHANGING A HOSTNAME USING HOSTNAMECTL

You can use the `hostnamectl` utility to update the hostname. By default, this utility sets the following hostname types:

- Static hostname: Stored in the `/etc/hostname` file. Typically, services use this name as the hostname.

- Pretty hostname: A descriptive name, such as `Proxy server in data center A`.

- Transient hostname: A fall-back value that is typically received from the network configuration.
Procedure

1. Optional: Display the current hostname setting:

   ```
   # hostnamectl status --static
   old-hostname.example.com
   ```

2. Set the new hostname:

   ```
   # hostnamectl set-hostname new-hostname.example.com
   ```

   This command sets the static, pretty, and transient hostname to the new value. To set only a specific type, pass the `--static`, `--pretty`, or `--transient` option to the command.

3. The `hostnamectl` utility automatically restarts the `systemd-hostnamed` to activate the new name. For the changes to take effect, reboot the host:

   ```
   # reboot
   ```

   Alternatively, if you know which services use the hostname:
   
   a. Restart all services that only read the hostname when the service starts:

      ```
      # systemctl restart <service_name>
      ```
   
   b. Active shell users must re-login for the changes to take effect.

Verification

- Display the hostname:

  ```
  # hostnamectl status --static
  new-hostname.example.com
  ```

Additional resources

- `hostnamectl(1)`

- `systemd-hostnamed.service(8)`
CHAPTER 19. CONFIGURING NETWORKMANAGER DHCP SETTINGS

NetworkManager provides different configuration options related to DHCP. For example, you can configure NetworkManager to use the built-in DHCP client (default) or an external client, and you can influence DHCP settings of individual profiles.

19.1. CHANGING THE DHCP CLIENT OF NETWORKMANAGER

By default, NetworkManager uses its internal DHCP client. However, if you require a DHCP client with features that the built-in client does not provide, you can alternatively configure NetworkManager to use `dhclient`.

Note that RHEL does not provide `dhcpcd` and, therefore, NetworkManager cannot use this client.

Procedure

1. Create the `/etc/NetworkManager/conf.d/dhcp-client.conf` file with the following content:

   ```
   [main]
   dhcp=dhclient
   ```

   You can set the `dhcp` parameter to `internal` (default) or `dhclient`.

2. If you set the `dhcp` parameter to `dhclient`, install the `dhcp-client` package:

   ```
   # dnf install dhcp-client
   ```

3. Restart NetworkManager:

   ```
   # systemctl restart NetworkManager
   ```

   Note that the restart temporarily interrupts all network connections.

Verification

- Search in the `/var/log/messages` log file for an entry similar to the following:

  ```
  Apr 26 09:54:19 server NetworkManager[27748]: <info> [1650959659.8483] dhcp-init: Using DHCP client 'dhclient'
  ```

  This log entry confirms that NetworkManager uses `dhclient` as DHCP client.

Additional resources

- `NetworkManager.conf(5)` man page

19.2. CONFIGURING THE DHCP BEHAVIOR OF A NETWORKMANAGER CONNECTION
A Dynamic Host Configuration Protocol (DHCP) client requests the dynamic IP address and corresponding configuration information from a DHCP server each time a client connects to the network.

When you configured a connection to retrieve an IP address from a DHCP server, the NetworkManager requests an IP address from a DHCP server. By default, the client waits 45 seconds for this request to be completed. When a DHCP connection is started, a dhcp client requests an IP address from a DHCP server.

Prerequisites

- A connection that uses DHCP is configured on the host.

Procedure

1. Set the `ipv4.dhcp-timeout` and `ipv6.dhcp-timeout` properties. For example, to set both options to 30 seconds, enter:

```bash
# nmcli connection modify connection_name ipv4.dhcp-timeout 30 ipv6.dhcp-timeout 30
```

Alternatively, set the parameters to `infinity` to configure that NetworkManager does not stop trying to request and renew an IP address until it is successful.

2. Optional: Configure the behavior if NetworkManager does not receive an IPv4 address before the timeout:

```bash
# nmcli connection modify connection_name ipv4.may-fail value
```

If you set the `ipv4.may-fail` option to:

- **yes**, the status of the connection depends on the IPv6 configuration:
  - If the IPv6 configuration is enabled and successful, NetworkManager activates the IPv6 connection and no longer tries to activate the IPv4 connection.
  - If the IPv6 configuration is disabled or not configured, the connection fails.
- **no**, the connection is deactivated. In this case:
  - If the `autoconnect` property of the connection is enabled, NetworkManager retries to activate the connection as many times as set in the `autoconnect-retries` property. The default is 4.
  - If the connection still cannot acquire a DHCP address, auto-activation fails. Note that after 5 minutes, the auto-connection process starts again to acquire an IP address from the DHCP server.

3. Optional: Configure the behavior if NetworkManager does not receive an IPv6 address before the timeout:

```bash
# nmcli connection modify connection_name ipv6.may-fail value
```

Additional resources

- `nm-settings(5)` man page
CHAPTER 20. RUNNING DHCLIENT EXIT HOOKS USING NETWORKMANAGER A DISPATCHER SCRIPT

You can use a NetworkManager dispatcher script to execute dhclient exit hooks.

20.1. THE CONCEPT OF NETWORKMANAGER DISPATCHER SCRIPTS

The NetworkManager-dispatcher service executes user-provided scripts in alphabetical order when network events happen. These scripts are typically shell scripts, but can be any executable script or application. You can use dispatcher scripts, for example, to adjust network-related settings that you cannot manage with NetworkManager.

You can store dispatcher scripts in the following directories:

- `/etc/NetworkManager/dispatcher.d/`: The general location for dispatcher scripts the root user can edit.
- `/usr/lib/NetworkManager/dispatcher.d/`: For pre-deployed immutable dispatcher scripts.

For security reasons, the NetworkManager-dispatcher service executes scripts only if the following conditions met:

- The script is owned by the root user.
- The script is only readable and writable by root.
- The setuid bit is not set on the script.

The NetworkManager-dispatcher service runs each script with two arguments:

1. The interface name of the device the operation happened on.
2. The action, such as up, when the interface has been activated.

The Dispatcher scripts section in the NetworkManager(8) man page provides an overview of actions and environment variables you can use in scripts.

The NetworkManager-dispatcher service runs one script at a time, but asynchronously from the main NetworkManager process. Note that, if a script is queued, the service will always run it, even if a later event makes it obsolete. However, the NetworkManager-dispatcher service runs scripts that are symbolic links referring to files in `/etc/NetworkManager/dispatcher.d/no-wait.d/` immediately, without waiting for the termination of previous scripts, and in parallel.

Additional resources

- NetworkManager(8) man page

20.2. CREATING A NETWORKMANAGER DISPATCHER SCRIPT THAT RUNS DHCLIENT EXIT HOOKS

When a DHCP server assigns or updates an IPv4 address, NetworkManager can run a dispatcher script stored in the `/etc/dhcp/dhclient-exit-hooks.d/` directory. This dispatcher script can then, for example, run dhclient exit hooks.
Prerequisites

- The `dhclient` exit hooks are stored in the `/etc/dhcp/dhclient-exit-hooks.d/` directory.

Procedure

1. Create the `/etc/NetworkManager/dispatcher.d/12-dhclient-down` file with the following content:

   ```bash
   #!/bin/bash
   # Run dhclient.exit-hooks.d scripts
   
   if [ -n "$DHCP4_DHCP_LEASE_TIME" ]; then
     if [ "$2" = "dhcp4-change" ] || [ "$2" = "up" ]; then
       if [ -d /etc/dhcp/dhclient-exit-hooks.d ] ; then
         for f in /etc/dhcp/dhclient-exit-hooks.d/*.sh ; do
           if [ -x "$f" ]; then
             $f
           fi
         done
       fi
     fi
   fi
   
   2. Set the root user as owner of the file:

      ```bash
      # chown root:root /etc/NetworkManager/dispatcher.d/12-dhclient-down
      ```

3. Set the permissions so that only the root user can execute it:

   ```bash
   # chmod 0700 /etc/NetworkManager/dispatcher.d/12-dhclient-down
   ```

4. Restore the SELinux context:

   ```bash
   # restorecon /etc/NetworkManager/dispatcher.d/12-dhclient-down
   ```

Additional resources

- `NetworkManager(8)` man page
CHAPTER 21. MANUALLY CONFIGURING THE /ETC/RESOLV.CONF FILE

By default, NetworkManager dynamically updates the /etc/resolv.conf file with the DNS settings from active NetworkManager connection profiles. However, you can disable this behavior and manually configure DNS settings in /etc/resolv.conf.

**NOTE**

Alternatively, if you require a specific order of DNS servers in /etc/resolv.conf, see Configuring the order of DNS servers.

21.1. DISABLING DNS PROCESSING IN THE NETWORKMANAGER CONFIGURATION

By default, NetworkManager manages DNS settings in the /etc/resolv.conf file, and you can configure the order of DNS servers. Alternatively, you can disable DNS processing in NetworkManager if you prefer to manually configure DNS settings in /etc/resolv.conf.

**Procedure**

1. As the root user, create the /etc/NetworkManager/conf.d/90-dns-none.conf file with the following content by using a text editor:

   ```
   [main]
   dns=none
   ```

2. Reload the NetworkManager service:

   ```
   # systemctl reload NetworkManager
   ```

   **NOTE**

   After you reload the service, NetworkManager no longer updates the /etc/resolv.conf file. However, the last contents of the file are preserved.

3. Optionally, remove the Generated by NetworkManager comment from /etc/resolv.conf to avoid confusion.

**Verification**

1. Edit the /etc/resolv.conf file and manually update the configuration.

2. Reload the NetworkManager service:

   ```
   # systemctl reload NetworkManager
   ```

3. Display the /etc/resolv.conf file:

   ```
   # cat /etc/resolv.conf
   ```
If you successfully disabled DNS processing, NetworkManager did not override the manually configured settings.

Troubleshooting

- Display the NetworkManager configuration to ensure that no other configuration file with a higher priority overrode the setting:

  ```
  # NetworkManager --print-config
  ...
  dns=none
  ...
  ```

Additional resources

- NetworkManager.conf(5) man page
- Configuring the order of DNS servers using NetworkManager

21.2. REPLACING /ETC/RESOLV.CONF WITH A SYMBOLIC LINK TO MANUALLY CONFIGURE DNS SETTINGS

By default, NetworkManager manages DNS settings in the `/etc/resolv.conf` file, and you can configure the order of DNS servers. Alternatively, you can disable DNS processing in NetworkManager if you prefer to manually configure DNS settings in `/etc/resolv.conf`. For example, NetworkManager does not automatically update the DNS configuration if `/etc/resolv.conf` is a symbolic link.

Prerequisites

- The NetworkManager `rc-manager` configuration option is not set to `file`. To verify, use the NetworkManager `--print-config` command.

Procedure

1. Create a file, such as `/etc/resolv.conf.manually-configured`, and add the DNS configuration for your environment to it. Use the same parameters and syntax as in the original `/etc/resolv.conf`.

2. Remove the `/etc/resolv.conf` file:

   ```
   # rm /etc/resolv.conf
   ```

3. Create a symbolic link named `/etc/resolv.conf` that refers to `/etc/resolv.conf.manually-configured`:

   ```
   # ln -s /etc/resolv.conf.manually-configured /etc/resolv.conf
   ```

Additional resources

- resolv.conf(5) man page
- NetworkManager.conf(5) man page
- Configuring the order of DNS servers using NetworkManager
CHAPTER 22. CONFIGURING THE ORDER OF DNS SERVERS

Most applications use the `getaddrinfo()` function of the `glibc` library to resolve DNS requests. By default, `glibc` sends all DNS requests to the first DNS server specified in the `/etc/resolv.conf` file. If this server does not reply, RHEL uses the next server in this file. NetworkManager enables you to influence the order of DNS servers in `/etc/resolv.conf`.

22.1. HOW NETWORKMANAGER ORDERS DNS SERVERS IN /ETC/RESOLV.CONF

NetworkManager orders DNS servers in the `/etc/resolv.conf` file based on the following rules:

- If only one connection profile exists, NetworkManager uses the order of IPv4 and IPv6 DNS server specified in that connection.

- If multiple connection profiles are activated, NetworkManager orders DNS servers based on a DNS priority value. If you set DNS priorities, the behavior of NetworkManager depends on the value set in the `dns` parameter. You can set this parameter in the `[main]` section in the `/etc/NetworkManager/NetworkManager.conf` file:
  
  - `dns=default` or if the `dns` parameter is not set:
    NetworkManager orders the DNS servers from different connections based on the `ipv4.dns-priority` and `ipv6.dns-priority` parameter in each connection.

    If you set no value or you set `ipv4.dns-priority` and `ipv6.dns-priority` to 0, NetworkManager uses the global default value. See Default values of DNS priority parameters.

  - `dns=dnsmasq` or `dns=systemd-resolved`:
    When you use one of these settings, NetworkManager sets either 127.0.0.1 for `dnsmasq` or 127.0.0.53 as `nameserver` entry in the `/etc/resolv.conf` file.

    Both the `dnsmasq` and `systemd-resolved` services forward queries for the search domain set in a NetworkManager connection to the DNS server specified in that connection, and forwards queries to other domains to the connection with the default route. When multiple connections have the same search domain set, `dnsmasq` and `systemd-resolved` forward queries for this domain to the DNS server set in the connection with the lowest priority value.

Default values of DNS priority parameters

NetworkManager uses the following default values for connections:

- **50** for VPN connections

- **100** for other connections

Valid DNS priority values:

You can set both the global default and connection-specific `ipv4.dns-priority` and `ipv6.dns-priority` parameters to a value between -2147483647 and 2147483647.

- A lower value has a higher priority.

- Negative values have the special effect of excluding other configurations with a greater value. For example, if at least one connection with a negative priority value exists, NetworkManager uses only the DNS servers specified in the connection profile with the lowest priority.
If multiple connections have the same DNS priority, NetworkManager prioritizes the DNS in the following order:

a. VPN connections
b. Connection with an active default route. The active default route is the default route with the lowest metric.

Additional resources

- **nm-settings(5) man page**
- **Using different DNS servers for different domains**

### 22.2. SETTING A NETWORKMANAGER-WIDE DEFAULT DNS SERVER PRIORITY VALUE

NetworkManager uses the following DNS priority default values for connections:

- **50** for VPN connections
- **100** for other connections

You can override these system-wide defaults with a custom default value for IPv4 and IPv6 connections.

**Procedure**

1. Edit the `/etc/NetworkManager/NetworkManager.conf` file:
   a. Add the `[connection]` section, if it does not exist:
      ```conf
      [connection]
      ```
   b. Add the custom default values to the `[connection]` section. For example, to set the new default for both IPv4 and IPv6 to **200**, add:
      ```conf
      ipv4.dns-priority=200
      ipv6.dns-priority=200
      ```
      
      You can set the parameters to a value between **-2147483647** and **2147483647**. Note that setting the parameters to **0** enables the built-in defaults (**50** for VPN connections and **100** for other connections).

2. Reload the NetworkManager service:
   ```shell
   # systemctl reload NetworkManager
   ```

Additional resources

- **NetworkManager.conf(5) man page**
22.3. SETTING THE DNS PRIORITY OF A NETWORKMANAGER CONNECTION

If you require a specific order of DNS servers you can set priority values in connection profiles. NetworkManager uses these values to order the servers when the service creates or updates the /etc/resolv.conf file.

Note that setting DNS priorities makes only sense if you have multiple connections with different DNS servers configured. If you have only one connection with multiple DNS servers configured, manually set the DNS servers in the preferred order in the connection profile.

Prerequisites

- The system has multiple NetworkManager connections configured.
- The system either has no dns parameter set in the /etc/NetworkManager/NetworkManager.conf file or the parameter is set to default.

Procedure

1. Optionally, display the available connections:

```bash
# nmcli connection show
NAME           UUID                                  TYPE      DEVICE
Example_con_1  d17ee488-4665-4de2-b28a-48befab0cd43  ethernet  enp1s0
Example_con_2  916e4f67-7145-3ffa-9f7b-e7cada8f6bf7  ethernet  enp7s0
...
```

2. Set the ipv4.dns-priority and ipv6.dns-priority parameters. For example, to set both parameters to 10 for the Example_con_1 connection:

```bash
# nmcli connection modify Example_con_1 ipv4.dns-priority 10 ipv6.dns-priority 10
```

3. Optionally, repeat the previous step for other connections.

4. Re-activate the connection you updated:

```bash
# nmcli connection up Example_con_1
```

Verification

- Display the contents of the /etc/resolv.conf file to verify that the DNS server order is correct:

```bash
# cat /etc/resolv.conf
```
CHAPTER 23. USING DIFFERENT DNS SERVERS FOR DIFFERENT DOMAINS

By default, Red Hat Enterprise Linux (RHEL) sends all DNS requests to the first DNS server specified in the `/etc/resolv.conf` file. If this server does not reply, RHEL uses the next server in this file. In environments where one DNS server cannot resolve all domains, administrators can configure RHEL to send DNS requests for a specific domain to a selected DNS server.

For example, you connect a server to a Virtual Private Network (VPN), and hosts in the VPN use the `example.com` domain. In this case, you can configure RHEL to process DNS queries in the following way:

- Send only DNS requests for `example.com` to the DNS server in the VPN network.
- Send all other requests to the DNS server that is configured in the connection profile with the default gateway.

23.1. USING DNMASQ IN NETWORKMANAGER TO SEND DNS REQUESTS FOR A SPECIFIC DOMAIN TO A SELECTED DNS SERVER

You can configure NetworkManager to start an instance of `dnsmasq`. This DNS caching server then listens on port 53 on the `loopback` device. Consequently, this service is only reachable from the local system and not from the network.

With this configuration, NetworkManager adds the `nameserver 127.0.0.1` entry to the `/etc/resolv.conf` file, and `dnsmasq` dynamically routes DNS requests to the corresponding DNS servers specified in the NetworkManager connection profiles.

Prerequisites

- The system has multiple NetworkManager connections configured.
- A DNS server and search domain are configured in the NetworkManager connection profile that is responsible for resolving a specific domain.
  For example, to ensure that the DNS server specified in a VPN connection resolves queries for the `example.com` domain, the VPN connection profile must contain the following settings:
    - A DNS server that can resolve `example.com`
    - A search domain set to `example.com` in the `ipv4.dns-search` and `ipv6.dns-search` parameters
- The `dnsmasq` service is not running or configured to listen on a different interface then `localhost`.

Procedure

1. Install the `dnsmasq` package:

   ```
   # dnf install dnsmasq
   ```

2. Edit the `/etc/NetworkManager/NetworkManager.conf` file, and set the following entry in the `[main]` section:
dns=dnsmasq

3. Reload the `NetworkManager` service:

```
# systemctl reload NetworkManager
```

**Verification**

1. Search in the `systemd` journal of the `NetworkManager` unit for which domains the service uses a different DNS server:

```
# journalctl -xeu NetworkManager
```

```
Jun 02 13:30:17 client_hostname dnsmasq[5298]: using nameserver 198.51.100.7#53 for domain example.com
```

2. Use the `tcpdump` packet sniffer to verify the correct route of DNS requests:

   a. Install the `tcpdump` package:

      ```
      # dnf install tcpdump
      ```

   b. On one terminal, start `tcpdump` to capture DNS traffic on all interfaces:

      ```
      # tcpdump -i any port 53
      ```

   c. On a different terminal, resolve host names for a domain for which an exception exists and another domain, for example:

      ```
      # host -t A www.example.com
      # host -t A www.redhat.com
      ```

   d. Verify in the `tcpdump` output that Red Hat Enterprise Linux sends only DNS queries for the `example.com` domain to the designated DNS server and through the corresponding interface:

      ```
      13:52:42.234533 tun0 Out IP server.43534 > 198.51.100.7.domain: 50121+ A? www.example.com. (33)
      ...
      ```

      Red Hat Enterprise Linux sends the DNS query for `www.example.com` to the DNS server on `198.51.100.7` and the query for `www.redhat.com` to `192.0.2.1`.

**Troubleshooting**

1. Verify that the `nameserver` entry in the `/etc/resolv.conf` file refers to `127.0.0.1`:
# cat /etc/resolv.conf
nameserver 127.0.0.1

If the entry is missing, check the `dns` parameter in the `/etc/NetworkManager/NetworkManager.conf` file.

2. Verify that the `dnsmasq` service listens on port 53 on the `loopback` device:

```bash
# ss -tulpn | grep "127.0.0.1:53"
udp  UNCONN 0  0    127.0.0.1:53   0.0.0.0:*    users:("dnsmasq",pid=7340,fd=18))
tcp  LISTEN 0  32   127.0.0.1:53   0.0.0.0:*    users:("dnsmasq",pid=7340,fd=19))
```

If the service does not listen on 127.0.0.1:53, check the journal entries of the `NetworkManager` unit:

```bash
# journalctl -u NetworkManager
```

## 23.2. USING SYSTEMD-RESOLVED IN NETWORKMANAGER TO SEND DNS REQUESTS FOR A SPECIFIC DOMAIN TO A SELECTED DNS SERVER

You can configure NetworkManager to start an instance of `systemd-resolved`. This DNS stub resolver then listens on port 53 on IP address `127.0.0.53`. Consequently, this stub resolver is only reachable from the local system and not from the network.

With this configuration, NetworkManager adds the `nameserver 127.0.0.53` entry to the `/etc/resolv.conf` file, and `systemd-resolved` dynamically routes DNS requests to the corresponding DNS servers specified in the NetworkManager connection profiles.

**IMPORTANT**

The `systemd-resolved` service is provided as a Technology Preview only. Technology Preview features are not supported with Red Hat production Service Level Agreements (SLAs), might not be functionally complete, and Red Hat does not recommend using them for production. These previews provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.


For a supported solution, see [Using dnsmasq in NetworkManager to send DNS requests for a specific domain to a selected DNS server](https://access.redhat.com/documentation/en-us/red_hat_enterprise_linux/8/html/networking_and_ip_v6/administration#using-dnsmasq).

**Prerequisites**

- The system has multiple NetworkManager connections configured.
- A DNS server and search domain are configured in the NetworkManager connection profile that is responsible for resolving a specific domain.

For example, to ensure that the DNS server specified in a VPN connection resolves queries for the `example.com` domain, the VPN connection profile must contain the following settings:
- A DNS server that can resolve example.com
- A search domain set to example.com in the ipv4.dns-search and ipv6.dns-search parameters

**Procedure**

1. Enable and start the **systemd-resolved** service:
   
   ```
   # systemctl --now enable systemd-resolved
   ```

2. Edit the `/etc/NetworkManager/NetworkManager.conf` file, and set the following entry in the `[main]` section:
   
   ```
   dns=systemd-resolved
   ```

3. Reload the **NetworkManager** service:
   
   ```
   # systemctl reload NetworkManager
   ```

**Verification**

1. Display the DNS servers **systemd-resolved** uses and for which domains the service uses a different DNS server:
   
   ```
   # resolvectl ...
   Link 2 (enp1s0)
   Current Scopes: DNS
   Protocols: +DefaultRoute ...
   Current DNS Server: 192.0.2.1
   DNS Servers: 192.0.2.1
   
   Link 3 (tun0)
   Current Scopes: DNS
   Protocols: -DefaultRoute ...
   Current DNS Server: 198.51.100.7
   DNS Servers: 198.51.100.7 203.0.113.19
   DNS Domain: example.com
   ```

   The output confirms that **systemd-resolved** uses different DNS servers for the example.com domain.

2. Use the **tcpdump** packet sniffer to verify the correct route of DNS requests:
   
   a. Install the **tcpdump** package:
      
      ```
      # dnf install tcpdump
      ```
   
   b. On one terminal, start **tcpdump** to capture DNS traffic on all interfaces:
      
      ```
      # tcpdump -i any port 53
      ```
c. On a different terminal, resolve host names for a domain for which an exception exists and another domain, for example:

```
# host -t A www.example.com
# host -t A www.redhat.com
```

d. Verify in the `tcpdump` output that Red Hat Enterprise Linux sends only DNS queries for the `example.com` domain to the designated DNS server and through the corresponding interface:

```
...  
13:52:42.234533 tun0   Out IP server.43534 > 198.51.100.7.domain: 50121+ A?
www.example.com. (33)
...  
13:52:57.753235 enp1s0 Out IP server.40864 > 192.0.2.1.domain: 6906+ A?
www.redhat.com. (33)
...
```

Red Hat Enterprise Linux sends the DNS query for `www.example.com` to the DNS server on `198.51.100.7` and the query for `www.redhat.com` to `192.0.2.1`.

### Troubleshooting

1. Verify that the `nameserver` entry in the `/etc/resolv.conf` file refers to `127.0.0.53`:

```
# cat /etc/resolv.conf
nameserver 127.0.0.53
```

If the entry is missing, check the `dns` parameter in the `/etc/NetworkManager/NetworkManager.conf` file.

2. Verify that the `systemd-resolved` service listens on port `53` on the local IP address `127.0.0.53`:

```
# ss -ulpn | grep "127.0.0.53"
udp UNCONN 0 0 127.0.0.53%lo:53 0.0.0.0:* users:(("systemd-resolve",pid=1050,fd=12))
tcp LISTEN 0 4096 127.0.0.53%lo:53 0.0.0.0:* users:(("systemd-resolve",pid=1050,fd=13))
```

If the service does not listen on `127.0.0.53:53`, check if the `systemd-resolved` service is running.
CHAPTER 24. MANAGING THE DEFAULT GATEWAY SETTING

The default gateway is a router that forwards network packets when no other route matches the destination of a packet. In a local network, the default gateway is typically the host that is one hop closer to the internet.

24.1. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING NMCLI

In most situations, administrators set the default gateway when they create a connection as explained in, for example, Configuring an Ethernet connection by using nmcli.

In most situations, administrators set the default gateway when they create a connection. However, you can also set or update the default gateway setting on a previously created connection using the nmcli utility.

Prerequisites

- At least one static IP address must be configured on the connection on which the default gateway will be set.
- If the user is logged in on a physical console, user permissions are sufficient. Otherwise, user must have root permissions.

Procedure

1. Set the IP address of the default gateway.
   For example, to set the IPv4 address of the default gateway on the example connection to 192.0.2.1:

   ```bash
   # nmcli connection modify example ipv4.gateway "192.0.2.1"
   ```

   For example, to set the IPv6 address of the default gateway on the example connection to 2001:db8:1::1:

   ```bash
   # nmcli connection modify example ipv6.gateway "2001:db8:1::1"
   ```

2. Restart the network connection for changes to take effect. For example, to restart the example connection using the command line:

   ```bash
   # nmcli connection up example
   ```

   **WARNING**
   All connections currently using this network connection are temporarily interrupted during the restart.

3. Optionally, verify that the route is active.
To display the IPv4 default gateway:

```
# ip -4 route
default via 192.0.2.1 dev example proto static metric 100
```

To display the IPv6 default gateway:

```
# ip -6 route
default via 2001:db8:1::1 dev example proto static metric 100 pref medium
```

### 24.2. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING THE NMCLI INTERACTIVE MODE

In most situations, administrators set the default gateway when they create a connection as explained in, for example, Configuring an Ethernet connection by using the nmcli interactive editor

In most situations, administrators set the default gateway when they create a connection. However, you can also set or update the default gateway setting on a previously created connection using the interactive mode of the *nmcli* utility.

**Prerequisites**

- At least one static IP address must be configured on the connection on which the default gateway will be set.
- If the user is logged in on a physical console, user permissions are sufficient. Otherwise, the user must have *root* permissions.

**Procedure**

1. Open the *nmcli* interactive mode for the required connection. For example, to open the *nmcli* interactive mode for the *example* connection:

   ```
   # nmcli connection edit example
   ```

2. Set the default gateway.
   For example, to set the IPv4 address of the default gateway on the *example* connection to 192.0.2.1:

   ```
   nmcli> set ipv4.gateway 192.0.2.1
   ```

   For example, to set the IPv6 address of the default gateway on the *example* connection to 2001:db8:1::1:

   ```
   nmcli> set ipv6.gateway 2001:db8:1::1
   ```

3. Optionally, verify that the default gateway was set correctly:

   ```
   nmcli> print
   ...
   ipv4.gateway: 192.0.2.1
   ```
4. Save the configuration:
   ```
   nmcli> save persistent
   ```

5. Restart the network connection for changes to take effect:
   ```
   nmcli> activate example
   ```

**WARNING**
- All connections currently using this network connection are temporarily interrupted during the restart.

6. Leave the `nmcli` interactive mode:
   ```
   nmcli> quit
   ```

7. Optionally, verify that the route is active.
   To display the IPv4 default gateway:
   ```
   # ip -4 route
   default via 192.0.2.1 dev example proto static metric 100
   ```
   To display the IPv6 default gateway:
   ```
   # ip -6 route
   default via 2001:db8:1::1 dev example proto static metric 100 pref medium
   ```

Additional resources
- Configuring an Ethernet connection by using the `nmcli` interactive editor

### 24.3. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING NM-CONNECTION-EDITOR

In most situations, administrators set the default gateway when they create a connection. However, you can also set or update the default gateway setting on a previously created connection using the `nm-connection-editor` application.

**Prerequisites**
- At least one static IP address must be configured on the connection on which the default gateway will be set.
Procedure

1. Open a terminal, and enter `nm-connection-editor`:

```
# nm-connection-editor
```

2. Select the connection to modify, and click the gear wheel icon to edit the existing connection.

3. Set the IPv4 default gateway. For example, to set the IPv4 address of the default gateway on the connection to `192.0.2.1`:
   a. Open the **IPv4 Settings** tab.
   b. Enter the address in the **gateway** field next to the IP range the gateway’s address is within:

<table>
<thead>
<tr>
<th>Addresses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td>Netmask</td>
<td>Gateway</td>
</tr>
<tr>
<td>192.0.2.123</td>
<td>24</td>
<td>192.0.2.1</td>
</tr>
</tbody>
</table>

4. Set the IPv6 default gateway. For example, to set the IPv6 address of the default gateway on the connection to `2001:db8:1::1`:
   a. Open the **IPv6** tab.
   b. Enter the address in the **gateway** field next to the IP range the gateway’s address is within:

<table>
<thead>
<tr>
<th>Addresses</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td>Prefix</td>
<td>Gateway</td>
</tr>
<tr>
<td>2001:db8:1:5</td>
<td>64</td>
<td>2001:db8:1:1</td>
</tr>
</tbody>
</table>

5. Click **OK**.

6. Click **Save**.

7. Restart the network connection for changes to take effect. For example, to restart the `example` connection using the command line:

```
# nmcli connection up example
```

**WARNING**

All connections currently using this network connection are temporarily interrupted during the restart.

8. Optionally, verify that the route is active.
   To display the IPv4 default gateway:
To display the IPv6 default gateway:

```
# ip -6 route
default via 2001:db8:1::1 dev example proto static metric 100 pref medium
```

Additional resources

- Configuring an Ethernet connection by using nm-connection-editor

### 24.4. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING CONTROL-CENTER

In most situations, administrators set the default gateway when they create a connection. However, you can also set or update the default gateway setting on a previously created connection using the `control-center` application.

**Prerequisites**

- At least one static IP address must be configured on the connection on which the default gateway will be set.

- The network configuration of the connection is open in the `control-center` application.

**Procedure**

1. Set the IPv4 default gateway. For example, to set the IPv4 address of the default gateway on the connection to `192.0.2.1`:
   
   a. Open the IPv4 tab.
   
   b. Enter the address in the `gateway` field next to the IP range the gateway’s address is within:

   ![Gateway Address Table]

2. Set the IPv6 default gateway. For example, to set the IPv6 address of the default gateway on the connection to `2001:db8:1::1`:
   
   a. Open the IPv6 tab.
   
   b. Enter the address in the `gateway` field next to the IP range the gateway’s address is within:

   ![Gateway Address Table]

3. Click **Apply**.
4. Back in the **Network** window, disable and re-enable the connection by switching the button for the connection to **Off** and back to **On** for changes to take effect.

![WARNING]

All connections currently using this network connection are temporarily interrupted during the restart.

5. Optionally, verify that the route is active.

   To display the IPv4 default gateway:
   
   ```
   $ ip -4 route
default via 192.0.2.1 dev example proto static metric 100
   ```

   To display the IPv6 default gateway:
   
   ```
   $ ip -6 route
default via 2001:db8:1::1 dev example proto static metric 100 pref medium
   ```

**Additional resources**

- Configuring an Ethernet connection by using control-center

**24.5. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING NMSTATECTL**

Use the `nmstatectl` utility to set the default gateway through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.

**Prerequisites**

- At least one static IP address must be configured on the connection on which the default gateway will be set.
- The `enp1s0` interface is configured, and the IP address of the default gateway is within the subnet of the IP configuration of this interface.
- The `nmstate` package is installed.

**Procedure**

1. Create a YAML file, for example `~/set-default-gateway.yml`, with the following content:

   ```yaml
   ---
   routes:
   config:
   ```
These settings define 192.0.2.1 as the default gateway, and the default gateway is reachable through the enp1s0 interface.

2. Apply the settings to the system:

```bash
# nmstatectl apply ~/set-default-gateway.yml
```

### Additional resources
- `nmstatectl(8)` man page
- `/usr/share/doc/nmstate/examples/` directory

## 24.6. SETTING THE DEFAULT GATEWAY ON AN EXISTING CONNECTION BY USING THE NETWORK RHEL SYSTEM ROLE

You can use the `network` RHEL system role to set the default gateway.

### IMPORTANT

When you run a play that uses the `network` RHEL system role and if the setting values do not match the values specified in the play, the role overrides the existing connection profile with the same name. To prevent resetting these values to their defaults, always specify the whole configuration of the network connection profile in the play, even if the configuration, for example the IP configuration, already exists.

Depending on whether it already exists, the procedure creates or updates the `enp1s0` connection profile with the following settings:

- A static IPv4 address - 198.51.100.20 with a /24 subnet mask
- A static IPv6 address - 2001:db8:1::1 with a /64 subnet mask
- An IPv4 default gateway - 198.51.100.254
- An IPv6 default gateway - 2001:db8:1::fffe
- An IPv4 DNS server - 198.51.100.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com

### Prerequisites
- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
**Procedure**

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure an Ethernet connection with static IP and default gateway
      ansible.builtin.include_role:
        name: rhel-system-roles.network
      vars:
        network_connections:
          - name: enp1s0
            type: ethernet
            autoconnect: yes
            ip:
              address:
                - 198.51.100.20/24
                - 2001:db8:1::1/64
              gateway4: 198.51.100.254
              gateway6: 2001:db8:1::fffe
            dns:
              - 198.51.100.200
              - 2001:db8:1::fffe
              dns_search:
                - example.com
            state: up
```

2. Validate the playbook syntax:

```bash
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```bash
$ ansible-playbook ~/playbook.yml
```

**Additional resources**

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory

**24.7. HOW NETWORKMANAGER MANAGES MULTIPLE DEFAULT GATEWAYS**

In certain situations, for example for fallback reasons, you set multiple default gateways on a host. However, to avoid asynchronous routing issues, each default gateway of the same protocol requires a separate metric value. Note that RHEL only uses the connection to the default gateway that has the lowest metric set.
You can set the metric for both the IPv4 and IPv6 gateway of a connection using the following command:

```bash
# nmcli connection modify connection-name ipv4.route-metric value ipv6.route-metric value
```

**IMPORTANT**

Do not set the same metric value for the same protocol in multiple connection profiles to avoid routing issues.

If you set a default gateway without a metric value, NetworkManager automatically sets the metric value based on the interface type. For that, NetworkManager assigns the default value of this network type to the first connection that is activated, and sets an incremented value to each other connection of the same type in the order they are activated. For example, if two Ethernet connections with a default gateway exist, NetworkManager sets a metric of 100 on the route to the default gateway of the connection that you activate first. For the second connection, NetworkManager sets 101.

The following is an overview of frequently-used network types and their default metrics:

<table>
<thead>
<tr>
<th>Connection type</th>
<th>Default metric value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VPN</td>
<td>50</td>
</tr>
<tr>
<td>Ethernet</td>
<td>100</td>
</tr>
<tr>
<td>MACsec</td>
<td>125</td>
</tr>
<tr>
<td>InfiniBand</td>
<td>150</td>
</tr>
<tr>
<td>Bond</td>
<td>300</td>
</tr>
<tr>
<td>Team</td>
<td>350</td>
</tr>
<tr>
<td>VLAN</td>
<td>400</td>
</tr>
<tr>
<td>Bridge</td>
<td>425</td>
</tr>
<tr>
<td>TUN</td>
<td>450</td>
</tr>
<tr>
<td>Wi-Fi</td>
<td>600</td>
</tr>
<tr>
<td>IP tunnel</td>
<td>675</td>
</tr>
</tbody>
</table>

**Additional resources**

- Configuring policy-based routing to define alternative routes
- Getting started with Multipath TCP
24.8. CONFIGURING NETWORKMANAGER TO AVOID USING A SPECIFIC PROFILE TO PROVIDE A DEFAULT GATEWAY

You can configure that NetworkManager never uses a specific profile to provide the default gateway. Follow this procedure for connection profiles that are not connected to the default gateway.

Prerequisites

- The NetworkManager connection profile for the connection that is not connected to the default gateway exists.

Procedure

1. If the connection uses a dynamic IP configuration, configure that NetworkManager does not use the connection as the default route for IPv4 and IPv6 connections:

   ```
   # nmcli connection modify connection_name ipv4.never-default yes ipv6.never-default yes
   ```

   Note that setting `ipv4.never-default` and `ipv6.never-default` to `yes`, automatically removes the default gateway's IP address for the corresponding protocol from the connection profile.

2. Activate the connection:

   ```
   # nmcli connection up connection_name
   ```

Verification

- Use the `ip -4 route` and `ip -6 route` commands to verify that RHEL does not use the network interface for the default route for the IPv4 and IPv6 protocol.

24.9. FIXING UNEXPECTED ROUTING BEHAVIOR DUE TO MULTIPLE DEFAULT GATEWAYS

There are only a few scenarios, such as when using multipath TCP, in which you require multiple default gateways on a host. In most cases, you configure only a single default gateway to avoid unexpected routing behavior or asynchronous routing issues.

NOTE

To route traffic to different internet providers, use policy-based routing instead of multiple default gateways.

Prerequisites

- The host uses NetworkManager to manage network connections, which is the default.
- The host has multiple network interfaces.
- The host has multiple default gateways configured.

Procedure
1. Display the routing table:
   - For IPv4, enter:
     ```sh
     # ip -4 route
     default via 192.0.2.1 dev enp1s0 proto static metric 101
     default via 198.51.100.1 dev enp7s0 proto static metric 102
     ...
     ```
   - For IPv6, enter:
     ```sh
     # ip -6 route
     default via 2001:db8:1::1 dev enp1s0 proto static metric 101 pref medium
     default via 2001:db8:2::1 dev enp7s0 proto static metric 102 pref medium
     ...
     ```

   Entries starting with `default` indicate a default route. Note the interface names of these entries displayed next to `dev`.

2. Use the following commands to display the NetworkManager connections that use the interfaces you identified in the previous step:
   ```sh
   # nmcli -f GENERAL.CONNECTION,IP4.GATEWAY,IP6.GATEWAY device show enp1s0
   GENERAL.CONNECTION:      Corporate-LAN
   IP4.GATEWAY:             192.0.2.1
   IP6.GATEWAY:             2001:db8:1::1
   ``
   ```sh
   # nmcli -f GENERAL.CONNECTION,IP4.GATEWAY,IP6.GATEWAY device show enp7s0
   GENERAL.CONNECTION:      Internet-Provider
   IP4.GATEWAY:             198.51.100.1
   IP6.GATEWAY:             2001:db8:2::1
   ``

   In these examples, the profiles named Corporate-LAN and Internet-Provider have the default gateways set. Because, in a local network, the default gateway is typically the host that is one hop closer to the internet, the rest of this procedure assumes that the default gateways in the Corporate-LAN are incorrect.

3. Configure that NetworkManager does not use the Corporate-LAN connection as the default route for IPv4 and IPv6 connections:
   ```sh
   # nmcli connection modify Corporate-LAN ipv4.never-default yes ipv6.never-default yes
   ```

   Note that setting `ipv4.never-default` and `ipv6.never-default` to `yes`, automatically removes the default gateway’s IP address for the corresponding protocol from the connection profile.

4. Activate the Corporate-LAN connection:
   ```sh
   # nmcli connection up Corporate-LAN
   ```

Verification
- Display the IPv4 and IPv6 routing tables and verify that only one default gateway is available for each protocol:
For IPv4, enter:

```
# ip -4 route
default via 192.0.2.1 dev enp1s0 proto static metric 101
```

For IPv6, enter:

```
# ip -6 route
default via 2001:db8:1::1 dev enp1s0 proto static metric 101 pref medium
```

Additional resources

- Configuring policy-based routing to define alternative routes
- Getting started with Multipath TCP
CHAPTER 25. CONFIGURING STATIC ROUTES

Routing ensures that you can send and receive traffic between mutually-connected networks. In larger environments, administrators typically configure services so that routers can dynamically learn about other routers. In smaller environments, administrators often configure static routes to ensure that traffic can reach from one network to the next.

You need static routes to achieve a functioning communication among multiple networks if all of these conditions apply:

- The traffic has to pass multiple networks.
- The exclusive traffic flow through the default gateways is not sufficient.

Section 25.1, "Example of a network that requires static routes" describes scenarios and how the traffic flows between different networks when you do not configure static routes.

25.1. EXAMPLE OF A NETWORK THAT REQUIRES STATIC ROUTES

You require static routes in this example because not all IP networks are directly connected through one router. Without the static routes, some networks cannot communicate with each other. Additionally, traffic from some networks flows only in one direction.

NOTE

The network topology in this example is artificial and only used to explain the concept of static routing. It is not a recommended topology in production environments.

For a functioning communication among all networks in this example, configure a static route to Raleigh (198.51.100.0/24) with next the hop Router 2 (203.0.113.10). The IP address of the next hop is the one of Router 2 in the data center network (203.0.113.0/24).

You can configure the static route as follows:

- For a simplified configuration, set this static route only on Router 1. However, this increases the traffic on Router 1 because hosts from the data center (203.0.113.0/24) send traffic to Raleigh (198.51.100.0/24) always through Router 1 to Router 2.

- For a more complex configuration, configure this static route on all hosts in the data center (203.0.113.0/24). All hosts in this subnet then send traffic directly to Router 2 (203.0.113.10) that is closer to Raleigh (198.51.100.0/24).

For more details between which networks traffic flows or not, see the explanations below the diagram.
In case that the required static routes are not configured, the following are the situations in which the communication works and when it does not:

- Hosts in the Berlin network (192.0.2.0/24):
  - Can communicate with other hosts in the same subnet because they are directly connected.
  - Can communicate with the internet because Router 1 is in the Berlin network (192.0.2.0/24) and has a default gateway, which leads to the internet.
  - Can communicate with the data center network (203.0.113.0/24) because Router 1 has interfaces in both the Berlin (192.0.2.0/24) and the data center (203.0.113.0/24) networks.
  - Cannot communicate with the Raleigh network (198.51.100.0/24) because Router 1 has no interface in this network. Therefore, Router 1 sends the traffic to its own default gateway (internet).

- Hosts in the data center network (203.0.113.0/24):
  - Can communicate with other hosts in the same subnet because they are directly connected.
  - Can communicate with the internet because they have their default gateway set to Router 1, and Router 1 has interfaces in both networks, the data center (203.0.113.0/24) and to the internet.
- Can communicate with the Berlin network (192.0.2.0/24) because they have their default gateway set to Router 1, and Router 1 has interfaces in both the data center (203.0.113.0/24) and the Berlin (192.0.2.0/24) networks.

- Cannot communicate with the Raleigh network (198.51.100.0/24) because the data center network has no interface in this network. Therefore, hosts in the data center network (203.0.113.0/24) send traffic to their default gateway (Router 1). Router 1 also has no interface in the Raleigh network (198.51.100.0/24) and, as a result, Router 1 sends this traffic to its own default gateway (internet).

- Hosts in the Raleigh network (198.51.100.0/24):
  - Can communicate with other hosts in the same subnet because they are directly connected.
  - Cannot communicate with hosts on the internet. Router 2 sends the traffic to Router 1 because of the default gateway settings. The actual behavior of Router 1 depends on the reverse path filter (rp_filter) system control (sysctl) setting. By default on RHEL, Router 1 drops the outgoing traffic instead of routing it to the internet. However, regardless of the configured behavior, communication is not possible without the static route.
  - Cannot communicate with the data center network (203.0.113.0/24). The outgoing traffic reaches the destination through Router 2 because of the default gateway setting. However, replies to packets do not reach the sender because hosts in the data center network (203.0.113.0/24) send replies to their default gateway (Router 1). Router 1 then sends the traffic to the internet.
  - Cannot communicate with the Berlin network (192.0.2.0/24). Router 2 sends the traffic to Router 1 because of the default gateway settings. The actual behavior of Router 1 depends on the rp_filter sysctl setting. By default on RHEL, Router 1 drops the outgoing traffic instead of sending it to the Berlin network (192.0.2.0/24). However, regardless of the configured behavior, communication is not possible without the static route.

**NOTE**

In addition to configuring the static routes, you must enable IP forwarding on both routers.

Additional resources

- Why cannot a server be pinged if net.ipv4.conf.all rp_filter is set on the server?
- Enabling IP forwarding

### 25.2. HOW TO USE THE NMCLi COMMAND TO CONFIGURE A STATIC ROUTE

To configure a static route, use the `nmcli` utility with the following syntax:

```bash
$ nmcli connection modify connection_name ipv4.routes "ip[/prefix] [next_hop] [metric] [attribute=value] [attribute=value] ...
```

The command supports the following route attributes:

- `cwnd=n`: Sets the congestion window (CWND) size, defined in number of packets.
- **lock-cwnd=true|false**: Defines whether or not the kernel can update the CWND value.

- **lock-mtu=true|false**: Defines whether or not the kernel can update the MTU to path MTU discovery.

- **lock-window=true|false**: Defines whether or not the kernel can update the maximum window size for TCP packets.

- **mtu=n**: Sets the maximum transfer unit (MTU) to use along the path to the destination.

- **onlink=true|false**: Defines whether the next hop is directly attached to this link even if it does not match any interface prefix.

- **scope=n**: For an IPv4 route, this attribute sets the scope of the destinations covered by the route prefix. Set the value as an integer (0-255).

- **src=address**: Sets the source address to prefer when sending traffic to the destinations covered by the route prefix.

- **table=table_id**: Sets the ID of the table the route should be added to. If you omit this parameter, NetworkManager uses the main table.

- **tos=n**: Sets the type of service (TOS) key. Set the value as an integer (0-255).

- **type=value**: Sets the route type. NetworkManager supports the *unicast*, *local*, *blackhole*, *unreachable*, *prohibit*, and *throw* route types. The default is *unicast*.

- **window=n**: Sets the maximal window size for TCP to advertise to these destinations, measured in bytes.

If you use the `ipv4.routes` sub-command, `nmcli` overrides all current settings of this parameter.

To add a route:

```bash
$ nmcli connection modify connection_name +ipv4.routes "<route>"
```

Similarly, to remove a specific route:

```bash
$ nmcli connection modify connection_name -ipv4.routes "<route>"
```

### 25.3. CONFIGURING A STATIC ROUTE BY USING NMCLI

You can add a static route to an existing NetworkManager connection profile using the `nmcli connection modify` command.

The procedure below configures the following routes:

- An IPv4 route to the remote **198.51.100.0/24** network. The corresponding gateway with the IP address **192.0.2.10** is reachable through the **example** connection.

- An IPv6 route to the remote **2001:db8:2::/64** network. The corresponding gateway with the IP address **2001:db8:1::10** is reachable through the **example** connection.

Prerequisites
The example connection profile exists and it configures this host to be in the same IP subnet as the gateways.

Procedure

1. Add the static IPv4 route to the example connection profile:

   ```bash
   # nmcli connection modify example +ipv4.routes "198.51.100.0/24 192.0.2.10"
   ```

   To set multiple routes in one step, pass the individual routes comma-separated to the command. For example, to add a route to the 198.51.100.0/24 and 203.0.113.0/24 networks, both routed through the 192.0.2.10 gateway, enter:

   ```bash
   # nmcli connection modify example +ipv4.routes "198.51.100.0/24 192.0.2.10,
   203.0.113.0/24 192.0.2.10"
   ```

2. Add the static IPv6 route to the example connection profile:

   ```bash
   # nmcli connection modify example +ipv6.routes "2001:db8:2::/64 2001:db8:1::10"
   ```

3. Re-activate the connection:

   ```bash
   # nmcli connection up example
   ```

Verification

1. Display the IPv4 routes:

   ```bash
   # ip -4 route
   ...
   198.51.100.0/24 via 192.0.2.10 dev enp1s0
   ```

2. Display the IPv6 routes:

   ```bash
   # ip -6 route
   ...
   2001:db8:2::/64 via 2001:db8:1::10 dev enp1s0 metric 1024 pref medium
   ```

Additional resources

- nmcli(1) man page
- nm-settings-nmcli(5) man page

**25.4. CONFIGURING A STATIC ROUTE BY USING NMTUI**

The nmtui application provides a text-based user interface for NetworkManager. You can use nmtui to configure static routes on a host without a graphical interface.

For example, the procedure below adds a route to the 192.0.2.0/24 network that uses the gateway running on 198.51.100.1, which is reachable through an existing connection profile.
NOTE

In `nmtui`:

- Navigate by using the cursor keys.
- Press a button by selecting it and hitting \texttt{Enter}.
- Select and deselect checkboxes by using \texttt{Space}.

Prerequisites

- The network is configured.
- The gateway for the static route must be directly reachable on the interface.
- If the user is logged in on a physical console, user permissions are sufficient. Otherwise, the command requires root permissions.

Procedure

1. Start `nmtui`:
   ```bash
   # nmtui
   ```

2. Select \textbf{Edit a connection}, and press \texttt{Enter}.

3. Select the connection profile through which you can reach the next hop to the destination network, and press \texttt{Enter}.

4. Depending on whether it is an IPv4 or IPv6 route, press the \textbf{Show} button next to the protocol’s configuration area.

5. Press the \textbf{Edit} button next to \textbf{Routing}. This opens a new window where you configure static routes:
   a. Press the \textbf{Add} button and fill in:
      - The destination network, including the prefix in Classless Inter-Domain Routing (CIDR) format
      - The IP address of the next hop
      - A metric value, if you add multiple routes to the same network and want to prioritize the routes by efficiency
   b. Repeat the previous step for every route you want to add and that is reachable through this connection profile.
   c. Press the \textbf{OK} button to return to the window with the connection settings.
6. Press the OK button to return to the nmtui main menu.

7. Select Activate a connection and press Enter.

8. Select the connection profile that you edited, and press Enter twice to deactivate and activate it again.

**IMPORTANT**

Skip this step if you run nmtui over a remote connection, such as SSH, that uses the connection profile you want to reactivate. In this case, if you would deactivate it in nmtui, the connection is terminated and, consequently, you cannot activate it again. To avoid this problem, use the nmcli connection connection_profile_name up command to reactivate the connection in the mentioned scenario.

9. Press the Back button to return to the main menu.

10. Select Quit, and press Enter to close the nmtui application.

**Verification**

- Verify that the route is active:

```
$ ip route
...
192.0.2.0/24 via 198.51.100.1 dev example proto static metric 100
```

**25.5. CONFIGURING A STATIC ROUTE BY USING CONTROL-CENTER**

You can use control-center in GNOME to add a static route to the configuration of a network connection.

The procedure below configures the following routes:

- An IPv4 route to the remote **198.51.100.0/24** network. The corresponding gateway has the IP address **192.0.2.10**.
- An IPv6 route to the remote **2001:db8:2::/64** network. The corresponding gateway has the IP address **2001:db8:1::10**.

**Prerequisites**

- The network is configured.
This host is in the same IP subnet as the gateways.

The network configuration of the connection is opened in the control-center application. See Configuring an Ethernet connection by using nm-connection-editor.

Procedure

1. On the IPv4 tab:
   a. Optional: Disable automatic routes by clicking the On button in the Routes section of the IPv4 tab to use only static routes. If automatic routes are enabled, Red Hat Enterprise Linux uses static routes and routes received from a DHCP server.
   b. Enter the address, netmask, gateway, and optionally a metric value of the IPv4 route:

<table>
<thead>
<tr>
<th>Routes</th>
<th>Automatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td>Netmask</td>
</tr>
<tr>
<td>198.51.100.0</td>
<td>24</td>
</tr>
</tbody>
</table>

2. On the IPv6 tab:
   a. Optional: Disable automatic routes by clicking the On button in the Routes section of the IPv4 tab to use only static routes.
   b. Enter the address, netmask, gateway, and optionally a metric value of the IPv6 route:

<table>
<thead>
<tr>
<th>Routes</th>
<th>Automatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Address</td>
<td>Prefix</td>
</tr>
<tr>
<td>2001:db8::</td>
<td>64</td>
</tr>
</tbody>
</table>

3. Click Apply.

4. Back in the Network window, disable and re-enable the connection by switching the button for the connection to Off and back to On for changes to take effect.

   **WARNING**

   Restarting the connection briefly disrupts connectivity on that interface.

Verification

1. Display the IPv4 routes:

   ```
   # ip -4 route
   ...
   198.51.100.0/24 via 192.0.2.10 dev enp1s0
   ```

2. Display the IPv6 routes:
25.6. CONFIGURING A STATIC ROUTE BY USING NM-CONNECTION-EDITOR

You can use the `nm-connection-editor` application to add a static route to the configuration of a network connection.

The procedure below configures the following routes:

- An IPv4 route to the remote 198.51.100.0/24 network. The corresponding gateway with the IP address 192.0.2.10 is reachable through the example connection.
- An IPv6 route to the remote 2001:db8:2::/64 network. The corresponding gateway with the IP address 2001:db8:1::10 is reachable through the example connection.

Prerequisites

- The network is configured.
- This host is in the same IP subnet as the gateways.

Procedure

1. Open a terminal, and enter `nm-connection-editor`:

   ```
   $ nm-connection-editor
   ```

2. Select the example connection profile, and click the gear wheel icon to edit the existing connection.

3. On the IPv4 Settings tab:
   a. Click the Routes button.
   b. Click the Add button and enter the address, netmask, gateway, and optionally a metric value.
   c. Click OK.

4. On the IPv6 Settings tab:
   a. Click the Routes button.
b. Click the **Add** button and enter the address, netmask, gateway, and optionally a metric value.

![Editing IPv6 routes for example](image)

<table>
<thead>
<tr>
<th>Address</th>
<th>Prefix</th>
<th>Gateway</th>
<th>Metric</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001:db8:2::</td>
<td>64</td>
<td>2001:db8:1::10</td>
<td></td>
</tr>
</tbody>
</table>

> Add  Delete


c. Click **OK**.

5. Click **Save**.

6. Restart the network connection for changes to take effect. For example, to restart the **example** connection using the command line:

```bash
# nmcli connection up example
```

**Verification**

1. Display the IPv4 routes:

```bash
# ip -4 route
... 198.51.100.0/24 via 192.0.2.10 dev enp1s0
```

2. Display the IPv6 routes:

```bash
# ip -6 route
... 2001:db8:2::/64 via 2001:db8:1::10 dev enp1s0 metric 1024 pref medium
```

**25.7. CONFIGURING A STATIC ROUTE BY USING THE NMCLI INTERACTIVE MODE**

You can use the interactive mode of the **nmcli** utility to add a static route to the configuration of a network connection.

The procedure below configures the following routes:

- An IPv4 route to the remote **198.51.100.0/24** network. The corresponding gateway with the IP address **192.0.2.10** is reachable through the **example** connection.

- An IPv6 route to the remote **2001:db8:2::/64** network. The corresponding gateway with the IP address **2001:db8:1::10** is reachable through the **example** connection.

**Prerequisites**

- The **example** connection profile exists and it configures this host to be in the same IP subnet as the gateways.
Procedure

1. Open the nmcli interactive mode for the example connection:

        # nmcli connection edit example

2. Add the static IPv4 route:

        nmcli> set ipv4.routes 198.51.100.0/24 192.0.2.10

3. Add the static IPv6 route:

        nmcli> set ipv6.routes 2001:db8:2::/64 2001:db8:1::10

4. Optionally, verify that the routes were added correctly to the configuration:

        nmcli> print

        ... 
        ipv4.routes: { ip = 198.51.100.0/24, nh = 192.0.2.10 }
        ... 
        ipv6.routes: { ip = 2001:db8:2::/64, nh = 2001:db8:1::10 }
        ...

        The **ip** attribute displays the network to route and the **nh** attribute the gateway (next hop).

5. Save the configuration:

        nmcli> save persistent

6. Restart the network connection:

        nmcli> activate example

7. Leave the nmcli interactive mode:

        nmcli> quit

Verification

1. Display the IPv4 routes:

        # ip -4 route

        ... 
        198.51.100.0/24 via 192.0.2.10 dev enp1s0

2. Display the IPv6 routes:

        # ip -6 route

        ... 
        2001:db8:2::/64 via 2001:db8:1::10 dev enp1s0 metric 1024 pref medium

Additional resources
25.8. CONFIGURING A STATIC ROUTE BY USING NMSTATECTL

Use the `nmstatectl` utility to configure a static route through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.

Prerequisites

- The `enp1s0` network interface is configured and is in the same IP subnet as the gateways.
- The `nmstate` package is installed.

Procedure

1. Create a YAML file, for example `~/add-static-route-to-enp1s0.yml`, with the following content:

```yaml
---
routes:
  config:
  - destination: 198.51.100.0/24
    next-hop-address: 192.0.2.10
    next-hop-interface: enp1s0
  - destination: 2001:db8:2::/64
    next-hop-address: 2001:db8:1::10
    next-hop-interface: enp1s0
```

These settings define the following static routes:

- An IPv4 route to the remote 198.51.100.0/24 network. The corresponding gateway with the IP address 192.0.2.10 is reachable through the `enp1s0` interface.

- An IPv6 route to the remote 2001:db8:2::/64 network. The corresponding gateway with the IP address 2001:db8:1::10 is reachable through the `enp1s0` interface.

2. Apply the settings to the system:

```bash
# nmstatectl apply ~/add-static-route-to-enp1s0.yml
```

Verification

1. Display the IPv4 routes:

```bash
# ip -4 route
... 198.51.100.0/24 via 192.0.2.10 dev enp1s0
```

2. Display the IPv6 routes:
# ip -6 route
...
2001:db8:2::/64 via 2001:db8:1::10 dev enp1s0 metric 1024 pref medium

Additional resources
- nmstatectl(8) man page
- /usr/share/doc/nmstate/examples/ directory

25.9. CONFIGURING A STATIC ROUTE BY USING THE NETWORK RHEL SYSTEM ROLE

You can use the network RHEL system role to configure static routes.

IMPORTANT
When you run a play that uses the network RHEL system role and if the setting values do not match the values specified in the play, the role overrides the existing connection profile with the same name. To prevent resetting these values to their defaults, always specify the whole configuration of the network connection profile in the play, even if the configuration, for example the IP configuration, already exists.

Prerequisites
- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has sudo permissions on them.

Procedure
1. Create a playbook file, for example ~/playbook.yml, with the following content:

```yaml
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure an Ethernet connection with static IP and additional routes
      ansible.builtin.include_role:
        name: rhel-system-roles.network
        vars:
          network_connections:
            - name: enp7s0
              type: ethernet
              autoconnect: yes
              ip:
                address:
                  - 192.0.2.1/24
                  - 2001:db8:1::1/64
              gateway4: 192.0.2.254
              gateway6: 2001:db8:1::ffe
dns:
```
Depending on whether it already exists, the procedure creates or updates the `enp7s0` connection profile with the following settings:

- A static IPv4 address - \texttt{192.0.2.1} with a \texttt{/24} subnet mask
- A static IPv6 address - \texttt{2001:db8:1::1} with a \texttt{/64} subnet mask
- An IPv4 default gateway - \texttt{192.0.2.254}
- An IPv6 default gateway - \texttt{2001:db8:1::fffe}
- An IPv4 DNS server - \texttt{192.0.2.200}
- An IPv6 DNS server - \texttt{2001:db8:1::ffbb}
- A DNS search domain - \texttt{example.com}
- Static routes:
  - \texttt{198.51.100.0/24} with gateway \texttt{192.0.2.10}
  - \texttt{2001:db8:2::/64} with gateway \texttt{2001:db8:1::10}

2. Validate the playbook syntax:

```
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```
$ ansible-playbook ~/playbook.yml
```

**Verification**

1. On the managed nodes:
   a. Display the IPv4 routes:

```
# ip -4 route
...
198.51.100.0/24 via 192.0.2.10 dev enp7s0
```
b. Display the IPv6 routes:

```
# ip -6 route
...
2001:db8:2::/64 via 2001:db8:1::10 dev enp7s0 metric 1024 pref medium
```

Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory
CHAPTER 26. CONFIGURING POLICY-BASED ROUTING TO DEFINE ALTERNATIVE ROUTES

By default, the kernel in RHEL decides where to forward network packets based on the destination address using a routing table. Policy-based routing enables you to configure complex routing scenarios. For example, you can route packets based on various criteria, such as the source address, packet metadata, or protocol.

NOTE
On systems that use NetworkManager, only the `nmcli` utility supports setting routing rules and assigning routes to specific tables.

26.1. ROUTING TRAFFIC FROM A SPECIFIC SUBNET TO A DIFFERENT DEFAULT GATEWAY BY USING NMCLI

You can use policy-based routing to configure a different default gateway for traffic from certain subnets. For example, you can configure RHEL as a router that, by default, routes all traffic to internet provider A using the default route. However, traffic received from the internal workstations subnet is routed to provider B.

The procedure assumes the following network topology:

Prerequisites

- The system uses NetworkManager to configure the network, which is the default.
- The RHEL router you want to set up in the procedure has four network interfaces:
  - The `enp7s0` interface is connected to the network of provider A. The gateway IP in the provider’s network is `198.51.100.2`, and the network uses a `/30` network mask.
  - The `enp1s0` interface is connected to the network of provider B. The gateway IP in the provider’s network is `192.0.2.2`, and the network uses a `/30` network mask.
  - The `enp8s0` interface is connected to the `10.0.0.0/24` subnet with internal workstations.
The `enp9s0` interface is connected to the `203.0.113.0/24` subnet with the company’s servers.

- Hosts in the internal workstations subnet use `10.0.0.1` as the default gateway. In the procedure, you assign this IP address to the `enp8s0` network interface of the router.

- Hosts in the server subnet use `203.0.113.1` as the default gateway. In the procedure, you assign this IP address to the `enp9s0` network interface of the router.

- The `firewalld` service is enabled and active.

**Procedure**

1. Configure the network interface to provider A:

   ```
   # nmcli connection add type ethernet con-name Provider-A ifname enp7s0
   ipv4.method manual ipv4.addresses 198.51.100.1/30 ipv4.gateway 198.51.100.2
   ipv4.dns 198.51.100.200 connection.zone external
   ```

   The `nmcli connection add` command creates a NetworkManager connection profile. The command uses the following options:

   - `type ethernet`: Defines that the connection type is Ethernet.
   - `con-name connection_name`: Sets the name of the profile. Use a meaningful name to avoid confusion.
   - `ifname network_device`: Sets the network interface.
   - `ipv4.method manual`: Enables to configure a static IP address.
   - `ipv4.addresses IP_address/subnet_mask`: Sets the IPv4 addresses and subnet mask.
   - `ipv4.gateway IP_address`: Sets the default gateway address.
   - `ipv4.dns IP_of_DNS_server`: Sets the IPv4 address of the DNS server.
   - `connection.zone firewalld_zone`: Assigns the network interface to the defined `firewalld` zone. Note that `firewalld` automatically enables masquerading for interfaces assigned to the `external` zone.

2. Configure the network interface to provider B:

   ```
   # nmcli connection add type ethernet con-name Provider-B ifname enp1s0
   ipv4.method manual ipv4.addresses 192.0.2.1/30 ipv4.routes "0.0.0.0/0 192.0.2.2 table=5000" connection.zone external
   ```

   This command uses the `ipv4.routes` parameter instead of `ipv4.gateway` to set the default gateway. This is required to assign the default gateway for this connection to a different routing table (5000) than the default. NetworkManager automatically creates this new routing table when the connection is activated.

3. Configure the network interface to the internal workstations subnet:
# nmcli connection add type ethernet con-name Internal-Workstations ifname enp8s0 ipv4.method manual ipv4.addresses 10.0.0.1/24 ipv4.routes "10.0.0.0/24 table=5000" ipv4.routing-rules "priority 5 from 10.0.0.0/24 table 5000" connection.zone trusted

This command uses the `ipv4.routes` parameter to add a static route to the routing table with ID `5000`. This static route for the `10.0.0.0/24` subnet uses the IP of the local network interface to provider B (`192.0.2.1`) as next hop.

Additionally, the command uses the `ipv4.routing-rules` parameter to add a routing rule with priority `5` that routes traffic from the `10.0.0.0/24` subnet to table `5000`. Low values have a high priority.

Note that the syntax in the `ipv4.routing-rules` parameter is the same as in an `ip rule add` command, except that `ipv4.routing-rules` always requires specifying a priority.

4. Configure the network interface to the server subnet:

```
# nmcli connection add type ethernet con-name Servers ifname enp9s0 ipv4.method manual ipv4.addresses 203.0.113.1/24 connection.zone trusted
```

Verification

1. On a RHEL host in the internal workstation subnet:
   a. Install the `traceroute` package:

```
# dnf install traceroute
```

   b. Use the `traceroute` utility to display the route to a host on the internet:

```
# traceroute redhat.com
traceroute to redhat.com (209.132.183.105), 30 hops max, 60 byte packets
  1  10.0.0.1 (10.0.0.1)     0.337 ms  0.260 ms  0.223 ms
  2  192.0.2.1 (192.0.2.1)   0.884 ms  1.066 ms  1.248 ms
...
```

The output of the command displays that the router sends packets over `192.0.2.1`, which is the network of provider B.

2. On a RHEL host in the server subnet:
   a. Install the `traceroute` package:

```
# dnf install traceroute
```

   b. Use the `traceroute` utility to display the route to a host on the internet:

```
# traceroute redhat.com
traceroute to redhat.com (209.132.183.105), 30 hops max, 60 byte packets
  1  203.0.113.1 (203.0.113.1)    2.179 ms  2.073 ms  1.944 ms
  2  198.51.100.2 (198.51.100.2)  1.868 ms  1.798 ms  1.549 ms
...
```
The output of the command displays that the router sends packets over 198.51.100.2, which is the network of provider A.

Troubleshooting steps

On the RHEL router:

1. Display the rule list:

```bash
# ip rule list
0: from all lookup local
5: from 10.0.0.0/24 lookup 5000
  32766: from all lookup main
  32767: from all lookup default
```

By default, RHEL contains rules for the tables local, main, and default.

2. Display the routes in table 5000:

```bash
# ip route list table 5000
0.0.0.0/0 via 192.0.2.2 dev enp1s0 proto static metric 100
10.0.0.0/24 dev enp8s0 proto static scope link src 192.0.2.1 metric 102
```

3. Display the interfaces and firewall zones:

```bash
# firewall-cmd --get-active-zones
external
  interfaces: enp1s0 enp7s0
trusted
  interfaces: enp8s0 enp9s0
```

4. Verify that the external zone has masquerading enabled:

```bash
# firewall-cmd --info-zone=external
external (active)
  target: default
  icmp-block-inversion: no
  interfaces: enp1s0 enp7s0
  sources:
  services: ssh
  ports:
  protocols:
    masquerade: yes
```

Additional resources

- `nm-settings(5)` man page
- `nmcli(1)` man page
- Is it possible to set up Policy Based Routing with NetworkManager in RHEL?
26.2. ROUTING TRAFFIC FROM A SPECIFIC SUBNET TO A DIFFERENT DEFAULT GATEWAY BY USING THE NETWORK RHEL SYSTEM ROLE

You can use policy-based routing to configure a different default gateway for traffic from certain subnets. For example, you can configure RHEL as a router that, by default, routes all traffic to internet provider A using the default route. However, traffic received from the internal workstations subnet is routed to provider B.

To configure policy-based routing remotely and on multiple nodes, you can use the network RHEL system role.

This procedure assumes the following network topology:

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has sudo permissions on them.
- The managed nodes uses the NetworkManager and firewalld services.
- The managed nodes you want to configure has four network interfaces:
  - The enp7s0 interface is connected to the network of provider A. The gateway IP in the provider’s network is 198.51.100.2, and the network uses a /30 network mask.
  - The enp1s0 interface is connected to the network of provider B. The gateway IP in the provider’s network is 192.0.1.2, and the network uses a /30 network mask.
  - The enp8s0 interface is connected to the 10.0.0.0/24 subnet with internal workstations.
  - The enp9s0 interface is connected to the 203.0.113.0/24 subnet with the company’s servers.

- Hosts in the internal workstations subnet use 10.0.0.1 as the default gateway. In the procedure, you assign this IP address to the enp8s0 network interface of the router.
Hosts in the server subnet use **203.0.113.1** as the default gateway. In the procedure, you assign this IP address to the **enp9s0** network interface of the router.

**Procedure**

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```yaml
---
- name: Configuring policy-based routing
  hosts: managed-node-01.example.com
  tasks:
    - name: Routing traffic from a specific subnet to a different default gateway
      ansible.builtin.include_role:
        name: rhel-system-roles.network
        vars:
          network_connections:
            - name: Provider-A
              interface_name: enp7s0
              type: ethernet
              autoconnect: True
              ip:
                address:
                  - 198.51.100.1/30
              gateway4: 198.51.100.2
              dns:
                - 198.51.100.200
              state: up
              zone: external
            - name: Provider-B
              interface_name: enp1s0
              type: ethernet
              autoconnect: True
              ip:
                address:
                  - 192.0.2.1/30
              route:
                - network: 0.0.0.0
                  prefix: 0
                  gateway: 192.0.2.2
                  table: 5000
              state: up
              zone: external
            - name: Internal-Workstations
              interface_name: enp8s0
              type: ethernet
              autoconnect: True
              ip:
                address:
                  - 10.0.0.1/24
              route:
                - network: 10.0.0.0
                  prefix: 24
                  table: 5000
              routing_rule:
```

---

Red Hat Enterprise Linux 9 Configuring and managing networking
2. Validate the playbook syntax:

```bash
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```bash
$ ansible-playbook ~/playbook.yml
```

Verification

1. On a RHEL host in the internal workstation subnet:
   a. Install the `traceroute` package:

   ```bash
   # dnf install traceroute
   ```

   b. Use the `traceroute` utility to display the route to a host on the internet:

   ```bash
   # traceroute redhat.com
   traceroute to redhat.com (209.132.183.105), 30 hops max, 60 byte packets
   1  10.0.0.1 (10.0.0.1)     0.337 ms  0.260 ms  0.223 ms
   2  192.0.2.1 (192.0.2.1)   0.884 ms  1.066 ms  1.248 ms
   ...
   
   The output of the command displays that the router sends packets over `192.0.2.1`, which is the network of provider B.

2. On a RHEL host in the server subnet:
   a. Install the `traceroute` package:

   ```bash
   # dnf install traceroute
   ```

   b. Use the `traceroute` utility to display the route to a host on the internet:
3. On the RHEL router that you configured using the RHEL system role:
   a. Display the rule list:

   ```bash
   # ip rule list
   0: from all lookup local
   5: from 10.0.0.0/24 lookup 5000
   32766: from all lookup main
   32767: from all lookup default
   
   By default, RHEL contains rules for the tables local, main, and default.
   
   b. Display the routes in table 5000:

   ```bash
   # ip route list table 5000
   0.0.0.0/0 via 192.0.2.2 dev enp1s0 proto static metric 100
   10.0.0.0/24 dev enp8s0 proto static scope link src 192.0.2.1 metric 102
   
   c. Display the interfaces and firewall zones:

   ```bash
   # firewall-cmd --get-active-zones
   external
     interfaces: enp1s0 enp7s0
   trusted
     interfaces: enp8s0 enp9s0
   
   d. Verify that the external zone has masquerading enabled:

   ```bash
   # firewall-cmd --info-zone=external
   external (active)
     target: default
     icmp-block-inversion: no
     interfaces: enp1s0 enp7s0
     sources:
     services: ssh
     ports:
     protocols:
     masquerade: yes
   ...
CHAPTER 27. REUSING THE SAME IP ADDRESS ON DIFFERENT INTERFACES

With Virtual routing and forwarding (VRF), administrators can use multiple routing tables simultaneously on the same host. For that, VRF partitions a network at layer 3. This enables the administrator to isolate traffic using separate and independent route tables per VRF domain. This technique is similar to virtual LANs (VLAN), which partitions a network at layer 2, where the operating system uses different VLAN tags to isolate traffic sharing the same physical medium.

One benefit of VRF over partitioning on layer 2 is that routing scales better considering the number of peers involved.

Red Hat Enterprise Linux uses a virtual vrt device for each VRF domain and adds routes to a VRF domain by adding existing network devices to a VRF device. Addresses and routes previously attached to the original device will be moved inside the VRF domain.

Note that each VRF domain is isolated from each other.

27.1. PERMANENTLY REUSING THE SAME IP ADDRESS ON DIFFERENT INTERFACES

You can use the virtual routing and forwarding (VRF) feature to permanently use the same IP address on different interfaces in one server.

IMPORTANT

To enable remote peers to contact both VRF interfaces while reusing the same IP address, the network interfaces must belong to different broadcasting domains. A broadcast domain in a network is a set of nodes, which receive broadcast traffic sent by any of them. In most configurations, all nodes connected to the same switch belong to the same broadcasting domain.

Prerequisites

- You are logged in as the root user.
- The network interfaces are not configured.

Procedure

1. Create and configure the first VRF device:
   a. Create a connection for the VRF device and assign it to a routing table. For example, to create a VRF device named vrf0 that is assigned to the 1001 routing table:

   ```
   # nmcli connection add type vrf ifname vrf0 con-name vrf0 table 1001 ipv4.method disabled ipv6.method disabled
   ```

   b. Enable the vrf0 device:

   ```
   # nmcli connection up vrf0
   ```
c. Assign a network device to the VRF just created. For example, to add the `enp1s0` Ethernet device to the `vrf0` VRF device and assign an IP address and the subnet mask to `enp1s0`, enter:

```
# nmcli connection add type ethernet con-name vrf.enp1s0 ifname enp1s0 master vrf0 ipv4.method manual ipv4.address 192.0.2.1/24
```

d. Activate the `vrf.enp1s0` connection:

```
# nmcli connection up vrf.enp1s0
```

2. Create and configure the next VRF device:

a. Create the VRF device and assign it to a routing table. For example, to create a VRF device named `vrf1` that is assigned to the `1002` routing table, enter:

```
# nmcli connection add type vrf ifname vrf1 con-name vrf1 table 1002 ipv4.method disabled ipv6.method disabled
```

b. Activate the `vrf1` device:

```
# nmcli connection up vrf1
```

c. Assign a network device to the VRF just created. For example, to add the `enp7s0` Ethernet device to the `vrf1` VRF device and assign an IP address and the subnet mask to `enp7s0`, enter:

```
# nmcli connection add type ethernet con-name vrf.enp7s0 ifname enp7s0 master vrf1 ipv4.method manual ipv4.address 192.0.2.1/24
```

d. Activate the `vrf.enp7s0` device:

```
# nmcli connection up vrf.enp7s0
```

27.2. TEMPORARILY REUSING THE SAME IP ADDRESS ON DIFFERENT INTERFACES

You can use the virtual routing and forwarding (VRF) feature to temporarily use the same IP address on different interfaces in one server. Use this procedure only for testing purposes, because the configuration is temporary and lost after you reboot the system.

IMPORTANT

To enable remote peers to contact both VRF interfaces while reusing the same IP address, the network interfaces must belong to different broadcasting domains. A broadcast domain in a network is a set of nodes which receive broadcast traffic sent by any of them. In most configurations, all nodes connected to the same switch belong to the same broadcasting domain.

Prerequisites

- You are logged in as the root user.
• The network interfaces are not configured.

Procedure

1. Create and configure the first VRF device:
   a. Create the VRF device and assign it to a routing table. For example, to create a VRF device named `blue` that is assigned to the **1001** routing table:

```
# ip link add dev blue type vrf table 1001
```

b. Enable the `blue` device:

```
# ip link set dev blue up
```

c. Assign a network device to the VRF device. For example, to add the `enp1s0` Ethernet device to the `blue` VRF device:

```
# ip link set dev enp1s0 master blue
```

d. Enable the `enp1s0` device:

```
# ip link set dev enp1s0 up
```

e. Assign an IP address and subnet mask to the `enp1s0` device. For example, to set it to `192.0.2.1/24`:

```
# ip addr add dev enp1s0 192.0.2.1/24
```

2. Create and configure the next VRF device:
   a. Create the VRF device and assign it to a routing table. For example, to create a VRF device named `red` that is assigned to the **1002** routing table:

```
# ip link add dev red type vrf table 1002
```

b. Enable the `red` device:

```
# ip link set dev red up
```

c. Assign a network device to the VRF device. For example, to add the `enp7s0` Ethernet device to the `red` VRF device:

```
# ip link set dev enp7s0 master red
```

d. Enable the `enp7s0` device:

```
# ip link set dev enp7s0 up
```

e. Assign the same IP address and subnet mask to the `enp7s0` device as you used for `enp1s0` in the `blue` VRF domain:
# ip addr add dev enp7s0 192.0.2.1/24

3. Optionally, create further VRF devices as described above.

27.3. ADDITIONAL RESOURCES

- `/usr/share/doc/kernel-doc-<kernel_version>/Documentation/networking/vrf.txt` from the `kernel-doc` package
CHAPTER 28. STARTING A SERVICE WITHIN AN ISOLATED VRF NETWORK

With virtual routing and forwarding (VRF), you can create isolated networks with a routing table that is different to the main routing table of the operating system. You can then start services and applications so that they have only access to the network defined in that routing table.

28.1. CONFIGURING A VRF DEVICE

To use virtual routing and forwarding (VRF), you create a VRF device and attach a physical or virtual network interface and routing information to it.

**WARNING**

To prevent that you lock out yourself out remotely, perform this procedure on the local console or remotely over a network interface that you do not want to assign to the VRF device.

**Prerequisites**

- You are logged in locally or using a network interface that is different to the one you want to assign to the VRF device.

**Procedure**

1. Create the `vrf0` connection with a same-named virtual device, and attach it to routing table 1000:

   ```sh
   # nmcli connection add type vrf ifname vrf0 con-name vrf0 table 1000 ipv4.method disabled ipv6.method disabled
   ```

2. Add the `enp1s0` device to the `vrf0` connection, and configure the IP settings:

   ```sh
   # nmcli connection add type ethernet con-name enp1s0 ifname enp1s0 master vrf0 ipv4.method manual ipv4.address 192.0.2.1/24 ipv4.gateway 192.0.2.254
   ```

   This command creates the `enp1s0` connection as a port of the `vrf0` connection. Due to this configuration, the routing information are automatically assigned to the routing table 1000 that is associated with the `vrf0` device.

3. If you require static routes in the isolated network:
   a. Add the static routes:

      ```sh
      # nmcli connection modify enp1s0 +ipv4.routes "198.51.100.0/24 192.0.2.2"
      ```

      This adds a route to the 198.51.100.0/24 network that uses 192.0.2.2 as the router.
   
   b. Activate the connection:
Verification

1. Display the IP settings of the device that is associated with vrf0:

   ```
   # nmcli connection up enp1s0
   ```

   ```
   # ip -br addr show vrf vrf0
   enp1s0   UP     192.0.2.1/24
   ```

2. Display the VRF devices and their associated routing table:

   ```
   # ip vrf show
   Name     Table
   -----------
   vrf0      1000
   ```

3. Display the main routing table:

   ```
   # ip route show
   default via 203.0.113.0/24 dev enp7s0 proto static metric 100
   ```

   The main routing table does not mention any routes associated with the device enp1s0 device or the 192.0.2.1/24 subnet.

4. Display the routing table 1000:

   ```
   # ip route show table 1000
   default via 192.0.2.254 dev enp1s0 proto static metric 101
   broadcast 192.0.2.255 dev enp1s0 proto kernel scope link src 192.0.2.1
   192.0.2.0/24 dev enp1s0 proto kernel scope link src 192.0.2.1 metric 101
   local 192.0.2.1 dev enp1s0 proto kernel scope host src 192.0.2.1
   broadcast 192.0.2.255 dev enp1s0 proto kernel scope link src 192.0.2.1
   198.51.100.0/24 via 192.0.2.2 dev enp1s0 proto static metric 101
   ```

   The default entry indicates that services that use this routing table, use 192.0.2.254 as their default gateway and not the default gateway in the main routing table.

5. Execute the traceroute utility in the network associated with vrf0 to verify that the utility uses the route from table 1000:

   ```
   # ip vrf exec vrf0 traceroute 203.0.113.1
   traceroute to 203.0.113.1 (203.0.113.1), 30 hops max, 60 byte packets
   1 192.0.2.254 (192.0.2.254) 0.516 ms 0.459 ms 0.430 ms
   ... 
   ```

   The first hop is the default gateway that is assigned to the routing table 1000 and not the default gateway from the system’s main routing table.

Additional resources

- ip-vrf(8) man page

28.2. STARTING A SERVICE WITHIN AN ISOLATED VRF NETWORK
You can configure a service, such as the Apache HTTP Server, to start within an isolated virtual routing and forwarding (VRF) network.

**IMPORTANT**

Services can only bind to local IP addresses that are in the same VRF network.

**Prerequisites**

- You configured the `vrf0` device.
- You configured Apache HTTP Server to listen only on the IP address that is assigned to the interface associated with the `vrf0` device.

**Procedure**

1. Display the content of the `httpd` systemd service:

   ```
   # systemctl cat httpd
   ...
   [Service]
   ExecStart=/usr/sbin/httpd $OPTIONS -DFOREGROUND
   ...
   
   You require the content of the **ExecStart** parameter in a later step to run the same command within the isolated VRF network.
   
   2. Create the `/etc/systemd/system/httpd.service.d/` directory:

   ```
   # mkdir /etc/systemd/system/httpd.service.d/
   
   3. Create the `/etc/systemd/system/httpd.service.d/override.conf` file with the following content:

   ```
   [Service]
   ExecStart=/usr/sbin/ip vrf exec vrf0 /usr/sbin/httpd $OPTIONS -DFOREGROUND
   
   To override the **ExecStart** parameter, you first need to unset it and then set it to the new value as shown.
   
   4. Reload systemd.

   ```
   # systemctl daemon-reload
   
   5. Restart the `httpd` service.

   ```
   # systemctl restart httpd
   
   **Verification**

   1. Display the process IDs (PID) of `httpd` processes:
# pidof -c httpd
1904 ...

2. Display the VRF association for the PIDs, for example:

```bash
# ip vrf identify 1904
vrf0
```

3. Display all PIDs associated with the **vrf0** device:

```bash
# ip vrf pids vrf0
1904  httpd
...
```

**Additional resources**

- [ip-vrf(8) man page](#)
CHAPTER 29. CONFIGURING ETHTOOL SETTINGS IN NETWORKMANAGER CONNECTION PROFILES

NetworkManager can configure certain network driver and hardware settings persistently. Compared to using the `ethtool` utility to manage these settings, this has the benefit of not losing the settings after a reboot.

You can set the following `ethtool` settings in NetworkManager connection profiles:

### Offload features
Network interface controllers can use the TCP offload engine (TOE) to offload processing certain operations to the network controller. This improves the network throughput.

### Interrupt coalesce settings
By using interrupt coalescing, the system collects network packets and generates a single interrupt for multiple packets. This increases the amount of data sent to the kernel with one hardware interrupt, which reduces the interrupt load, and maximizes the throughput.

### Ring buffers
These buffers store incoming and outgoing network packets. You can increase the ring buffer sizes to reduce a high packet drop rate.

#### 29.1. CONFIGURING AN ETHTOOL OFFLOAD FEATURE BY USING NMCLI

You can use NetworkManager to enable and disable `ethtool` offload features in a connection profile.

**Procedure**

1. For example, to enable the RX offload feature and disable TX offload in the `enp1s0` connection profile, enter:

   ```
   # nmcli con modify enp1s0 ethtool.feature-rx on ethtool.feature-tx off
   ```

   This command explicitly enables RX offload and disables TX offload.

2. To remove the setting of an offload feature that you previously enabled or disabled, set the feature's parameter to a null value. For example, to remove the configuration for TX offload, enter:

   ```
   # nmcli con modify enp1s0 ethtool.feature-tx ""
   ```

3. Reactivate the network profile:

   ```
   # nmcli connection up enp1s0
   ```

**Verification**

- Use the `ethtool -k` command to display the current offload features of a network device:

  ```
  # ethtool -k network_device
  ```
Additional resources

- nm-settings-nmcli(5) man page

29.2. CONFIGURING AN ETHTOOL OFFLOAD FEATURE BY USING THE NETWORK RHEL SYSTEM ROLE

You can use the network RHEL system role to configure ethtool features of a NetworkManager connection.

IMPORTANT

When you run a play that uses the network RHEL system role and if the setting values do not match the values specified in the play, the role overrides the existing connection profile with the same name. To prevent resetting these values to their defaults, always specify the whole configuration of the network connection profile in the play, even if the configuration, for example the IP configuration, already exists.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has sudo permissions on them.

Procedure

1. Create a playbook file, for example ~/playbook.yml, with the following content:

```yaml
---
- name: Configure the network
  hosts: managed-node-01.example.com
  tasks:
    - name: Configure an Ethernet connection with ethtool features
      ansible.builtin.include_role:
        name: rhel-system-roles.network
        vars:
          network_connections:
            - name: enp1s0
              type: ethernet
              autoconnect: yes
              ip:
                address:
                  - 198.51.100.20/24
                  - 2001:db8:1::1/64
                gateway4: 198.51.100.254
                gateway6: 2001:db8:1::fffe
                dns:
                  - 198.51.100.200
                  - 2001:db8:1::ffbb
                dns_search:
                  - example.com
              ethtool:
                features:
```
This playbook creates the `enp1s0` connection profile with the following settings, or updates it if the profile already exists:

- A static IPv4 address - **198.51.100.20** with a **/24** subnet mask
- A static IPv6 address - **2001:db8:1::1** with a **/64** subnet mask
- An IPv4 default gateway - **198.51.100.254**
- An IPv6 default gateway - **2001:db8:1::fffe**
- An IPv4 DNS server - **198.51.100.200**
- An IPv6 DNS server - **2001:db8:1::ffbb**
- A DNS search domain - **example.com**
- `ethtool` features:
  - Generic receive offload (GRO): disabled
  - Generic segmentation offload (GSO): enabled
  - TX stream control transmission protocol (SCTP) segmentation: disabled

2. Validate the playbook syntax:

```sh
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```sh
$ ansible-playbook ~/playbook.yml
```

Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory

### 29.3. Configuring an `ethtool` Coalesce Settings by Using NMCLI

You can use NetworkManager to set `ethtool` coalesce settings in connection profiles.

**Procedure**
1. For example, to set the maximum number of received packets to delay to 128 in the `enp1s0` connection profile, enter:

```
# nmcli connection modify enp1s0 ethtool.coalesce-rx-frames 128
```

2. To remove a coalesce setting, set it to a null value. For example, to remove the `ethtool.coalesce-rx-frames` setting, enter:

```
# nmcli connection modify enp1s0 ethtool.coalesce-rx-frames ""
```

3. To reactivate the network profile:

```
# nmcli connection up enp1s0
```

**Verification**

1. Use the `ethtool -c` command to display the current offload features of a network device:

```
# ethtool -c network_device
```

**Additional resources**

- `nm-settings-nmcli(5)` man page

### 29.4. Configuring an Ethtool Coalesce Settings by Using the Network RHEL System Role

You can use the `network` RHEL system role to configure `ethtool` coalesce settings of a NetworkManager connection.

**IMPORTANT**

When you run a play that uses the `network` RHEL system role and if the setting values do not match the values specified in the play, the role overrides the existing connection profile with the same name. To prevent resetting these values to their defaults, always specify the whole configuration of the network connection profile in the play, even if the configuration, for example the IP configuration, already exists.

**Prerequisites**

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.

**Procedure**

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

```
---
  - name: Configure the network
    ```
This playbook creates the `enp1s0` connection profile with the following settings, or updates it if the profile already exists:

- A static IPv4 address - `198.51.100.20` with a `/24` subnet mask
- A static IPv6 address - `2001:db8:1::1` with a `/64` subnet mask
- An IPv4 default gateway - `198.51.100.254`
- An IPv6 default gateway - `2001:db8:1::fffe`
- An IPv4 DNS server - `198.51.100.200`
- An IPv6 DNS server - `2001:db8:1::ffbb`
- A DNS search domain - `example.com`
- `ethtool` coalesce settings:
  - RX frames: `128`
  - TX frames: `128`

2. Validate the playbook syntax:

```bash
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.
3. Run the playbook:

```
$ ansible-playbook ~/playbook.yml
```

Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
- `/usr/share/doc/rhel-system-roles/network/` directory

## 29.5. INCREASING THE RING BUFFER SIZE TO REDUCE A HIGH PACKET DROP RATE BY USING NMCLI

Increase the size of an Ethernet device’s ring buffers if the packet drop rate causes applications to report a loss of data, timeouts, or other issues.

Receive ring buffers are shared between the device driver and network interface controller (NIC). The card assigns a transmit (TX) and receive (RX) ring buffer. As the name implies, the ring buffer is a circular buffer where an overflow overwrites existing data. There are two ways to move data from the NIC to the kernel, hardware interrupts and software interrupts, also called SoftIRQs.

The kernel uses the RX ring buffer to store incoming packets until the device driver can process them. The device driver drains the RX ring, typically by using SoftIRQs, which puts the incoming packets into a kernel data structure called an `sk_buff` or `skb` to begin its journey through the kernel and up to the application that owns the relevant socket.

The kernel uses the TX ring buffer to hold outgoing packets which should be sent to the network. These ring buffers reside at the bottom of the stack and are a crucial point at which packet drop can occur, which in turn will adversely affect network performance.

### Procedure

1. Display the packet drop statistics of the interface:

   ```
   # ethtool -S enp1s0
   ...
   rx_queue_0_drops: 97326
   rx_queue_1_drops: 63783
   ...
   ```

   Note that the output of the command depends on the network card and the driver.

   High values in `discard` or `drop` counters indicate that the available buffer fills up faster than the kernel can process the packets. Increasing the ring buffers can help to avoid such loss.

2. Display the maximum ring buffer sizes:

   ```
   # ethtool -g enp1s0
   Ring parameters for `enp1s0`:
   Pre-set maximums:
   RX:  4096
   RX Mini:  0
   RX Jumbo: 16320
   TX:  4096
   ```
Current hardware settings:
RX: 255
RX Mini: 0
RX Jumbo: 0
TX: 255

If the values in the Pre-set maximums section are higher than in the Current hardware settings section, you can change the settings in the next steps.

3. Identify the NetworkManager connection profile that uses the interface:

```bash
# nmcli connection show
NAME    UUID                                      TYPE    DEVICE
Example-Connection a5eb6490-cc20-3668-81f8-0314a27f3f75 ethernet enp1s0
```

4. Update the connection profile, and increase the ring buffers:

- To increase the RX ring buffer, enter:

  ```bash
  # nmcli connection modify Example-Connection ethtool.ring-rx 4096
  ```

- To increase the TX ring buffer, enter:

  ```bash
  # nmcli connection modify Example-Connection ethtool.ring-tx 4096
  ```

5. Reload the NetworkManager connection:

```bash
# nmcli connection up Example-Connection
```

**IMPORTANT**

Depending on the driver your NIC uses, changing in the ring buffer can shortly interrupt the network connection.

Additional resources

- ifconfig and ip commands report packet drops
- Should I be concerned about a 0.05% packet drop rate?
- ethtool(8) man page

**29.6. INCREASING THE RING BUFFER SIZE TO REDUCE A HIGH PACKET DROP RATE BY USING THE NETWORK RHEL SYSTEM ROLE**

Increase the size of an Ethernet device’s ring buffers if the packet drop rate causes applications to report a loss of data, timeouts, or other issues.

Ring buffers are circular buffers where an overflow overwrites existing data. The network card assigns a transmit (TX) and receive (RX) ring buffer. Receive ring buffers are shared between the device driver and the network interface controller (NIC). Data can move from NIC to the kernel through either hardware interrupts or software interrupts, also called SoftIRQs.
The kernel uses the RX ring buffer to store incoming packets until the device driver can process them. The device driver drains the RX ring, typically by using SoftIRQs, which puts the incoming packets into a kernel data structure called an `sk_buff` or `skb` to begin its journey through the kernel and up to the application that owns the relevant socket.

The kernel uses the TX ring buffer to hold outgoing packets which should be sent to the network. These ring buffers reside at the bottom of the stack and are a crucial point at which packet drop can occur, which in turn will adversely affect network performance.

**IMPORTANT**

When you run a play that uses the `network` RHEL system role and if the setting values do not match the values specified in the play, the role overrides the existing connection profile with the same name. To prevent resetting these values to their defaults, always specify the whole configuration of the network connection profile in the play, even if the configuration, for example the IP configuration, already exists.

**Prerequisites**

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has `sudo` permissions on them.
- You know the maximum ring buffer sizes that the device supports.

**Procedure**

1. Create a playbook file, for example `~/playbook.yml`, with the following content:

   ```yaml
   ---
   - name: Configure the network
     hosts: managed-node-01.example.com
     tasks:
       - name: Configure an Ethernet connection with increased ring buffer sizes
         ansible.builtin.include_role:
           name: rhel-system-roles.network
         vars:
           network_connections:
             - name: enp1s0
               type: ethernet
               autoconnect: yes
               ip:
                 address:
                   - 198.51.100.20/24
                   - 2001:db8:1::/64
               gateway4: 198.51.100.254
               gateway6: 2001:db8:1::fffe
               dns:
                 - 198.51.100.200
                 - 2001:db8:1::ffbb
               dns_search:
                 - example.com
               ethtool:
                 ring:
   ```
This playbook creates the `enp1s0` connection profile with the following settings, or updates it if the profile already exists:

- A static IPv4 address - \textbf{198.51.100.20} with a /24 subnet mask
- A static IPv6 address - \textbf{2001:db8:1::1} with a /64 subnet mask
- An IPv4 default gateway - \textbf{198.51.100.254}
- An IPv6 default gateway - \textbf{2001:db8:1::fffe}
- An IPv4 DNS server - \textbf{198.51.100.200}
- An IPv6 DNS server - \textbf{2001:db8:1::ffbb}
- A DNS search domain - \textbf{example.com}
- Maximum number of ring buffer entries:
  - Receive (RX): 4096
  - Transmit (TX): 4096

2. Validate the playbook syntax:

```
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```
$ ansible-playbook ~/playbook.yml
```

Additional resources

- \texttt{/usr/share/ansible/roles/rhel-system-roles.network/README.md} file
- \texttt{/usr/share/doc/rhel-system-roles/network/} directory
CHAPTER 30. INTRODUCTION TO NETWORKMANAGER DEBUGGING

Increasing the log levels for all or certain domains helps to log more details of the operations that NetworkManager performs. You can use this information to troubleshoot problems. NetworkManager provides different levels and domains to produce logging information. The /etc/NetworkManager/NetworkManager.conf file is the main configuration file for NetworkManager. The logs are stored in the journal.

30.1. INTRODUCTION TO NETWORKMANAGER REAPPLY METHOD

The NetworkManager service uses a profile to manage the connection settings of a device. Desktop Bus (D-Bus) API can create, modify, and delete these connection settings. For any changes in a profile, D-Bus API clones the existing settings to the modified settings of a connection. Despite cloning, changes do not apply to the modified settings. To make it effective, reactivate the existing settings of a connection or use the reapply() method.

The reapply() method has the following features:

1. Updating modified connection settings without deactivation or restart of a network interface.

2. Removing pending changes from the modified connection settings. As NetworkManager does not revert the manual changes, you can reconfigure the device and revert external or manual parameters.

3. Creating different modified connection settings than that of the existing connection settings.

Also, reapply() method supports the following attributes:

- bridge.ageing-time
- bridge.forward-delay
- bridge.group-address
- bridge.group-forward-mask
- bridge.hello-time
- bridge.max-age
- bridge.multicast-hash-max
- bridge.multicast-last-member-count
- bridge.multicast-last-member-interval
- bridge.multicast-membership-interval
- bridge.multicast-querier
- bridge.multicast-querier-interval
- bridge.multicast-query-interval
- bridge.multicast-query-response-interval
CHAPTER 30. INTRODUCTION TO NETWORKMANAGER DEBUGGING

- bridge.multicast-query-use-ifaddr
- bridge.multicast-router
- bridge.multicast-snooping
- bridge.multicast-startup-query-count
- bridge.multicast-startup-query-interval
- bridge.priority
- bridge.stp
- bridge.VLAN-filtering
- bridge.VLAN-protocol
- bridge.VLANs
- 802-3-ethernet.accept-all-mac-addresses
- 802-3-ethernet.cloned-mac-address
- IPv4.addresses
- IPv4.dhcp-client-id
- IPv4.dhcp-iaid
- IPv4.dhcp-timeout
- IPv4.DNS
- IPv4.DNS-priority
- IPv4.DNS-search
- IPv4.gateway
- IPv4.ignore-auto-DNS
- IPv4.ignore-auto-routes
- IPv4.may-fail
- IPv4.method
- IPv4.never-default
- IPv4.route-table
- IPv4.routes
- IPv4.routing-rules
- IPv6.addr-gen-mode
30.2. SETTING THE NETWORKMANAGER LOG LEVEL

By default, all the log domains are set to record the \texttt{INFO} log level. Disable rate-limiting before collecting debug logs. With rate-limiting, \texttt{systemd-journald} drops messages if there are too many of them in a short time. This can occur when the log level is \texttt{TRACE}.

This procedure disables rate-limiting and enables recording debug logs for the all (ALL) domains.

\textbf{Procedure}

1. To disable rate-limiting, edit the \texttt{/etc/systemd/journald.conf} file, uncomment the \texttt{RateLimitBurst} parameter in the \texttt{[Journal]} section, and set its value as \texttt{0}:

   \begin{verbatim}
   RateLimitBurst=0
   \end{verbatim}

2. Restart the \texttt{systemd-journald} service.

   \begin{verbatim}
   # systemctl restart systemd-journald
   \end{verbatim}
3. Create the /etc/NetworkManager/conf.d/95-nm-debug.conf file with the following content:

```
[logging]
domains=ALL:TRACE
```

The `domains` parameter can contain multiple comma-separated `domain:level` pairs.

4. Restart the NetworkManager service.

```
# systemctl restart NetworkManager
```

**Verification**

- Query the `systemd` journal to display the journal entries of the `NetworkManager` unit:

```
# journalctl -u NetworkManager
```

```
Jun 30 15:24:32 server NetworkManager[164187]: <debug> [1656595472.4939] active-connection[0x5565143c80a0]: update activation type from assume to managed
Jun 30 15:24:32 server NetworkManager[164187]: <trace> [1656595472.4939] l3cfg[4281fdf43e356454,ifindex=3]: commit type register (type "update", source "device", existing a369f23014b9ede3) -> a369f23014b9ede3
Jun 30 15:24:32 server NetworkManager[164187]: <info> [1656595472.4940] manager: NetworkManager state is now CONNECTED_SITE

```

**30.3. TEMPORARILY SETTING LOG LEVELS AT RUN TIME USING NMCLI**

You can change the log level at run time using `nmcli`. However, Red Hat recommends to enable debugging using configuration files and restart NetworkManager. Updating debugging `levels` and `domains` using the `.conf` file helps to debug boot issues and captures all the logs from the initial state.

**Procedure**

1. Optional: Display the current logging settings:

```
# nmcli general logging
```

```
LEVEL   DOMAINS
INFO
PLATFORM,RFKILL,ETHER,WIFI,BT,MPI,MB,DHCP4,DHCP6,PPP,WIFI_SCAN,IP4,IP6,AUTOIP4,DNS,VPN,SHARING,SUPPLICANT,AGENTS,SETTINGS,SUSPEND,CORE,DEVICE,OLPC,WIMAX,INFINIBAND,FIREWALL,ADSL,BOND,VLAN,BRIDGE,DBUS_PROPS,TEAM,CONC,HECK,DC,B,DISPATCH
```

2. To modify the logging level and domains, use the following options:

- To set the log level for all domains to the same `LEVEL`, enter:
To change the level for specific domains, enter:

```
# nmcli general logging level LEVEL domains DOMAINS
```

Note that updating the logging level using this command disables logging for all the other domains.

To change the level of specific domains and preserve the level of all other domains, enter:

```
# nmcli general logging level KEEP domains DOMAIN:LEVEL,DOMAIN:LEVEL
```

### 30.4. VIEWING NETWORKMANAGER LOGS

You can view the NetworkManager logs for troubleshooting.

**Procedure**

- To view the logs, enter:

  ```
  # journalctl -u NetworkManager -b
  ```

**Additional resources**

- NetworkManager.conf(5) man page
- journalctl(1) man page

### 30.5. DEBUGGING LEVELS AND DOMAINS

You can use the **levels** and **domains** parameters to manage the debugging for NetworkManager. The level defines the verbosity level, whereas the domains define the category of the messages to record the logs with given severity (**level**).

<table>
<thead>
<tr>
<th>Log levels</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OFF</td>
<td>Does not log any messages about NetworkManager</td>
</tr>
<tr>
<td>ERR</td>
<td>Logs only critical errors</td>
</tr>
<tr>
<td>WARN</td>
<td>Logs warnings that can reflect the operation</td>
</tr>
<tr>
<td>INFO</td>
<td>Logs various informational messages that are useful for tracking state and operations</td>
</tr>
<tr>
<td>DEBUG</td>
<td>Enables verbose logging for debugging purposes</td>
</tr>
<tr>
<td>TRACE</td>
<td>Enables more verbose logging than the DEBUG level</td>
</tr>
</tbody>
</table>
Note that subsequent levels log all messages from earlier levels. For example, setting the log level to INFO also logs messages contained in the ERR and WARN log level.

Additional resources

- NetworkManager.conf(5) man page
CHAPTER 31. USING LLDP TO DEBUG NETWORK CONFIGURATION PROBLEMS

You can use the Link Layer Discovery Protocol (LLDP) to debug network configuration problems in the topology. This means that LLDP can report configuration inconsistencies with other hosts or routers and switches.

31.1. DEBUGGING AN INCORRECT VLAN CONFIGURATION USING LLDP INFORMATION

If you configured a switch port to use a certain VLAN and a host does not receive these VLAN packets, you can use the Link Layer Discovery Protocol (LLDP) to debug the problem. Perform this procedure on the host that does not receive the packets.

Prerequisites

- The `nmstate` package is installed.
- The switch supports LLDP.
- LLDP is enabled on neighbor devices.

Procedure

1. Create the `~/enable-LLDP-enp1s0.yml` file with the following content:

   ```yaml
   interfaces:
   - name: enp1s0
     type: ethernet
     lldp:
       enabled: true
   
```

2. Use the `~/enable-LLDP-enp1s0.yml` file to enable LLDP on interface `enp1s0`:

   ```bash
   # nmstatectl apply ~/enable-LLDP-enp1s0.yml
   
```

3. Display the LLDP information:

   ```bash
   # nmstatectl show enp1s0
   - name: enp1s0
     type: ethernet
     state: up
     ipv4:
       enabled: false
       dhcp: false
     ipv6:
       enabled: false
       autoconf: false
       dhcp: false
     lldp:
       enabled: true
     neighbors:
       - - type: 5
   ```
system-name: Summit300-48
- type: 6
  system-description: Summit300-48 - Version 7.4e.1 (Build 5)
  05/27/05 04:53:11
- type: 7
  system-capabilities:
  - MAC Bridge component
  - Router
  - type: 1
    _description: MAC address
    chassis-id: 00:01:30:F9:AD:A0
    chassis-id-type: 4
- type: 2
  _description: Interface name
  port-id: 1/1
  port-id-type: 5
- type: 127
  ieee-802-1-vlans:
  - name: v2-0488-03-0505
    vid: 488
    oui: 00:80:c2
    subtype: 3
- type: 127
  ieee-802-3-mac-phy-conf:
    autoneg: true
    operational-mau-type: 16
    pmd-autoneg-cap: 27648
    oui: 00:12:0f
    subtype: 1
- type: 127
  ieee-802-1-ppvids:
  - 0
    oui: 00:80:c2
    subtype: 2
- type: 8
  management-addresses:
  - address: 00:01:30:F9:AD:A0
    address-subtype: MAC
    interface-number: 1001
    interface-number-subtype: 2
- type: 127
  ieee-802-3-max-frame-size: 1522
  oui: 00:12:0f
  subtype: 4
  mac-address: 82:75:BE:6F:8C:7A
  mtu: 1500

4. Verify the output to ensure that the settings match your expected configuration. For example, the LLDP information of the interface connected to the switch shows that the switch port this host is connected to uses VLAN ID 448:

- type: 127
  ieee-802-1-vlans:
  - name: v2-0488-03-0505
    vid: 488
If the network configuration of the enp1s0 interface uses a different VLAN ID, change it accordingly.

Additional resources

Configuring VLAN tagging
CHAPTER 32. LINUX TRAFFIC CONTROL

Linux offers tools for managing and manipulating the transmission of packets. The Linux Traffic Control (TC) subsystem helps in policing, classifying, shaping, and scheduling network traffic. TC also mangles the packet content during classification by using filters and actions. The TC subsystem achieves this by using queuing disciplines (qdisc), a fundamental element of the TC architecture.

The scheduling mechanism arranges or rearranges the packets before they enter or exit different queues. The most common scheduler is the First-In-First-Out (FIFO) scheduler. You can do the qdiscs operations temporarily using the tc utility or permanently using NetworkManager.

In Red Hat Enterprise Linux, you can configure default queueing disciplines in various ways to manage the traffic on a network interface.

32.1. OVERVIEW OF QUEUING DISCIPLINES

Queuing disciplines (qdiscs) help with queuing up and, later, scheduling of traffic transmission by a network interface. A qdisc has two operations;

- enqueue requests so that a packet can be queued up for later transmission and
- dequeue requests so that one of the queued-up packets can be chosen for immediate transmission.

Every qdisc has a 16-bit hexadecimal identification number called a handle, with an attached colon, such as 1: or abcd:. This number is called the qdisc major number. If a qdisc has classes, then the identifiers are formed as a pair of two numbers with the major number before the minor, <major>:<minor>, for example abcd:1. The numbering scheme for the minor numbers depends on the qdisc type. Sometimes the numbering is systematic, where the first-class has the ID <major>:1, the second one <major>:2, and so on. Some qdiscs allow the user to set class minor numbers arbitrarily when creating the class.

Classful qdiscs

Different types of qdiscs exist and help in the transfer of packets to and from a networking interface. You can configure qdiscs with root, parent, or child classes. The point where children can be attached are called classes. Classes in qdisc are flexible and can always contain either multiple children classes or a single child, qdisc. There is no prohibition against a class containing a classful qdisc itself, this facilitates complex traffic control scenarios.

Classful qdiscs do not store any packets themselves. Instead, they enqueue and dequeue requests down to one of their children according to criteria specific to the qdisc. Eventually, this recursive packet passing ends up where the packets are stored (or picked up from in the case of dequeuing).

Classless qdiscs

Some qdiscs contain no child classes and they are called classless qdiscs. Classless qdiscs require less customization compared to classful qdiscs. It is usually enough to attach them to an interface.

Additional resources

- tc(8) man page
- tc-actions(8) man page

32.2. INTRODUCTION TO CONNECTION TRACKING

At a firewall, the Netfilter framework filters packets from an external network. After a packet arrives,
Netfilter assigns a connection tracking entry. Connection tracking is a Linux kernel networking feature for logical networks that tracks connections and identifies packet flow in those connections. This feature filters and analyzes every packet, sets up the connection tracking table to store connection status, and updates the connection status based on identified packets. For example, in the case of FTP connection, Netfilter assigns a connection tracking entry to ensure all packets of FTP connection work in the same manner. The connection tracking entry stores a Netfilter mark and tracks the connection state information in the memory table in which a new packet tuple maps with an existing entry. If the packet tuple does not map with an existing entry, the packet adds a new connection tracking entry that groups packets of the same connection.

You can control and analyze traffic on the network interface. The tc traffic controller utility uses the qdisc discipline to configure the packet scheduler in the network. The qdisc kernel-configured queuing discipline enqueues packets to the interface. By using qdisc, Kernel catches all the traffic before a network interface transmits it. Also, to limit the bandwidth rate of packets belonging to the same connection, use the tc qdisc command.

To retrieve data from connection tracking marks into various fields, use the tc utility with the ctinfo module and the connmark functionality. For storing packet mark information, the ctinfo module copies the Netfilter mark and the connection state information into a socket buffer (skb) mark metadata field.

Transmitting a packet over a physical medium removes all the metadata of a packet. Before the packet loses its metadata, the ctinfo module maps and copies the Netfilter mark value to a specific value of the Diffserv code point (DSCP) in the packet’s IP field.

Additional resources
- tc(8) and tc-ctinfo(8) man pages

32.3. INSPECTING QDISCS OF A NETWORK INTERFACE USING THE TC UTILITY

By default, Red Hat Enterprise Linux systems use fq_codel qdisc. You can inspect the qdisc counters using the tc utility.

Procedure
1. Optional: View your current qdisc:

   ```
   # tc qdisc show dev enp0s1
   ```

2. Inspect the current qdisc counters:

   ```
   # tc -s qdisc show dev enp0s1
   qdisc fq_codel 0: root refcnt 2 limit 10240p flows 1024 quantum 1514 target 5.0ms interval 100.0ms memory_limit 32Mb ecn
   Sent 1008193 bytes 5559 pkt (dropped 233, overlimits 55 requeues 77)
   backlog 0b 0p requeues 0
   ```

   - **dropped** - the number of times a packet is dropped because all queues are full
   - **overlimits** - the number of times the configured link capacity is filled
   - **sent** - the number of dequeues
32.4. UPDATING THE DEFAULT QDISC

If you observe networking packet losses with the current qdisc, you can change the qdisc based on your network-requirements.

Procedure

1. View the current default qdisc:

   ```
   # sysctl -a | grep qdisc
   net.core.default_qdisc = fq_codel
   ```

2. View the qdisc of current Ethernet connection:

   ```
   # tc -s qdisc show dev enp0s1
   qdisc fq_codel 0: root refcnt 2 limit 10240p flows 1024 quantum 1514 target 5.0ms interval 100.0ms memory_limit 32Mb ecn
   Sent 0 bytes 0 pkt (dropped 0, overlimits 0 requeues 0)
   backlog 0b 0p requeues 0
   maxpacket 0 drop_overlimit 0 new_flow_count 0 ecn_mark 0
   new_flows_len 0 old_flows_len 0
   ```

3. Update the existing qdisc:

   ```
   # sysctl -w net.core.default_qdisc=pfifo_fast
   ```

4. To apply the changes, reload the network driver:

   ```
   # modprobe -r NETWORKDRIVERNAME
   # modprobe NETWORKDRIVERNAME
   ```

5. Start the network interface:

   ```
   # ip link set enp0s1 up
   ```

Verification

- View the qdisc of the Ethernet connection:

   ```
   # tc -s qdisc show dev enp0s1
   qdisc pfifo_fast 0: root refcnt 2 bands 3 priomap 1 2 2 1 2 0 0 1 1 1 1 1 1 1 1
   Sent 373186 bytes 5333 pkt (dropped 0, overlimits 0 requeues 0)
   backlog 0b 0p requeues 0
   ....
   ```

Additional resources

- How to set sysctl variables on Red Hat Enterprise Linux

32.5. TEMPORARILY SETTING THE CURRENT QDISC OF A NETWORK INTERFACE USING THE TC UTILITY
You can update the current `qdisc` without changing the default one.

**Procedure**

1. Optional: View the current `qdisc`:
   ```bash
   # tc -s qdisc show dev enp0s1
   ``

2. Update the current `qdisc`:
   ```bash
   # tc qdisc replace dev enp0s1 root htb
   ``

**Verification**

- View the updated current `qdisc`:
  ```bash
  # tc -s qdisc show dev enp0s1
  qdisc htb 8001: root refcnt 2 r2q 10 default 0 direct_packets_stat 0 direct_qlen 1000
  Sent 0 bytes 0 pkt (dropped 0, overlimits 0 requeues 0)
  backlog 0b 0p requeues 0
  ```

**32.6. PERMANENTLY SETTING THE CURRENT QDISC OF A NETWORK INTERFACE USING NETWORKMANAGER**

You can update the current `qdisc` value of a NetworkManager connection.

**Procedure**

1. Optional: View the current `qdisc`:
   ```bash
   # tc qdisc show dev enp0s1
   qdisc fq_codel 0: root refcnt 2
   ``

2. Update the current `qdisc`:
   ```bash
   # nmcli connection modify enp0s1 tc.qdiscs 'root pfifo_fast'
   ```

3. Optional: To add another `qdisc` over the existing `qdisc`, use the `+tc.qdisc` option:
   ```bash
   # nmcli connection modify enp0s1 +tc.qdisc 'ingress handle ffff:'
   ```

4. Activate the changes:
   ```bash
   # nmcli connection up enp0s1
   ```

**Verification**

- View current `qdisc` the network interface:
Additional resources

- nm-settings(5) man page

### 32.7. CONFIGURING THE RATE LIMITING OF PACKETS BY USING THE TC-CTINFO UTILITY

You can limit network traffic and prevent the exhaustion of resources in the network by using rate limiting. With rate limiting, you can also reduce the load on servers by limiting repetitive packet requests in a specific time frame. In addition, you can manage bandwidth rate by configuring traffic control in the kernel with the `tc-ctinfo` utility.

The connection tracking entry stores the Netfilter mark and connection information. When a router forwards a packet from the firewall, the router either removes or modifies the connection tracking entry from the packet. The connection tracking information (ctinfo) module retrieves data from connection tracking marks into various fields. This kernel module preserves the Netfilter mark by copying it into socket buffer (skb) mark metadata field.

**Prerequisites**

- The `iperf3` utility is installed on a server and a client.

**Procedure**

1. Perform the following steps on the server:
   a. Add a virtual link to the network interface:

   ```
   # ip link add name ifb4eth0 numtxqueues 48 numrxqueues 48 type ifb
   
   This command has the following parameters:
   ```

   ```
   name ifb4eth0
   Sets new virtual device interface.
   numtxqueues 48
   Sets the number of transmit queues.
   numrxqueues 48
   Sets the number of receive queues.
   type ifb
   Sets the type of the new device.
   ```

   b. Change the state of the interface:

   ```
   # ip link set dev ifb4eth0 up
   ```

   c. Add the `qdisc` attribute on the physical network interface and apply it to the incoming traffic:
In the `handle ffff:` option, the `handle` parameter assigns the major number `fff:` as a default value to a classful `qdisc` on the `enp1s0` physical network interface, where `qdisc` is a queueing discipline parameter to analyze traffic control.

d. Add a filter on the physical interface of the `ip` protocol to classify packets:

```
# tc filter add dev enp1s0 parent ffff: protocol ip u32 match u32 0 0 action ctinfo
cpmark 100 action mirred egress redirect dev ifb4eth0
```

This command has the following attributes:

**parent ffff:**
- Sets major number `fff:` for the parent `qdisc`.

**u32 match u32 0 0**
- Sets the `u32` filter to `match` the IP headers of `u32` pattern. The first `0` represents the second byte of IP header while the other `0` is for the mask match telling the filter which bits to match.

**action ctinfo**
- Sets action to retrieve data from the connection tracking mark into various fields.

**cpmark 100**
- Copies the connection tracking mark (connmark) 100 into the packet IP header field.

**action mirred egress redirect dev ifb4eth0**
- Sets `action mirred` to redirect the received packets to the `ifb4eth0` destination interface.

e. Add a classful `qdisc` to the interface:

```
# tc qdisc add dev ifb4eth0 root handle 1: htb default 1000
```

This command sets the major number `1` to root `qdisc` and uses the `htb` hierarchy token bucket with classful `qdisc` of minor-id `1000`.

f. Limit the traffic on the interface to 1 Mbit/s with an upper limit of 2 Mbit/s:

```
# tc class add dev ifb4eth0 parent 1:1 classid 1:100 htb ceil 2mbit rate 1mbit prio 100
```

This command has the following parameters:

**parent 1:1**
- Sets `parent` with `classid` as `1` and `root` as `1`.

**classid 1:100**
- Sets `classid` as `1:100` where `1` is the number of parent `qdisc` and `100` is the number of classes of the parent `qdisc`.

**htb ceil 2mbit**
- The `htb` classful `qdisc` allows upper limit bandwidth of `2 Mbit/s` as the `ceil` rate limit.
g. Apply the Stochastic Fairness Queuing (sfq) of classless qdisc to interface with a time interval of 60 seconds to reduce queue algorithm perturbation:

```
# tc qdisc add dev ifb4eth0 parent 1:100 sfq perturb 60
```

h. Add the firewall mark (fw) filter to the interface:

```
# tc filter add dev ifb4eth0 parent 1:0 protocol ip prio 100 handle 100 fw classid 1:100
```

i. Restore the packet meta mark from the connection mark (CONNMARK):

```
# nft add rule ip mangle PREROUTING counter meta mark set ct mark
```

In this command, the nft utility has a mangle table with the PREROUTING chain rule specification that alters incoming packets before routing to replace the packet mark with CONNMARK.

j. If no nft table and chain exist, create a table and add a chain rule:

```
# nft add table ip mangle
# nft add chain ip mangle PREROUTING {type filter hook prerouting priority mangle \\}
```

k. Set the meta mark on tcp packets that are received on the specified destination address 192.0.2.3:

```
# nft add rule ip mangle PREROUTING ip daddr 192.0.2.3 counter meta mark set 0x64
```

l. Save the packet mark into the connection mark:

```
# nft add rule ip mangle PREROUTING counter ct mark set mark
```

m. Run the iperf3 utility as the server on a system by using the -s parameter and the server then waits for the response of the client connection:

```
# iperf3 -s
```

2. On the client, run iperf3 as a client and connect to the server that listens on IP address 192.0.2.3 for periodic HTTP request–response timestamp:

```
# iperf3 -c 192.0.2.3 -t TCP_STREAM | tee rate
```

192.0.2.3 is the IP address of the server while 192.0.2.4 is the IP address of the client.

3. Terminate the iperf3 utility on the server by pressing Ctrl+C:
4. Terminate the **iperf3** utility on the client by pressing **Ctrl+C**:

```
Connecting to host 192.0.2.3, port 5201
local 192.0.2.4 port 52130 connected to 192.0.2.3 port 5201
```

```
[5] Interval        Transfer  Bitrate      Retr  Cwnd
[5]   0.00-1.00   sec   481 KBytes  3.94 Mbits/sec 0   76.4 KBytes
[5]   1.00-2.00   sec   223 KBytes  1.83 Mbits/sec 0   82.0 KBytes
```

```
iperf3: error - the server has terminated
```

**Verification**

1. Display the statistics about packet counts of the **htb** and **sfq** classes on the interface:

```
# tc -s qdisc show dev ifb4eth0
```

```
qdisc htb 1: root
... Sent 26611455 bytes 3054 pkt (dropped 76, overlimits 4887 requeues 0)
... qdisc sfq 8001: parent
... Sent 26535030 bytes 2296 pkt (dropped 76, overlimits 0 requeues 0)
```

2. Display the statistics of packet counts for the **mirred** and **ctinfo** actions:

```
# tc -s filter show dev enp1s0 ingress
```

```
filter parent ffff: protocol ip pref 49152 u32 chain 0
filter parent ffff: protocol ip pref 49152 u32 chain 0 fh 800: ht divisor 1
filter parent ffff: protocol ip pref 49152 u32 chain 0 fh 800::800 order 2048 key ht 800 bkt 0
terminal flowid not_in_hw (rule hit 8075 success 8075)
match 00000000/00000000 at 0 (success 8075 )
action order 1 : ctinfo zone 0 pipe
    index 1 ref 1 bind 1 cpmark 0x00000064 installed 3105 sec firstused 3105 sec DSCP set 0 error 0
    CPMARK set 7712
Action statistics:
Sent 25891504 bytes 3137 pkt (dropped 0, overlimits 0 requeues 0)
backlog 0b 0p requeues 0

action order 2: mirred (Egress Redirect to device ifb4eth0) stolen
    index 1 ref 1 bind 1 installed 3105 sec firstused 3105 sec
```
3. Display the statistics of the htb rate-limiter and its configuration:

```
# tc -s class show dev ifb4eth0
class htb 1:100 root leaf 8001: prio 7 rate 1Mbit ceil 2Mbit burst 1600b cburst 1600b
Sent 26541716 bytes 2373 pkt (dropped 61, overlimits 4887 requeues 0)
backlog 0b 0p requeues 0
lended: 7248 borrowed: 0 giants: 0
tokens: 187250 ctokens: 93625
```

Additional resources

- `tc(8)` and `tc-ctinfo(8)` man page
- `nft(8)` man page

### 32.8. AVAILABLE QDISCS IN RHEL

Each qdisc addresses unique networking-related issues. The following is the list of qdiscs available in RHEL. You can use any of the following qdisc to shape network traffic based on your networking requirements.

<table>
<thead>
<tr>
<th>qdisc name</th>
<th>Included in</th>
<th>Offload support</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asynchronous Transfer Mode (ATM)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Class-Based Queueing</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Credit-Based Shaper</td>
<td>kernel-modules-extra</td>
<td>Yes</td>
</tr>
<tr>
<td>CHOose and Keep for responsive flows, CHOose and Kill for unresponsive flows (CHOKE)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Controlled Delay (CoDel)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>Deficit Round Robin (DRR)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Differentiated Services marker (DSMARK)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Enhanced Transmission Selection (ETS)</td>
<td>kernel-modules-extra</td>
<td>Yes</td>
</tr>
<tr>
<td>Fair Queue (FQ)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>qdisc name</td>
<td>Included in</td>
<td>Offload support</td>
</tr>
<tr>
<td>---------------------------------------------------------</td>
<td>-------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Fair Queuing Controlled Delay (FQ_CODel)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>Generalized Random Early Detection (GRED)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Hierarchical Fair Service Curve (HSFC)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>Heavy-Hitter Filter (HHF)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>Hierarchy Token Bucket (HTB)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>INGRESS</td>
<td>kernel-core</td>
<td>Yes</td>
</tr>
<tr>
<td>Multi Queue Priority (MQPrio)</td>
<td>kernel-modules-extra</td>
<td>Yes</td>
</tr>
<tr>
<td>Multiqueue (MULTIQ)</td>
<td>kernel-modules-extra</td>
<td>Yes</td>
</tr>
<tr>
<td>Network Emulator (NETEM)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Proportional Integral-controller Enhanced (PIE)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>PLUG</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>Quick Fair Queueing (QFQ)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Random Early Detection (RED)</td>
<td>kernel-modules-extra</td>
<td>Yes</td>
</tr>
<tr>
<td>Stochastic Fair Blue (SFB)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
<tr>
<td>Stochastic Fairness Queueing (SFQ)</td>
<td>kernel-core</td>
<td></td>
</tr>
<tr>
<td>Token Bucket Filter (TBF)</td>
<td>kernel-core</td>
<td>Yes</td>
</tr>
<tr>
<td>Trivial Link Equalizer (TEQL)</td>
<td>kernel-modules-extra</td>
<td></td>
</tr>
</tbody>
</table>

**IMPORTANT**

The `qdisc` offload requires hardware and driver support on NIC.

**Additional resources**

- `tc(8)` man page
CHAPTER 33. AUTHENTICATING A RHEL CLIENT TO THE NETWORK BY USING THE 802.1X STANDARD WITH A CERTIFICATE STORED ON THE FILE SYSTEM

Administrators frequently use port-based Network Access Control (NAC) based on the IEEE 802.1X standard to protect a network from unauthorized LAN and Wi-Fi clients. To enable a client to connect to such networks, you must configure 802.1X authentication on this clients.

33.1. CONFIGURING 802.1X NETWORK AUTHENTICATION ON AN EXISTING ETHERNET CONNECTION BY USING NMCLI

You can use the `nmcli` utility to configure an Ethernet connection with 802.1X network authentication on the command line.

Prerequisites

- The network supports 802.1X network authentication.
- The Ethernet connection profile exists in NetworkManager and has a valid IP configuration.
- The following files required for TLS authentication exist on the client:
  - The client key stored is in the `/etc/pki/tls/private/client.key` file, and the file is owned and only readable by the `root` user.
  - The client certificate is stored in the `/etc/pki/tls/certs/client.crt` file.
  - The Certificate Authority (CA) certificate is stored in the `/etc/pki/tls/certs/ca.crt` file.
- The `wpa_supplicant` package is installed.

Procedure

1. Set the Extensible Authentication Protocol (EAP) to `tls` and the paths to the client certificate and key file:

   ```shell
   # nmcli connection modify enp1s0 802-1x.eap tls 802-1x.client-cert /etc/pki/tls/certs/client.crt 802-1x.private-key /etc/pki/tls/certs/certs/client.key
   ```

   Note that you must set the `802-1x.eap`, `802-1x.client-cert`, and `802-1x.private-key` parameters in a single command.

2. Set the path to the CA certificate:

   ```shell
   # nmcli connection modify enp1s0 802-1x.ca-cert /etc/pki/tls/certs/ca.crt
   ```

3. Set the identity of the user used in the certificate:

   ```shell
   # nmcli connection modify enp1s0 802-1x.identity user@example.com
   ```

4. Optionally, store the password in the configuration:

   ```shell
   # nmcli connection modify enp1s0 802-1x.private-key-password password
   ```
IMPORTANT

By default, NetworkManager stores the password in clear text in the
/etc/sysconfig/network-scripts/keys-connection_name file, that is readable
only by the root user. However, clear text passwords in a configuration file can be
a security risk.

To increase the security, set the 802-1x.password-flags parameter to 0x1. With
this setting, on servers with the GNOME desktop environment or the nm-applet
running, NetworkManager retrieves the password from these services. In other
cases, NetworkManager prompts for the password.

5. Activate the connection profile:

   # nmcli connection up enp1s0

Verification

- Access resources on the network that require network authentication.

Additional resources

- Configuring an Ethernet connection
- nm-settings(5) man page
- nmcli(1) man page

33.2. CONFIGURING A STATIC ETHERNET CONNECTION WITH 802.1X
NETWORK AUTHENTICATION BY USING NMSTATECTL

Use the nmstatectl utility to configure an Ethernet connection with 802.1X network authentication
through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result
matches the configuration file. If anything fails, nmstatectl automatically rolls back the changes to avoid
leaving the system in an incorrect state.

NOTE

The nmstate library only supports the TLS Extensible Authentication Protocol (EAP) method.

Prerequisites

- The network supports 802.1X network authentication.
- The managed node uses NetworkManager.
- The following files required for TLS authentication exist on the client:
  - The client key stored is in the /etc/pki/tls/private/client.key file, and the file is owned and
    only readable by the root user.
The client certificate is stored in the `/etc/pki/tls/certs/client.crt` file.

The Certificate Authority (CA) certificate is stored in the `/etc/pki/tls/certs/ca.crt` file.

**Procedure**

1. Create a YAML file, for example `~/create-ethernet-profile.yml`, with the following content:

```yaml
---
interfaces:
  - name: enp1s0
    type: ethernet
    state: up
    ipv4:
      enabled: true
      address:
        - ip: 192.0.2.1
          prefix-length: 24
        dhcp: false
    ipv6:
      enabled: true
      address:
        - ip: 2001:db8:1::1
          prefix-length: 64
      autoconf: false
      dhcp: false
    802.1x:
      ca-cert: /etc/pki/tls/certs/ca.crt
      client-cert: /etc/pki/tls/certs/client.crt
      eap-methods:
        - tls
      identity: client.example.org
      private-key: /etc/pki/tls/private/client.key
      private-key-password: password
    routes:
      config:
        - destination: 0.0.0.0/0
          next-hop-address: 192.0.2.254
          next-hop-interface: enp1s0
        - destination: ::/0
          next-hop-address: 2001:db8:1::ffe
          next-hop-interface: enp1s0
    dns-resolver:
      config:
        search:
          - example.com
        server:
          - 192.0.2.200
          - 2001:db8:1::ffbb
```

These settings define an Ethernet connection profile for the `enp1s0` device with the following settings:

- A static IPv4 address – `192.0.2.1` with a `/24` subnet mask
- A static IPv6 address – `2001:db8:1::1` with a `/64` subnet mask
- An IPv4 default gateway - 192.0.2.254
- An IPv6 default gateway - 2001:db8:1::fffe
- An IPv4 DNS server - 192.0.2.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com
- 802.1X network authentication using the TLS EAP protocol

2. Apply the settings to the system:

   # nmstatectl apply ~/create-ethernet-profile.yml

Verification

- Access resources on the network that require network authentication.

33.3. CONFIGURING A STATIC ETHERNET CONNECTION WITH 802.1X NETWORK AUTHENTICATION BY USING THE NETWORK RHEL SYSTEM ROLE

You can remotely configure an Ethernet connection with 802.1X network authentication by using the network RHEL system role.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has sudo permissions on them.
- The network supports 802.1X network authentication.
- The managed nodes uses NetworkManager.
- The following files required for TLS authentication exist on the control node:
  - The client key is stored in the /srv/data/client.key file.
  - The client certificate is stored in the /srv/data/client.crt file.
  - The Certificate Authority (CA) certificate is stored in the /srv/data/ca.crt file.

Procedure

1. Create a playbook file, for example ~/playbook.yml, with the following content:

   ---
   - name: Configure an Ethernet connection with 802.1X authentication
     hosts: managed-node-01.example.com
     tasks:
- name: Copy client key for 802.1X authentication
  ansible.builtin.copy:
    src: "/srv/data/client.key"
    dest: "/etc/pki/tls/private/client.key"
    mode: 0600

- name: Copy client certificate for 802.1X authentication
  ansible.builtin.copy:
    src: "/srv/data/client.crt"
    dest: "/etc/pki/tls/certs/client.crt"

- name: Copy CA certificate for 802.1X authentication
  ansible.builtin.copy:
    src: "/srv/data/ca.crt"
    dest: "/etc/pki/ca-trust/source/anchors/ca.crt"

- name: Configure connection
  ansible.builtin.include_role:
    name: rhel-system-roles.network
    vars:
      network_connections:
        - name: enp1s0
          type: ethernet
          autoconnect: yes
          ip:
            address:
              - 192.0.2.1/24
              - 2001:db8:1::1/64
          gateway4: 192.0.2.254
          gateway6: 2001:db8:1::ffe
          dns:
            - 192.0.2.200
            - 2001:db8:1::ffbb
          dns_search:
            - example.com
          ieee802_1x:
            identity: user_name
            eap: tls
            private_key: "/etc/pki/tls/private/client.key"
            private_key_password: "password"
            client_cert: "/etc/pki/tls/certs/client.crt"
            ca_cert: "/etc/pki/ca-trust/source/anchors/ca.crt"
            domain_suffix_match: example.com
          state: up

These settings define an Ethernet connection profile for the **enp1s0** device with the following settings:

- A static IPv4 address – **192.0.2.1** with a /24 subnet mask
- A static IPv6 address – **2001:db8:1::1** with a /64 subnet mask
- An IPv4 default gateway – **192.0.2.254**
- An IPv6 default gateway – **2001:db8:1::ffe**
- An IPv4 DNS server - 192.0.2.200
- An IPv6 DNS server - 2001:db8:1::ffbb
- A DNS search domain - example.com
- 802.1X network authentication using the TLS Extensible Authentication Protocol (EAP)

2. Validate the playbook syntax:

```
$ ansible-playbook --syntax-check ~/playbook.yml
```

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

```
$ ansible-playbook ~/playbook.yml
```

Additional resources

- /usr/share/ansible/roles/rhel-system-roles.network/README.md file
- /usr/share/doc/rhel-system-roles/network/ directory

33.4. CONFIGURING A WIFI CONNECTION WITH 802.1X NETWORK AUTHENTICATION BY USING THE NETWORK RHEL SYSTEM ROLE

Using RHEL system role, you can automate the creation of a wifi connection. For example, you can remotely add a wireless connection profile for the wlp1s0 interface using an Ansible Playbook. The created profile uses the 802.1X standard to authenticate the client to a wifi network. The playbook configures the connection profile to use DHCP. To configure static IP settings, adapt the parameters in the ip dictionary accordingly.

Prerequisites

- You have prepared the control node and the managed nodes
- You are logged in to the control node as a user who can run playbooks on the managed nodes.
- The account you use to connect to the managed nodes has sudo permissions on them.
- The network supports 802.1X network authentication.
- You installed the wpa_supplicant package on the managed node.
- DHCP is available in the network of the managed node.
- The following files required for TLS authentication exist on the control node:
  - The client key is stored in the /srv/data/client.key file.
  - The client certificate is stored in the /srv/data/client.crt file.
  - The CA certificate is stored in the /srv/data/ca.crt file.
**Procedure**

1. Create a playbook file, for example ~/playbook.yml, with the following content:

```yaml
---
- name: Configure a wifi connection with 802.1X authentication
  hosts: managed-node-01.example.com
  tasks:
    - name: Copy client key for 802.1X authentication
      ansible.builtin.copy:
        src: "/srv/data/client.key"
        dest: "/etc/pki/tls/private/client.key"
        mode: 0400
    - name: Copy client certificate for 802.1X authentication
      ansible.builtin.copy:
        src: "/srv/data/client.crt"
        dest: "/etc/pki/tls/certs/client.crt"
    - name: Copy CA certificate for 802.1X authentication
      ansible.builtin.copy:
        src: "/srv/data/ca.crt"
        dest: "/etc/pki/ca-trust/source/anchors/ca.crt"
    - block:
      - ansible.builtin.import_role:
        name: rhel-system-roles.network
        vars:
          network_connections:
            - name: Configure the Example-wifi profile
              interface_name: wlp1s0
              state: up
              type: wireless
              autoconnect: yes
              ip:
                dhcp4: true
                auto6: true
              wireless:
                ssid: "Example-wifi"
                key_mgmt: "wpa-eap"
                ieee802_1x:
                  identity: "user_name"
                  eap: tls
                  private_key: "/etc/pki/tls/client.key"
                  private_key_password: "password"
                  private_key_password_flags: none
                  client_cert: "/etc/pki/tls/client.pem"
                  ca_cert: "/etc/pki/tls/cacert.pem"
                  domain_suffix_match: "example.com"
```

These settings define a wifi connection profile for the **wlp1s0** interface. The profile uses 802.1X standard to authenticate the client to the wifi network. The connection retrieves IPv4 addresses, IPv6 addresses, default gateway, routes, DNS servers, and search domains from a DHCP server and IPv6 stateless address autoconfiguration (SLAAC).

2. Validate the playbook syntax:
$ ansible-playbook --syntax-check ~/playbook.yml

Note that this command only validates the syntax and does not protect against a wrong but valid configuration.

3. Run the playbook:

$ ansible-playbook ~/playbook.yml

Additional resources

- /usr/share/ansible/roles/rhel-system-roles.network/README.md file
- /usr/share/doc/rhel-system-roles/network/ directory
CHAPTER 34. SETTING UP AN 802.1X NETWORK AUTHENTICATION SERVICE FOR LAN CLIENTS BY USING HOSTAPD WITH FREERADIUS BACKEND

The IEEE 802.1X standard defines secure authentication and authorization methods to protect networks from unauthorized clients. By using the `hostapd` service and FreeRADIUS, you can provide network access control (NAC) in your network.

In this documentation, the RHEL host acts as a bridge to connect different clients with an existing network. However, the RHEL host grants only authenticated clients access to the network.

34.1. PREREQUISITES

- A clean installation of FreeRADIUS.
  
  If the `freeradius` package is already installed, remove the `/etc/raddb/` directory, uninstall and then install the package again. Do not reinstall the package by using the `dnf reinstall` command, because the permissions and symbolic links in the `/etc/raddb/` directory are then different.

34.2. SETTING UP THE BRIDGE ON THE AUTHENTICATOR

A network bridge is a link-layer device which forwards traffic between hosts and networks based on a table of MAC addresses. If you set up RHEL as an 802.1X authenticator, add both the interfaces on which to perform authentication and the LAN interface to the bridge.

**Prerequisites**

- The server has multiple Ethernet interfaces.

**Procedure**

1. Create the bridge interface:

   ```bash
   # nmcli connection add type bridge con-name br0 ifname br0
   ```

2. Assign the Ethernet interfaces to the bridge:

   ```bash
   # nmcli connection add type ethernet slave-type bridge con-name br0-port1 ifname enp1s0 master br0
   # nmcli connection add type ethernet slave-type bridge con-name br0-port2 ifname 
   ```
3. Enable the bridge to forward extensible authentication protocol over LAN (EAPOL) packets:

   # nmcli connection modify br0 group-forward-mask 8

4. Configure the connection to automatically activate the ports:

   # nmcli connection modify br0 connection.autoconnect-slaves 1

5. Activate the connection:

   # nmcli connection up br0

**Verification**

1. Display the link status of Ethernet devices that are ports of a specific bridge:

   # ip link show master br0

   3: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel master
   br0 state UP mode DEFAULT group default qlen 1000
   link/ether 52:54:00:62:61:0e brd ff:ff:ff:ff:ff:
   
   ...  

2. Verify if forwarding of EAPOL packets is enabled on the *br0* device:

   # cat /sys/class/net/br0/bridge/group_fwd_mask

   0x8

   If the command returns 0x8, forwarding is enabled.

**Additional resources**

- [nm-settings(5) man page](#)

**34.3. CERTIFICATE REQUIREMENTS BY FREERADIUS**

For a secure FreeRADIUS service, you require TLS certificates for different purposes:

- A TLS server certificate for encrypted connections to the server. Use a trusted certificate authority (CA) to issue the certificate.
  The server certificate requires the extended key usage (EKU) field set to **TLS Web Server Authentication**.

- Client certificates issued by the same CA for extended authentication protocol transport layer security (EAP-TLS). EAP-TLS provides certificate-based authentication and is enabled by default.
  The client certificates require their EKU field set to **TLS Web Client Authentication**.
### 34.4. CREATING A SET OF CERTIFICATES ON A FREERADIUS SERVER FOR TESTING PURPOSES

For testing purposes, the `freeradius` package installs scripts and configuration files in the `/etc/raddb/certs/` directory to create your own certificate authority (CA) and issue certificates.

**IMPORTANT**

If you use the default configuration, certificates generated by these scripts expire after 60 days and keys use an insecure password ("whatever"). However, you can customize the CA, server, and client configuration.

After you perform the procedure, the following files, which you require later in this documentation, are created:

- `/etc/raddb/certs/ca.pem`: CA certificate
- `/etc/raddb/certs/server.key`: Private key of the server certificate
- `/etc/raddb/certs/server.pem`: Server certificate
- `/etc/raddb/certs/client.key`: Private key of the client certificate
- `/etc/raddb/certs/client.pem`: Client certificate

**Prerequisites**

- You installed the `freeradius` package.

**Procedure**

1. Change into the `/etc/raddb/certs/` directory:

   ```
   # cd /etc/raddb/certs/
   ```

2. Optional: Customize the CA configuration in the `/etc/raddb/certs/ca.cnf` file:

   ```
   [ req ]
   default_bits = 2048
   input_password = ca_password
   output_password = ca_password
   ...
   ```
3. Optional: Customize the server configuration in the `/etc/raddb/certs/server.cnf` file::

```bash
[ CA_default ]
default_days = 730

[ req ]
distinguished_name = server
default_bits = 2048
input_password = key_password
output_password = key_password

[ server ]
countryName = US
stateOrProvinceName = North Carolina
localityName = Raleigh
organizationName = Example Inc.
emailAddress = admin@example.org
commonName = "Example Server Certificate"
```

4. Optional: Customize the client configuration in the `/etc/raddb/certs/client.cnf` file::

```bash
[ CA_default ]
default_days = 365

[ req ]
distinguished_name = client
default_bits = 2048
input_password = password_on_private_key
output_password = password_on_private_key

[ client ]
countryName = US
stateOrProvinceName = North Carolina
localityName = Raleigh
organizationName = Example Inc.
emailAddress = user@example.org
commonName = user@example.org
```

5. Create the certificates:

```
# make all
```
6. Change the group on the `/etc/raddb/certs/server.pem` file to `radiusd`:

```bash
# chgrp radiusd /etc/raddb/certs/server.pem
```

Additional resources

- `/etc/raddb/certs/README.md`

### 34.5. CONFIGURING FREERADIUS TO AUTHENTICATE NETWORK CLIENTS SECURESLY BY USING EAP

FreeRADIUS supports different methods of the Extensible authentication protocol (EAP). However, for a secure network, configure FreeRADIUS to support only the following secure EAP authentication methods:

- **EAP-TLS** (transport layer security) uses a secure TLS connection to authenticate clients by using certificates. To use EAP-TLS, you need TLS client certificates for each network client and a server certificate for the server. Note that the same certificate authority (CA) must have issued the certificates. Always use your own CA to create certificates, because all client certificates issued by the CA you use can authenticate to your FreeRADIUS server.

- **EAP-TTLS** (tunneled transport layer security) uses a secure TLS connection and authenticates clients by using mechanisms, such as password authentication protocol (PAP) or challenge handshake authentication protocol (CHAP). To use EAP-TTLS, you need a TLS server certificate.

- **EAP-PEAP** (protected extensible authentication protocol) uses a secure TLS connection as the outer authentication protocol to set up the tunnel. The authenticator authenticates the certificate of the RADIUS server. Afterwards, the supplicant authenticates through the encrypted tunnel by using Microsoft challenge handshake authentication protocol version 2 (MS-CHAPv2) or other methods.

**NOTE**

The default FreeRADIUS configuration files serve as documentation and describe all parameters and directives. If you want to disable certain features, comment them out instead of removing the corresponding parts in the configuration files. This enables you to preserve the structure of the configuration files and the included documentation.

### Prerequisites

- You installed the `freeradius` package.
- The configuration files in the `/etc/raddb/` directory are unchanged and as provided by the `freeradius` package.
- The following files exist on the server:
  - TLS private key of the FreeRADIUS host: `/etc/raddb/certs/server.key`
  - TLS server certificate of the FreeRADIUS host: `/etc/raddb/certs/server.pem`
  - TLS CA certificate: `/etc/raddb/certs/ca.pem`
If you store the files in a different location or if they have different names, set the `private_key_file`, `certificate_file`, and `ca_file` parameters in the `/etc/raddb/mods-available/eap` file accordingly.

**Procedure**

1. If the `/etc/raddb/certs/dh` with Diffie–Hellman (DH) parameters does not exist, create one. For example, to create a DH file with a 2048 bits prime, enter:

   ```
   # openssl dhparam -out /etc/raddb/certs/dh 2048
   ```

   For security reasons, do not use a DH file with less than a 2048 bits prime. Depending on the number of bits, the creation of the file can take several minutes.

2. Set secure permissions on the TLS private key, server certificate, CA certificate, and the file with DH parameters:

   ```
   # chmod 640 /etc/raddb/certs/server.key /etc/raddb/certs/server.pem
   /etc/raddb/certs/ca.pem /etc/raddb/certs/dh
   # chown root:radiusd /etc/raddb/certs/server.key /etc/raddb/certs/server.pem
   /etc/raddb/certs/ca.pem /etc/raddb/certs/dh
   ```

3. Edit the `/etc/raddb/mods-available/eap` file:

   a. Set the password of the private key in the `private_key_password` parameter:

   ```
   eap {
   ...
   tls-config tls-common {
   ...
   private_key_password = key_password
   ...
   }
   }
   ```

   b. Depending on your environment, set the `default_eap_type` parameter in the `eap` directive to your primary EAP type you use:

   ```
   eap {
   ...
   default_eap_type = ttls
   ...
   }
   ```

   For a secure environment, use only `ttls`, `tls`, or `peap`.

   c. Comment out the `md5` directives to disable the insecure EAP-MD5 authentication method:

   ```
   eap {
   ...
   # md5 {
   # }
   ...
   }
   ```
Note that, in the default configuration file, other insecure EAP authentication methods are commented out by default.

4. Edit the `/etc/raddb/sites-available/default` file, and comment out all authentication methods other than `eap`:

```plaintext
authenticate {
...
# Auth-Type PAP {
#   pap
# }

# Auth-Type CHAP {
#   chap
# }

# Auth-Type MS-CHAP {
#   mschap
# }

# mschap

# digest
...
}
```

This leaves only EAP enabled and disables plain-text authentication methods.

5. Edit the `/etc/raddb/clients.conf` file:

a. Set a secure password in the `localhost` and `localhost_ipv6` client directives:

```plaintext
client localhost {
  ipaddr = 127.0.0.1
...
  secret = client_password
...
}

client localhost_ipv6 {
  ipv6addr = ::1
  secret = client_password
}
```

b. If RADIUS clients, such as network authenticators, on remote hosts should be able to access the FreeRADIUS service, add corresponding client directives for them:

```plaintext
client hostapd.example.org {
  ipaddr = 192.0.2.2/32
  secret = client_password
}
```

The `ipaddr` parameter accepts IPv4 and IPv6 addresses, and you can use the optional classless inter-domain routing (CIDR) notation to specify ranges. However, you can set only one value in this parameter. For example, to grant access to an IPv4 and IPv6 address, add two client directives.
Use a descriptive name for the client directive, such as a hostname or a word that describes where the IP range is used.

6. If you want to use EAP-TTLS or EAP-PEAP, add the users to the `/etc/raddb/users` file:

   ```
   example_user     Cleartext-Password := "user_password"
   ```

   For users who should use certificate-based authentication (EAP-TLS), do not add any entry.

7. Verify the configuration files:

   ```
   # radiusd -XC
   ...
   Configuration appears to be OK
   ```

8. Enable and start the `radiusd` service:

   ```
   # systemctl enable --now radiusd
   ```

Verification

- Testing EAP-TTLS authentication against a FreeRADIUS server or authenticator
- Testing EAP-TLS authentication against a FreeRADIUS server or authenticator

Troubleshooting

1. Stop the `radiusd` service:

   ```
   # systemctl stop radiusd
   ```

2. Start the service in debug mode:

   ```
   # radiusd -X
   ...
   Ready to process requests
   ```

3. Perform authentication tests on the FreeRADIUS host, as referenced in the Verification section.

Next steps

- Disable no longer required authentication methods and other features you do not use.

34.6. CONFIGURING HOSTAPD AS AN AUTHENTICATOR IN A WIRED NETWORK

The host access point daemon (`hostapd`) service can act as an authenticator in a wired network to provide 802.1X authentication. For this, the `hostapd` service requires a RADIUS server that authenticates the clients.
The hostapd service provides an integrated RADIUS server. However, use the integrated RADIUS server only for testing purposes. For production environments, use FreeRADIUS server, which supports additional features, such as different authentication methods and access control.

**IMPORTANT**

The hostapd service does not interact with the traffic plane. The service acts only as an authenticator. For example, use a script or service that uses the hostapd control interface to allow or deny traffic based on the result of authentication events.

Prerequisites

- You installed the hostapd package.
- The FreeRADIUS server has been configured, and it is ready to authenticate clients.

Procedure

1. Create the `/etc/hostapd/hostapd.conf` file with the following content:

```plaintext
# General settings of hostapd
# =========================================

# Control interface settings
ctrl_interface=/var/run/hostapd
ctrl_interface_group=wheel

# Enable logging for all modules
logger_syslog=-1
logger_stdout=-1

# Log level
logger_syslog_level=2
logger_stdout_level=2

# Wired 802.1X authentication
# =========================================

# Driver interface type
driver=wired

# Enable IEEE 802.1X authorization
ieee8021x=1

# Use port access entry (PAE) group address
# (01:80:c2:00:00:03) when sending EAPOL frames
use_pae_group_addr=1

# Network interface for authentication requests
interface=br0

# RADIUS client configuration
```
# Local IP address used as NAS-IP-Address
own_ip_addr=192.0.2.2

# Unique NAS-Identifier within scope of RADIUS server
nas_identifier=hostapd.example.org

# RADIUS authentication server
auth_server_addr=192.0.2.1
auth_server_port=1812
auth_server_shared_secret=client_password

# RADIUS accounting server
acct_server_addr=192.0.2.1
acct_server_port=1813
acct_server_shared_secret=client_password

For further details about the parameters used in this configuration, see their descriptions in the /usr/share/doc/hostapd/hostapd.conf example configuration file.

2. Enable and start the hostapd service:

```bash
# systemctl enable --now hostapd
```

Verification

- See:
  - Testing EAP-TTLS authentication against a FreeRADIUS server or authenticator
  - Testing EAP-TLS authentication against a FreeRADIUS server or authenticator

Troubleshooting

1. Stop the hostapd service:

```bash
# systemctl stop hostapd
```

2. Start the service in debug mode:

```bash
# hostapd -d /etc/hostapd/hostapd.conf
```

3. Perform authentication tests on the FreeRADIUS host, as referenced in the Verification section.

Additional resources

- hostapd.conf(5) man page
- /usr/share/doc/hostapd/hostapd.conf file
34.7. TESTING EAP-TTLS AUTHENTICATION AGAINST A FREERADIUS SERVER OR AUTHENTICATOR

To test if authentication by using extensible authentication protocol (EAP) over tunneled transport layer security (EAP-TTLS) works as expected, run this procedure:

- After you set up the FreeRADIUS server
- After you set up the `hostapd` service as an authenticator for 802.1X network authentication.

The output of the test utilities used in this procedure provide additional information about the EAP communication and help you to debug problems.

Prerequisites

- When you want to authenticate to:
  - A FreeRADIUS server:
    - The `eapol_test` utility, provided by the `hostapd` package, is installed.
    - The client, on which you run this procedure, has been authorized in the FreeRADIUS server’s client databases.
  - An authenticator, the `wpa_supplicant` utility, provided by the same-named package, is installed.
  - You stored the certificate authority (CA) certificate in the `/etc/pki/tls/certs/ca.pem` file.

Procedure

1. Create the `/etc/wpa_supplicant/wpa_supplicant-TTLS.conf` file with the following content:

   ```
   ap_scan=0
   network={
     eap=TTLS
     eapol_flags=0
     key_mgmt=IEEE8021X
     # Anonymous identity (sent in unencrypted phase 1)
     # Can be any string
     anonymous_identity="anonymous"
     # Inner authentication (sent in TLS-encrypted phase 2)
     phase2="auth=PAP"
     identity="example_user"
     password="user_password"
     # CA certificate to validate the RADIUS server’s identity
     ca_cert="/etc/pki/tls/certs/ca.pem"
   }
   ```

2. To authenticate to:

- A FreeRADIUS server, enter:
The `-a` option defines the IP address of the FreeRADIUS server, and the `-s` option specifies the password for the host on which you run the command in the FreeRADIUS server’s client configuration.

- An authenticator, enter:

```bash
# wpa_supplicant -c /etc/wpa_supplicant/wpa_supplicant-TTLS.conf -D wired -i enp0s31f6
```

The `-i` option specifies the network interface name on which `wpa_supplicant` sends out extended authentication protocol over LAN (EAPOL) packets.

For more debugging information, pass the `-d` option to the command.

Additional resources

- `/usr/share/doc/wpa_supplicant/wpa_supplicant.conf` file

### 34.8. TESTING EAP-TLS AUTHENTICATION AGAINST A FREERADIUS SERVER OR AUTHENTICATOR

To test if authentication by using extensible authentication protocol (EAP) transport layer security (EAP-TLS) works as expected, run this procedure:

- After you set up the FreeRADIUS server

- After you set up the `hostapd` service as an authenticator for 802.1X network authentication.

The output of the test utilities used in this procedure provide additional information about the EAP communication and help you to debug problems.

Prerequisites

- When you want to authenticate to:
  - A FreeRADIUS server:
    - The `eapol_test` utility, provided by the `hostapd` package, is installed.
    - The client, on which you run this procedure, has been authorized in the FreeRADIUS server’s client databases.
- An authenticator, the `wpa_supplicant` utility, provided by the same-named package, is installed.

- You stored the certificate authority (CA) certificate in the `/etc/pki/tls/certs/ca.pem` file.

- The CA that issued the client certificate is the same that issued the server certificate of the FreeRADIUS server.

- You stored the client certificate in the `/etc/pki/tls/certs/client.pem` file.

- You stored the private key of the client in the `/etc/pki/tls/private/client.key` file.

### Procedure

1. Create the `/etc/wpa_supplicant/wpa_supplicant-TLS.conf` file with the following content:

   ```
ap_scan=0

network={
    eap=TLS
    eapol_flags=0
    key_mgmt=IEEE8021X

    identity="user@example.org"
    client_cert="/etc/pki/tls/certs/client.pem"
    private_key="/etc/pki/tls/private/client.key"
    private_key_passwd="password_on_private_key"

    # CA certificate to validate the RADIUS server’s identity
cert_certs="/etc/pki/tls/certs/ca.pem"
}
```

2. To authenticate to:

   - A FreeRADIUS server, enter:

     ```
     # eaplo_test -c /etc/wpa_supplicant/wpa_supplicant-TLS.conf -a 192.0.2.1 -s client_password
     ...
     EAP: Status notification: remote certificate verification (param=success)
     ...
     CTRL-EVENT-EAP-SUCCESS EAP authentication completed successfully
     ...
     SUCCESS
     ```

     The `-a` option defines the IP address of the FreeRADIUS server, and the `-s` option specifies the password for the host on which you run the command in the FreeRADIUS server’s client configuration.

   - An authenticator, enter:

     ```
     # wpa_supplicant -c /etc/wpa_supplicant/wpa_supplicant-TLS.conf -D wired -i enp0s31f6
     ...
     enp0s31f6: CTRL-EVENT-EAP-SUCCESS EAP authentication completed successfully
     ...
     ```
The `-i` option specifies the network interface name on which `wpa_supplicant` sends out extended authentication protocol over LAN (EAPOL) packets.

For more debugging information, pass the `-d` option to the command.

Additional resources

- `/usr/share/doc/wpa_supplicant/wpa_supplicant.conf` file

### 34.9. BLOCKING AND ALLOWING TRAFFIC BASED ON HOSTAPD AUTHENTICATION EVENTS

The `hostapd` service does not interact with the traffic plane. The service acts only as an authenticator. However, you can write a script to allow and deny traffic based on the result of authentication events.

**IMPORTANT**

This procedure is not supported and is no enterprise-ready solution. It only demonstrates how to block or allow traffic by evaluating events retrieved by `hostapd_cli`.

When the `802-1x-tr-mgmt` systemd service starts, RHEL blocks all traffic on the listen port of `hostapd` except extensible authentication protocol over LAN (EAPOL) packets and uses the `hostapd_cli` utility to connect to the `hostapd` control interface. The `/usr/local/bin/802-1x-tr-mgmt` script then evaluates events. Depending on the different events received by `hostapd_cli`, the script allows or blocks traffic for MAC addresses. Note that, when the `802-1x-tr-mgmt` service stops, all traffic is automatically allowed again.

Perform this procedure on the `hostapd` server.

**Prerequisites**

- The `hostapd` service has been configured, and the service is ready to authenticate clients.

**Procedure**

1. Create the `/usr/local/bin/802-1x-tr-mgmt` file with the following content:

```bash
#!/bin/sh

if [ "$1" == "xblock_all" ]
then

    nft delete table bridge tr-mgmt-br0 2>/dev/null || true
    nft -f - << EOF
    table bridge tr-mgmt-br0 {
        set allowed_macs {
            type ether_addr
        }
    }

    chain accesscontrol {
        ether saddr @allowed_macs accept
        ether daddr @allowed_macs accept
drop
    }

EOF

else
    nft delete table bridge tr-mgmt-br0 2>/dev/null || true
    nft -f - << EOF
    table bridge tr-mgmt-br0 {
        set allowed_macs {
            type ether_addr
        }
    }

    chain accesscontrol {
        ether saddr @allowed_macs accept
        ether daddr @allowed_macs accept
drop
    }

EOF

fi
```
2. Create the `/etc/systemd/system/802-1x-tr-mgmt@.service` systemd service file with the following content:

```plaintext
[Unit]
Description=Example 802.1x traffic management for hostapd
After=hostapd.service
After=sys-devices-virtual-net-%i.device

[Service]
Type=simple
ExecStartPre=/bin/sh -c '/usr/sbin/tc qdisc del dev %i ingress > /dev/null 2>&1'
ExecStartPre=/bin/sh -c '/usr/sbin/tc qdisc del dev %i clsact > /dev/null 2>&1'
ExecStartPre=/usr/sbin/tc qdisc add dev %i clsact
ExecStartPre=/usr/sbin/tc filter add dev %i ingress pref 10000 protocol 0x888e matchall action ok index 100
ExecStartPre=/usr/sbin/tc filter add dev %i ingress pref 10001 protocol all matchall action drop index 101
ExecStart=/usr/sbin/hostapd_cli -i %i -a /usr/local/bin/802-1x-tr-mgmt
ExecStopPost=/usr/sbin/tc qdisc del dev %i clsact

[Install]
WantedBy=multi-user.target
```
3. Reload systemd:

   # systemctl daemon-reload

4. Enable and start the 802-1x-tr-mgmt service with the interface name hostapd is listening on:

   # systemctl enable --now 802-1x-tr-mgmt@br0.service

Verification

- Authenticate with a client to the network. See:
  - Testing EAP-TTLS authentication against a FreeRADIUS server or authenticator
  - Testing EAP-TLS authentication against a FreeRADIUS server or authenticator

Additional resources

- systemctl.service(5) man page
CHAPTER 35. GETTING STARTED WITH MULTIPATH TCP

Transmission Control Protocol (TCP) ensures reliable delivery of the data through the internet and automatically adjusts its bandwidth in response to network load. Multipath TCP (MPTCP) is an extension to the original TCP protocol (single-path). MPTCP enables a transport connection to operate across multiple paths simultaneously, and brings network connection redundancy to user endpoint devices.

35.1. UNDERSTANDING MPTCP

The Multipath TCP (MPTCP) protocol allows for simultaneous usage of multiple paths between connection endpoints. The protocol design improves connection stability and also brings other benefits compared to the single-path TCP.

NOTE

In MPTCP terminology, links are considered as paths.

The following are some of the advantages of using MPTCP:

- It allows a connection to simultaneously use multiple network interfaces.
- In case a connection is bound to a link speed, the usage of multiple links can increase the connection throughput. Note, that in case of the connection is bound to a CPU, the usage of multiple links causes the connection slowdown.
- It increases the resilience to link failures.

For more details about MPTCP, we highly recommend you review the Additional resources.

Additional resources

- Understanding Multipath TCP: High availability for endpoints and the networking highway of the future
- RFC8684: TCP Extensions for Multipath Operation with Multiple Addresses

35.2. PREPARING RHEL TO ENABLE MPTCP SUPPORT

By default the MPTCP support is disabled in RHEL. Enable MPTCP so that applications that support this feature can use it. Additionally, you have to configure user space applications to force use MPTCP sockets if those applications have TCP sockets by default.

Prerequisites

The following packages are installed:

- iperf3
- mptcpd
- systemtap

Procedure
1. Enable MPTCP sockets in the kernel:

```
# echo "net.mptcp.enabled=1" > /etc/sysctl.d/90-enable-MPTCP.conf
# sysctl -p /etc/sysctl.d/90-enable-MPTCP.conf
```

2. Start the `iperf3` server, and force it to create MPTCP sockets instead of TCP sockets:

```
# mptcpize run iperf3 -s
Server listening on 5201
```

3. Connect the client to the server, and force it to create MPTCP sockets instead of TCP sockets:

```
# mptcpize iperf3 -c 127.0.0.1 -t 3
```

4. After the connection is established, verify the `ss` output to see the subflow-specific status:

```
# ss -nti '( dport :5201 )'
```

```
State Recv-Q Send-Q Local Address:Port Peer Address:Port Process
ESTAB 0 0 127.0.0.1:41842 127.0.0.1:5201
advmss:65483 cwnd:7 data_segs_out:142 bytes_sent:393050505bps lastsnd:2813 lastrcv:2772
lastack:2772 pacing_rate 785946640bps delivery_rate 10944000000bps delivered:4
busy:41ms rcv_space:43690 rcv_ssthresh:43690 minrtt:0.008 tcp-ulp-mptcp flags:Mmec
token:0000(id:0)/2ff053ec(id:0) seq:3e2cbea12d7673d4 sfseq:3 ssnoff:ad3d00f4 maplen:2
```

5. Verify MPTCP counters:

```
# nstat MPTcp
```

```
#kernel
MPTcpExtMPCapableSYNRX 2 0.0
MPTcpExtMPCapableSYNTX 2 0.0
MPTcpExtMPCapableSYNACKRX 2 0.0
MPTcpExtMPCapableACKRX 2 0.0
```

**Additional resources**

- `tcp(7)` man page
- `mptcpize(8)` man page

### 35.3. USING IPROUTE2 TO TEMPORARILY CONFIGURE AND ENABLE MULTIPLE PATHS FOR MPTCP APPLICATIONS

Each MPTCP connection uses a single subflow similar to plain TCP. To get the MPTCP benefits, specify a higher limit for maximum number of subflows for each MPTCP connection. Then configure additional endpoints to create those subflows.
IMPORTANT

The configuration in this procedure will not persist after rebooting your machine.

Note that MPTCP does not yet support mixed IPv6 and IPv4 endpoints for the same socket. Use endpoints belonging to the same address family.

Prerequisites

- The `mptcpd` package is installed
- The `iperf3` package is installed
- Server network interface settings:
  - `enp4s0`: **192.0.2.1/24**
  - `enp1s0`: **198.51.100.1/24**
- Client network interface settings:
  - `enp4s0f0`: **192.0.2.2/24**
  - `enp4s0f1`: **198.51.100.2/24**

Procedure

1. Configure the client to accept up to 1 additional remote address, as provided by the server:
   ```
   # ip mptcp limits set add_addr_accepted 1
   ```

2. Add IP address **198.51.100.1** as a new MPTCP endpoint on the server:
   ```
   # ip mptcp endpoint add 198.51.100.1 dev enp1s0 signal
   ```
   The `signal` option ensures that the `ADD_ADDR` packet is sent after the three-way-handshake.

3. Start the `iperf3` server, and force it to create MPTCP sockets instead of TCP sockets:
   ```
   # mptcpize run iperf3 -s
   Server listening on 5201
   ```

4. Connect the client to the server, and force it to create MPTCP sockets instead of TCP sockets:
   ```
   # mptcpize iperf3 -c 192.0.2.1 -t 3
   ```

Verification

1. Verify the connection is established:
   ```
   # ss -nti '( sport :5201 )'
   ```

2. Verify the connection and IP address limit:
# ip mptcp limit show

3. Verify the newly added endpoint:

# ip mptcp endpoint show

4. Verify MPTCP counters by using the `nstat MPTcp*` command on a server:

```bash
# nstat MPTcp*
#kernel
MPTcpExtMPCapableSYNRX  2  0.0
MPTcpExtMPCapableACKRX  2  0.0
MPTcpExtMPJoinSynRx     2  0.0
MPTcpExtMPJoinAckRx     2  0.0
MPTcpExtEchoAdd         2  0.0
```

Additional resources

- `ip-mptcp(8)` man page
- `mptcpize(8)` man page

### 35.4. PERMANENTLY CONFIGURING MULTIPLE PATHS FOR MPTCP APPLICATIONS

You can configure MultiPath TCP (MPTCP) using the `nmcli` command to permanently establish multiple subflows between a source and destination system. The subflows can use different resources, different routes to the destination, and even different networks. Such as Ethernet, cellular, wifi, and so on. As a result, you achieve combined connections, which increase network resilience and throughput.

The server uses the following network interfaces in our example:

- `enp4s0`: 192.0.2.1/24
- `enp1s0`: 198.51.100.1/24
- `enp7s0`: 192.0.2.3/24

The client uses the following network interfaces in our example:

- `enp4s0f0`: 192.0.2.2/24
- `enp4s0f1`: 198.51.100.2/24
- `enp6s0`: 192.0.2.5/24

**Prerequisites**

- You configured the default gateway on the relevant interfaces.

**Procedure**

1. Enable MPTCP sockets in the kernel:
2. Optional: The RHEL kernel default for subflow limit is 2. If you require more:
   a. Create the `/etc/systemd/system/set_mptcp_limit.service` file with the following content:

   ```
   [Unit]
   Description=Set MPTCP subflow limit to 3
   After=network.target
   
   [Service]
   ExecStart=ip mptcp limits set subflows 3
   Type=oneshot
   
   [Install]
   WantedBy=multi-user.target
   ```

   The `oneshot` unit executes the `ip mptcp limits set subflows 3` command after your network (`network.target`) is operational during every boot process.

   The `ip mptcp limits set subflows 3` command sets the maximum number of `additional` subflows for each connection, so 4 in total. It is possible to add maximally 3 additional subflows.

   b. Enable the `set_mptcp_limit` service:

   ```
   # systemctl enable --now set_mptcp_limit
   ```

3. Enable MPTCP on all connection profiles that you want to use for connection aggregation:

   ```
   # nmcli connection modify <profile_name> connection.mptcp-flags signal,subflow,also-without-default-route
   ```

   The `connection.mptcp-flags` parameter configures MPTCP endpoints and the IP address flags. If MPTCP is enabled in a NetworkManager connection profile, the setting will configure the IP addresses of the relevant network interface as MPTCP endpoints.

   By default, NetworkManager does not add MPTCP flags to IP addresses if there is no default gateway. If you want to bypass that check, you need to use also the `also-without-default-route` flag.

**Verification**

1. Verify that you enabled the MPTCP kernel parameter:

   ```
   # sysctl net.mptcp.enabled
   net.mptcp.enabled = 1
   ```

2. Verify that you set the subflow limit correctly, in case the default was not enough:

   ```
   # ip mptcp limit show
   add_addr_accepted 2 subflows 3
   ```
3. Verify that you configured the per-address MPTCP setting correctly:

```bash
# ip mptcp endpoint show
192.0.2.1 id 1 subflow dev enp4s0
198.51.100.1 id 2 subflow dev enp1s0
192.0.2.3 id 3 subflow dev enp7s0
192.0.2.4 id 4 subflow dev enp3s0
...```

Additional resources

- `nm-settings-nmcli(5)`
- `ip-mptcp(8)`
- Section 35.1, "Understanding MPTCP"
- Understanding Multipath TCP: High availability for endpoints and the networking highway of the future
- RFC8684: TCP Extensions for Multipath Operation with Multiple Addresses
- Using Multipath TCP to better survive outages and increase bandwidth

### 35.5. MONITORING MPTCP SUB-FLOWS

The life cycle of a multipath TCP (MPTCP) socket can be complex: The main MPTCP socket is created, the MPTCP path is validated, one or more sub-flows are created and eventually removed. Finally, the MPTCP socket is terminated.

The MPTCP protocol allows monitoring MPTCP-specific events related to socket and sub-flow creation and deletion, using the `ip` utility provided by the `iproute` package. This utility uses the `netlink` interface to monitor MPTCP events.

This procedure demonstrates how to monitor MPTCP events. For that, it simulates a MPTCP server application, and a client connects to this service. The involved clients in this example use the following interfaces and IP addresses:

- **Server:** 192.0.2.1
- **Client (Ethernet connection):** 192.0.2.2
- **Client (WiFi connection):** 192.0.2.3

To simplify this example, all interfaces are within the same subnet. This is not a requirement. However, it is important that routing has been configured correctly, and the client can reach the server via both interfaces.

**Prerequisites**

- A RHEL client with two network interfaces, such as a laptop with Ethernet and WiFi
- The client can connect to the server via both interfaces
- A RHEL server
Both the client and the server run RHEL 9.0 or later

You installed the **mptcpd** package on both the client and the server

**Procedure**

1. Set the per connection additional subflow limits to 1 on both client and server:

   ```
   # ip mptcp limits set add_addr_accepted 0 subflows 1
   ```

2. On the server, to simulate a MPTCP server application, start `netcat (nc)` in listen mode with enforced MPTCP sockets instead of TCP sockets:

   ```
   # mptcpize run nc -l -k -p 12345
   ```
   The `-k` option causes that `nc` does not close the listener after the first accepted connection. This is required to demonstrate the monitoring of sub-flows.

3. On the client:

   a. Identify the interface with the lowest metric:

   ```
   # ip -4 route
   192.0.2.0/24 dev enp1s0 proto kernel scope link src 192.0.2.2 metric 100
   192.0.2.0/24 dev wlp1s0 proto kernel scope link src 192.0.2.3 metric 600
   ```

   The **enp1s0** interface has a lower metric than **wlp1s0**. Therefore, RHEL uses **enp1s0** by default.

   b. On the first terminal, start the monitoring:

   ```
   # ip mptcp monitor
   ```

   c. On the second terminal, start a MPTCP connection to the server:

   ```
   # mptcpize run nc 192.0.2.1 12345
   ```

   RHEL uses the **enp1s0** interface and its associated IP address as a source for this connection.

   On the monitoring terminal, the `ip mptcp monitor` command now logs:

   ```
   [       CREATED] token=63c070d2 remid=0 locid=0 saddr4=192.0.2.2 daddr4=192.0.2.1
   sport=36444 dport=12345
   ```

   The token identifies the MPTCP socket as an unique ID, and later it enables you to correlate MPTCP events on the same socket.

   d. On the terminal with the running `nc` connection to the server, press **Enter**. This first data packet fully establishes the connection. Note that, as long as no data has been sent, the connection is not established.

   On the monitoring terminal, `ip mptcp monitor` now logs:
e. Optional: Display the connections to port 12345 on the server:

```
# ss -taunp | grep ":12345"
tcp ESTAB  0  0         192.0.2.2:36444 192.0.2.1:12345
```

At this point, only one connection to the server has been established.

f. On a third terminal, create another endpoint:

```
# ip mptcp endpoint add dev wlp1s0 192.0.2.3 subflow
```

This command sets the name and IP address of the WiFi interface of the client in this command.

On the monitoring terminal, `ip mptcp monitor` now logs:

```
[SF_ESTABLISHED] token=63c070d2 remid=0 locid=2 saddr4=192.0.2.3
daddr4=192.0.2.1 sport=53345 dport=12345 backup=0 ifindex=3
```

The `locid` field displays the local address ID of the new sub-flow and identifies this sub-flow even if the connection uses network address translation (NAT). The `saddr4` field matches the endpoint’s IP address from the `ip mptcp endpoint add` command.

g. Optional: Display the connections to port 12345 on the server:

```
# ss -taunp | grep ":12345"
tcp ESTAB  0  0         192.0.2.2:36444 192.0.2.1:12345
tcp ESTAB  0  0  192.0.2.3%wlp1s0:53345 192.0.2.1:12345
```

The command now displays two connections:

- The connection with source address 192.0.2.2 corresponds to the first MPTCP sub-flow that you established previously.
- The connection from the sub-flow over the `wlp1s0` interface with source address 192.0.2.3.

h. On the third terminal, delete the endpoint:

```
# ip mptcp endpoint delete id 2
```

Use the ID from the `locid` field from the `ip mptcp monitor` output, or retrieve the endpoint ID using the `ip mptcp endpoint show` command.

On the monitoring terminal, `ip mptcp monitor` now logs:

```
[     SF_CLOSED] token=63c070d2 remid=0 locid=2 saddr4=192.0.2.3
daddr4=192.0.2.1 sport=53345 dport=12345 backup=0 ifindex=3
```

i. On the first terminal with the `nc` client, press Ctrl+C to terminate the session.

On the monitoring terminal, `ip mptcp monitor` now logs:
Additional resources

- ip-mptcp(1) man page
- How NetworkManager manages multiple default gateways

35.6. DISABLING MULTIPATH TCP IN THE KERNEL

You can explicitly disable the MPTCP option in the kernel.

Procedure

- Disable the `mptcp.enabled` option.

```
# echo "net.mptcp.enabled=0" > /etc/sysctl.d/90-enable-MPTCP.conf
# sysctl -p /etc/sysctl.d/90-enable-MPTCP.conf
```

Verification

- Verify whether the `mptcp.enabled` is disabled in the kernel.

```
# sysctl -a | grep mptcp.enabled
net.mptcp.enabled = 0
```
CHAPTER 36. MANAGING THE MPTCPD SERVICE

This section describes the basic management of the mptcpd service. The mptcpd package provides the mptcpize tool, which switches on the mptcp protocol in the TCP environment.

36.1. CONFIGURING MPTCPD

The mptcpd service is a component of the mptcp protocol which provides an instrument to configure mptcp endpoints. The mptcpd service creates a subflow endpoint for each address by default. The endpoint list is updated dynamically according to IP addresses modification on the running host. The mptcpd service creates the list of endpoints automatically. It enables multiple paths as an alternative to using the ip utility.

Prerequisites

- The mptcp package installed

Procedure

1. Enable mptcp.enabled option in the kernel with the following command:

```
# echo "net.mptcp.enabled=1" > /etc/sysctl.d/90-enable-MPTCP.conf
# sysctl -p /etc/sysctl.d/90-enable-MPTCP.conf
```

2. Start the mptcpd service:

```
# systemctl start mptcp.service
```

3. Verify endpoint creation:

```
# ip mptcp endpoint
```

4. To stop the mptcpd service, use the following command:

```
# systemctl stop mptcp.service
```

5. To configure mptcpd service manually, modify the /etc/mptcpd/mptcpd.conf configuration file.

Note, that the endpoint, which mptcpd service creates, lasts till the host shutdown.

Additional resources

- mptcpd(8) man page.

36.2. MANAGING APPLICATIONS WITH MPTCPIZE TOOL

Using the mptcpize tool manage applications and services.

The instruction below shows how to use the mptcpize tool to manage applications in the TCP environment.
Assuming, you need to run the `iperf3` utility with the enabled MPTCP socket. You can achieve this goal by following the procedure below.

**Prerequisites**
- The `mptcpd` package is installed
- The `iperf3` package is installed

**Procedure**
- Start `iperf3` utility with MPTCP sockets enabled:
  ```
  # mptcpize run iperf3 -s &
  ```

### 36.3. ENABLING MPTCP SOCKETS FOR A SERVICES USING THE MPTCPIZE UTILITY

The following set of commands instruct you how to manage services using the `mptcpize` tool. You can enable or disable the `mptcp` socket for a service.

Assuming, you need to manage `mptcp` socket for the `nginx` service. You can achieve this goal by following the procedure below.

**Prerequisites**
- The `mptcpd` package is installed
- The `nginx` package is installed

**Procedure**
1. Enable MPTCP sockets for a service:
   ```
   # mptcpize enable nginx
   ```
2. Disable the MPTCP sockets for a service:
   ```
   # mptcpize disable nginx
   ```
3. Restart the service to make the changes to take effect:
   ```
   # systemctl restart nginx
   ```
CHAPTER 37. NETWORKMANAGER CONNECTION PROFILES IN KEYFILE FORMAT

By default, NetworkManager in Red Hat Enterprise Linux 9 and later stores connection profiles in keyfile format. Unlike the deprecated `ifcfg` format, the keyfile format supports all connection settings that NetworkManager provides.

37.1. THE KEYFILE FORMAT OF NETWORKMANAGER PROFILES

The keyfile format is similar to the INI format. For example, the following is an Ethernet connection profile in keyfile format:

```
[connection]
id=example_connection
uuid=82c6272d-1ff7-4d56-9c7c-0eb27c300029
type=ethernet
autoconnect=true

[ipv4]
method=auto

[ipv6]
method=auto

[ethernet]
mac-address=00:53:00:8f:fa:66
```

Each section corresponds to a NetworkManager setting name as described in the `nm-settings(5)` and `nm-settings-keyfile(5)` man pages. Each key-value-pair in a section is one of the properties listed in the settings specification of the man page.

Most variables in NetworkManager keyfiles have a one-to-one mapping. This means that a NetworkManager property is stored in the keyfile as a variable of the same name and in the same format. However, there are exceptions, mainly to make the keyfile syntax easier to read. For a list of these exceptions, see the `nm-settings-keyfile(5)` man page.

IMPORTANT

For security reasons, because connection profiles can contain sensitive information, such as private keys and passphrases, NetworkManager uses only configuration files owned by the root user and that are only readable and writable by root.

Depending on the purpose of the connection profile, save it in one of the following directories:

- `/etc/NetworkManager/system-connections/`: The location of persistent profiles. If you modify a persistent profile by using the NetworkManager API, NetworkManager writes and overwrites files in this directory.
- `/run/NetworkManager/system-connections/`: For temporary profiles that are automatically removed when you reboot the system.
- `/usr/lib/NetworkManager/system-connections/`: For pre-deployed immutable profiles. When you edit such a profile using the NetworkManager API, NetworkManager copies this profile to either the persistent or temporary storage.
NetworkManager does not automatically reload profiles from disk. When you create or update a connection profile in keyfile format, use the **nmcli connection reload** command to inform NetworkManager about the changes.

### 37.2. USING NMCLI TO CREATE KEYFILE CONNECTION PROFILES IN OFFLINE MODE

Red Hat recommends using NetworkManager utilities, such as **nmcli**, the **network** RHEL system role, or the **nmstate** API to manage NetworkManager connections, to create and update configuration files. However, you can also create various connection profiles in the keyfile format in offline mode using the **nmcli --offline connection add** command.

The offline mode ensures that **nmcli** operates without the **NetworkManager** service to produce keyfile connection profiles through standard output. This feature can be useful if:

- You want to create your connection profiles that need to be pre-deployed somewhere. For example in a container image, or as an RPM package.
- You want to create your connection profiles in an environment where the **NetworkManager** service is not available. For example when you want to use the **chroot** utility. Alternatively, when you want to create or modify the network configuration of the RHEL system to be installed through the Kickstart **%post** script.

You can create the following connection profile types:

- static Ethernet connection
- dynamic Ethernet connection
- network bond
- network bridge
- VLAN or any kind of supported connections

**Procedure**

1. Create a new connection profile in the keyfile format. For example, for a connection profile of an Ethernet device that does not use DHCP, run a similar **nmcli** command:

   ```
   # nmcli --offline connection add type ethernet con-name Example-Connection
   ipv4.addresses 192.0.2.1/24 ipv4.dns 192.0.2.200 ipv4.method manual >
   /etc/NetworkManager/system-connections/output.nmconnection
   ```

**NOTE**

The connection name you specified with the **con-name** key is saved into the **id** variable of the generated profile. When you use the **nmcli** command to manage this connection later, specify the connection as follows:

- When the **id** variable is not omitted, use the connection name, for example **Example-Connection**.
- When the **id** variable is omitted, use the file name without the **.nmconnection** suffix, for example **output**.
2. Set permissions to the configuration file so that only the root user can read and update it:

```
# chmod 600 /etc/NetworkManager/system-connections/output.nmconnection
# chown root:root /etc/NetworkManager/system-connections/output.nmconnection
```

3. Start the NetworkManager service:

```
# systemctl start NetworkManager.service
```

4. If you set the autoconnect variable in the profile to false, activate the connection:

```
# nmcli connection up Example-Connection
```

Verification

1. Verify that the NetworkManager service is running:

```
# systemctl status NetworkManager.service
● NetworkManager.service - Network Manager
  Loaded: loaded (/usr/lib/systemd/system/NetworkManager.service; enabled; vendor preset: enabled)
  Active: active (running) since Wed 2022-08-03 13:08:32 CEST; 1min 40s ago
...
```

2. Verify that NetworkManager can read the profile from the configuration file:

```
# nmcli -f TYPE,FILENAME,NAME connection
TYPE      FILENAME                                                    NAME
ethernet  /etc/NetworkManager/system-connections/output.nmconnection Example-Connection
ethernet  /etc/sysconfig/network-scripts/ifcfg-enp1s0                 enp1s0
...
```

If the output does not show the newly created connection, verify that the keyfile permissions and the syntax you used are correct.

3. Display the connection profile:

```
# nmcli connection show Example-Connection
connection.id:                          Example-Connection
connection.uuid:                        232290ce-5225-422a-9228-cb83b22056b4
connection.stable-id:                   --
connection.type:                        802-3-ethernet
connection.interface-name:              --
connection.autoconnect:                 yes
...
```

Additional resources

- nmcli(1)
- nm-settings-keyfile(5)
37.3. MANUALLY CREATING A NETWORKMANAGER PROFILE IN KEYFILE FORMAT

You can manually create a NetworkManager connection profile in keyfile format.

**NOTE**

Manually creating or updating the configuration files can result in an unexpected or non-functional network configuration. As an alternative, you can use `nmcli` in offline mode.

See [Using nmcli to create keyfile connection profiles in offline mode](#).

**Procedure**

1. If you create a profile for a hardware interface, such as Ethernet, display the MAC address of this interface:

   ```
   # ip address show enp1s0
   2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
      link/ether 00:53:00:8f:fa:66 brd ff:ff:ff:ff:ff:ff
   
   2: enp1s0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default qlen 1000
      link/ether 00:53:00:8f:fa:66 brd ff:ff:ff:ff:ff:ff
   ```

2. Create a connection profile. For example, for a connection profile of an Ethernet device that uses DHCP, create the `/etc/NetworkManager/system-connections/example.nmconnection` file with the following content:

   ```
   [connection]
   id=example_connection
   type=ethernet
   autoconnect=true

   [ipv4]
   method=auto

   [ipv6]
   method=auto

   [ethernet]
   mac-address=00:53:00:8f:fa:66
   ```
NOTE

You can use any file name with a .nmconnection suffix. However, when you later use nmcli commands to manage the connection, you must use the connection name set in the id variable when you refer to this connection. When you omit the id variable, use the file name without the .nmconnection to refer to this connection.

3. Set permissions on the configuration file so that only the root user can read and update it:

```bash
# chown root:root /etc/NetworkManager/system-connections/example.nmconnection
# chmod 600 /etc/NetworkManager/system-connections/example.nmconnection
```

4. Reload the connection profiles:

```bash
# nmcli connection reload
```

5. Verify that NetworkManager read the profile from the configuration file:

```bash
# nmcli -f NAME,UUID,FILENAME connection
NAME                UUID                                  FILENAME
example-connection  86da2486-068d-4d05-9ac7-957ec118afba
/etc/NetworkManager/system-connections/example.nmconnection
```

If the command does not show the newly added connection, verify that the file permissions and the syntax you used in the file are correct.

6. If you set the autoconnect variable in the profile to false, activate the connection:

```bash
# nmcli connection up example_connection
```

Verification

1. Display the connection profile:

```bash
# nmcli connection show example_connection
```

Additional resources

- nm-settings-keyfile (5)

37.4. THE DIFFERENCES IN INTERFACE RENAMING WITH PROFILES IN IF CFG AND KEYFILE FORMAT

You can define custom network interface names, such as provider or lan to make interface names more descriptive. In this case, the udev service renames the interfaces. The renaming process works differently depending on whether you use connection profiles in ifcfg or keyfile format.

The interface renaming process when using a profile ifcfg format

The helper utility searches for the `HWADDR` parameter in `/etc/sysconfig/network-scripts/ifcfg-*` files.

If the value set in the variable matches the MAC address of an interface, the helper utility renames the interface to the name set in the `DEVICE` parameter of the file.

The interface renaming process when using a profile in keyfile format

1. Create a `systemd` link file or a `udev` rule to rename an interface.
2. Use the custom interface name in the `interface-name` property of a NetworkManager connection profile.

Additional resources

- How the udev device manager renames network interfaces
- Configuring user-defined network interface names by using udev rules
- Configuring user-defined network interface names by using systemd link files

### 37.5. MIGRATING NETWORKMANAGER PROFILES FROM IFCFG TO KEYFILE FORMAT

If you still use connection profiles in the deprecated `ifcfg` format, you can convert them to the keyfile format.

**NOTE**

If an `ifcfg` file contains the `NM_CONTROLLED=no` setting, NetworkManager does not control this profile and, consequently the migration process ignores it.

**Prerequisites**

- You have connection profiles in `ifcfg` format in the `/etc/sysconfig/network-scripts/` directory.
- If the connection profiles contain a `DEVICE` variable that is set to a custom device name, such as `provider` or `lan`, you created a `systemd` link file or a `udev` rule for each of the custom device names.

**Procedure**

- Migrate the connection profiles:

```bash
# nmcli connection migrate
Connection 'enp1s0' (43ed18ab-f0c4-4934-af3d-2b3333948e45) successfully migrated.
Connection 'enp2s0' (883333e8-1b87-4947-8ceb-1f8812a80a9b) successfully migrated.
...
```

**Verification**

- Optionally, you can verify that you successfully migrated all your connection profiles:
Additional resources

- nm-settings-keyfile(5)
- nm-settings-ifcfg-rh(5)
- How the udev device manager renames network interfaces
CHAPTER 38. SYSTEMD NETWORK TARGETS AND SERVICES

NetworkManager configures the network during the system boot process. However, when booting with a remote root (/), such as if the root directory is stored on an iSCSI device, the network settings are applied in the initial RAM disk (initrd) before RHEL is started. For example, if the network configuration is specified on the kernel command line using `rd.neednet=1` or a configuration is specified to mount remote file systems, then the network settings are applied on initrd.

RHEL uses the network and network-online targets and the NetworkManager-wait-online service while applying network settings. Also, you can configure systemd services to start after the network is fully available if these services cannot dynamically reload.

38.1. DIFFERENCES BETWEEN THE NETWORK AND NETWORK-ONLINE SYSTEMD TARGET

Systemd maintains the network and network-online target units. The special units such as NetworkManager-wait-online.service, have WantedBy=network-online.target and Before=network-online.target parameters. If enabled, these units get started with network-online.target and delay the target to be reached until some form of network connectivity is established. They delay the network-online target until the network is connected.

The network-online target starts a service, which adds substantial delays to further execution. Systemd automatically adds dependencies with Wants and After parameters for this target unit to all the System V (SysV) init script service units with a Linux Standard Base (LSB) header referring to the $network facility. The LSB header is metadata for init scripts. You can use it to specify dependencies. This is similar to the systemd target.

The network target does not significantly delay the execution of the boot process. Reaching the network target means that the service that is responsible for setting up the network has started. However, it does not mean that a network device was configured. This target is important during the shutdown of the system. For example, if you have a service that was ordered after the network target during bootup, then this dependency is reversed during the shutdown. The network does not get disconnected until your service has been stopped. All mount units for remote network file systems automatically start the network-online target unit and order themselves after it.

NOTE

The network-online target unit is only useful during the system starts. After the system has completed booting up, this target does not track the online state of the network. Therefore, you cannot use network-online to monitor the network connection. This target provides a one–time system startup concept.

38.2. OVERVIEW OF NETWORKMANAGER-WAIT-ONLINE

The NetworkManager-wait-online service waits with a timeout for the network to be configured. This network configuration involves plugging-in an Ethernet device, scanning for a Wi-Fi device, and so forth. NetworkManager automatically activates suitable profiles that are configured to start automatically. The failure of the automatic activation process due to a DHCP timeout or similar event might keep NetworkManager busy for an extended period of time. Depending on the configuration, NetworkManager retries activating the same profile or a different profile.

When the startup completes, either all profiles are in a disconnected state or are successfully activated. You can configure profiles to auto-connect. The following are a few examples of parameters that set timeouts or define when the connection is considered active:
- **connection.wait-device-timeout** - sets the timeout for the driver to detect the device
- **ipv4.may-fail** and **ipv6.may-fail** - sets activation with one IP address family ready, or whether a particular address family must have completed configuration.
- **ipv4.gateway-ping-timeout** - delays activation.

Additional resources
- [nm-settings(5)](man) man page

### 38.3. Configuring a Systemd Service to Start After the Network Has Been Started

Red Hat Enterprise Linux installs *systemd* service files in the `/usr/lib/systemd/system/` directory. This procedure creates a drop-in snippet for a service file in `/etc/systemd/system/service_name.service.d/` that is used together with the service file in `/usr/lib/systemd/system/` to start a particular service after the network is online. It has a higher priority if settings in the drop-in snippet overlap with the ones in the service file in `/usr/lib/systemd/system/`.

#### Procedure

1. To open the service file in the editor, enter:

   ```
   # systemctl edit service_name
   ```

2. Enter the following, and save the changes:

   ```
   [Unit]
   After=network-online.target
   ```

3. Reload the *systemd* service.

   ```
   # systemctl daemon-reload
   ```
CHAPTER 39. INTRODUCTION TO NMSTATE

CHAPTER 39. INTRODUCTION TO NMSTATE
Nmstate is a declarative network manager API. The nmstate package provides the libnmstate Python
library and a command-line utility, nmstatectl, to manage NetworkManager on RHEL. When you use
Nmstate, you describe the expected networking state using YAML or JSON-formatted instructions.
Nmstate has many benefits. For example, it:
Provides a stable and extensible interface to manage RHEL network capabilities
Supports atomic and transactional operations at the host and cluster level
Supports partial editing of most properties and preserves existing settings that are not specified
in the instructions
Provides plug-in support to enable administrators to use their own plug-ins

39.1. USING THE LIBNMSTATE LIBRARY IN A PYTHON APPLICATION
The libnmstate Python library enables developers to use Nmstate in their own application
To use the library, import it in your source code:
import libnmstate
Note that you must install the nmstate package to use this library.
Example 39.1. Querying the network state using the libnmstate library
The following Python code imports the libnmstate library and displays the available network
interfaces and their state:
import json
import libnmstate
from libnmstate.schema import Interface
net_state = libnmstate.show()
for iface_state in net_state[Interface.KEY]:
print(iface_state[Interface.NAME] + ": "
+ iface_state[Interface.STATE])

39.2. UPDATING THE CURRENT NETWORK CONFIGURATION USING
NMSTATECTL
You can use the nmstatectl utility to store the current network configuration of one or all interfaces in a
file. You can then use this file to:
Modify the configuration and apply it to the same system.
Copy the file to a different host and configure the host with the same or modified settings.

For example, you can export the settings of the enp1s0 interface to a file, modify the configuration, and

331


For example, you can export the settings of the `enp1s0` interface to a file, modify the configuration, and apply the settings to the host.

**Prerequisites**

- The **nmstate** package is installed.

**Procedure**

1. Export the settings of the `enp1s0` interface to the `~/network-config.yml` file:

   ```
   # nmstatectl show enp1s0 > ~/network-config.yml
   ```

   This command stores the configuration of `enp1s0` in YAML format. To store the output in JSON format, pass the `--json` option to the command.

   If you do not specify an interface name, `nmstatectl` exports the configuration of all interfaces.

2. Modify the `~/network-config.yml` file using a text editor to update the configuration.

3. Apply the settings from the `~/network-config.yml` file:

   ```
   # nmstatectl apply ~/network-config.yml
   ```

   If you exported the settings in JSON format, pass the `--json` option to the command.

### 39.3. THE NMSTATE SYSTEMD SERVICE

You can automatically apply new network settings when the Red Hat Enterprise Linux system boots by configuring the **nmstate systemd** service.

With the **nmstate** package installed, you can store `.yml` files with Nmstate instructions in the `/etc/nmstate/` directory. The **nmstate** service then automatically applies the files on the next reboot or when you manually restart the service. After Nmstate successfully applies a file, it renames the file’s `.yml` suffix to `.applied` to prevent the service from processing the same file again.

The **nmstate** service is a **oneshot systemd** service. Consequently, **systemd** executes it only when the system boots and when you manually restart the service.

**NOTE**

By default, the **nmstate** service is disabled. Use the **systemctl enable nmstate** command to enable it. Afterwards, **systemd** executes this service each time when the system starts.

### 39.4. NETWORK STATES FOR THE NETWORK RHEL SYSTEM ROLE

The **network** RHEL system role supports state configurations in playbooks to configure the devices. For this, use the **network_state** variable followed by the state configurations.

Benefits of using the **network_state** variable in a playbook:

- Using the declarative method with the state configurations, you can configure interfaces, and the NetworkManager creates a profile for these interfaces in the background.
With the `network_state` variable, you can specify the options that you require to change, and all the other options will remain the same as they are. However, with the `network_connections` variable, you must specify all settings to change the network connection profile.

For example, to create an Ethernet connection with dynamic IP address settings, use the following `vars` block in your playbook:

```
vars:
  network_state:
    interfaces:
      - name: enp7s0
        type: ethernet
        state: up
        ipv4:
          enabled: true
          auto-dns: true
          auto-gateway: true
          auto-routes: true
          dhcp: true
        ipv6:
          enabled: true
          auto-dns: true
          auto-gateway: true
          auto-routes: true
          autoconf: true
          dhcp: true
```

For example, to only change the connection status of dynamic IP address settings that you created as above, use the following `vars` block in your playbook:

```
vars:
  network_state:
    interfaces:
      - name: enp7s0
        type: ethernet
        state: down
```

```yaml
vars:
  network_connections:
    - name: enp7s0
      interface_name: enp7s0
      type: ethernet
      autoconnect: yes
      ip:
        dhcp4: yes
        auto6: yes
      state: up
```

### Additional resources

- `/usr/share/ansible/roles/rhel-system-roles.network/README.md` file
39.5. ADDITIONAL RESOURCES

- /usr/share/doc/rhel-system-roles/network/ directory
- /usr/share/doc/nmstate/README.md
- /usr/share/doc/nmstate/examples/
CHAPTER 40. CAPTURING NETWORK PACKETS

To debug network issues and communications, you can capture network packets. The following sections provide instructions and additional information about capturing network packets.

40.1. USING XDPDUMP TO CAPTURE NETWORK PACKETS INCLUDING PACKETS DROPPED BY XDP PROGRAMS

The `xdpdump` utility captures network packets. Unlike the `tcpdump` utility, `xdpdump` uses an extended Berkeley Packet Filter (eBPF) program for this task. This enables `xdpdump` to also capture packets dropped by Express Data Path (XDP) programs. User-space utilities, such as `tcpdump`, are not able to capture these dropped packages, as well as original packets modified by an XDP program.

You can use `xdpdump` to debug XDP programs that are already attached to an interface. Therefore, the utility can capture packets before an XDP program is started and after it has finished. In the latter case, `xdpdump` also captures the XDP action. By default, `xdpdump` captures incoming packets at the entry of the XDP program.

**IMPORTANT**

On other architectures than AMD and Intel 64-bit, the `xdpdump` utility is provided as a Technology Preview only. Technology Preview features are not supported with Red Hat production Service Level Agreements (SLAs), might not be functionally complete, and Red Hat does not recommend using them for production. These previews provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.

See [Technology Preview Features Support Scope](#) on the Red Hat Customer Portal for information about the support scope for Technology Preview features.

Note that `xdpdump` has no packet filter or decode capabilities. However, you can use it in combination with `tcpdump` for packet decoding.

**Prerequisites**

- A network driver that supports XDP programs.
- An XDP program is loaded to the `enp1s0` interface. If no program is loaded, `xdpdump` captures packets in a similar way `tcpdump` does, for backward compatibility.

**Procedure**

1. To capture packets on the `enp1s0` interface and write them to the `/root/capture.pcap` file, enter:

   ```
   # xdpdump -i enp1s0 -w /root/capture.pcap
   ```

2. To stop capturing packets, press **Ctrl+C**.

**Additional resources**

- `xdpdump(8)` man page
If you are a developer and you are interested in the source code of `xdpdump`, download and install the corresponding source RPM (SRPM) from the Red Hat Customer Portal.

### 40.2. ADDITIONAL RESOURCES

- How to capture network packets with `tcpdump`?
CHAPTER 41. UNDERSTANDING THE EBPF NETWORKING FEATURES IN RHEL 9

The extended Berkeley Packet Filter (eBPF) is an in-kernel virtual machine that allows code execution in the kernel space. This code runs in a restricted sandbox environment with access only to a limited set of functions.

In networking, you can use eBPF to complement or replace kernel packet processing. Depending on the hook you use, eBPF programs have, for example:

- Read and write access to packet data and metadata
- Can look up sockets and routes
- Can set socket options
- Can redirect packets

41.1. OVERVIEW OF NETWORKING EBPF FEATURES IN RHEL 9

You can attach extended Berkeley Packet Filter (eBPF) networking programs to the following hooks in RHEL:

- eXpress Data Path (XDP): Provides early access to received packets before the kernel networking stack processes them.
- tc eBPF classifier with direct-action flag: Provides powerful packet processing on ingress and egress.
- Control Groups version 2 (cgroup v2): Enables filtering and overriding socket-based operations performed by programs in a control group.
- Socket filtering: Enables filtering of packets received from sockets. This feature was also available in the classic Berkeley Packet Filter (cBPF), but has been extended to support eBPF programs.
- Stream parser: Enables splitting up streams to individual messages, filtering, and redirecting them to sockets.
- SO_REUSEPORT socket selection: Provides a programmable selection of a receiving socket from a reuseport socket group.
- Flow dissector: Enables overriding the way the kernel parses packet headers in certain situations.
- TCP congestion control callbacks: Enables implementing a custom TCP congestion control algorithm.
- Routes with encapsulation: Enables creating custom tunnel encapsulation.

XDP

You can attach programs of the BPF_PROG_TYPE_XDP type to a network interface. The kernel then executes the program on received packets before the kernel network stack starts processing them. This allows fast packet forwarding in certain situations, such as fast packet dropping to prevent distributed denial of service (DDoS) attacks and fast packet redirects for load balancing scenarios.
You can also use XDP for different forms of packet monitoring and sampling. The kernel allows XDP programs to modify packets and to pass them for further processing to the kernel network stack.

The following XDP modes are available:

- **Native (driver) XDP**: The kernel executes the program from the earliest possible point during packet reception. At this moment, the kernel did not parse the packet and, therefore, no metadata provided by the kernel is available. This mode requires that the network interface driver supports XDP but not all drivers support this native mode.

- **Generic XDP**: The kernel network stack executes the XDP program early in the processing. At that time, kernel data structures have been allocated, and the packet has been pre-processed. If a packet should be dropped or redirected, it requires a significant overhead compared to the native mode. However, the generic mode does not require network interface driver support and works with all network interfaces.

- **Offloaded XDP**: The kernel executes the XDP program on the network interface instead of on the host CPU. Note that this requires specific hardware, and only certain eBPF features are available in this mode.

On RHEL, load all XDP programs using the `libxdp` library. This library enables system-controlled usage of XDP.

**NOTE**
Currently, there are some system configuration limitations for XDP programs. For example, you must disable certain hardware offload features on the receiving interface. Additionally, not all features are available with all drivers that support the native mode.

In RHEL 9, Red Hat supports the XDP features only if you use the `libxdp` library to load the program into the kernel.

**AF_XDP**
Using an XDP program that filters and redirects packets to a given AF_XDP socket, you can use one or more sockets from the AF_XDP protocol family to quickly copy packets from the kernel to the user space.

**Traffic Control**
The Traffic Control (tc) subsystem offers the following types of eBPF programs:

- **BPF_PROG_TYPE_SCHED_CLS**
- **BPF_PROG_TYPE_SCHED_ACT**

These types enable you to write custom tc classifiers and tc actions in eBPF. Together with the parts of the tc ecosystem, this provides the ability for powerful packet processing and is the core part of several container networking orchestration solutions.

In most cases, only the classifier is used, as with the direct-action flag, the eBPF classifier can execute actions directly from the same eBPF program. The clsact Queueing Discipline (qdisc) has been designed to enable this on the ingress side.

Note that using a flow dissector eBPF program can influence operation of some other qdiscs and tc classifiers, such as flower.

**Socket filter**
Several utilities use or have used the classic Berkeley Packet Filter (cBPF) for filtering packets received on a socket. For example, the `tcpdump` utility enables the user to specify expressions, which `tcpdump` then translates into cBPF code.

As an alternative to cBPF, the kernel allows eBPF programs of the `BPF_PROG_TYPE_SOCKET_FILTER` type for the same purpose.

Control Groups
In RHEL, you can use multiple types of eBPF programs that you can attach to a cgroup. The kernel executes these programs when a program in the given cgroup performs an operation. Note that you can use only cgroups version 2.

The following networking-related cgroup eBPF programs are available in RHEL:

- **BPF_PROG_TYPE_SOCK_OPS**: The kernel calls this program on various TCP events. The program can adjust the behavior of the kernel TCP stack, including custom TCP header options, and so on.

- **BPF_PROG_TYPE_CGROUP_SOCK_ADDR**: The kernel calls this program during `connect`, `bind`, `sendto`, `recvmsg`, `getpeername`, and `gethostname` operations. This program allows changing IP addresses and ports. This is useful when you implement socket-based network address translation (NAT) in eBPF.

- **BPF_PROG_TYPE_CGROUP_SOCKOPT**: The kernel calls this program during `setsockopt` and `getsockopt` operations and allows changing the options.

- **BPF_PROG_TYPE_CGROUP_SOCK**: The kernel calls this program during socket creation, socket releasing, and binding to addresses. You can use these programs to allow or deny the operation, or only to inspect socket creation for statistics.

- **BPF_PROG_TYPE_CGROUP_SKB**: This program filters individual packets on ingress and egress, and can accept or reject packets.

- **BPF_PROG_TYPE_CGROUP_SYSCTL**: This program allows filtering of access to system controls (`sysctl`).

Stream Parser
A stream parser operates on a group of sockets that are added to a special eBPF map. The eBPF program then processes packets that the kernel receives or sends on those sockets.

The following stream parser eBPF programs are available in RHEL:

- **BPF_PROG_TYPE_SK_SKB**: An eBPF program parses packets received from the socket into individual messages, and instructs the kernel to drop those messages or send them to another socket in the group.

- **BPF_PROG_TYPE_SK_MSG**: This program filters egress messages. An eBPF program parses the packets into individual messages and either approves or rejects them.

SO_REUSEPORT socket selection
Using this socket option, you can bind multiple sockets to the same IP address and port. Without eBPF, the kernel selects the receiving socket based on a connection hash. With the `BPF_PROG_TYPE_SK_REUSEPORT` program, the selection of the receiving socket is fully programmable.

Flow dissector
When the kernel needs to process packet headers without going through the full protocol decode, they
are dissected. For example, this happens in the tc subsystem, in multipath routing, in bonding, or when calculating a packet hash. In this situation the kernel parses the packet headers and fills internal structures with the information from the packet headers. You can replace this internal parsing using the BPF_PROG_TYPE_FLOW_DISSECTOR program. Note that you can only dissect TCP and UDP over IPv4 and IPv6 in eBPF in RHEL.

TCP Congestion Control
You can write a custom TCP congestion control algorithm using a group of BPF_PROG_TYPE_STRUCT_OPS programs that implement struct tcp_congestion_oops callbacks. An algorithm that is implemented this way is available to the system alongside the built-in kernel algorithms.

Routes with encapsulation
You can attach one of the following eBPF program types to routes in the routing table as a tunnel encapsulation attribute:

- BPF_PROG_TYPE_LWT_IN
- BPF_PROG_TYPE_LWT_OUT
- BPF_PROG_TYPE_LWT_XMIT

The functionality of such an eBPF program is limited to specific tunnel configurations and does not allow creating a generic encapsulation or decapsulation solution.

Socket lookup
To bypass limitations of the bind system call, use an eBPF program of the BPF_PROG_TYPE_SK_LOOKUP type. Such programs can select a listening socket for new incoming TCP connections or an unconnected socket for UDP packets.

41.2. OVERVIEW OF XDP FEATURES IN RHEL 9 BY NETWORK CARDS

The following is an overview of XDP–enabled network cards and the XDP features you can use with them:

<table>
<thead>
<tr>
<th>Network card</th>
<th>Driver</th>
<th>Basic Redirect Target HW offlow Zero- copy Larg e MTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon Elastic Network Adapter</td>
<td>ena</td>
<td>yes   yes   [a] no   no   no</td>
</tr>
<tr>
<td>aQuantia AQtion Ethernet card</td>
<td>atlantic</td>
<td>yes   yes   no   no   no   no</td>
</tr>
<tr>
<td>Broadcom NetXtreme–C/E 10/25/40/50 gigabit Ethernet</td>
<td>bnxt_en</td>
<td>yes   yes   yes   no   no   yes</td>
</tr>
<tr>
<td>Cavium Thunder Virtual function</td>
<td>nicvf</td>
<td>yes   no    no   no   no   no</td>
</tr>
<tr>
<td>Google Virtual NIC (gVNIC) support</td>
<td>gve</td>
<td>yes   yes   yes   no   yes   no</td>
</tr>
</tbody>
</table>

[a] Indicates functionality that is experimental.
<table>
<thead>
<tr>
<th>Network card</th>
<th>Driver</th>
<th>Basic</th>
<th>Redirect</th>
<th>Target</th>
<th>HW offload</th>
<th>Zero copy</th>
<th>Large MTU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intel® 10GbE PCI Express Virtual Function Ethernet</td>
<td>ixgbevf</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Intel® 10GbE PCI Express adapters</td>
<td>ixgbe</td>
<td>yes</td>
<td>yes</td>
<td>yes [a]</td>
<td>no</td>
<td>yes</td>
<td>yes [b]</td>
</tr>
<tr>
<td>Intel® Ethernet Connection E800 Series</td>
<td>ice</td>
<td>yes</td>
<td>yes</td>
<td>yes [a]</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Intel® Ethernet Controller I225-LM/I225-V family</td>
<td>igc</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes [b]</td>
</tr>
<tr>
<td>Intel® PCI Express Gigabit adapters</td>
<td>igb</td>
<td>yes</td>
<td>yes</td>
<td>yes [a]</td>
<td>no</td>
<td>no</td>
<td>yes [b]</td>
</tr>
<tr>
<td>Intel® Ethernet Controller XL710 Family</td>
<td>i40e</td>
<td>yes</td>
<td>yes</td>
<td>yes [a][c]</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Marvell OcteonTX2</td>
<td>rvu_nicpf</td>
<td>yes</td>
<td>yes</td>
<td>yes [a][c]</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Mellanox 5th generation network adapters (ConnectX series)</td>
<td>mlx5_core</td>
<td>yes</td>
<td>yes</td>
<td>yes [c]</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Mellanox Technologies 1/10/40Gbit Ethernet</td>
<td>mlx4_en</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Microsoft Azure Network Adapter</td>
<td>mana</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Microsoft Hyper-V virtual network</td>
<td>hv_netvsc</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Netronome® NFP4000/NFP6000 NIC [d]</td>
<td>nfp</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>QEMU Virtio network</td>
<td>virtio_net</td>
<td>yes</td>
<td>yes</td>
<td>yes [a]</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>QLogic QED 25/40/100Gb Ethernet NIC</td>
<td>qede</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>STMicroelectronics Multi-Gigabit Ethernet</td>
<td>stmmac</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Network card</td>
<td>Driver</td>
<td>Basic</td>
<td>Redirect</td>
<td>Target</td>
<td>HW offload</td>
<td>Zero-copy</td>
<td>Large MTU</td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>-------</td>
<td>----------</td>
<td>--------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>Solarflare</td>
<td>sfc</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>SFC9000/SFC9100/EF100-family</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Universal TUN/TAP device</td>
<td>tun</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Virtual Ethernet pair device</td>
<td>veth</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>VMware VMXNET3 ethernet driver</td>
<td>vmxnet3</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Xen paravirtual network device</td>
<td>xen-netfront</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>no</td>
<td>no</td>
</tr>
</tbody>
</table>

[a] Only if an XDP program is loaded on the interface.

[b] Transmitting side only. Cannot receive large packets through XDP.

[c] Requires several XDP TX queues allocated that is larger or equal to the largest CPU index.

[d] Some of the listed features are not available for the Netronome® NFP3800 NIC.

Legend:

- Basic: Supports basic return codes: **DROP**, **PASS**, **ABORTED**, and **TX**.
- Redirect: Supports the **XDP_REDIRECT** return code.
- Target: Can be a target of a **XDP_REDIRECT** return code.
- HW offload: Supports XDP hardware offload.
- Zero-copy: Supports the zero-copy mode for the **AF_XDP** protocol family.
- Large MTU: Supports packets larger than page size.
CHAPTER 42. NETWORK TRACING USING THE BPF COMPILER COLLECTION

BPF Compiler Collection (BCC) is a library, which facilitates the creation of the extended Berkeley Packet Filter (eBPF) programs. The main utility of eBPF programs is analyzing the operating system performance and network performance without experiencing overhead or security issues.

BCC removes the need for users to know deep technical details of eBPF, and provides many out-of-the-box starting points, such as the bcc-tools package with pre-created eBPF programs.

NOTE

The eBPF programs are triggered on events, such as disk I/O, TCP connections, and process creations. It is unlikely that the programs should cause the kernel to crash, loop or become unresponsive because they run in a safe virtual machine in the kernel.

42.1. INSTALLING THE BCC-TOOLS PACKAGE

Install the bcc-tools package, which also installs the BPF Compiler Collection (BCC) library as a dependency.

Procedure

1. Install bcc-tools.

   ```
   # dnf install bcc-tools
   ```

   The BCC tools are installed in the `/usr/share/bcc/tools/` directory.

2. Optionally, inspect the tools:

   ```
   # ll /usr/share/bcc/tools/
   ...
   -rwxr-xr-x. 1 root root  4198 Dec 14 17:53 dcsnoop
   -rwxr-xr-x. 1 root root  3931 Dec 14 17:53 dcstat
   -rwxr-xr-x. 1 root root 20040 Dec 14 17:53 deadlock_detector
   -rw-r--r--. 1 root root  7105 Dec 14 17:53 deadlock_detector.c
   drwxr-xr-x. 3 root root  8192 Mar 11 10:28 doc
   -rwxr-xr-x. 1 root root 7588 Dec 14 17:53 execsnoop
   -rwxr-xr-x. 1 root root  6373 Dec 14 17:53 ext4dist
   -rwxr-xr-x. 1 root root 10401 Dec 14 17:53 ext4slower
   ...
   ```

   The doc directory in the listing above contains documentation for each tool.

42.2. DISPLAYING TCP CONNECTIONS ADDED TO THE KERNEL’S ACCEPT QUEUE

After the kernel receives the ACK packet in a TCP 3-way handshake, the kernel moves the connection from the SYN queue to the accept queue after the connection’s state changes to ESTABLISHED. Therefore, only successful TCP connections are visible in this queue.

The tcpaccept utility uses eBPF features to display all connections the kernel adds to the accept
The utility is lightweight because it traces the `accept()` function of the kernel instead of capturing packets and filtering them. For example, use `tcpaccept` for general troubleshooting to display new connections the server has accepted.

**Procedure**

1. Enter the following command to start the tracing the kernel `accept` queue:

   ```bash
   # /usr/share/bcc/tools/tcpaccept
   PID  COMM     IP   RADDR       RPORT   LADDR     LPORT
   843  sshd      4  192.0.2.17 50598  192.0.2.1  22
   1107 ns-slapd   4  198.51.100.6 38772  192.0.2.1  389
   1107 ns-slapd   4  203.0.113.85 38774  192.0.2.1  389
   ...
   ```

   Each time the kernel accepts a connection, `tcpaccept` displays the details of the connections.

2. Press `Ctrl+C` to stop the tracing process.

**Additional resources**

- `tcpaccept(8)` man page
- `/usr/share/bcc/tools/doc/tcpaccept_example.txt` file

## 42.3. TRACING OUTGOING TCP CONNECTION ATTEMPTS

The `tcpconnect` utility uses eBPF features to trace outgoing TCP connection attempts. The output of the utility also includes connections that failed.

The `tcpconnect` utility is lightweight because it traces, for example, the `connect()` function of the kernel instead of capturing packets and filtering them.

**Procedure**

1. Enter the following command to start the tracing process that displays all outgoing connections:

   ```bash
   # /usr/share/bcc/tools/tcpconnect
   PID  COMM       IP   SADDR      DADDR       DPORT
   31346 curl      4  192.0.2.1  198.51.100.16  80
   31348 telnet    4  192.0.2.1  203.0.113.231  23
   31361 isc-worker00 4  192.0.2.1  192.0.2.254  53
   ...
   ```

   Each time the kernel processes an outgoing connection, `tcpconnect` displays the details of the connections.

2. Press `Ctrl+C` to stop the tracing process.

**Additional resources**

- `tcpconnect(8)` man page
- `/usr/share/bcc/tools/doc/tcpconnect_example.txt` file
42.4. MEASURING THE LATENCY OF OUTGOING TCP CONNECTIONS

The TCP connection latency is the time taken to establish a connection. This typically involves the kernel TCP/IP processing and network round trip time, and not the application runtime.

The `tcpconnlat` utility uses eBPF features to measure the time between a sent SYN packet and the received response packet.

**Procedure**

1. Start measuring the latency of outgoing connections:

   ```
   # /usr/share/bcc/tools/tcpconnlat
   PID    COMM         IP SADDR      DADDR          DPORT LAT(ms)
   32151  isc-worker00 4  192.0.2.1  192.0.2.254    53    0.60
   32155  ssh          4  192.0.2.1  203.0.113.190  22    26.34
   32319  curl         4  192.0.2.1  198.51.100.59  443   188.96
   ...  
   ```

   Each time the kernel processes an outgoing connection, `tcpconnlat` displays the details of the connection after the kernel receives the response packet.

2. Press Ctrl+C to stop the tracing process.

**Additional resources**

- `tcpconnlat(8)` man page
- `/usr/share/bcc/tools/doc/tcpconnlat_example.txt` file

42.5. DISPLAYING DETAILS ABOUT TCP PACKETS AND SEGMENTS THAT WERE DROPPED BY THE KERNEL

The `tcpdrop` utility enables administrators to display details about TCP packets and segments that were dropped by the kernel. Use this utility to debug high rates of dropped packets that can cause the remote system to send timer-based retransmits. High rates of dropped packets and segments can impact the performance of a server.

Instead of capturing and filtering packets, which is resource-intensive, the `tcpdrop` utility uses eBPF features to retrieve the information directly from the kernel.

**Procedure**

1. Enter the following command to start displaying details about dropped TCP packets and segments:

   ```
   # /usr/share/bcc/tools/tcpdrop
   TIME     PID    IP SADDR:SPORT       > DADDR:DPORT   STATE (FLAGS)
   13:28:39 32253  4  192.0.2.85:51616  > 192.0.2.1:22  CLOSE_WAIT (FIN|ACK)
   b'tcp_drop+0x1'
   b'tcp_data_queue+0x2b9'
   ...  
   13:28:39 1      4  192.0.2.85:51616  > 192.0.2.1:22  CLOSE (ACK)
   ```
b'tcp_drop+0x1'
b'tcp_rcv_state_process+0xe2'
...

Each time the kernel drops TCP packets and segments, tcpdrop displays the details of the connection, including the kernel stack trace that led to the dropped package.

2. Press **Ctrl+C** to stop the tracing process.

Additional resources

- tcpdrop(8) man page
- /usr/share/bcc/tools/doc/tcpdrop_example.txt file

### 42.6. TRACING TCP SESSIONS

The tcplife utility uses eBPF to trace TCP sessions that open and close, and prints a line of output to summarize each one. Administrators can use tcplife to identify connections and the amount of transferred traffic.

For example, you can display connections to port **22** (SSH) to retrieve the following information:

- The local process ID (PID)
- The local process name
- The local IP address and port number
- The remote IP address and port number
- The amount of received and transmitted traffic in KB.
- The time in milliseconds the connection was active

**Procedure**

1. Enter the following command to start the tracing of connections to the local port **22**:

   ```
   /usr/share/bcc/tools/tcplife -L 22
   PID   COMM    LADDR      LPORT RADDR       RPORT TX_KB  RX_KB      MS
   19392 sshd    192.0.2.1  22    192.0.2.17  43892    53     52 6681.95
   19431 sshd    192.0.2.1  22    192.0.2.245 43902    81 249381 7585.09
   19487 sshd    192.0.2.1  22    192.0.2.121 43970  6998     7 16740.35
   ...
   ```

   Each time a connection is closed, tcplife displays the details of the connections.

2. Press **Ctrl+C** to stop the tracing process.

Additional resources

- tcplife(8) man page
- /usr/share/bcc/tools/doc/tcplife_example.txt file
42.7. TRACING TCP RETRANSMISSIONS

The `tcpretrans` utility displays details about TCP retransmissions, such as the local and remote IP address and port number, as well as the TCP state at the time of the retransmissions.

The utility uses eBPF features and, therefore, has a very low overhead.

**Procedure**

1. Use the following command to start displaying TCP retransmission details:

   ```
   # /usr/share/bcc/tools/tcpretrans
   TIME    PID  IP  LADDR:LPORT   RADDR:RPORT   STATE
   00:23:02 0    4  192.0.2.1:22  R> 198.51.100.0:26788  ESTABLISHED
   00:23:02 0    4  192.0.2.1:22  R> 198.51.100.0:26788  ESTABLISHED
   00:45:43 0    4  192.0.2.1:22  R> 198.51.100.0:17634  ESTABLISHED
   ...
   
   Each time the kernel calls the TCP retransmit function, `tcpretrans` displays the details of the connection.
   
   2. Press **Ctrl+C** to stop the tracing process.

**Additional resources**

- `tcpretrans(8)` man page
- `/usr/share/bcc/tools/doc/tcpretrans_example.txt` file

42.8. DISPLAYING TCP STATE CHANGE INFORMATION

During a TCP session, the TCP state changes. The `tcpstates` utility uses eBPF functions to trace these state changes, and prints details including the duration in each state. For example, use `tcpstates` to identify if connections spend too much time in the initialization state.

**Procedure**

1. Use the following command to start tracing TCP state changes:

   ```
   # /usr/share/bcc/tools/tcpstates
   SKADDR           C-PID C-COMM     LADDR     LPORT RADDR       RPORT OLDSTATE    ->
                   NEWSTATE   MS
   ffff9cd377b3af80 0     swapper/1  0.0.0.0   22    0.0.0.0     0     LISTEN      -> SYN_RECV
   0.000
   ffff9cd377b3af80 0     swapper/1  192.0.2.1 22    192.0.2.45  53152 SYN_RECV    ->
   ESTABLISHED 0.067
   ffff9cd377b3af80 818   sssd_nss   192.0.2.1 22    192.0.2.45  53152 ESTABLISHED ->
   CLOSE_WAIT 65636.773
   ffff9cd377b3af80 1432  sshd       192.0.2.1 22    192.0.2.45  53152 CLOSE_WAIT  ->
   LAST_ACK 24.409
   ffff9cd377b3af80 1267  pulseaudio 192.0.2.1 22    192.0.2.45  53152 LAST_ACK    ->
   CLOSE 0.376
   ...
   ```
Each time a connection changes its state, **tcpstates** displays a new line with updated connection details.

If multiple connections change their state at the same time, use the socket address in the first column (**SKADDR**) to determine which entries belong to the same connection.

2. Press **Ctrl+C** to stop the tracing process.

**Additional resources**

- **tcpstates**(8) man page
- **/usr/share/bcc/tools/doc/tcpstates_example.txt** file

### 42.9. SUMMARIZING AND AGGREGATING TCP TRAFFIC SENT TO SPECIFIC SUBNETS

The **tcpsubnet** utility summarizes and aggregates IPv4 TCP traffic that the local host sends to subnets and displays the output on a fixed interval. The utility uses eBPF features to collect and summarize the data to reduce the overhead.

By default, **tcpsubnet** summarizes traffic for the following subnets:

- **127.0.0.1/32**
- **10.0.0.0/8**
- **172.16.0.0/12**
- **192.0.2.0/24/16**
- **0.0.0.0/0**

Note that the last subnet (**0.0.0.0/0**) is a catch-all option. The **tcpsubnet** utility counts all traffic for subnets different than the first four in this catch-all entry.

Follow the procedure to count the traffic for the **192.0.2.0/24** and **198.51.100.0/24** subnets. Traffic to other subnets will be tracked in the **0.0.0.0/0** catch-all subnet entry.

**Procedure**

1. Start monitoring the amount of traffic send to the **192.0.2.0/24**, **198.51.100.0/24**, and other subnets:

   ```bash
   # /usr/share/bcc/tools/tcpsubnet 192.0.2.0/24,198.51.100.0/24,0.0.0.0/0
   Tracing... Output every 1 secs. Hit Ctrl-C to end
   [02/21/20 10:04:50]
   192.0.2.0/24  856
   198.51.100.0/24  7467
   [02/21/20 10:04:51]
   192.0.2.0/24  1200
   198.51.100.0/24  8763
   0.0.0.0/0  673
   ...
   ```
This command displays the traffic in bytes for the specified subnets once per second.

2. Press `Ctrl+C` to stop the tracing process.

Additional resources

- `tcpsubnet(8)` man page
- `/usr/share/bcc/tools/doc/tcpsubnet.txt` file

### 42.10. DISPLAYING THE NETWORK THROUGHPUT BY IP ADDRESS AND PORT

The `tcptop` utility displays TCP traffic the host sends and receives in kilobytes. The report automatically refreshes and contains only active TCP connections. The utility uses eBPF features and, therefore, has only a very low overhead.

**Procedure**

1. To monitor the sent and received traffic, enter:

```bash
#/usr/share/bcc/tools/tcptop
13:46:29 loadavg: 0.10 0.03 0.01 1/215 3875

PID    COMM         LADDR           RADDR              RX_KB   TX_KB
3853   3853         192.0.2.1:22    192.0.2.165:41838  32     102626
1285   sshd         192.0.2.1:22    192.0.2.45:39240   0           0
...```

The output of the command includes only active TCP connections. If the local or remote system closes a connection, the connection is no longer visible in the output.

2. Press `Ctrl+C` to stop the tracing process.

Additional resources

- `tcptop(8)` man page
- `/usr/share/bcc/tools/doc/tcptop.txt` file

### 42.11. TRACING ESTABLISHED TCP CONNECTIONS

The `tcptracer` utility traces the kernel functions that connect, accept, and close TCP connections. The utility uses eBPF features and, therefore, has a very low overhead.

**Procedure**

1. Use the following command to start the tracing process:

```bash
#/usr/share/bcc/tools/tcptracer
Tracing TCP established connections. Ctrl-C to end.
T  PID    COMM        IP SADDR        DADDR       SPORT  DPORT
A  1088   ns-slapd    4  192.0.2.153  192.0.2.1   0      65535```

CHAPTER 42. NETWORK TRACING USING THE BPF COMPILER COLLECTION
Each time the kernel connects, accepts, or closes a connection, tcptracer displays the details of the connections.

2. Press Ctrl+C to stop the tracing process.

Additional resources

- tcptracer(8) man page
- /usr/share/bcc/tools/doc/tcptracer_example.txt file

42.12. TRACING IPV4 AND IPV6 LISTEN ATTEMPTS

The solisten utility traces all IPv4 and IPv6 listen attempts. It traces the listen attempts including that ultimately fail or the listening program that does not accept the connection. The utility traces function that the kernel calls when a program wants to listen for TCP connections.

Procedure

1. Enter the following command to start the tracing process that displays all listen TCP attempts:

```
# /usr/share/bcc/tools/solisten
```

```
PID    COMM           PROTO         BACKLOG     PORT     ADDR
3643   nc             TCPv4         1           4242     0.0.0.0
3659   nc             TCPv6         1           4242     2001:db8:1::1
4221   redis-server   TCPv6         128         6379     ::
4221   redis-server   TCPv4         128         6379     0.0.0.0
....
```

2. Press Ctrl+C to stop the tracing process.

Additional resources

- solisten(9) man page
- /usr/share/bcc/tools/doc/solisten_example.txt file

42.13. SUMMARIZING THE SERVICE TIME OF SOFT INTERRUPTS

The softirqs utility summarizes the time spent servicing soft interrupts (soft IRQs) and shows this time as either totals or histogram distributions. This utility uses the irq:softirq_enter and irq:softirq_exit kernel tracepoints, which is a stable tracing mechanism.

Procedure

1. Enter the following command to start the tracing soft irq event time:

```
# /usr/share/bcc/tools/softirqs
```

```
Tracing soft irq event time... Hit Ctrl-C to end.
^C
```
### Additional resources

- **softirqs(8) man page**
- **/usr/share/bcc/tools/doc/softirqs_example.txt file**
- **mpstat(1) man page**

### 42.14. SUMMARIZING PACKETS SIZE AND COUNT ON A NETWORK INTERFACE

The netqtop utility displays statistics about the attributes of received (RX) and transmitted (TX) packets on each network queue of a particular network interface. The statistics include:

- Bytes per second (BPS)
- Packets per second (PPS)
- The average packet size
- Total number of packets

To generate these statistics, netqtop traces the kernel functions that perform events of transmitted packets `net_dev_start_xmit` and received packets `netif_receive_skb`.

#### Procedure

1. Display the number of packets within the range of bytes size of the time interval of 2 seconds:

```bash
# /usr/share/bcc/tools/netqtop -n enp1s0 -i 2
Fri Jan 31 18:08:55 2023
TX
QueueID avg_size [0, 64) [64, 512) [512, 2K) [2K, 16K) [16K, 64K)
0 0 0 0 0
Total 0 0 0 0 0
RX
QueueID avg_size [0, 64) [64, 512) [512, 2K) [2K, 16K) [16K, 64K)
0 38.0 1 0 0
Total 38.0 1 0 0

Fri Jan 31 18:08:57 2023
TX
QueueID avg_size [0, 64) [64, 512) [512, 2K) [2K, 16K) [16K, 64K)
```
2. Press Ctrl+C to stop netqtop.

Additional resources

- netqtop(8) man page
- /usr/share/bcc/tools/doc/netqtop_example.txt

42.15. ADDITIONAL RESOURCES

- /usr/share/doc/bcc/README.md
CHAPTER 43. CONFIGURING NETWORK DEVICES TO ACCEPT TRAFFIC FROM ALL MAC ADDRESSES

Network devices usually intercept and read packets that their controller is programmed to receive. You can configure the network devices to accept traffic from all MAC addresses in a virtual switch or at the port group level.

You can use this network mode to:

- Diagnose network connectivity issues
- Monitor network activity for security reasons
- Intercept private data-in-transit or intrusion in the network

You can enable this mode for any kind of network device, except InfiniBand.

43.1. TEMPORARILY CONFIGURING A DEVICE TO ACCEPT ALL TRAFFIC

You can use the ip utility to temporary configure a network device to accept all traffic regardless of the MAC addresses.

Procedure

1. Optional: Display the network interfaces to identify the one for which you want to receive all traffic:

```
# ip address show
1: enp1s0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc fq_codel state DOWN group default qlen 1000
   link/ether 98:fa:9b:a4:34:09 brd ff:ff:ff:ff:ff:ff
   ...```

2. Modify the device to enable or disable this property:

   - To enable the accept-all-mac-addresses mode for enp1s0:
     ```
     # ip link set enp1s0 promisc on
     ```
   - To disable the accept-all-mac-addresses mode for enp1s0:
     ```
     # ip link set enp1s0 promisc off
     ```

Verification

- Verify that the accept-all-mac-addresses mode is enabled:

```
# ip link show enp1s0
1: enp1s0: <NO-CARRIER,BROADCAST,MULTICAST,PROMISC,UP> mtu 1500 qdisc fq_codel state DOWN mode DEFAULT group default qlen 1000
   link/ether 98:fa:9b:a4:34:09 brd ff:ff:ff:ff:ff:ff
   ...```
The **PROMISC** flag in the device description indicates that the mode is enabled.

## 43.2. PERMANENTLY CONFIGURING A NETWORK DEVICE TO ACCEPT ALL TRAFFIC USING NMCLI

You can use the `nmcli` utility to permanently configure a network device to accept all traffic regardless of the MAC addresses.

### Procedure

1. Optional: Display the network interfaces to identify the one for which you want to receive all traffic:

   ```bash
   # ip address show
   1: enp1s0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc fq_codel state DOWN group default qlen 1000
   link/ether 98:fa:9b:a4:34:09 brd ff:ff:ff:ff:ff:ff
   ...
   ```

   You can create a new connection, if you do not have any.

2. Modify the network device to enable or disable this property.

   - To enable the `ethernet.accept-all-mac-addresses` mode for `enp1s0`:
     ```bash
     # nmcli connection modify enp1s0 ethernet.accept-all-mac-addresses yes
     ```

   - To disable the `accept-all-mac-addresses` mode for `enp1s0`:
     ```bash
     # nmcli connection modify enp1s0 ethernet.accept-all-mac-addresses no
     ```

3. Apply the changes, reactivate the connection:

   ```bash
   # nmcli connection up enp1s0
   ```

### Verification

- Verify that the `ethernet.accept-all-mac-addresses` mode is enabled:

  ```bash
  # nmcli connection show enp1s0
  ...  
  802-3-ethernet.accept-all-mac-addresses:1 (true)
  ```

The `802-3-ethernet.accept-all-mac-addresses: true` indicates that the mode is enabled.

## 43.3. PERMANENTLY CONFIGURING A NETWORK DEVICE TO ACCEPT ALL TRAFFIC USING NMSTATECTL

Use the `nmstatectl` utility to configure a device to accept all traffic regardless of the MAC addresses through the Nmstate API. The Nmstate API ensures that, after setting the configuration, the result matches the configuration file. If anything fails, `nmstatectl` automatically rolls back the changes to avoid leaving the system in an incorrect state.
Prerequisites

- The `nmstate` package is installed.
- The `enp1s0.yml` file that you used to configure the device is available.

Procedure

1. Edit the existing `enp1s0.yml` file for the `enp1s0` connection and add the following content to it:

   ```yaml
   ---
   interfaces:
      - name: enp1s0
        type: ethernet
        state: up
        accept-all-mac-address: true
   ```

   These settings configure the `enp1s0` device to accept all traffic.

2. Apply the network settings:

   ```bash
   # nmstatectl apply ~/enp1s0.yml
   ```

Verification

- Verify that the `802-3-ethernet.accept-all-mac-addresses` mode is enabled:

   ```bash
   # nmstatectl show enp1s0
   interfaces:
      - name: enp1s0
        type: ethernet
        state: up
        accept-all-mac-addresses: true
   ...
   ```

   The `802-3-ethernet.accept-all-mac-addresses: true` indicates that the mode is enabled.

Additional resources

- `nmstatectl(8)` man page
- `/usr/share/doc/nmstate/examples/` directory
CHAPTER 44. MIRRORING A NETWORK INTERFACE BY USING NMCLI

Network administrators can use port mirroring to replicate inbound and outbound network traffic being communicated from one network device to another. Mirroring traffic of an interface can be helpful in the following situations:

- To debug networking issues and tune the network flow
- To inspect and analyze the network traffic
- To detect an intrusion

Prerequisites

- A network interface to mirror the network traffic to.

Procedure

1. Add a network connection profile that you want to mirror the network traffic from:

   ```
   # nmcli connection add type ethernet ifname enp1s0 con-name enp1s0 autoconnect no
   ```

2. Attach a `prio qdisc` to `enp1s0` for the egress (outgoing) traffic with the `10:` handle:

   ```
   # nmcli connection modify enp1s0 +tc.qdisc "root prio handle 10:"
   ```

   The `prio qdisc` attached without children allows attaching filters.

3. Add a `qdisc` for the ingress traffic, with the `ffff:` handle:

   ```
   # nmcli connection modify enp1s0 +tc.qdisc "ingress handle ffff:"
   ```

4. Add the following filters to match packets on the ingress and egress `qdiscs`, and to mirror them to `enp7s0`:

   ```
   # nmcli connection modify enp1s0 +tc.tfilter "parent ffff: matchall action mirred egress mirror dev enp7s0"
   # nmcli connection modify enp1s0 +tc.tfilter "parent 10: matchall action mirred egress mirror dev enp7s0"
   ```

   The `matchall` filter matches all packets, and the `mirred` action redirects packets to destination.

5. Activate the connection:

   ```
   # nmcli connection up enp1s0
   ```

Verification

1. Install the `tcpdump` utility:
# dnf install tcpdump

2. Display the traffic mirrored on the target device (**enp7s0**):

    # tcpdump -i **enp7s0**

Additional resources

- How to capture network packets using **tcpdump**
CHAPTER 45. USING NMSTATE-AUTOCONF TO AUTOMATICALLY CONFIGURE THE NETWORK STATE USING LLDP

Network devices can use the Link Layer Discovery Protocol (LLDP) to advertise their identity, capabilities, and neighbors in a LAN. The `nmstate-autoconf` utility can use this information to automatically configure local network interfaces.

**IMPORTANT**

The `nmstate-autoconf` utility is provided as a Technology Preview only. Technology Preview features are not supported with Red Hat production Service Level Agreements (SLAs), might not be functionally complete, and Red Hat does not recommend using them for production. These previews provide early access to upcoming product features, enabling customers to test functionality and provide feedback during the development process.

See [Technology Preview Features Support Scope](#) on the Red Hat Customer Portal for information about the support scope for Technology Preview features.

45.1. USING NMSTATE-AUTOCONF TO AUTOMATICALLY CONFIGURE NETWORK INTERFACES

The `nmstate-autoconf` utility uses LLDP to identify the VLAN settings of interfaces connected to a switch to configure local devices.

This procedure assumes the following scenario and that the switch broadcasts the VLAN settings using LLDP:

- The `enp1s0` and `enp2s0` interfaces of the RHEL server are connected to switch ports that are configured with VLAN ID 100 and VLAN name `prod-net`.
- The `enp3s0` interface of the RHEL server is connected to a switch port that is configured with VLAN ID 200 and VLAN name `mgmt-net`.

The `nmstate-autoconf` utility then uses this information to create the following interfaces on the server:

- `bond100` - A bond interface with `enp1s0` and `enp2s0` as ports.
- `prod-net` - A VLAN interface on top of `bond100` with VLAN ID 100.
- `mgmt-net` - A VLAN interface on top of `enp3s0` with VLAN ID 200

If you connect multiple network interfaces to different switch ports for which LLDP broadcasts the same VLAN ID, `nmstate-autoconf` creates a bond with these interfaces and, additionally, configures the common VLAN ID on top of it.

**Prerequisites**

- The `nmstate` package is installed.
- LLDP is enabled on the network switch.
- The Ethernet interfaces are up.
Procedure

1. Enable LLDP on the Ethernet interfaces:
   a. Create a YAML file, for example `~/enable-lldp.yml`, with the following content:

   ```yaml
   interfaces:
   - name: enp1s0
     type: ethernet
     lldp:
       enabled: true
   - name: enp2s0
     type: ethernet
     lldp:
       enabled: true
   - name: enp3s0
     type: ethernet
     lldp:
       enabled: true
   
   b. Apply the settings to the system:

   ```bash
   # nmstatectl apply ~/enable-lldp.yml
   ```

2. Configure the network interfaces using LLDP:
   a. Optional, start a dry-run to display and verify the YAML configuration that `nmstate-autoconf` generates:

   ```bash
   # nmstate-autoconf -d enp1s0,enp2s0,enp3s0
   
   ---
   interfaces:
   - name: prod-net
     type: vlan
     state: up
     vlan:
       base-iface: bond100
       id: 100
   - name: mgmt-net
     type: vlan
     state: up
     vlan:
       base-iface: enp3s0
       id: 200
   - name: bond100
     type: bond
     state: up
     link-aggregation:
       mode: balance-rr
       port:
       - enp1s0
       - enp2s0
   ```
   b. Use `nmstate-autoconf` to generate the configuration based on information received from LLDP, and apply the settings to the system:

   ```bash
   #
   ```
# nmstate-autoconf enp1s0, enp2s0, enp3s0

## Next steps

- If there is no DHCP server in your network that provides the IP settings to the interfaces, configure them manually. For details, see:
  - Configuring an Ethernet connection
  - Configuring network bonding

## Verification

1. Display the settings of the individual interfaces:

   # nmstatectl show <interface_name>

## Additional resources

- nmstate-autoconf(8) man page
CHAPTER 46. CONFIGURING 802.3 LINK SETTINGS

Auto-negotiation is a feature of the IEEE 802.3u Fast Ethernet protocol. It targets the device ports to provide an optimal performance of speed, duplex mode, and flow control for information exchange over a link. Using the auto-negotiation protocol, you have optimal performance of data transfer over the Ethernet.

NOTE

To utilize maximum performance of auto-negotiation, use the same configuration on both sides of a link.

46.1. CONFIGURING 802.3 LINK SETTINGS USING THE NMCLI UTILITY

To configure the 802.3 link settings of an Ethernet connection, modify the following configuration parameters:

- 802-3-ethernet.auto-negotiate
- 802-3-ethernet.speed
- 802-3-ethernet.duplex

Procedure

1. Display the current settings of the connection:

```
# nmcli connection show Example-connection
...
802-3-ethernet.speed: 0
802-3-ethernet.duplex: --
802-3-ethernet.auto-negotiate: no
...
```

You can use these values if you need to reset the parameters in case of any problems.

2. Set the speed and duplex link settings:

```
# nmcli connection modify Example-connection 802-3-ethernet.auto-negotiate yes 802-3-ethernet.speed 10000 802-3-ethernet.duplex full
```

This command enables auto-negotiation and sets the speed of the connection to 10000 Mbit full duplex.

3. Reactivate the connection:

```
# nmcli connection up Example-connection
```

Verification

- Use the `ethtool` utility to verify the values of Ethernet interface `enp1s0`:

```
# ethtool enp1s0
```
Settings for enp1s0:
...
Speed: 10000 Mb/s
Duplex: Full
Auto-negotiation: on
...
Link detected: yes

Additional resources

- **nm-settings(5)** man page
CHAPTER 47. GETTING STARTED WITH DPDK

The data plane development kit (DPDK) provides libraries and network drivers to accelerate packet processing in user space.

Administrators use DPDK, for example, in virtual machines to use Single Root I/O Virtualization (SR-IOV) to reduce latencies and increase I/O throughput.

NOTE

Red Hat does not support experimental DPDK APIs.

47.1. INSTALLING THE DPDK PACKAGE

To use DPDK, install the dpdk package.

Procedure

- Use the dnf utility to install the dpdk package:

  ```
  # dnf install dpdk
  ```

47.2. ADDITIONAL RESOURCES

- Network Adapter Fast Datapath Feature Support Matrix
CHAPTER 48. GETTING STARTED WITH TIPC

Transparent Inter-process Communication (TIPC), which is also known as Cluster Domain Sockets, is an Inter-process Communication (IPC) service for cluster-wide operation.

Applications that are running in a high-available and dynamic cluster environment have special needs. The number of nodes in a cluster can vary, routers can fail, and, due to load balancing considerations, functionality can be moved to different nodes in the cluster. TIPC minimizes the effort by application developers to deal with such situations, and maximizes the chance that they are handled in a correct and optimal way. Additionally, TIPC provides a more efficient and fault-tolerant communication than general protocols, such as TCP.

48.1. THE ARCHITECTURE OF TIPC

TIPC is a layer between applications using TIPC and a packet transport service (bearer), and spans the level of transport, network, and signaling link layers. However, TIPC can use a different transport protocol as bearer, so that, for example, a TCP connection can serve as a bearer for a TIPC signaling link.

TIPC supports the following bearers:

- Ethernet
- InfiniBand
- UDP protocol

TIPC provides a reliable transfer of messages between TIPC ports, that are the endpoints of all TIPC communication.

The following is a diagram of the TIPC architecture:

48.2. LOADING THE TIPC MODULE WHEN THE SYSTEM BOOTS

Before you can use the TIPC protocol, you must load the tipc kernel module. You can configure Red Hat Enterprise Linux to automatically load this kernel module automatically when the system boots.

Procedure

1. Create the `/etc/modules-load.d/tipc.conf` file with the following content:
2. Restart the `systemd-modules-load` service to load the module without rebooting the system:

```
# systemctl start systemd-modules-load
```

**Verification**

1. Use the following command to verify that RHEL loaded the `tipc` module:

```
# lsmod | grep tipc
```

```
tipc 311296 0
```

If the command shows no entry for the `tipc` module, RHEL failed to load it.

**Additional resources**

- `modules-load.d(5)` man page

### 48.3. CREATING A TIPC NETWORK

To create a TIPC network, perform this procedure on each host that should join the TIPC network.

**IMPORTANT**

The commands configure the TIPC network only temporarily. To permanently configure TIPC on a node, use the commands of this procedure in a script, and configure RHEL to execute that script when the system boots.

**Prerequisites**

- The `tipc` module has been loaded. For details, see [Loading the `tipc` module when the system boots](#).

**Procedure**

1. Optional: Set a unique node identity, such as a UUID or the node’s host name:

```
# tipc node set identity host_name
```

   The identity can be any unique string consisting of a maximum 16 letters and numbers.

   You cannot set or change an identity after this step.

2. Add a bearer. For example, to use Ethernet as media and `enp0s1` device as physical bearer device, enter:

```
# tipc bearer enable media eth device enp1s0
```

3. Optional: For redundancy and better performance, attach further bearers using the command from the previous step. You can configure up to three bearers, but not more than two on the same media.
4. Repeat all previous steps on each node that should join the TIPC network.

Verification

1. Display the link status for cluster members:

```
# tipc link list
broadcast-link: up
5254006b74be:enp1s0-525400df55d1:enp1s0: up
```

This output indicates that the link between bearer `enp1s0` on node `5254006b74be` and bearer `enp1s0` on node `525400df55d1` is up.

2. Display the TIPC publishing table:

```
# tipc nametable show
Type  Lower      Upper      Scope    Port       Node
0     1795222054 1795222054 cluster  0          5254006b74be
0     3741353223 3741353223 cluster  0          525400df55d1
1     1           1           node     2399405586 5254006b74be
2     3741353223 3741353223 node     0          5254006b74be
```

- The two entries with service type `0` indicate that two nodes are members of this cluster.
- The entry with service type `1` represents the built-in topology service tracking service.
- The entry with service type `2` displays the link as seen from the issuing node. The range limit `3741353223` represents the peer endpoint’s address (a unique 32-bit hash value based on the node identity) in decimal format.

Additional resources

- `tipc-bearer(8)` man page
- `tipc-namespace(8)` man page

48.4. ADDITIONAL RESOURCES

- Red Hat recommends to use other bearer level protocols to encrypt the communication between nodes based on the transport media. For example:
  - MACSec: See Using MACsec to encrypt layer 2 traffic
  - IPsec: See Configuring a VPN with IPsec
- For examples of how to use TIPC, clone the upstream GIT repository using the `git clone git://git.code.sf.net/p/tipc/tipcutils` command. This repository contains the source code of demos and test programs that use TIPC features. Note that this repository is not provided by Red Hat.
CHAPTER 49. AUTOMATICALLY CONFIGURING NETWORK INTERFACES IN PUBLIC CLOUDS USING NM-CLOUD-SETUP

Usually, a virtual machine (VM) has only one interface that is configurable by DHCP. However, DHCP cannot configure VMs with multiple network entities, such as interfaces, IP subnets, and IP addresses. Additionally, you cannot apply settings when the VM instance is running. To solve this runtime configuration issue, the `nm-cloud-setup` utility automatically retrieves configuration information from the metadata server of the cloud service provider and updates the network configuration of the host. The utility automatically picks up multiple network interfaces, multiple IP addresses, or IP subnets on one interface and helps to reconfigure the network of the running VM instance.

49.1. CONFIGURING AND PRE-DEPLOYING NM-CLOUD-SETUP

To enable and configure network interfaces in public clouds, run `nm-cloud-setup` as a timer and service.

**NOTE**

On Red Hat Enterprise Linux On Demand and AWS golden images, `nm-cloud-setup` is already enabled and no action is required.

**Prerequisite**

- A network connection exists.
- The connection uses DHCP.
  
  By default, NetworkManager creates a connection profile which uses DHCP. If no profile was created because you set the `no-auto-default` parameter in `/etc/NetworkManager/NetworkManager.conf`, create this initial connection manually.

**Procedure**

1. Install the `nm-cloud-setup` package:

   ```bash
   # dnf install NetworkManager-cloud-setup
   ```

2. Create and run the snap-in file for the `nm-cloud-setup` service:
   
   a. Use the following command to start editing the snap-in file:

      ```bash
      # systemctl edit nm-cloud-setup.service
      ```

      It is important to either start the service explicitly or reboot the system to make configuration settings effective.

   b. Use the `systemd` snap-in file to configure the cloud provider in `nm-cloud-setup`. For example, to use Amazon EC2, enter:

      ```
      [Service]
      Environment=NM_CLOUD_SETUP_EC2=yes
      ```

      You can set the following environment variables to enable the cloud provider you use:

      - `NM_CLOUD_SETUP_AZURE` for Microsoft Azure
• **NM_CLOUD_SETUP_EC2** for Amazon EC2 (AWS)
• **NM_CLOUD_SETUP_GCP** for Google Cloud Platform (GCP)
• **NM_CLOUD_SETUP_ALIYUN** for Alibaba Cloud (Aliyun)

c. Save the file and quit the editor.

3. **Reload the** **systemd** **configuration:**

```
# systemctl daemon-reload
```

4. **Enable and start the** **nm-cloud-setup** **service:**

```
# systemctl enable --now nm-cloud-setup.service
```

5. **Enable and start the** **nm-cloud-setup** **timer:**

```
# systemctl enable --now nm-cloud-setup.timer
```

**Additional resources**

- **nm-cloud-setup(8)** man page
- **Configuring an Ethernet connection**

**49.2. UNDERSTANDING THE ROLE OF IMDSV2 AND NM-CLOUD-SETUP IN THE RHEL EC2 INSTANCE**

The instance metadata service (IMDS) in Amazon EC2 allows you to manage permissions to access instance metadata of a running Red Hat Enterprise Linux (RHEL) EC2 instance. The RHEL EC2 instance uses IMDS version 2 (IMDSv2), a session-oriented method. By using the **nm-cloud-setup** utility, administrators can reconfigure the network and automatically update the configuration of running RHEL EC2 instances. The **nm-cloud-setup** utility handles IMDSv2 API calls by using IMDSv2 tokens without any user intervention.

- IMDS runs on a link-local address **169.254.169.254** for providing access to native applications on a RHEL EC2 instance.

- After you have specified and configured IMDSv2 for each RHEL EC2 instance for applications and users, you can no longer access IMDSv1.

- By using IMDSv2, the RHEL EC2 instance maintains metadata without using the IAM role while remaining accessible through the IAM role.

- When the RHEL EC2 instance boots, the **nm-cloud-setup** utility automatically runs to fetch the EC2 instance API access token for using the RHEL EC2 instance API.

**NOTE**

Use the IMDSv2 token as an HTTP header to check the details of the EC2 environment.

**Additional resources**
• **nm-cloud-setup(8)** man page