Red Hat OpenStack Platform 10

Network Functions Virtualization Product Guide

Overview of the Network Functions Virtualization (NFV)
Overview of the Network Functions Virtualization (NFV)

OpenStack Team
rhos-docs@redhat.com
Abstract

This guide introduces Network Functions Virtualization (NFV), its advantages, supported configurations, architecture and components, installation and integration information.
# Table of Contents

PREFACE .................................................................................................................. 3.

CHAPTER 1. UNDERSTANDING RED HAT NETWORK FUNCTIONS VIRTUALIZATION (NFV) .......... 4.
  1.1. ADVANTAGES OF NFV 4
  1.2. SUPPORTED CONFIGURATIONS FOR NFV DEPLOYMENTS 4
  1.3. ETSI NFV ARCHITECTURE 5
  1.4. SUBSCRIPTIONS 5

CHAPTER 2. SOFTWARE ............................................................................................. 7.
  2.1. ARCHITECTURE AND COMPONENTS 7
  2.2. INTEGRATION 8
  2.3. INSTALLATION SUMMARY 8

CHAPTER 3. HARDWARE ............................................................................................ 10.

CHAPTER 4. DATAPATH SUPPORTABILITY MATRIX .................................................. 11.

CHAPTER 5. NFV PERFORMANCE CONSIDERATIONS ................................................. 13.
  5.1. CPUS AND NUMA NODES 13
    5.1.1. NUMA Node Example 13
  5.2. CPU PINNING 14
  5.3. HUGEPAGES 15

CHAPTER 6. FINDING MORE INFORMATION .............................................................. 16.
PREFACE

Red Hat OpenStack Platform provides the foundation to build a private or public Infrastructure-as-a-Service (IaaS) cloud on top of Red Hat Enterprise Linux. It offers a massively scalable, fault-tolerant platform for the development of cloud-enabled workloads.

Network Functions Virtualization (NFV) uses virtualization to move network node functions into building blocks that interconnect to create communication services. NFV is a new way to define, create, and manage networks by replacing dedicated hardware appliances with software and automation.

This guide briefly discusses Red Hat’s effort to accelerate NFV deployments using the Red Hat OpenStack Platform.
CHAPTER 1. UNDERSTANDING RED HAT NETWORK FUNCTIONS VIRTUALIZATION (NFV)

Network Functions Virtualization (NFV) is a software-based solution that helps the Communication Service Providers (CSPs) to move beyond the traditional, proprietary hardware to achieve greater efficiency and agility while reducing the operational costs.

NFV virtualizes network functions on general-purpose, cloud-based infrastructure to provide more agility, flexibility, simplicity, efficiency, and scalability than legacy infrastructure, while also reducing costs and allowing greater innovation.

An NFV environment allows for IT and network convergence by providing a virtualised infrastructure using the standard virtualization technologies that run on standard hardware devices such as switches, routers, and storage to virtualize network functions. The management and orchestration logic deploys and sustains these services. NFV also includes a Systems Administration, Automation and Life-Cycle Management thereby reducing the manual work necessary. It makes sure that people use modern tools, for example, devops, to work faster and at a larger scale.

1.1. ADVANTAGES OF NFV

The main advantages of implementing NFV are as follows:

- It accelerates the time-to-market by allowing quick deployment of new networking services because you do not need to install specialized new hardware to support changing business requirements. NFV allows Communication Service Providers (CSPs) to try and develop services to meet the growing customer demands, thus reducing the risk associated with new services.

- It delivers agility and flexibility by allowing you to quickly scale services to address changing demands, and supports innovation by enabling service developers to self-manage their resources and prototype using the same platform that will be used in production.

- It addresses customer demands in hours or minutes instead of weeks or days, without sacrificing security or performance.

- It reduces capital expenditures because it uses commodity-off-the-shelf hardware instead of expensive tailor-made equipment.

- It reduces operational costs by streamlined operations and automation that optimizes day-to-day tasks and improves employee productivity.

1.2. SUPPORTED CONFIGURATIONS FOR NFV DEPLOYMENTS

Red Hat OpenStack Platform 10 supports NFV deployments with the inclusion of automated OVS-DPDK and SR-IOV configuration. Furthermore, customers looking for a Hyper-converged Infrastructure (HCI) solution can now co-locate the Compute sub-system with the Red Hat Ceph Storage nodes. This hyper-converged model delivers lower cost of entry, smaller initial deployment footprints, maximized capacity utilization, and more efficient management in NFV use cases.

With the previous releases of Red Hat OpenStack Platform director, overcloud consists of predefined nodes, namely, Controller, Compute, Storage and so on. Each node consisted of a set of services defined in the core heat template collection on the director node. With the Red Hat OpenStack Platform 10, you can create custom deployment roles, using the composable roles feature, adding or removing services from each role. With Red Hat OpenStack Platform 10 release, the director offers many configuration options to deploy the latest OpenStack and Ceph versions. It also supports, with limitations,
the deployment of Technology Preview services such as OpenDaylight, Containerized Compute node or the Real-Time KVM hypervisor. For more information on the support scope for features marked as technology previews, see Technology Preview.

For more information on the Composable Roles, see Composable Roles and Services.

1.3. ETSI NFV ARCHITECTURE

The European Telecommunications Standards Institute (ETSI) is an independent standardization group that develops standards for information and communications technologies (ICT) in Europe.

Network functions virtualization (NFV) focuses on addressing problems involved in using proprietary hardware devices. With NFV, the necessity to install network-specific equipment is reduced, depending upon the use case requirements and economic benefits. The ETSI Industry Specification Group for Network Functions Virtualization (ETSI ISG NFV) sets the requirements, reference architecture, and the infrastructure specifications necessary to ensure virtualized functions are supported.

Red Hat is offering an open-source based cloud-optimized solution to help the Communication Service Providers (CSP) to achieve IT and network convergence, by adding the NFV features like SR-IOV, DPDK-OVS to existing open source projects like OpenStack.

1.4. SUBSCRIPTIONS

To install Red Hat OpenStack Platform, you must register all systems in the OpenStack environment either through the Red Hat Content Delivery Network, or through Red Hat Satellite 6. If using a Red Hat Satellite Server, synchronize the required repositories to your Red Hat OpenStack Platform environment. Subscribing to the right channels allows you access to certain repositories and each repository allows you to download the necessary packages required for installing and configuring Red Hat OpenStack Platform. The following is a list of CDN channels that you need to subscribe for installing Red Hat OpenStack Platform 10 with NFV components:

Table: Red Hat OpenStack Platform Repositories

<table>
<thead>
<tr>
<th>Name</th>
<th>Repository</th>
<th>Description of Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Hat Enterprise Linux 7 Server (RPMs)</td>
<td>rhel-7-server-rpms</td>
<td>Base operating system repository.</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux 7 Server - Extras (RPMs)</td>
<td>rhel-7-server-extras-rpms</td>
<td>Contains Red Hat OpenStack Platform dependencies</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux 7 Server - RH Common (RPMs)</td>
<td>rhel-7-server-rh-common-rpms</td>
<td>Contains tools for deploying and configuring Red Hat OpenStack Platform.</td>
</tr>
<tr>
<td>Red Hat Enterprise Linux High Availability (for RHEL 7 Server) (RPMs)</td>
<td>rhel-ha-for-rhel-7-server-rpms</td>
<td>High availability tools for Red Hat Enterprise Linux. Used for Controller node high availability.</td>
</tr>
<tr>
<td>Product Description</td>
<td>Repository Name</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Red Hat OpenStack Platform 10 for RHEL 7 (RPMs)</td>
<td><code>rhel-7-server-openstack-10-rpms</code></td>
<td>Core Red Hat OpenStack Platform repository. Also contains packages for Red Hat OpenStack Platform Director.</td>
</tr>
<tr>
<td>Red Hat Ceph Storage OSD 2 for Red Hat Enterprise Linux 7 Server (RPMs)</td>
<td><code>rhel-7-server-rhceph-2-osd-rpms</code></td>
<td>(For Ceph Storage Nodes) Repository for Ceph Storage Object Storage daemon. Installed on Ceph Storage nodes.</td>
</tr>
<tr>
<td>Red Hat Ceph Storage MON 2 for Red Hat Enterprise Linux 7 Server (RPMs)</td>
<td><code>rhel-7-server-rhceph-2-mon-rpms</code></td>
<td>(For Ceph Storage Nodes) Repository for Ceph Storage Monitor daemon. Installed on Controller nodes in OpenStack environments using Ceph Storage nodes.</td>
</tr>
<tr>
<td>Red Hat Ceph Storage Tools 2 for Red Hat Enterprise Linux 7 Workstation (RPMs)</td>
<td><code>rhel-7-server-rhceph-2-tools-rpms</code></td>
<td>(For Ceph Storage Nodes) Provides Rados REST gateway required for Ceph object storage.</td>
</tr>
</tbody>
</table>

For more information on the steps to subscribe to the channels, see Subscription Basics.
In general, NFV platform has the following components:

- **Virtualized Network Functions (VNFs)** - the software implementation of network functions, such as routers, firewalls, load balancers, broadband gateways, mobile packet processors, servicing nodes, signalling, location services, and so on.

- **NFV Infrastructure (NFVI)** - the physical resources (compute, storage, network) and the virtualization layer that make up the infrastructure. The network includes the datapath for forwarding packets between virtual machines and across hosts. This allows you to install VNFs without being concerned about the details of the underlying hardware. NFVI forms the foundation of the NFV stack. NFVI supports multi-tenancy and is managed by the Virtual Infrastructure Manager (VIM). Enhanced Platform Awareness (EPA) allows Red Hat OpenStack Platform to improve the virtual machine packet forwarding performance (throughput, latency, jitter) by exposing low-level CPU and NIC acceleration components to the VNF.

- **NFV Management and Orchestration (MANO)** - the management and orchestration layer focuses on all the service management tasks required throughout the lifecycle of the VNF. The main goals of MANO is to allow service definition, automation, error-correlation, monitoring and lifecycle of the network functions offered by the operator to its customers, decoupled from the physical infrastructure. This decoupling requires additional layer of management, provided by the Virtual Network Function Manager (VNFM). VNFM manages the lifecycle of the virtual...
machines and VNFs by either interacting directly with them or through the Element Management System (EMS) provided by the VNF vendor. The other important component defined by MANO is the Orchestrator also known as NFVO. NFVO interfaces with various databases and systems including Operations/Business Support Systems (OSS/BSS) on the top and with the VNFM on the bottom. If the NFVO wants to create a new service for a customer, it asks the VNFM to trigger the instantiation of a VNF (which may result in multiple virtual machines).

- **Operations and Business Support Systems (OSS/BSS)** - provides the essential business function applications, for example, operations support and billing. The OSS/BSS needs to be adapted to NFV, integrating with both legacy systems and the new MANO components. The BSS systems set policies based on service subscriptions and manage reporting and billing.

- **Systems Administration, Automation and Life-Cycle Management** - manages system administration, automation of the infrastructure components and life-cycle of the NFVI platform.

### 2.2. INTEGRATION

Red Hat’s solution for NFV includes a range of products that can act as the different components of the NFV framework in the ETSI model. The following products from the Red Hat portfolio will have NFV features:

- **Red Hat OpenStack Platform**, where a service provider can run IT and NFV workloads. The Enhanced Platform Awareness (EPA) features deliver deterministic performance improvements through OpenStack features like CPU Pinning, Huge pages, Non-Uniform Memory Access (NUMA) affinity and network adaptors (NICs) supporting SR-IOV and OVS-DPDK.

- **Red Hat Enterprise Linux** and **Red Hat Enterprise Linux Atomic Host** allow the creation of virtual machines and containers as VNFs.

- **Red Hat Ceph Storage** as the unified elastic and high-performance storage layer for all the needs of the service provider workloads.

- **Red Hat JBoss Middleware** and **OpenShift Enterprise** by Red Hat can be optionally used to modernize the OSS/BSS components.

- **Red Hat CloudForms** acts as the VNF manager and presents data from multiple sources like the VIM and the NFVi in a unified display.

- **Red Hat Satellite** and **Ansible** by Red Hat as optional components to provide enhanced Systems Administration, Automation and Life-cycle Management.

### 2.3. INSTALLATION SUMMARY

The Red Hat OpenStack Platform director is a toolset for installing and managing a complete OpenStack environment. It is based primarily on the upstream OpenStack project TripleO, which is an abbreviation for “OpenStack-On-OpenStack”. This project takes advantage of the OpenStack components to install a fully operational OpenStack environment; this includes a minimal OpenStack node called the undercloud that provisions and controls the bare metal systems to be use as the production OpenStack nodes, called the overcloud. This provides a simple method for installing a complete Red Hat OpenStack Platform environment that is both lean and robust.

For more information on installing the undercloud and overcloud, see [Red Hat OpenStack Platform Director Installation and Usage](#).

To install the NFV features, you need to configure the following additional steps:
● Include SR-IOV, and PCI Passthrough parameters to your `network-environment.yaml` file, update the `first-boot.yaml` file to set the Compute kernel arguments, modify the `compute.yaml` file and run the `overcloud_deploy.sh` script to deploy the overcloud.

● Install the DPDK libraries and drivers for fast packets processing by polling data directly from the NICs. Include the DPDK parameters to your `network-environment.yaml` file, update the `first-boot.yaml` files to set Compute kernel arguments, update the `compute.yaml` file to set the bridge with DPDK port, update the `controller.yaml` file to set the bridge and an interface with VLAN configured and run the `overcloud_deploy.sh` script to deploy the overcloud.

For detailed information on the configuration procedure steps, see Network Function Virtualization Configuration Guide.
CHAPTER 3. HARDWARE

You can use Red Hat Technologies Ecosystem to check for a list of certified hardware, software, cloud provider, component by choosing the category and then selecting the product version.

For a complete list of the certified hardware for Red Hat OpenStack Platform, see Red Hat OpenStack Platform certified hardware.

For a list of tested NICs for NFV, see Network Adapter Support. (Customer Portal login required.)
CHAPTER 4. DATAPATH SUPPORTABILITY MATRIX

With the introduction of NFV, more networking vendors are starting to implement their traditional devices as VNFs. While the majority of them are looking into virtual machines (VMs), some are also looking at container-based approach, per design choice. OpenStack-based solution should be rich and flexible due to two primary reasons:

- **Application readiness** - Network vendors are currently in the process of transforming their devices into VNFs. So different VNFs in the market have different maturity levels; common barriers to this readiness include enabling RESTful interfaces in their APIs, evolving their data models to become stateless, and providing automated management operations. OpenStack should provide a common platform for all.

- **Broad use-cases** - NFV includes a broad range of applications that serve different use-cases. For example, Virtual Customer Premise Equipment (vCPE) aims at providing a number of network functions such as routing, firewall, VPN, and NAT at customer premises. Virtual Evolved Packet Core (vEPC), is a cloud architecture that provides a cost-effective platform for the core components of Long-Term Evolution (LTE) network, allowing dynamic provisioning of gateways and mobile endpoints to sustain the increased volumes of data traffic from smartphones and other devices.

These use-cases, by nature, are implemented using different network applications and protocols, and require different connectivity, isolation and performance characteristics from the infrastructure. It is also common to separate between control plane interfaces and protocols and the actual forwarding plane. It is necessary for OpenStack to be flexible enough to offer different datapath connectivity options.

In principle, there are two common approaches for providing data plane connectivity to virtual machines:

- **Direct hardware access** bypasses the linux kernel and provides secure direct memory access (DMA) to the physical Network Interface Card (NIC) using technologies such as PCI Passthrough (denominated SR-IOV PF in OpenStack now) or single root I/O virtualization (SR-IOV) for both Virtual Function (VF) and Physical Function (PF) pass-through.

- **Using a virtual switch (vswitch)**, implemented as a software service of the hypervisor. Virtual machines are connected to the vSwitch using virtual interfaces (vNICs), and the vSwitch is capable of forwarding traffic between virtual machines as well as between virtual machines and the physical network.

Some of the common datapath options are as follows:

- **Single Root I/O Virtualization (SR-IOV)** - is a standard that makes a single PCI hardware device appear as multiple virtual PCI devices. It works by introducing Physical Functions (PFs) which are the full featured PCIe functions representing the physical hardware ports and Virtual Functions (VFs) that are lightweight functions that can be assigned to the virtual machines. The virtual machines see the VF as a regular NIC that communicates directly with the hardware. NICs support multiple VFs.

- **Open vSwitch (OVS)** - is an open source software switch that is designed to be used as a virtual switch within a virtualized server environment. OVS supports the capabilities of a regular L2-L3 switch and also offers support to the SDN protocols such as OpenFlow to create user-defined overlay networks (for example, VXLAN). OVS uses linux kernel networking to switch packets between virtual machines and across hosts using physical NIC. OVS now supports connection tracking (Conntrack) and built-in firewall capability which avoids overhead of linux bridge with iptables/ebtables. Open vSwitch for Red Hat OpenStack Platform environments offer out-of-box OpenStack Networking (neutron) integration with OVS.
- **Data Plane Development Kit (DPDK)** - consists of a set of libraries and poll mode drivers (PMD) for fast packet processing. It is designed to run mostly in the user-space, enabling applications to perform their own packet processing directly from/to the NIC. DPDK reduces latency and allows more packets to be processed. DPDK Poll Mode Drivers (PMDs) run in busy loop, constantly scanning the NIC ports on host and vNIC ports in guest for arrival of packets.

- **DPDK accelerated Open vSwitch (OVS-DPDK)** - is Open vSwitch bundled with DPDK for high performance user-space solution with Linux kernel bypass and direct memory access (DMA) to physical NICs. The idea is to replace the standard OVS kernel datapath with a DPDK-based datapath, creating a user-space vSwitch on the host, which uses DPDK internally for its packet forwarding. The advantage of this architecture is that it is mostly transparent to users as the basic OVS features as well as the interfaces it exposes (such as OpenFlow, OVSDB, the command line) remain mostly the same.
CHAPTER 5. NFV PERFORMANCE CONSIDERATIONS

For an NFV solution to be useful, its virtualized functions must meet or exceed the performance of physical implementations. Red Hat’s virtualization technologies are based on the high-performance Kernel-based Virtual Machine (KVM) hypervisor, common in OpenStack and cloud deployments.

5.1. CPUs AND NUMA NODES

Previously, all memory on x86 systems was equally accessible to all CPUs in the system. This resulted in memory access times that were the same regardless of which CPU in the system was performing the operation and was referred to as Uniform Memory Access (UMA).

In Non-Uniform Memory Access (NUMA), system memory is divided into zones called nodes, which are allocated to particular CPUs or sockets. Access to memory that is local to a CPU is faster than memory connected to remote CPUs on that system. Normally, each socket on a NUMA system has a local memory node whose contents can be accessed faster than the memory in the node local to another CPU or the memory on a bus shared by all CPUs.

Similarly, physical NICs are placed in PCI slots on the Compute node hardware. These slots connect to specific CPU sockets which are associated to a particular NUMA node. For optimum performance, connect your datapath NICs to the same NUMA nodes in your configuration (SR-IOV or OVS-DPDK).

The performance impact of NUMA misses are significant, generally starting at a 10% performance hit or higher. Each CPU socket can have multiple CPU cores which are treated as individual CPUs for virtualization purposes.

OpenStack Compute makes smart scheduling and placement decisions when launching instances. Administrators who want to take advantage of these features can create customized performance flavors to target specialized workloads including NFV and High Performance Computing (HPC).

TIP

Background information about NUMA is available in the following article: What is NUMA and how does it work on Linux?

5.1.1. NUMA Node Example

The following diagram provides an example of a two-node NUMA system and the way the CPU cores and memory pages are made available:
NOTE
Remote memory available via Interconnect is accessed only if VM1 from NUMA node 0 has a CPU core in NUMA node 1. In this case, the memory of NUMA node 1 will act as local for the third CPU core of VM1 (for example, if VM1 is allocated with CPU 4 in the diagram above), but at the same time, it will act as remote memory for the other CPU cores of the same VM.

WARNING
At present, it is impossible to migrate an instance which has been configured to use CPU pinning. For more information about this issue, see the following solution: Instance migration fails when using cpu-pinning from a numa-cell and flavor-property "hw:cpu_policy=dedicated".

5.2. CPU PINNING
CPU pinning is the ability to run a specific virtual machine's virtual CPU on a specific physical CPU, in a given host. vCPU pinning provides similar advantages to task pinning on bare-metal systems. Since virtual machines run as user space tasks on the host operating system, pinning increases cache efficiency.
5.3. HUGEPAGES

Physical memory is segmented into contiguous regions called pages. For efficiency, the system retrieves memory by accessing entire pages instead of individual bytes of memory. To perform this translation, the system looks in the Translation Lookaside Buffers (TLB) which contain the physical to virtual address mappings for the most recently or frequently used pages. When a mapping being searched for is not in the TLB, the processor must iterate through all the page tables to determine the address mappings. This causes a performance penalty. It is therefore preferable to optimise the TLB so as to ensure the target process can avoid the TLB misses as much as possible.

The typical page size in an x86 system is 4KB, with other larger page sizes available. Larger page sizes mean that there are fewer pages overall, and therefore increases the amount of system memory that can have its virtual to physical address translation stored in the TLB, and as a result lowers the potential for TLB misses, which increases performance. With larger page sizes, there is an increased potential for memory to be wasted as processes must allocate in pages, but not all of the memory is likely required. As a result, choosing a page size is a trade off between providing faster access times by using larger pages and ensuring maximum memory utilization by using smaller pages.
CHAPTER 6. FINDING MORE INFORMATION

The following table includes additional Red Hat documentation for reference:

The Red Hat OpenStack Platform documentation suite can be found here: Red Hat OpenStack Platform 10 Documentation Suite

Table 6.1. List of Available Documentation

<table>
<thead>
<tr>
<th>Component</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red Hat Enterprise Linux</td>
<td>Red Hat OpenStack Platform is supported on Red Hat Enterprise Linux 7.3. For information on installing Red Hat Enterprise Linux, see the corresponding installation guide at: Red Hat Enterprise Linux Documentation Suite.</td>
</tr>
<tr>
<td>Red Hat OpenStack Platform</td>
<td>To install OpenStack components and their dependencies, use the Red Hat OpenStack Platform director. The director uses a basic OpenStack installation as the undercloud to install, configure and manage the OpenStack nodes in the final overcloud. Be aware that you will need one extra host machine for the installation of the undercloud, in addition to the environment necessary for the deployed overcloud. For detailed instructions, see Red Hat OpenStack Platform Director Installation and Usage. For information on configuring advanced features for a Red Hat OpenStack Platform enterprise environment using the Red Hat OpenStack Platform director such as network isolation, storage configuration, SSL communication, and general configuration method, see Advanced Overcloud Customization. You can also manually install the Red Hat OpenStack Platform components, see Manual Installation Procedures.</td>
</tr>
<tr>
<td>NFV Documentation</td>
<td>For more details on planning your Red Hat OpenStack Platform deployment with NFV, see Network Function Virtualization Planning Guide. For information on configuring SR-IOV and OVS-DPDK with Red Hat OpenStack Platform 10 director, see the Network Functions Virtualization Configuration Guide.</td>
</tr>
</tbody>
</table>