



Red Hat Enterprise Linux 6

Guía de administración de energía

Manejo del consumo de energía en Red Hat Enterprise Linux 6
Edición 1.0

Red Hat Enterprise Linux 6 Guía de administración de energía

Manejo del consumo de energía en Red Hat Enterprise Linux 6

Edición 1.0

Don Domingo

Red Hat Servicios de contenido de Ingeniería

Rüdiger Landmann

Red Hat Servicios de contenido de Ingeniería

r.landmann@redhat.com

Red Hat Inc.

Legal Notice

Copyright © 2010 Red Hat Inc..

This document is licensed by Red Hat under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](#). If you distribute this document, or a modified version of it, you must provide attribution to Red Hat, Inc. and provide a link to the original. If the document is modified, all Red Hat trademarks must be removed.

Red Hat, as the licensor of this document, waives the right to enforce, and agrees not to assert, Section 4d of CC-BY-SA to the fullest extent permitted by applicable law.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, OpenShift, Fedora, the Infinity logo, and RHCE are trademarks of Red Hat, Inc., registered in the United States and other countries.

Linux ® is the registered trademark of Linus Torvalds in the United States and other countries.

Java ® is a registered trademark of Oracle and/or its affiliates.

XFS ® is a trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

MySQL ® is a registered trademark of MySQL AB in the United States, the European Union and other countries.

Node.js ® is an official trademark of Joyent. Red Hat Software Collections is not formally related to or endorsed by the official Joyent Node.js open source or commercial project.

The OpenStack ® Word Mark and OpenStack logo are either registered trademarks/service marks or trademarks/service marks of the OpenStack Foundation, in the United States and other countries and are used with the OpenStack Foundation's permission. We are not affiliated with, endorsed or sponsored by the OpenStack Foundation, or the OpenStack community.

All other trademarks are the property of their respective owners.

Resumen

Este documento explica cómo manejar efectivamente el consumo de energía en sistemas de Red Hat Enterprise Linux 6. Las secciones a continuación describen diferentes técnicas para disminuir el consumo de energía (para servidores y portátiles), y la forma como cada una de estas técnicas afecta el rendimiento general de su sistema.

Table of Contents

CAPÍTULO 1. VISTA PRELIMINAR	3
1.1. IMPORTANCIA DE ADMINISTRACIÓN DE ENERGÍA	3
1.2. FUNDAMENTOS DE ADMINISTRACIÓN DE ENERGÍA	4
CAPÍTULO 2. HERRAMIENTAS DE ADMINISTRACIÓN Y AUDITORÍA DE ENERGÍA	6
2.1. VISTA PRELIMINAR DE AUDITORÍA Y ANÁLISIS	6
2.2. POWERTOP	6
2.3. DISKDEVSTAT Y NETDEVSTAT	8
2.4. KIT DE HERRAMIENTAS DE VIDA DE BATERÍA	12
2.5. TUNED Y KTUNE	13
2.5.1. El archivo tuned.conf	14
2.5.2. Tuned-adm	16
2.6. DEVICEKIT DE ENERGÍA Y DEVKIT DE ENERGÍA	18
2.7. EL GESTOR DE ENERGÍA DE GNOME	19
2.8. OTROS MEDIOS DE AUDITORÍA	20
CAPÍTULO 3. INFRAESTRUCTURA CENTRAL Y MECÁNICA	21
3.1. ESTADOS DE CPU INACTIVOS	21
3.2. USO DE GOBERNADORES CPUFREQ	21
3.2.1. Tipos de gobernadores CPUfreq	22
3.2.2. Configuración de CPUfreq	23
3.2.3. Ajuste de la política CPUfreq y Velocidad	24
3.3. SUSPENDER Y REANUDAR	25
3.4. KERNEL SIN INTERVALOS	25
3.5. ADMINISTRADOR DE ENERGÍA DE ESTADO ACTIVO	26
3.6. ADMINISTRACIÓN DE ENERGÍA DE ENLACE AGRESIVO	26
3.7. OPTIMIZACIÓN DE ACCESO DE UNIDAD DE RELATIME	27
3.8. LÍMITE DE ENERGÍA	28
3.9. ADMINISTRACIÓN DE ENERGÍA EN GRÁFICAS MEJORADAS	29
3.10. RFKILL	29
3.11. OPTIMIZACIONES EN ESPACIO DE USUARIO	30
CAPÍTULO 4. CASOS DE USO	32
4.1. EJEMPLO — SERVIDOR	32
4.2. EJEMPLO — PORTÁTIL	33
APÉNDICE A. CONSEJOS PARA DESARROLLADORES	36
A.1. USO DE HILOS	36
A.2. DESPERTADORES	37
A.3. FSYNC	37
APÉNDICE B. HISTORIAL DE REVISIONES	39

CAPÍTULO 1. VISTA PRELIMINAR

La administración de energía ha sido uno de los puntos de interés de mejoras para Red Hat Enterprise Linux 6. La administración de energía es uno de los aspectos de *Tecnología de informática verde* (informática ecológica), una serie de consideraciones que también incluye aspectos tales como el uso de materiales reciclables, el impacto ambiental de la producción de hardware y conciencia ambiental en el diseño e implementación de sistemas. En este documento, proporcionamos orientación e información sobre la administración de energía de los sistemas que ejecutan Red Hat Enterprise Linux 6.

1.1. IMPORTANCIA DE ADMINISTRACIÓN DE ENERGÍA

En el centro de la administración de energía se entiende cómo optimizar de forma eficaz el consumo de energía de cada componente del sistema. Esto conlleva a estudiar las diferentes tareas que su sistema realiza y a configurar cada componente para garantizar que el rendimiento sea el correcto para el trabajo.

La razón principal para administrar energía es:

- reducción general de consumo de energía para ahorrar costo

El uso correcto de la administración de energía resulta en:

- reducción de calor para servidores y centros de cómputo
- reducción de costos secundarios que incluyen enfriamiento, espacio, cables, generadores y *Sistema de alimentación ininterrumpida*(UPS).
- extensión de vida de baterías para portátiles
- baja salida de dióxido de carbono
- cumplimiento de regulaciones gubernamentales o requisitos legales sobre Informática verde, por ejemplo Energy Star
- cumplimiento de reglamentación empresarial para nuevos sistemas

Como regla general, la reducción del consumo de energía de un componente específico del sistema (o del sistema como un todo) conducirá a reducir el calor y por ende el rendimiento. Consecuentemente, se deberá estudiar a fondo y probar la reducción en rendimiento permitida por las configuraciones que se realicen, especialmente para sistemas de misión crítica.

Si estudia las diferentes tareas que realiza su sistema y configura cada componente para garantizar que su rendimiento sea apenas suficiente para el trabajo, podrá ahorrar energía, generar menos calor y optimizar la vida de la batería para portátiles. Muchos de los principios para análisis y ajuste de un sistema con respecto al consumo de energía son similares a los del ajuste de rendimiento. En cierto grado, la administración energética y el ajuste de rendimiento son enfoques opuestos a la configuración del sistema porque generalmente los sistemas, se optimizan hacia el rendimiento o hacia la energía.

Red Hat Enterprise Linux 6 ya se distribuye con nuevas funcionalidades de manejo de energía que están habilitadas por defecto. Fueron escogidas de forma selectiva para no impactar el rendimiento de un servidor típico o el caso de escritorio. Sin embargo, para casos más específicos de máximo rendimiento, se requiere latencia más baja o rendimiento de CPU superior, una revisión de esos predeterminados puede ser necesaria.

Para decidir si debe optimizar sus máquinas mediante las técnicas descritas en este documento, hágase las siguientes preguntas:

P: ¿Debo optimizar?

R: La importancia de optimizar energía depende de si su compañía tiene lineamientos que necesiten seguirse o si hay algunas reglas que deben cumplirse.

P: ¿Cuánto necesito optimizar?

R: Varias de las técnicas que presentamos no requieren que usted vaya a través de todo el proceso de auditoría y análisis de su máquina, sino que ofrecen un set de optimizaciones generales que típicamente mejoran el uso de energía. Aquellas, por supuesto, no son tan buenas como un sistema auditado y optimizado manualmente aunque proporcionan un buen compromiso.

P: ¿La optimización reducirá el rendimiento del sistema a un nivel inaceptable?

R: La mayoría de las técnicas descritas en este documento impactan notablemente el rendimiento de su sistema. Si opta por implementar la administración de energía por fuera de los determinados en Red Hat Enterprise Linux 6, debe monitorizar el rendimiento del sistema después de optimizar la energía y decidir si la pérdida de rendimiento es aceptable.

P: ¿El tiempo y los recursos utilizados para optimizar el sistema superarán las ganancias obtenidas?

R: Optimizar manualmente un solo sistema siguiendo todo el proceso, no vale la pena, ya que el tiempo empleado y costo para hacerlo es hasta el momento superior al beneficio que se obtendría de una sola máquina. Por otra parte, podría ser una buena idea, si por ejemplo, usted despliega 10000 sistemas de escritorio para sus oficinas todas con la misma configuración y montaje y luego crea una configuración optimizada y la aplica a todas las máquinas.

Las siguientes secciones explicarán cómo el óptimo rendimiento de hardware beneficia a su sistema en términos de consumo de energía.

1.2. FUNDAMENTOS DE ADMINISTRACIÓN DE ENERGÍA

La administración de energía efectiva se construye sobre los siguiente principios:

Una CPU inactiva solo debe despertarse cuando sea necesario

El kernel de Red Hat Enterprise Linux 5, utilizaba un temporizador periódico para cada CPU. Dicho temporizador evita que la CPU quede en realidad inactiva (lo cual sucede cada pocos milisegundos, según la configuración), independientemente de si se está ejecutando algún proceso o no. Una gran parte de la administración de energía efectiva implica la reducción de la frecuencia en la cual se crean los despertadores de CPU.

Debido a lo anterior, el kernel de Linux en Red Hat Enterprise Linux 6 elimina el temporizador periódico: como resultado, el estado de inactividad de una CPU es ahora *sin intervalo*. Así, se evita el consumo de energía innecesario de la CPU cuando está inactiva. No obstante, los beneficios de esta función se pueden alterar si el sistema tiene aplicaciones que crean eventos de temporizador innecesarios. Los eventos de polling (tales como los controles de cambios de volumen, movimiento del ratón o similares) son ejemplos de dichos eventos.

Red Hat Enterprise Linux 5 incluye herramientas con las cuales se pueden identificar aplicaciones audit en la base del uso de su CPU. Consulte el [Capítulo 2, Herramientas de administración y auditoría de energía](#) para obtener mayor información.

Hardware y dispositivos no utilizados deben desactivarse completamente.

Esto es cierto especialmente para dispositivos que tienen partes móviles (tales como los discos duros). Además, algunas aplicaciones pueden dejar un dispositivo sin utilizar pero habilitado "abierto"; cuando esto se presenta, el kernel asume que el dispositivo está en uso, lo que puede impedir que el dispositivo entre en un estado de ahorro de energía.

Actividad baja debe traducirse en bajo vataje

Sin embargo, en muchos casos esto depende de un hardware moderno y de una configuración de BIOS correcta. Los componentes viejos del sistema, por lo general, no soportan las nuevas funciones que ahora podemos soportar en Red Hat Enterprise Linux 6. Asegúrese de utilizar el firmware oficial más reciente para sus sistemas y que estén habilitadas en la administración de energía o secciones de configuración de dispositivo de BIOS. Algunas funciones que se deben buscar son:

- SpeedStep
- PowerNow!
- Cool'n'Quiet
- ACPI (C state)
- Smart

Si su hardware soporta estas funciones y están habilitadas en el BIOS, Red Hat Enterprise Linux 6 las utilizará como predeterminadas.

Formas diferentes de estados de CPU y sus efectos

Las CPU modernas junto con la *Configuración avanzada e Interfaz de Energía (ACPI)* proporcionan diferentes estados de energía. Los tres estados diferentes son:

- Dormido (Estados-C)
- Frecuencia (estados-P)
- La salida de calor (estados-T o "thermal states")

Una CPU que se ejecute en el estado dormido más bajo posible, consumirá la menor cantidad de vatios, pero también utilizará más tiempo para despertarse de su estado cuando se necesite. En varios casos esto conlleva a que la CPU tenga que despertarse al poco tiempo de quedarse dormida. Esta situación produce una CPU ocupada permanentemente y pierde parte de su potencial energético si otro estado había sido utilizado.

Una máquina apagada utiliza la menor cantidad de energía

Obviamente, apagar los sistemas es una de las mejores formas de ahorrar energía. Por ejemplo, su compañía puede desarrollar una cultura corporativa enfocada a una conciencia de tecnología informática "verde" con una línea directriz para apagar las máquinas durante las horas de almuerzo, receso, o cuando termina la jornada laboral. También se podrían consolidar varios servidores físicos en un servidor más grande y virtualizarlos mediante la tecnología de virtualización que distribuimos con Red Hat Enterprise Linux 6.

CAPÍTULO 2. HERRAMIENTAS DE ADMINISTRACIÓN Y AUDITORÍA DE ENERGÍA

2.1. VISTA PRELIMINAR DE AUDITORÍA Y ANÁLISIS

La auditoría manual detallada, el análisis y ajuste de un sistema único suelen ser la excepción, porque el tiempo y costo utilizado para hacerlo, por lo general, superan los beneficios obtenidos de estas últimas partes de ajuste del sistema. Sin embargo, puede ser muy útil, si estas tareas se realizan una vez para un gran número de sistemas casi idénticos en los que se puede reutilizar la misma configuración para todos los sistemas. Por ejemplo, considere la implementación de miles de sistemas de escritorio o un cluster HPC donde las máquinas son casi idénticas. Otra razón para hacer auditoría y análisis es proporcionar una base de comparación con la cual pueda identificar regresiones o cambios en conducta del sistema en el futuro. Los resultados de este análisis pueden ser muy útiles en casos donde se suceden regularmente actualizaciones de hardware, BIOS, o de software y no se quieren tener sorpresas con respecto al consumo de energía. Por lo general, la auditoría y análisis le da una mejor idea de lo que realmente está pasando en su sistema.

La auditoría y análisis de un sistema con respecto al consumo de energía es relativamente difícil, incluso con los sistemas más modernos disponibles. La mayoría de los sistemas no proporcionan los medios necesarios para medir el uso de energía a través de software. Aunque existen excepciones: la consola de administración ILO de sistemas de servidor de Hewlett Packard tiene un módulo de administración que se puede acceder a través de la red. IBM proporciona una solución parecida en su módulo de administración de energía BladeCenter. En algunos sistemas Dell, el Asistente de Tecnología Informática también ofrece opciones de monitorización de energía. Otros proveedores probablemente ofrecen opciones similares para sus plataformas de servidor, pero como se puede ver no hay una solución única disponible que sea soportada por todos los proveedores. Si su sistema no tiene mecanismos incorporados para medir el consumo de energía, solo existen unas pocas opciones. Podría instalar una fuente de alimentación para su sistema que ofrezca información de consumo de energía a través de USB. La fuente de alimentación Gigabyte Odin GT 550 W PC , es un ejemplo de ello y el software para leer esos valores bajo Linux está disponible externamente en <http://mgmt.sth.sze.hu/~andras/dev/gopsu/>. Como último recurso, algunos medidores de vatios externos como Watts up? PRO tienen un conector USB.

Suelen necesitarse medidas directas de consumo de energía para maximizar el ahorro hasta donde sea posible. Afortunadamente, otros medios están disponibles para medir si los cambios se efectúan y cómo se está comportando el sistema. Este capítulo describe las herramientas necesarias.

2.2. POWERTOP

La introducción del kernel sin marca en Red Hat Enterprise Linux 6 (consulte la [Sección 3.4, “Kernel sin intervalos”](#)) permite a la CPU entrar en estado de inactividad con más frecuencia, reduciendo el consumo de energía y mejorando la administración de energía. La nueva herramienta **PowerTOP** identifica componentes específicos de kernel y aplicaciones de espacio de usuario que frecuentemente despiertan a la CPU. **PowerTOP** servía en desarrollo para realizar las auditorías descritas en la [Sección 3.11, “Optimizaciones en espacio de usuario”](#), las cuales dieron lugar a numerosas aplicaciones que se está ajustando en esta versión, reduciendo así despertadores de CPU innecesarios en un factor de 10.

Instale **PowerTOP** con el comando:

```
yum install powertop
```

Ejecute **PowerTOP** con el comando:

■

power top

Tenga en cuenta que necesitará ejecutar **PowerTOP** con privilegios de root para que la aplicación pueda hacer algo útil.

Al ejecutarse, **PowerTOP** reúne estadísticas del sistema y presenta una lista de los componentes que están enviando con mayor frecuencia despertadores a la CPU. **PowerTOP** también hace sugerencias sobre ajustes al sistema para reducir consumo de energía. Dichas sugerencias aparecen en la parte inferior de la pantalla, y especifican la tecla para aceptar la sugerencia de **PowerTOP**. Puesto que **PowerTOP** se actualiza periódicamente, otras sugerencias aparecerán. Observe en la [Figura 2.1, “PowerTOP en operación”](#), la sugerencia para **aumentar el tiempo de reescritura sucia de VM**, y la tecla (**w**) para aceptar la sugerencia.

Cuando se ejecuta, **PowerTOP** reúne estadísticas desde el sistema y le presenta varias listas importantes de información. En la parte superior de la lista está la lista de cuánto tiempo sus núcleos de CPU han estado disponibles en estados C y P. Entre más CPU permanezca en los estados superiores C o P es mejor (**C4** es superior a **C3**) y es un buen indicador de qué tan bien se ajusta el sistema al uso de CPU. Su meta debe ser la residencia de 90% o más en los estados superiores C o P cuando el sistema esté inactivo.

La segunda parte de la información es un resumen de los despertadores por segundo reales de la máquina. El número de despertadores por segundo da una medida de qué tan bien los servicios u otros dispositivos y controladores del kernel están ejecutándose con respecto al uso de energía en su sistema. Entre más despertadores por segundo tenga, habrá más consumo de energía, por lo tanto entre más baja es mejor aquí.

Luego, **PowerTOP** proporciona una aproximación del uso de energía del sistema, si está disponible. Espere a que **PowerTOP** reporte este dato en los portátiles cuando estén en modo de batería.

Cualquier estimación del uso de energía está acompañada por una lista detallada de todos los componentes que envían con más frecuencia despertadores a la CPU. En la parte superior de la lista están todos los componentes que debe investigar de cerca para optimizar su sistema y reducir el uso de energía. Si se trata de componentes de kernel, (indicados por el nombre del componente que se lista en <>) entonces los despertadores estarán siempre asociados con un controlador específico que los genera. El ajuste de controladores suele requerir cambios de kernel que van más allá del objetivo de este documento. Sin embargo, los procesos de espacio de usuario que envían despertadores se administran con mayor facilidad. Primero, identifique si este servicio o aplicación debe ejecutarse en este sistema. Si no, simplemente, desactívelo. Para apagar este servicio de forma permanente, ejecute:

```
chkconfig nombre_de_servicio off
```

Para obtener mayores detalles sobre lo que realmente este componente hace, ejecute:

```
ps -awux | grep nombre_de_componente
strace -p processid
```

Si el rastro parece que se está repitiendo, entonces probablemente ha encontrado un bucle ocupado. Para corregir esto se requeriría cambiar el código en ese componente, sin embargo este tema va más allá del tópico de este documento.

Por último, **PowerTOP** también hace sugerencias sobre ajustes del sistema para reducir el consumo de energía. Dichas sugerencias aparecen en la parte inferior de la pantalla, y especifican la tecla para aceptar la sugerencia de **PowerTOP**. Puesto que **PowerTOP** se actualiza periódicamente, otras sugerencias aparecerán. Observe en la [Figura 2.1, “PowerTOP en operación”](#), la sugerencia para **aumentar el tiempo de reescritura sucia de VM** y la tecla (**w**) para aceptar la sugerencia.

Estos cambios solo se activarán después del reinicio. Para ayudarle a hacer los cambios permanentes, **PowerTOP** desplegará el comando exacto que ejecuta para realizar esta optimización. Añada el comando a su archivo `/etc/rc.local` con su editor preferido para que el cambio se efectúe cada vez que el computador arranque.

```

PowerTOP version 1.11      (C) 2007 Intel Corporation

Cn          Avg residency      P-states (frequencies)
C0 (cpu running)      ( 4.4%)          2.81 Ghz   3.2%
polling          0.1ms ( 0.0%)          2.81 Ghz   0.2%
C1 mwait          0.0ms ( 0.0%)          2.14 Ghz   0.1%
C2 mwait          0.5ms ( 1.1%)          1.60 Ghz   0.4%
C4 mwait          4.3ms (94.5%)          800 Mhz    96.2%

Wakeup-from-idle per second : 245.5      interval: 15.0s
no ACPI power usage estimate available

Top causes for wakeups:
38.3% (163.7)  <kernel core> : hrtimer_start_range_ns (tick_sched_timer)
 8.8% ( 37.8)  <interrupt> : iwlgagn
 8.6% ( 36.9)  <kernel IPI> : Rescheduling interrupts
 7.9% ( 33.9)  <interrupt> : extra timer interrupt
 7.9% ( 33.7)  firefox : hrtimer_start_range_ns (hrtimer_wakeup)
 4.6% ( 19.9)  popfile.pl : hrtimer_start_range_ns (hrtimer_wakeup)
 3.2% ( 13.8)  <kernel core> : hrtimer_start (tick_sched_timer)
 2.7% ( 11.7)  <interrupt> : i915
 2.6% ( 11.2)  <interrupt> : ahci
 2.2% (  9.5)  <interrupt> : ehci_hcd:usb1
 2.2% (  9.5)  USB device 1-5.1.2 : Microsoft 3-Button Mouse with IntelliEye(TM) (Microsoft)
 2.1% (  9.0)  <kernel core> : __mod_timer (ehci_watchdog)
 1.5% (  6.5)  thunderbird-bin : hrtimer_start_range_ns (hrtimer_wakeup)
 1.3% (  5.5)  simpress.bin : hrtimer_start_range_ns (hrtimer_wakeup)
 1.3% (  5.5)  plasma-desktop : hrtimer_start_range_ns (hrtimer_wakeup)
 1.2% (  5.3)  <interrupt> : eth0
 0.9% (  4.0)  <kernel core> : __mod_timer (rh_timer_func)
 0.2% (  1.0)  klipper : hrtimer_start_range_ns (hrtimer_wakeup)
 0.2% (  1.0)  httpd : hrtimer_start_range_ns (hrtimer_wakeup)
 0.2% (  0.9)  konversation : hrtimer_start_range_ns (hrtimer_wakeup)

Suggestion: increase the VM dirty writeback time from 5.00 to 15 seconds with:
echo 1500 > /proc/sys/vm/dirty_writeback_centisecs
This wakes the disk up less frequently for background VM activity
Q - Quit      R - Refresh      W - Increase Writeback time
    
```

Figura 2.1. PowerTOP en operación

La página *Less Watts* publica una lista de aplicaciones que **PowerTOP** ha identificado para mantener activas las CPU. Consulte <http://www.lesswatts.org/projects/powertop/known.php>.

2.3. DISKDEVSTAT Y NETDEVSTAT

Diskdevstat y **netdevstat** son herramientas de **SystemTap** que colectan información detallada sobre la actividad de disco y de la red de todas las aplicaciones que se ejecutan en un sistema. Estas herramientas fueron inspiradas por **PowerTOP**, la cual muestra el número de despertadores de CPU por cada aplicación por segundo (consulte la [Sección 2.2, “PowerTOP”](#)). La estadística que recogen estas herramientas le permiten identificar aplicaciones que desperdician energía con operaciones de E/S muy pequeñas más que con las más grandes. Otras herramientas de monitorización que solo miden velocidades de transferencia no ayudan a identificar este tipo de uso.

Instalar estas herramientas con **SystemTap** con el comando:

```
yum install systemtap tuned-utils kernel-debuginfo
```

Ejecute las herramientas con el comando:

```
diskdevstat
```

o el comando:

```
netdevstat
```

Ambos comandos pueden tomar hasta tres parámetros, así:

```
diskdevstat update_interval duración_total mostrar_histograma
```

```
netdevstat update_interval duración_total mostrar_histograma
```

actualizar_intervalo

El tiempo en segundos entre actualizaciones de la pantalla. Predeterminado: **5**

duración_total

El tiempo en segundos para el total. Predeterminado: **86400** (1 día)

mostrar_histograma

Indica si hacer histograma para todos los datos recolectados al final de la ejecución.

La salida se parece a la de **PowerTOP**. A continuación, una muestra de una ejecución más larga de **diskdevstat** en un sistema que ejecuta Fedora 10 KDE 4.2:

```

  PID  UID DEV      WRITE_CNT WRITE_MIN WRITE_MAX WRITE_AVG  READ_CNT
  READ_MIN READ_MAX READ_AVG COMMAND
  2789 2903 sda1      854      0.000   120.000   39.836     0
0.000  0.000  0.000 plasma
 15494 0 sda1      0        0.000   0.000     0.000     758
0.000  0.012  0.000 0logwatch
 15520 0 sda1      0        0.000   0.000     0.000     140
0.000  0.009  0.000 perl
 15549 0 sda1      0        0.000   0.000     0.000     140
0.000  0.009  0.000 perl
 15585 0 sda1      0        0.000   0.000     0.000     108
0.001  0.002  0.000 perl
  2573 0 sda1      63      0.033  3600.015  515.226     0
0.000  0.000  0.000 auditd
 15429 0 sda1      0        0.000   0.000     0.000     62
0.009  0.009  0.000 crond
 15379 0 sda1      0        0.000   0.000     0.000     62
0.008  0.008  0.000 crond
 15473 0 sda1      0        0.000   0.000     0.000     62
0.008  0.008  0.000 crond
 15415 0 sda1      0        0.000   0.000     0.000     62
0.008  0.008  0.000 crond
 15433 0 sda1      0        0.000   0.000     0.000     62
0.008  0.008  0.000 crond
 15425 0 sda1      0        0.000   0.000     0.000     62
0.007  0.007  0.000 crond
 15375 0 sda1      0        0.000   0.000     0.000     62
0.008  0.008  0.000 crond
 15477 0 sda1      0        0.000   0.000     0.000     62

```

0.007	0.007	0.000	crond				
15469	0 sda1	0	0.000	0.000	0.000		62
0.007	0.007	0.000	crond				
15419	0 sda1	0	0.000	0.000	0.000		62
0.008	0.008	0.000	crond				
15481	0 sda1	0	0.000	0.000	0.000		61
0.000	0.001	0.000	crond				
15355	0 sda1	0	0.000	0.000	0.000		37
0.000	0.014	0.001	laptop_mode				
2153	0 sda1	26	0.003	3600.029	1290.730		0
0.000	0.000	0.000	rsyslogd				
15575	0 sda1	0	0.000	0.000	0.000		16
0.000	0.000	0.000	cat				
15581	0 sda1	0	0.000	0.000	0.000		12
0.001	0.002	0.000	perl				
15582	0 sda1	0	0.000	0.000	0.000		12
0.001	0.002	0.000	perl				
15579	0 sda1	0	0.000	0.000	0.000		12
0.000	0.001	0.000	perl				
15580	0 sda1	0	0.000	0.000	0.000		12
0.001	0.001	0.000	perl				
15354	0 sda1	0	0.000	0.000	0.000		12
0.000	0.170	0.014	sh				
15584	0 sda1	0	0.000	0.000	0.000		12
0.001	0.002	0.000	perl				
15548	0 sda1	0	0.000	0.000	0.000		12
0.001	0.014	0.001	perl				
15577	0 sda1	0	0.000	0.000	0.000		12
0.001	0.003	0.000	perl				
15519	0 sda1	0	0.000	0.000	0.000		12
0.001	0.005	0.000	perl				
15578	0 sda1	0	0.000	0.000	0.000		12
0.001	0.001	0.000	perl				
15583	0 sda1	0	0.000	0.000	0.000		12
0.001	0.001	0.000	perl				
15547	0 sda1	0	0.000	0.000	0.000		11
0.000	0.002	0.000	perl				
15576	0 sda1	0	0.000	0.000	0.000		11
0.001	0.001	0.000	perl				
15518	0 sda1	0	0.000	0.000	0.000		11
0.000	0.001	0.000	perl				
15354	0 sda1	0	0.000	0.000	0.000		10
0.053	0.053	0.005	lm_lid.sh				

Las columnas son:

PID

El ID de proceso de la aplicación

UID

El ID de usuario bajo el cual se ejecutan las aplicaciones

DEV

El dispositivo en el cual la E/S tienen lugar

WRITE_CNT

El número total de operaciones de escritura

WRITE_MIN

El tiempo mínimo para dos escritos consecutivos (en segundos)

WRITE_MAX

El tiempo máximo para dos escritos consecutivos (en segundos)

WRITE_AVG

El tiempo promedio para dos escritos consecutivos (en segundos)

READ_CNT

El número total de operaciones de lectura

READ_MIN

El tiempo mínimo para dos lecturas consecutivas (en segundos)

READ_MAX

El máximo tiempo para dos lecturas consecutivas (en segundos)

READ_AVG

El tiempo promedio para dos lecturas consecutivas (en segundos)

COMMAND

El nombre del proceso

En este ejemplo, tres aplicaciones muy obvias llaman la atención:

```

PID   UID DEV      WRITE_CNT WRITE_MIN WRITE_MAX WRITE_AVG  READ_CNT
READ_MIN READ_MAX READ_AVG COMMAND
2789  2903 sda1      854      0.000    120.000   39.836     0
0.000  0.000  0.000 plasma
2573   0 sda1       63      0.033   3600.015  515.226     0
0.000  0.000  0.000 auditd
2153   0 sda1       26      0.003   3600.029 1290.730     0
0.000  0.000  0.000 rsyslogd

```

Estas tres aplicaciones tienen un **WRITE_CNT** mayor que **0**, lo cual significa que realizan un forma de escritura durante la medición. De ellas, **plasma** fue el peor transgresor en un alto grado: realizó la mayoría de operaciones de escritura y por lo tanto el promedio de tiempo entre escrituras fue el más bajo. Por esa razón **Plasma** sería el mejor candidato para investigar si está interesado en aplicaciones ineficientes de energía.

Utilice los comandos **strace** y **ltrace** para examinar aplicaciones de una forma más cercana al rastrear todas las llamadas de sistema del ID de proceso determinado. En este ejemplo, puede ejecutar:

```
strace -p 2789
```

En este ejemplo, la salida de **strace** contenía un patrón repetitivo cada 45 segundos que abría para escribir el archivo cache del icono de KDE del usuario para escribir y luego lo cerraba inmediatamente. Esto llevaba necesariamente a una escritura física al disco duro cuando los metadatos del archivo (específicamente, el tiempo de modificación) habían cambiado. La corrección final era evitar llamadas innecesarias cuando no se presentaban actualizaciones de iconos.

2.4. KIT DE HERRAMIENTAS DE VIDA DE BATERÍA

Red Hat Enterprise Linux 6 introduce el **Kit de herramientas de vida de batería** (BLTK), un paquete de prueba que simula y analiza la vida y rendimiento de la batería. BLTK lo hace mediante la realización de un set de tareas que simulan grupos de usuario específicos y reportan los resultados. Aunque BLTK está diseñado específicamente para probar rendimiento de notebook, también puede reportar el rendimiento de computadores de escritorio cuando se inicia con la opción **-a**.

BLTK le permite generar cargas de trabajo reproducibles que son comparables al uso real de una máquina. Por ejemplo, la carga de trabajo de **office** escribe un texto, corrige cosas en él y hace lo mismo para una hoja de cálculo. Al ejecutar BLTK junto con **PowerTOP** o cualquiera otra herramienta de auditoría o análisis podrá probar si las optimizaciones que usted realizó tienen algún efecto cuando la máquina está activa en uso en lugar de cuando solo está en inactividad. Puesto que la misma carga de trabajo se puede ejecutar varias veces para diferentes configuraciones, usted podrá comparar resultados para diferentes configuraciones

Instalar BLTK con el comando:

```
yum install bltk
```

Ejecutar BLTK con el comando:

```
bltk carga_de_trabajo opciones
```

Por ejemplo, para ejecutar la carga de trabajo **idle** por 120 segundos:

```
bltk -I -T 120
```

Las cargas predeterminadas disponibles son:

-I, --idle

El sistema está inactivo, usar una línea de base para comparar con otras cargas de trabajo

-R, --reader

Simula lectura de documentos (por defecto, con **Firefox**)

-P, --player

Simula ver archivos de multimedia desde una unidad de CD o DVD (por defecto, con **mplayer**)

-O, --office

Simula la edición de documentos con el paquete de **OpenOffice.org**

Otras opciones le permiten especificar:

-a, --ac-ignore

Ignorar si energía de Corriente alterna está disponible (necesaria para usar escritorio)

-T número_de_segundos, --time número_de_segundos

El tiempo (en segundos) sobre el cual ejecutar la prueba; use esta opción con la carga de trabajo **idle**

-F nombre_de_archivo, --file nombre_de_archivo

Especifica un archivo para ser usado por una carga determinada, por ejemplo, un archivo para la carga de trabajo del **reproductor** para que se reproduzca en lugar de acceder a la unidad de CD o DVD

-W aplicación, --prog aplicación

Especifica una aplicación para ser utilizada por una carga de trabajo particular, por ejemplo, un navegador diferente a **Firefox** para la carga de trabajo de **lector**

BLTK soporta una gran cantidad de opciones especializadas. Para obtener mayor información, consulte la página man **bltk**.

BLTK guarda los resultados que genera en un directorio especificado por defecto en el archivo de configuración `/etc/bltk.conf`, `~/.bltk/workload.results.number/`. Por ejemplo, el directorio `~/.bltk/reader.results.002/` guarda los resultados de la tercera prueba con la carga de trabajo de **lector** (la primera prueba no está numerada). Los resultados están esparcidos a través de varios archivos de texto. Para condensar estos resultados dentro de un formato fácil de leer, ejecute:

```
bltk_report ruta_a_directorio_de_resultados
```

Los resultados aparecen ahora en un archivo de texto llamado **Report** en el directorio de resultados. Para ver los resultados en un emulador de terminal, utilice la opción **-o**:

```
bltk_report -o ruta_a_directorio_de_resultados
```

2.5. TUNED Y KTUNE

Tuned es un daemon que monitoriza el uso de los componentes del sistema y ajusta en forma dinámica la configuración del sistema con base en esa información de monitorización. El ajuste dinámico se encarga de la forma en que varios componentes del sistema se utilizan de forma diferente a través del uptime para cualquier sistema dado. Por ejemplo, el disco duro se utiliza bastante durante el arranque y el inicio de sesión, pero escasamente se utiliza con aplicaciones tales como OpenOffice o clientes de correo-e. Igualmente, los dispositivos de CPU y red se utilizan de forma diferente en varios momentos. **Tuned** monitoriza la actividad de dichos componentes y reacciona ante los cambios en su uso.

A manera de ejemplo, considere una estación de trabajo típica. La mayoría del tiempo, la interfaz de red de Ethernet estará muy inactiva. solo unos pocos correos-e entrarán y saldrán se cargarán algunas páginas web de vez en cuando. Para esa clase de cargas, la interfaz de red no tiene que ejecutarse a toda la velocidad todo el tiempo, como lo hace en forma predeterminada. **Tuned** tiene un complemento de monitorización y ajuste para dispositivos de red que puede detectar esa baja actividad y luego automáticamente bajar la velocidad de esa interfaz, resultando en un uso de energía más bajo. Si la actividad en la interfaz aumenta drásticamente por un periodo de tiempo mayor, por ejemplo debido a una imagen de DVD que se está descargando o un correo-e que está abriendo un archivo adjunto **tuned**

lo detectará y establecerá la velocidad de interfaz al máximo para ofrecer el mejor rendimiento cuando el nivel de actividad es tan alto. Este principio también se aplica para otras conexiones de CPU y discos duros.

Los dispositivos de red no están configurados para comportarse de esa manera porque los cambios de velocidad pueden tardar varios segundos en efectuarse y por lo tanto impactarían visiblemente la experiencia del usuario. Consideraciones similares se aplican para conexiones de ajuste de CPU y disco duro. Cuando un disco duro girado hacia abajo, puede tardar varios segundos para que vuelva a girar, lo que produce una falta de respuesta del sistema durante ese periodo. El efecto secundario de latencia es menor para la conexión de CPU, pero al menos aún se puede medir, aunque es casi imperceptible al usuario.

Ahora junto con **tuned** ofrecemos **ktune**. **Ktune** fue presentada en Red Hat Enterprise Linux 5.3 como marco de trabajo y servicio para optimizar el rendimiento de la máquina para casos de uso específico. Desde entonces, **ktune** ha mejorado en tal grado que ahora la usamos como parte estática de nuestro marco de trabajo de ajustes generales. Se utiliza principalmente en los perfiles predefinidos en [Sección 2.5.2, “Tuned-adm”](#).

Instale el paquete **tuned** y sus scripts de **systemtap** asociados con el comando:

```
yum install tuned
```

Al instalar el paquete **tuned** también establece un archivo de configuración en **/etc/tuned.conf** y activa el perfil predeterminado.

Inicie **tuned** al ejecutar:

```
service tuned start
```

Para iniciar **tuned** cada vez que la máquina arranque, ejecute:

```
chkconfig tuned on
```

Tuned por sí misma tiene opciones adicionales que usted puede ejecutar manualmente. Las opciones disponibles son:

-d, --daemon

Inicia **tuned** como daemon en lugar de en primer plano.

-c, --conf file

Usa un archivo de configuración con el nombre y ruta especificados, por ejemplo, **--conf file=/etc/tuned2.conf**. El predeterminado es **/etc/tuned.conf**.

-D, --debug

Usa el nivel superior de registro.

2.5.1. El archivo **tuned.conf**

El archivo **tuned.conf** contiene las configuraciones de **tuned**. Por defecto, se localiza en **/etc/tuned.conf**, pero puede especificar un nombre y dirección diferentes si inicia **tuned** con la opción **--conf file**.

El archivo config siempre debe contener una sección **[main]** que define los parámetros generales para **tuned**. El archivo contiene una sección para cada complemento.

La sección **[main]** contiene las siguientes opciones:

interval

El intervalo en el que **tuned** debe monitorizar y ajustar el sistema, en segundos. El predeterminado es el valor **10**.

verbose

Especifica si la salida debe ser detallada. El valor predeterminado es **Falso**.

logging

Especifica la prioridad mínima de mensajes que deben ser registrados. En orden descendente, los valores permitidos son: **crítico**, **error**, **advertencia**, **info** y **depuración**. El valor predeterminado es **info**.

logging_disable

Especifica la prioridad máxima de mensajes a ser registrados; cualquier mensaje con esta prioridad o menor no será registrado. En orden descendente, los valores permitidos son: **crítico**, **error**, **advertencia**, **info**, y **depuración**. El valor **notset** inhabilita esta opción.

Cada complemento tiene su propia sección, especificada con el nombre del complemento en paréntesis cuadrados; por ejemplo: **[CPUtuning]**. Cada complemento puede tener sus propias opciones, pero lo siguiente se aplica a todos los complementos:

enabled

Especifica si el complemento está habilitado o no. El valor predeterminado es **Verdadero**.

verbose

Especifica si la salida debe ser detallada. Si no se establece para este complemento, el valor se heredará de **[main]**.

logging

Especifica la prioridad mínima de mensajes a ser registrados. Si no se establece para este complemento, el valor se heredará de **[main]**.

A continuación, una muestra del archivo config:

```
[main]
interval=10
pidfile=/var/run/tuned.pid
logging=info
logging_disable=notset

# Disk monitoring section

[DiskMonitor]
enabled=True
```

```
logging=debug

# Disk tuning section

[DiskTuning]
enabled=True
hdparm=False
alpm=False
logging=debug

# Net monitoring section

[NetMonitor]
enabled=True
logging=debug

# Net tuning section

[NetTuning]
enabled=True
logging=debug

# CPU monitoring section

[CPUMonitor]
# Enabled or disable the plugin. Default is True. Any other value
# disables it.
enabled=True

# CPU tuning section

[CPUTuning]
# Enabled or disable the plugin. Default is True. Any other value
# disables it.
enabled=True
```

2.5.2. Tuned-adm

Por lo general, una auditoría y análisis de un sistema consume mucho tiempo y probablemente no vale la pena el ahorro de unos pocos vatios adicionales al hacerla. Anteriormente la única alternativa era usar los predeterminados. Ahora, Red Hat Enterprise Linux 6 incluye perfiles independientes para casos de uso específicos como una alternativa entre esos extremos, junto con la herramienta **tuned-adm** que le permite cambiar entre perfiles fácilmente en la línea de comandos. Red Hat Enterprise Linux 6 incluye un número de perfiles predefinidos para casos de uso típicos que usted puede seleccionar y activar con el comando **tuned-adm**, pero también usted mismo puede crear, modificar o borrar perfiles.

Para listar todos los perfiles disponibles e identificar el perfil actual activo, ejecute:

```
tuned-adm list
```

Para mostrar el perfil actualmente activo, ejecute:

```
tuned-adm active
```

Para cambiar uno de los perfiles, ejecute:

```
tuned-adm profile profile_name
```

por ejemplo:

```
tuned-adm profile server-powersave
```

Para desactivar todos los ajustes:

```
tuned-adm off
```

Cuando instale por primera vez **tuned**, el perfil **predeterminado** se activará. Red Hat Enterprise Linux 6 también incluye los siguiente perfiles predeterminados:

default

Es el perfil predeterminado de ahorro de energía. Tiene el más bajo impacto en ahorro de energía de los perfiles disponibles y solo habilita CPU y complementos de disco de **tuned**.

desktop-powersave

Es un perfil de ahorro de energía dirigido a sistemas de escritorio. Habilita ahorro de energía ALPM para adaptadores de host SATA (consulte la [Sección 3.6, “Administración de energía de enlace agresivo”](#)) y complementos de CPU, Ethernet, y disco de **tuned**.

server-powersave

Es un perfil de ahorro de energía dirigido a los sistemas de servidores. Habilita el ahorro de energía ALPM para adaptadores de host SATA, inhabilita polling de CD-ROM a través de **HAL** (consulte la página man hal-disable-polling) y activa la CPU y complementos de disco de **tuned**.

laptop-ac-powersave

Es un perfil de ahorro de energía de medio impacto dirigido a portátiles en AC. Habilita ahorro de energía ALPM para adaptadores de host SATA, ahorro de energía WIFI, como también complementos de CPU, Ethernet y disco de **tuned**.

laptop-battery-powersave

Es un perfil de ahorro de energía de alto impacto dirigido a portátiles que se ejecutan con batería. Activa todos los mecanismos de ahorro de energía de los perfiles anteriores y habilita el programador de ahorro de energía multi-núcleos para sistemas despertadores bajos,. Además, se asegura de que el gobernador ondemand esté activo y que el ahorro de energía AC97 esté habilitado. Este perfil se puede usar para ahorrar la máxima cantidad de energía en cualquier clase de sistema, no solo portátiles con batería. El sacrificio a cambio es un notable impacto en rendimiento, específicamente en latencia de disco y E/S de red.

throughput-performance

Es un perfil de servidor para ajuste de rendimiento. Inhabilita los mecanismos de ahorro de energía **tuned** y **ktune**, habilita la configuración de **sysctl** que mejora el rendimiento de procesamiento de su E/S de red y disco y cambia al **programador de fecha límite**.

latency-performance

Es un perfil de servidor para ajuste de rendimiento de latencia típica. Inhabilita los mecanismos de ahorro de energía de **tuned** y **ktune** habilita la configuración de **sysctl** para mejorar el rendimiento de latencia de su E/S de red.

Todos los perfiles están almacenados en subdirectorios independientes bajo **/etc/tune-profiles**. Por lo tanto, **/etc/tune-profiles/desktop-powersave** contiene todos los archivos y configuraciones necesarios para ese perfil. Cada uno de estos directorios contiene hasta cuatro archivos:

tuned.conf

la configuración para el servicio ajustado como activo para este perfil.

sysctl.ktune

la configuración **sysctl** utilizada por **ktune**. El formato es idéntico al archivo **/etc/sysconfig/sysctl** (consulte las páginas man **sysctl** y **sysctl.conf**).

ktune.sysconfig

el archivo de configuración de **ktune**, típicamente **/etc/sysconfig/ktune**.

ktune.sh

un script de shell estilo **init** utilizado por **ktune** el cual puede ejecutar comandos específicos durante el inicio del sistema para ajustar el sistema.

La forma más fácil de iniciar un nuevo perfil es copiando uno existente. El perfil **laptop-battery-powersave** ya contiene un set de ajustes muy variados y es por lo tanto, útil como punto de partida. Simplemente copie todo el directorio al nuevo nombre de perfil así:

```
cp -a /etc/tune-profiles/laptop-battery-powersave/ /etc/tune-profiles/myprofile
```

Modifique los archivos en el nuevo perfil para que coincidan con sus requerimientos personales. Por ejemplo, si usted requiere la detección de cambios de CD, puede desactivar esa optimización quitando el comentario de la línea apropiada en **ktune.sh** script:

```
# Disable HAL polling of CDROMS
# for i in /dev/scd*; do hal-disable-polling --device $i; done > /dev/null
2>&1
```

2.6. DEVICEKIT DE ENERGÍA Y DEVKIT DE ENERGÍA

En Red Hat Enterprise Linux 6 el **DeviceKit de energía** asume las funciones de administrador de energía que hacían parte de **HAL** y algunas de las funciones que hacían parte del **Gestor de energía de GNOME** en los lanzamientos anteriores de Red Hat Enterprise Linux (consulte también la [Sección 2.7, “El Gestor de energía de GNOME”](#)). El **DeviceKit de energía** proporciona un daemon, un API, y un set de herramientas de línea de comandos. Cada fuente de alimentación en el sistema está representada como un dispositivo, ya sea un dispositivo físico o no. Por ejemplo, una batería de portátil y una fuente de alimentación de corriente alterna(CA) y ambos representados como dispositivos.

Puede acceder a las herramientas de línea de comandos con el comando **devkit -power** y las siguientes opciones:

--enumerate, -e

Presenta una ruta de objeto para cada dispositivo de energía en el sistema, por ejemplo:

```
/org/freedesktop/DeviceKit/power/devices/line_power_AC
/org/freedesktop/UPower/DeviceKit/power/battery_BAT0
```

--dump, -d

Presenta los parámetros para todos los dispositivos de energía en el sistema.

--wakeups, -w

Presenta los despertadores de CPU en el sistema.

--monitor, -m

Monitoriza el sistema para los cambios a los dispositivos de energía, por ejemplo, la conexión o desconexión de una fuente de alimentación de CA o el agotamiento de una pila. Presione **Ctrl+C** para detener la monitorización del sistema.

--monitor-detail

Monitoriza el sistema para cambios a dispositivos de energía, por ejemplo, la conexión o desconexión de una fuente de alimentación de CA o el agotamiento de una pila. La opción **--monitor-detail** presenta más información que la opción **--monitor**. Presione **Ctrl+C** para detener la monitorización del sistema.

--show-info *object_path*, -i *object_path*

Presenta toda la información disponible para una ruta de objeto determinada. Por ejemplo, para obtener información sobre la batería de su sistema representada por la ruta de objeto **/org/freedesktop/UPower/DeviceKit/power/battery_BAT0**, ejecute:

```
devkit-power -i /org/freedesktop/UPower/DeviceKit/power/battery_BAT0
```

2.7. EL GESTOR DE ENERGÍA DE GNOME

El **Gestor de energía de GNOME** es un daemon que está instalado como parte del escritorio de GNOME. La mayoría de la funciones de administración de energía que el **Gestor de energía de GNOME** proporcionaba en versiones anteriores de Red Hat Enterprise Linux se convirtió en parte de **DeviceKit-power** en Red Hat Enterprise Linux 6 (consulte, la [Sección 2.6, “DeviceKit de energía y devkit de energía”](#)). Sin embargo, el **Gestor de energía de GNOME** es aún una interfaz para la funcionalidad. A través de una mini-aplicación en la bandeja del sistema, el **Gestor de energía de GNOME** notifica sus cambios en el estatus de energía del sistema; por ejemplo, un cambio de la batería a la energía de corriente alterna. También reporta el estado de la batería y le avisa cuando la energía de la batería está baja.

El **Gestor de energía de GNOME** también le permite establecer algunas configuraciones de administración de energía. Para acceder a estas configuraciones, haga clic en el icono **Gestor de energía de GNOME** en la bandeja del sistema, luego haga clic en **Preferencias**

La pantalla de **Preferencias de administración de energía** contiene las siguientes pestañas:

- En energía de CA
- En energía de batería
- General

Use las pestañas **En energía CA** y **En energía de batería** para especificar cuánto tiempo debe pasar antes de que la pantalla se apaga en un sistema inactivo, cuánto tiempo debe pasar antes de que un sistema inactivo se duerma, y si el sistema debe dejar de girar los discos duros cuando no estén en uso. La pestaña **En energía de batería** también le permite establecer el brillo de la pantalla y elegir una conducta para un sistema con una batería críticamente baja. Por ejemplo, el **Gestor de energía de GNOME**, por defecto, hace al sistema invernarse cuando su nivel de pila alcanza un nivel bajo crítico. Use la pestaña **General** para establecer las conductas para el botón de energía (física) y el botón de suspender en su sistema y especifique las circunstancias bajo las cuales el icono del **Gestor de energía de GNOME** deberá aparecer en la bandeja del sistema.

2.8. OTROS MEDIOS DE AUDITORÍA

Red Hat Enterprise Linux 6 ofrece algunas herramientas más con las cuales puede realizar auditorías y análisis. La mayoría de ellas pueden utilizarse como fuentes suplementarias de información en caso de que desee verificar lo que ya ha descubierto o en caso de que necesite información en profundidad sobre algunas partes. La mayoría de estas herramientas también se utilizan para ajustar el rendimiento. Ellas incluyen:

vmstat

vmstat proporciona información detallada sobre procesos, memoria, búsqueda, E/S de bloque, trampas y actividad de CPU. Utilícelo para ver más de cerca lo que el sistema global hace y dónde está ocupado.

iostat

iostat es similar a **vmstat**, pero solo para dispositivos de bloque de E/S. También proporciona estadísticas y una salida más detallada.

blktrace

blktrace es un programa de rastreo de E/S de bloque. Analiza la información en bloques únicos asociados con aplicaciones. Es muy útil en combinación con **diskdevstat**.

CAPÍTULO 3. INFRAESTRUCTURA CENTRAL Y MECÁNICA

3.1. ESTADOS DE CPU INACTIVOS

Las CPU con arquitectura x86 soportan estados en los que partes de la CPU están desactivadas o se ejecutan en configuraciones de rendimiento inferior. Dichos estados, conocidos como *estados-C*, permiten a los sistemas ahorrar energía al desactivar parcialmente las CPU que no están en uso. Los estados-C se enumeran de C0 en adelante, con números superiores que representan funcionalidad de CPU disminuida y mayor ahorro de energía. Los estados-C de un determinado número son bastante similares a través de procesadores, aunque los detalles exactos de la configuración de una función específica del estado puede variar entre familias de procesadores. Los estados-C se definen de así:

C0

El estado en operación o en ejecución. En este estado, la CPU está funcionando y no esta inactiva de ninguna manera.

C1, Halt

Un estado en el que el procesador no está ejecutando instrucciones, pero típicamente no está en un estado de energía bajo. La CPU puede continuar procesando sin ninguna demora. Todos los procesadores que ofrecen estados-C necesitan soportar dicho estado. Los procesadores Pentium 4 soportan un estado C1 mejorado llamado C1E que en efecto es un estado de bajo consumo energético.

C2, Stop-Clock

Un estado opcional en el que el reloj está congelado para este procesador, pero que mantiene un estado completo para sus registradores y memorias cache, por lo tanto, después de reiniciar el reloj puede inmediatamente reiniciar el procesamiento. Éste es un estado opcional.

C3, Dormido

Un estado opcional en el que el procesador se va a dormir y no necesita mantener su memoria cache actualizada. Despertar de este estado toma un tiempo considerablemente mayor que desde C2.

Las CPU Intel recientes con la microarquitectura "Nehalem" representan un nuevo estado-C, C6, el cual puede reducir a cero el voltaje proporcionado de una CPU, pero por lo general, reduce el consumo de energía entre el 80 y 90 por ciento. El kernel en Red Hat Enterprise Linux 6 incluye optimizaciones para este nuevo estado-C.

3.2. USO DE GOBERNADORES CPUFREQ

Una de las formas más efectivas para reducir el consumo de energía y la salida de calor en su sistema es utilizar CPUfreq. CPUfreq — también se conoce como escalabilidad de velocidad de CPU — permite el ajuste inmediato de la velocidad del reloj del procesador. De esta manera, el sistema puede ejecutar a una velocidad reducida de reloj para ahorrar energía. El gobernador de CPUfreq define las reglas de cambio de frecuencias, ya sea a una velocidad de reloj menor o mayor, y cuándo cambiar las frecuencias.

El gobernador define las características de energía de la CPU del sistema que a su vez afectan el rendimiento de la CPU. Cada gobernador tiene su propia conducta, propósito e idoneidad en términos de carga de trabajo. Esta sección describe cómo elegir y configurar un gobernador de CPUfreq, las

características de cada gobernador y la clase de carga de trabajo para la cual cada gobernador es adecuado.

3.2.1. Tipos de gobernadores CPUfreq

Esta sección lista y describe los distintos tipos de gobernadores CPUfreq disponibles en Red Hat Enterprise Linux 6.

cpufreq_performance

El Gobernador de Rendimiento fuerza a la CPU a usar la frecuencia de reloj más alta posible. Esta frecuencia se establecerá de forma estática y no cambiará. Como tal, este gobernador *no ofrece beneficio de ahorro de energía*. Solamente es apropiado para horas de carga de trabajo pesado y aún así solo cuando la CPU está rara vez (o nunca) inactiva.

cpufreq_powersave

En contraste, el gobernador Powersave fuerza a la CPU a usar la frecuencia de reloj más baja posible. Esta frecuencia se establecerá de forma estática y no cambiará. Como tal, este gobernador ofrece ahorro máximo de energía, pero a expensas del *rendimiento de CPU más bajo*.

El término "powersave" puede algunas veces ser engañoso, ya que (en principio) una CPU lenta con carga completa consume más energía que una CPU rápida que no esté cargada. Entonces, aunque puede ser recomendable establecer la CPU para utilizar el gobernador Powersave durante los periodos en que se espera baja actividad, cualquier carga alta inesperada durante ese tiempo puede hacer que el sistema consuma más energía.

El gobernador Powersave es más que una "limitante de velocidad" para la CPU y un "ahorrador de energía". Es más útil en sistemas y entornos donde el sobrecalentamiento puede ser un problema.

cpufreq_ondemand

El gobernador Ondemand es un gobernador dinámico que permite a la CPU alcanzar la máxima frecuencia de reloj cuando el sistema está inactivo. Aunque esto permite al sistema ajustar el consumo de energía según la carga del sistema, lo hace a expensas de *la latencia entre el cambio de frecuencia*. Como resultado, la latencia puede alterar los beneficios de ahorro de energía /rendimiento ofrecidos por el gobernador Ondemand si el sistema cambia, con demasiada frecuencia, entre inactivo y carga de trabajo pesada.

Para la mayoría de sistemas, el gobernador Ondemand puede proporcionar el mejor compromiso entre emisión de calor, consumo de energía, rendimiento y gestión. Cuando el sistema solo está ocupado en momentos del día específicos, el gobernador Ondemand cambiará automáticamente entre frecuencia máxima y mínima según la carga sin ninguna otra intervención.

cpufreq_userspace

El gobernador Userspace le permite establecer la frecuencia a programas de espacio de usuario (o cualquier proceso que esté ejecutándose como root). Este gobernador se utiliza junto con el daemon **cpuspeed**. De todos los gobernadores, Userspace es el más adaptable; y dependiendo de cómo se configure, puede ofrecer el mejor balance entre rendimiento y consumo para su sistema.

cpufreq_conservative

Al igual que el gobernador Ondemand, el gobernador Conservative ajusta la frecuencia de reloj según el uso. Sin embargo, mientras el gobernador Ondemand lo hace de una manera agresiva (es decir, desde lo máximo a lo mínimo y viceversa), el gobernador Conservative cambia de frecuencias gradualmente.

Es decir que el gobernador Conservative se ajustará a una frecuencia de reloj que estime correcta para la carga, en lugar de elegir simplemente entre máxima y mínima. Aunque esto podría proporcionar un

significativo ahorro en consumo de energía, lo hace siempre a una *latencia mayor* que la del gobernador Ondemand.



NOTA

Puede habilitar al gobernador mediante trabajos **cron**. Esto le permite automáticamente establecer gobernadores específicos durante horas específicas del día. Como consecuencia, puede especificar un gobernador de baja frecuencia durante tiempos de inactividad (por ejemplo después de horas de trabajo) y retornar a un gobernador de frecuencia más alta durante horas de carga de trabajo pesada.

Para obtener instrucciones sobre cómo habilitar un determinado gobernador, consulte el [Procedimiento 3.2, “Cómo habilitar un gobernador CPUfreq”](#) en la [Sección 3.2.2, “Configuración de CPUfreq”](#).

3.2.2. Configuración de CPUfreq

Antes de seleccionar y configurar un gobernador CPUfreq, se necesita añadir el controlador de CPUfreq.

Procedimiento 3.1. Cómo añadir el controlador CPUfreq

1. Utilice el siguiente comando para ver cuáles controladores de CPUfreq están disponibles en su sistema:

```
ls /lib/modules/[versión de
kernel]/kernel/arch/[arquitectura]/kernel/cpu/cpufreq/
```

2. Use **modprobe** para añadir el controlador CPUfreq apropiado.

```
modprobe [controlador de CPUfreq]
```

Al usar el comando anterior, asegúrese de eliminar el sufijo de nombre de archivo **.ko**



IMPORTANTE

Al escoger un controlador CPUfreq apropiado, siempre elija **acpi-cpufreq** sobre **p4-clockmod**. Aunque el uso del controlador **p4-clockmod** reduce la frecuencia del reloj de una CPU, no reduce el voltaje. Por otra parte, **acpi-cpufreq** reduce el voltaje junto con la frecuencia del reloj de la CPU, permitiendo menos consumo de energía y salida de calor para cada reducción de unidad en rendimiento..

3. Una vez el controlador CPUfreq esté configurado, se podrá ver el sistema actual que está gobernando, mediante:

```
cat /sys/devices/system/cpu/cpu0/cpufreq/scaling_governor
```

También puede ver qué gobernadores están disponibles para una CPU específica, mediante:

```
cat /sys/devices/system/cpu/[cpu ID]/cpufreq/scaling_available_governors
```

Algunos gobernadores de CPUfreq pueden no estar disponibles para su uso. En este caso, utilice **modprobe** para añadir los módulos de kernel necesarios para habilitar el gobernador CPUfreq específico que desee utilizar. Dichos módulos de kernel están disponibles en `/lib/modules/[versión de kernel]/kernel/drivers/cpufreq/`.

Procedimiento 3.2. Cómo habilitar un gobernador CPUfreq

1. Si el gobernador específico no aparece en la lista como disponible, utilice **modprobe** para habilitar el gobernador que desee utilizar. Por ejemplo, si el gobernador **ondemand** no está disponible para su CPU, use el siguiente comando:

```
modprobe cpufreq_ondemand
```

2. Cuando un gobernador esté listado como disponible para su CPU, usted podrá habilitarlo mediante:

```
echo [gobernador] > /sys/devices/system/cpu/cpu0/cpufreq/scaling_governor
```

3.2.3. Ajuste de la política CPUfreq y Velocidad

Una vez haya escogido el gobernador CPUfreq, podrá ajustar la velocidad de cada una de las CPU con los ajustables que se hallan en `/sys/devices/system/cpu/[cpu ID]/cpufreq/`. Dichos ajustadores son:

- **cpufreq_info_min_freq** — Muestra la frecuencia operativa mínima disponible de CPU' (en KHz).
- **cpufreq_info_max_freq** — Muestra la frecuencia operativa máxima disponible de CPU' (en KHz).
- **cpufreq_scaling_driver** — Muestra qué controlador CPUfreq se utiliza para establecer la frecuencia en esta CPU.
- **cpufreq_scaling_available_governors** — Muestra los gobernadores de CPUfreq disponibles en este kernel. Si desea utilizar un gobernador de CPUfreq que no esté listado en este archivo, consulte el [Procedimiento 3.2, “Cómo habilitar un gobernador CPUfreq”](#) en la [Sección 3.2.2, “Configuración de CPUfreq”](#) para obtener instrucciones sobre cómo hacerlo.
- **cpufreq_scaling_governor** — Muestra el gobernador CPUfreq que está en uso. Para utilizar un gobernador diferente, utilice `echo [governor] > /sys/devices/system/cpu/[cpu ID]/cpufreq/scaling_governor` (consulte el [Procedimiento 3.2, “Cómo habilitar un gobernador CPUfreq”](#) en la [Sección 3.2.2, “Configuración de CPUfreq”](#) para obtener mayor información).
- **cpufreq_info_cur_freq** — Muestra la velocidad actual de la CPU (en KHz).
- **cpufreq_scaling_available_frequencies** — Lista las frecuencias disponibles para la CPU, en KHz.
- **cpufreq_scaling_min_freq** y **cpufreq_scaling_max_freq** — Establecen los *límites de política* de la CPU, en KHz.



IMPORTANTE

Al establecer límites de políticas, debe configurar **scaling_max_freq** antes de **scaling_min_freq**.

- **affected_cpus** — Lista las CPU que requieren software de coordinación de frecuencia.
- **scaling_setspeed** — Utilizado para cambiar la velocidad del reloj de la CPU, en KHz. Solo puede establecer una velocidad entre los límites de políticas de las CPU (como en **scaling_min_freq** y **scaling_max_freq**).

Para ver el valor actual de cada ajustable, utilice **cat [tunable]**. Por ejemplo, para ver la velocidad actual de `cpu0` (en KHz), use:

```
cat /sys/devices/system/cpu/cpu0/cpufreq/cpuinfo_cur_freq.
```

Para cambiar el valor de cada ajustable, use **echo [value]> /sys/devices/system/cpu/[cpu ID]/cpufreq/[tunable]**. Por ejemplo, para establecer la mínima velocidad de `cpu0` a 360 KHz, use:

```
echo 360000 > /sys/devices/system/cpu/cpu0/cpufreq/scaling_min_freq
```

3.3. SUSPENDER Y REANUDAR

Cuando se suspende un sistema, el kernel le pide a los controladores almacenar sus estados y luego los descarga. Cuando se reanuda el sistema, este recarga dichos controladores, los cuales intentan reprogramar sus dispositivos. La habilidad de los controladores para realizar esta tarea, determina si el sistema se puede reanudar con éxito.

Los controladores de vídeo son en particular problemáticos en este respecto, porque la especificación de la *Configuración avanzada e Interfaz de energía* (ACPI) no requiere que firmware sea capaz de reprogramar el hardware de vídeo. Por lo tanto, a menos que los controladores de vídeo puedan programar hardware desde un estado sin inicializar completamente, pueden impedir al sistema reanudar.

Red Hat Enterprise Linux 6 incluye mayor soporte para nuevos chipsets gráficos, lo cual garantiza que la suspensión y la reanudación funcionen en un número mayor de plataformas. El soporte para chipsets de NVIDIA ha mejorado bastante; en particular para las series GeForce 8800.

3.4. KERNEL SIN INTERVALOS

Anteriormente, el kernel de Linux interrumpía periódicamente cada CPU en un sistema a una frecuencia predeterminada — 100 Hz, 250 Hz, ó 1000 Hz, según la plataforma. El kernel le preguntaba a la CPU por los procesos que estaba ejecutando y utilizaba los resultados para contabilidad de procesos y balance de cargas. Conocido como el *temporizador de intervalo*, el kernel realizaba esta interrupción independientemente del estado de energía de la CPU. Por lo tanto, incluso una CPU inactiva respondía hasta 1000 de estas solicitudes por segundo. En sistemas que implementaban medidas de ahorro de energía para CPU inactivas, el temporizador de intervalos evitaba que la CPU permaneciera inactiva lo suficiente como para que el sistema se beneficiara de estos ahorros de energía

El kernel en Red Hat Enterprise Linux 6 ejecuta *sin intervalos*: es decir, reemplaza las interrupciones periódicas antiguas por las interrupciones en demanda. Por lo tanto, a las CPU inactivas se les permite permanecer inactivas hasta que haya una nueva tarea en cola para procesar y las CPU que han entrado a estados de energía inferiores, pueden permanecer en esos estados por más tiempo.

3.5. ADMINISTRADOR DE ENERGÍA DE ESTADO ACTIVO

El *Administrador de energía de estado activo* (ASPM) ahorra energía en el subsistema de *Interconexión de componentes periféricos Express* (PCI Express o PCIe) al establecer el estado de baja energía para enlaces PCIe cuando los dispositivos que ellos conectan no están en uso. ASPM controla el estado de energía en ambos extremos del enlace y ahorra energía en el enlace aún cuando el dispositivo en el extremo del enlace esté en estado de energía Total.

Cuando ASPM está habilitado, la latencia de dispositivo aumenta debido al tiempo requerido para la transición del enlace entre diferentes estados de energía. ASPM tiene tres políticas para determinar los tres estados de energía:

default

establece los estados de enlace de energía de PCIe según los predeterminados especificados por firmware en el sistema (por ejemplo, BIOS). Este es el estado predeterminado para ASPM.

powersave

Configura ASPM para ahorrar energía cuando sea posible, independiente del costo de rendimiento.

performance

Inhabilita ASPM para permitir que enlaces PCIe operen con el máximo rendimiento

Las políticas ASPM se establecen en `/sys/module/pcie_aspm/parameters/policy`, pero también pueden especificarse en el momento de arranque con el parámetro de kernel `pcie_aspm`, donde `pcie_aspm=off` inhabilita a ASPM y `pcie_aspm=force` habilita a ASPM, incluso en dispositivos que no soportan ASPM.



AVISO

Si se establece `pcie_aspm=force`, el hardware que no soporte ASPM puede hacer que el sistema deje de responder. Antes de configurar `pcie_aspm=force`, verifique si todo el hardware PCIe en el sistema soporta ASPM.

3.6. ADMINISTRACIÓN DE ENERGÍA DE ENLACE AGRESIVO

Administración de energía de enlace agresivo (ALPM) es una técnica de ahorro de energía que ayuda al disco a ahorrar energía al establecer un enlace SATA al disco a baja energía durante el tiempo de inactividad (es decir, cuando no hay E/S). ALPM restablece automáticamente el enlace SATA a un estado de energía activo una vez que las solicitudes de E/S estén en la cola para ese enlace.

El ahorro de energía introducido por ALPM viene a expensas de la latencia de disco. Como tal, solo debe usar ALPM si espera que el sistema presente largos periodos de tiempo de inactividad de E/S.

ALPM solo está disponible en controladores SATA que utilizan la *Interfaz de controlador de host avanzado* (AHCI). Para obtener mayor información sobre AHCI, consulte <http://www.intel.com/technology/serialata/ahci.htm>.

Cuando está disponible, ALPM está habilitado por defecto. ALPM tiene tres modos:

min_power

Este modo establece el enlace a su mínimo estado de energía, (LETARGO) cuando no hay E/S en el disco. Dicho modo es útil para cuando se espera un periodo extendido de tiempo de inactividad.

medium_power

Este modo establece el enlace en el segundo estado más bajo de energía (PARCIAL) cuando no hay E/S en el disco. Dicho modo está diseñado para permitir transiciones en estados de enlace de energía (por ejemplo durante los momentos de alta carga de E/S intermitente y tiempo de inactividad de E/S) con el menor impacto posible en rendimiento.

El modo **medium_power** permite el enlace a la transición entre estados PARCIAL y TOTAL (es decir "ACTIVO"), dependiendo de la carga. Observe que no es posible la transición directa de un enlace PARCIAL a SLUMBER y viceversa; en este caso, ningún estado de energía puede pasar al otro sin haber hecho primero la transición al estado ACTIVO.

max_performance

ALPM está inhabilitado; el enlace no entra a ningún estado de baja cuando no hay E/S en el disco.

Para verificar si sus adaptadores de host SATA en realidad soportan ALPM, puede chequear si el archivo `/sys/class/scsi_host/host*/link_power_management_policy` existe.

**IMPORTANTE**

Al establecer ALPM a **min_power** o a **medium_power** se desactivará automáticamente la función "Conexión en caliente".

3.7. OPTIMIZACIÓN DE ACCESO DE UNIDAD DE RELATIME

El estándar POSIX requiere que los sistemas operativos mantengan metadatos de sistema de archivos que registren la última vez que cada archivo ha sido accedido. Esta marca de tiempo se denomina **atime**, y el mantenimiento de ésta requiere una serie constante de operaciones de escritura para almacenaje. Dichas escrituras mantienen dispositivos de almacenamiento y sus enlaces ocupados y encendidos. Puesto que pocas aplicaciones utilizan datos **atime**, esta actividad de dispositivo de almacenamiento desperdicia energía. La escritura a almacenamiento se presenta incluso si el archivo no fue leído desde el almacenaje, sino desde la memoria cache. Por algún tiempo, el kernel de Linux ha soportado la opción **noatime** para **mount** y no escribiría datos **atime** a sistemas de archivos montados con esta opción. Sin embargo, la desactivación de esta funcionalidad es problemática, puesto que algunas aplicaciones dependen de los datos de **atime** y fallarán si no está disponible.

El kernel utilizado en Red Hat Enterprise Linux 6 soporta otra alternativa — **relatime**. **Relatime** mantiene datos **atime**, pero no para cada vez que se acceda a un archivo. Con esta opción habilitada, los datos **atime** se escriben al disco únicamente si el archivo ha sido modificado desde la última vez que los datos **atime** fueron actualizados (**mtime**), o si el archivo fue accedido por un periodo de tiempo mayor (por defecto, un día).

Por defecto, todos los sistemas de archivos ahora se montan con **relatime** habilitada. Para suprimir esta función a través de todo el sistema, use el parámetro de arranque **default_relatime=0**. Si **relatime** está habilitada por defecto en un sistema, podrá suprimirla para un determinado sistema de archivos montando ese sistema de archivos con la opción **norelatime**. Por último, para variar la longitud predeterminada antes de que el sistema actualice los datos **atime** de un archivo, use el parámetro de arranque **relatime_interval=**, especificando el periodo en segundos. El valor predeterminado es **86400**.

3.8. LÍMITE DE ENERGÍA

Red Hat Enterprise Linux 6 es compatible con las funciones de límite de energía de hardware reciente, tales como tecnología *Para límite de energía dinámica* (DPC) de HP y el Administrador de nodos de Intel (NM). El límite de energía permite a los administradores limitar la energía consumida por los servidores, pero también permite a los administradores planificar los centros de datos de forma más eficiente, porque el riesgo de sobrecarga de fuentes de alimentación existente se disminuye considerablemente. Los administradores pueden colocar más servidores dentro del mismo espacio físico y tener confianza en que si el consumo de energía de servidor está limitado, la demanda de energía durante la carga pesada no superará la energía disponible.

Límite de energía dinámico de HP

El límite de energía dinámico es una función disponible en la selección de servidores ProLiant y BladeSystem que permite a los administradores de sistema limitar el consumo de energía de un servidor o un grupo de servidores. El límite es un límite definitivo que el servidor no superará, independientemente de su carga de trabajo actual. El límite no surte ningún efecto hasta que el servidor alcanza su límite de consumo de energía. En ese momento, el procesador de gestión ajusta estados de CPU P y la aceleración del reloj para limitar la energía consumida.

El límite de energía dinámico modifica el comportamiento de CPU independientemente del sistema operativo, sin embargo, firmware *integrated Lights-Out 2* (iLO2) de HP, permite el acceso de sistemas operativos al procesador de gestión y por lo tanto, las aplicaciones en el espacio de usuario pueden consultar el procesador de gestión. El kernel de Red Hat Enterprise Linux 6 incluye un controlador para firmware de HP iLO e iLO2, el cual permite que los programas consulten procesadores de gestión en `/hpilo/dev/d X ccb N`. El kernel también incluye una extensión de la interfaz de `sysfs hwmon` que soporta funciones de límite de energía y un controlador `hwmon` para medidores de energía ACPI 4.0 que utilizan la interfaz `sysfs`. Estas características permiten que el sistema operativo y las herramientas de espacio de usuario lean el valor configurado para el límite de energía, junto con el uso actual de energía del sistema.

Para obtener mayor información sobre el Límite de energía dinámica de HP, consulte *HP Power Capping and HP Dynamic Power Capping for ProLiant Servers*, disponible en <http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01549455/c01549455.pdf>

Administrador de Nodos Intel

El Administrador de nodos de Intel ,impone un límite de energía en sistemas, mediante el procesador de estados- P y estados-T para limitar el rendimiento de la CPU y, por tanto, el consumo de energía. Al establecer una política de gestión de energía, los administradores pueden configurar sistemas para consumir menos energía en momentos en que las cargas de sistema son bajas, por ejemplo, por la noche o los fines de semana.

El administrador nodos Intel ajusta el rendimiento de la CPU mediante la *Administración de energía y configuración dirigida del Sistema operativo* (OSPM) a través de la *Interfaz de energía y configuración avanzada* estándar. Cuando el Administrador de nodos de Intel notifica al controlador de OSPM sobre los cambios a estados-T, el controlador hace los cambios correspondientes al procesador de estados-P. Igualmente, cuando el Administrador de nodos de Intel notifica al controlador de OSPM de los cambios de estados -P, el controlador cambia a estados-T respectivamente. Estos cambios suceden automáticamente y no requieren más entrada del sistema operativo. Los administradores configuran y monitorizan al Administrador de nodos de Intel con software *Administrador de centro de datos Intel* (DCM).

Para obtener mayor información sobre el Administrador de nodos de Intel, consulte *Node Manager — A Dynamic Approach To Managing Power in The Data Center*, disponible en <http://communities.intel.com/docs/DOC-4766>

3.9. ADMINISTRACIÓN DE ENERGÍA EN GRÁFICAS MEJORADAS

Red Hat Enterprise Linux 6 ahorra energía en gráficas y muestra dispositivos al eliminar varias fuentes de consumo innecesario.

LVDS reclocking

La *Señal diferencial de bajo voltaje* (LVDS) es un sistema para llevar señales electrónicas en alambre de cobre. Una aplicación importante del sistema es la transmisión de información pixel a *pantalla de cristal líquido* (LCD) en computadores notebook. Todas las pantallas tienen *tasa de actualización* — la tasa en la cual reciben datos actualizados desde un controlador gráfico y redirigen la imagen en la pantalla. Por lo general, la pantalla recibe datos actualizados sesenta veces por segundo (una frecuencia de 60 Hz). Cuando una pantalla y controlador gráfico están conectados por LVDS, el sistema LVDS utiliza energía en cada ciclo de actualización. Cuando está inactivo, la tasa de actualización de varias pantallas LCD pueden caer a 30 Hz sin ningún efecto perceptible (a diferencia de los monitores de *tubo de rayos catódicos* (CRT), en los cuales la reducción en la tasa de actualizaciones produce un parpadeo característico). El controlador Intel para adaptadores gráficos incorporado en el kernel utilizado en Red Hat Enterprise Linux 6 realiza automáticamente este *downclocking* y ahorra cerca de 0.5 W cuando la pantalla está inactiva.

Habilitando memoria de auto-actualización

Memoria de acceso aleatorio dinámico y sincrónico (SDRAM) — utilizada para memoria de vídeo en adaptadores gráficos — se recarga miles de veces por segundo para que las células de memoria individual retengan los datos que se almacenan en ellos. Aparte de su función principal de administrar datos cuando entran y salen de la memoria, el controlador de memoria es responsable de iniciar estos ciclos de actualización. Sin embargo, SDRAM también tiene un modo de *auto actualización* de baja energía. En dicho modo, la memoria utiliza un temporizador interno para generar su propios ciclos de actualización, lo cual permite al sistema apagar el controlador de memoria sin dañar los datos que están actualmente en memoria. El kernel utilizado en Red Hat Enterprise Linux 6 puede reactivar la auto-actualización de memoria en los adaptadores gráficos Intel cuando están en inactividad, lo cual ahorra cerca de 0.8 W.

Reducción de GPU de relojes

Las unidades de procesamiento gráfico típicas (GPU) contienen relojes internos que gobiernan varias partes de su circuito interno. El kernel utilizado en Red Hat Enterprise Linux 6 puede reducir la frecuencia de algunos de los relojes internos en GPU Intel y ATI. Al reducir el número de ciclos que los componentes de GPU realizan en un tiempo determinado, se ahorra la energía que hubieran consumido en los ciclos que no tenían que actuar. El kernel reduce automáticamente la velocidad de estos relojes cuando la GPU está inactiva y la incrementa cuando la actividad de GPU aumenta. Al reducir los ciclos de reloj de GPU se pueden ahorrar hasta 5 W.

GPU powerdown

Los controladores gráficos Intel y ATI en Red Hat Enterprise Linux 6 pueden detectar cuando no hay un monitor conectado a un adaptador y por lo tanto, apagar completamente la GPU. Esta función es en particular importante para servidores que no tienen monitores conectados regularmente a ellos.

3.10. RFKILL

Muchos sistemas informáticos contienen radiotransmisores, que incluyen dispositivos Wi-Fi, Bluetooth, y 3G. Dichos dispositivos consumen energía, la cual se pierda si no se utiliza.

RFKill es un subsistema en el kernel de Linux que proporciona una interfaz a través de la cual radiotransmisores en un sistema informático se pueden solicitar, activar o desactivar. Cuando los transmisores se desactivan, pueden ser localizados en un estado en el que software puede reactivarlos (un *bloque blando*) o en el que software no pueda reactivarlos (un *bloque duro*).

El núcleo RFKill proporciona la interfaz de programación de aplicación (API) para el subsistema. Los controladores de Kernel que han sido diseñados para soportar RFKill usan esta API para registrarse al kernel, e incluyen métodos para habilitar e inhabilitar el dispositivo. Además, el núcleo RFKill proporciona notificaciones que las aplicaciones de usuario pueden interpretar y formas para que aplicaciones de usuario soliciten estados de transmisor.

La interfaz RFKill está localizada en `/dev/rfkill`, el cual contiene el estado actual de todos los radiotransmisores en el sistema. Cada dispositivo tiene su estado RFKill actual registrado en `sysfs`. Además, RFKill genera `uevents` para cada cambio de estado en un dispositivo RFKill-habilitado.

Rfkill es una herramienta de línea de comando con la cual se puedes solicitar y cambiar dispositivos de RFKill-habilitado en el sistema. Para obtener la herramienta, instale el paquete `rfkill`.

Use el comando `rfkill list` para obtener una lista de dispositivos, cada uno de los cuales tiene un *número de índice* asociado a este, iniciando en `0`. Se puede usar este número de índice para decirle a **rfkill** que bloquee o desbloquee un dispositivo, por ejemplo:

```
rfkill block 0
```

bloquea el primer dispositivo RFKill-habilitado en el sistema.

También puede usar **rfkill** para bloquear algunas categorías de dispositivos o todos los dispositivos de RFKill-habilitado. Por ejemplo:

```
rfkill block wifi
```

bloquea todos los dispositivos Wi-Fi en el sistema. Para bloquear todos los dispositivos RFKill-habilitado, ejecute:

```
rfkill block all
```

Para desbloquear dispositivos, ejecute `rfkill unblock` en lugar de `rfkill block`. Para obtener una lista completa de las categorías de dispositivos que **rfkill** puede bloquear, ejecute: `rfkill help`

3.11. OPTIMIZACIONES EN ESPACIO DE USUARIO

La reducción de la cantidad de trabajo realizada por hardware del sistema es fundamental para ahorrar energía. Por lo tanto, aunque los cambios descritos en [Capítulo 3, Infraestructura central y mecánica](#) permiten que el sistema opere en varios estados de consumo de energía reducido, las aplicaciones en espacio de usuario que solicitan trabajo innecesario del hardware del sistema evitan que el hardware entre en esos estados. Durante el desarrollo de Red Hat Enterprise Linux 6, se realizaron auditorías en las siguientes áreas para reducir demandas innecesarias de hardware:

Despertadores reducidos

Red Hat Enterprise Linux 6 utiliza un *kernel sin intervalo* (consulte la [Sección 3.4, “Kernel sin intervalos”](#)), el cual permite a la CPU permanecer en los estados más profundos de inactividad por más tiempo. Sin embargo, el *temporizador de intervalo* no es la única fuente de despertadores de CPU excesivos y las llamadas de función desde aplicaciones pueden también evitar que la CPU entre o permanezca en estados de inactividad. Las llamadas de función innecesarias se redujeron en más de 50 aplicaciones.

Almacenaje reducido y red de E/S

La entrada o salida (E/S) para dispositivos de almacenamiento y las interfaces de red obligan a

dispositivos a consumir energía. En dispositivos de almacenamiento y de red esa función reducía los estados de energía cuando estaban inactivos (por ejemplo, ALPM o ASPM), este tráfico puede impedir que el dispositivo entrar o permanecer en un Estado de inactividad y puede impedir que unidades de disco duro de girar hacia abajo cuando no esté en uso. Las demandas excesivas e innecesarias de almacenamiento de información han sido minimizadas en varias aplicaciones. En particular, aquellas demandas que impedían a unidades de disco duro girar hacia abajo.

Auditoría Initscript

Los servicios que se inicien automáticamente si se requieren o no tienen gran potencial para residuos de recursos del sistema. En su lugar, los servicios deben por defecto estar en "off" o "en demand" siempre que sea posible. Por ejemplo, el servicio de **BlueZ** que permite el soporte de Bluetooth previamente se ejecutaba automáticamente cuando se iniciaba el sistema, si el hardware de Bluetooth estaba presente o no. Initscript de **BlueZ** ahora comprueban que hardware Bluetooth está presente en el sistema antes de iniciar el servicio.

CAPÍTULO 4. CASOS DE USO

Este capítulo describe dos tipos de casos de uso para ilustrar los métodos de análisis y configuración descritos en esta guía. El primer ejemplo, observa servidores típicos y el segundo observa un portátil típico.

4.1. EJEMPLO — SERVIDOR

Un servidor estándar normal viene con todas las funciones necesarias de hardware soportadas en Red Hat Enterprise Linux 6. Lo primero que hay que tener en cuenta es la clase d

Independientemente del tipo de servidor, el rendimiento gráfico no se requiere por lo general. Por lo tanto, el ahorro de energía GPU puede dejarse encendida.

Servidor de Web

Un servidor de Web necesita E/S de red y disco. Según la velocidad de conexión externa 100 Mbit/s podría ser suficiente. Si la máquina sirve en su mayoría páginas estáticas, el rendimiento de la CPU podría no ser importante. Las opciones de administración de energía por lo tanto incluirían:

- no hay complementos de disco o de red para **tuned**.
- ALPM encendido.
- Gobernador **ondemand** encendido.
- tarjeta de red limitada a 100 Mbit/s.

Servidor de cómputo

Un servidor de cómputo principalmente necesita CPU. Las opciones de administración de energía podrían incluir:

- Según los trabajos y dónde se suceda el almacenamiento de datos, complementos de disco y red para **tuned**; o para sistemas de modo-batch, completamente activos **tuned**.
- según el uso, quizás el gobernador de **rendimiento**.

Servidor de correo

Un servidor de correo necesita la mayoría de las opciones de E/S de disco y CPU. Las opciones de administración de energía podrían incluir:

- Gobernador **ondemand** encendido, puesto que el último porcentaje de rendimiento de CPU no es importante.
- no hay complementos de disco o de red para **tuned**.
- la velocidad de red no debe ser limitada, puesto que el correo a menudo es interno y puede, por lo tanto, beneficiarse de un enlace de 1 Gbit/s o 10 Gbit/s.

Servidor de archivos

Los requisitos del Servidor de archivos son similares a los del servidor de correo, pero dependiendo del protocolo utilizado, podría requerir más rendimiento de CPU. Por lo general, servidores basados en Samba, requieren más CPU que NFS y, NFS requiere más de iSCSI. Incluso con esto, debería poder utilizar el gobernador **ondemand**.

Servidor de directorio

Un servidor de directorio generalmente tiene requisitos más bajos para E/S de disco, específicamente si está equipado con suficiente RAM. La latencia de red es importante aunque la E/S de red sea menor. Debería considerar el ajuste de latencia de red con una velocidad de enlace inferior, pero debería probarlo con cuidado en su red particular.

4.2. EJEMPLO — PORTÁTIL

Los portátiles son un lugar bastante común donde el ahorro y la administración de energía pueden realmente hacer la diferencia. Como los portátiles por diseño normalmente consumen drásticamente menos energía que las estaciones de trabajo o servidores, el potencial para ahorros absolutos es menor que para otras máquinas. Cuando está en modo de batería, sin embargo, cualquier ahorro puede ayudar a alargar unos minutos más la vida de batería de un portátil. Aunque esta sección se centra en portátiles en modo de batería, seguramente también puede aún utilizar alguno o todos estos ajustes mientras ejecuta en energía de corriente alterna.

Los ahorros para componentes únicos suelen hacer una diferencia relativamente mayor en los portátiles que en las estaciones de trabajo. Por ejemplo, una interfaz de red de 1 Gbit/s ejecutándose a 100 Mbits/s ahorra cerca de 3–4 vatios. Para un servidor típico con un consumo de cerca de 400 vatios, este ahorro es aproximadamente de 1 %. En un portátil con un consumo total de energía alrededor de 40 vatios, el ahorro de energía en solo este componente llega a 10 % del total.

Las optimizaciones de ahorro de energía en un portátil incluyen:

- Configure el BIOS del sistema para inhabilitar todo el hardware que no utilice. Por ejemplo, los puertos seriales y paralelos, los lectores de tarjeta, cámaras web, WiFi y Bluetooth entre otros candidatos.
- Atenúe el monitor en entornos más oscuros donde no sea necesaria la plena iluminación para leer la pantalla cómodamente. Utilice **Sistema+Preferencias** → **Administración de energía** en el escritorio de GNOME, **Lanzador de aplicaciones Kickoff+Equipo+Configuración del sistema+Avanzado** → **Administración de energía** en el escritorio de KDE o **Gestor de energía de GNOME** o **xbacklight** en la línea de comandos; o las teclas de función de su portátil.
- Use el perfil de **laptop-battery-powersave** de **tuned-adm** para habilitar todo un set de mecanismos de ahorro de energía. Observe que el rendimiento y la latencia para el disco duro y la interfaz de red se ven afectados.

Además o como otra alternativa, puede realizar varios ajustes pequeños a varias configuraciones del sistema:

- usar el gobernador **ondemand** (activado por defecto en Red Hat Enterprise Linux 6)
- activar modo portátil (parte del perfil de **laptop-battery-powersave**):

```
echo 5 > /proc/sys/vm/laptop_mode
```

- aumentar tiempo de descarga a disco (parte del perfil **laptop-battery-powersave**):

```
echo 1500 > /proc/sys/vm/dirty_writeback_centisecs
```

- inhabilitar nmi watchdog (parte del perfil **laptop-battery-powersave**):

```
echo 0 > /proc/sys/kernel/nmi_watchdog
```

- habilitar AC97 ahorro de energía de audio (habilitado por defecto en Red Hat Enterprise Linux 6):

```
echo Y > /sys/module/snd_ac97_codec/parameters/power_save
```

- habilitar ahorro de energía multi-núcleos (parte el perfil **laptop-battery-powersave**):

```
echo 1 > /sys/devices/system/cpu/sched_mc_power_savings
```

- habilitar auto-suspensión de USB:

```
for i in /sys/bus/usb/devices/*/power/autosuspend; do echo 1 > $i; done
```

Observe que la auto-suspensión de USB no funciona correctamente en todos los dispositivos USB.

- habilitar mínima configuración de energía para ALPM (parte del perfil **laptop-battery-powersave**):

```
echo min_power > /sys/class/scsi_host/host*/link_power_management_policy
```

- montar sistema de archivos mediante relatime (por defecto en Red Hat Enterprise Linux 6):

```
mount -o remount,relatime mountpoint
```

- activar el mejor modo de ahorro de energía para discos duros (parte del perfil **laptop-battery-powersave**):

```
hdparm -B 1 -S 200 /dev/sd*
```

- desactivar CD-ROM de polling (parte del perfil de **laptop-battery-powersave**):

```
hal-disable-polling --device /dev/scd*
```

- reducir brillo de pantalla a **50** o menos, por ejemplo:

```
xbacklight -set 50
```

- activar DPMS para inactividad de pantalla:

```
xset +dpms; xset dpms 0 0 300
```

- reducir niveles de energía Wi-Fi (parte del perfil **laptop-battery-powersave** profile):

```
for i in /sys/bus/pci/devices/*/power_level ; do echo 5 > $i ; done
```

- desactivar Wi-Fi:

```
echo 1 > /sys/bus/pci/devices/*/rf_kill
```

- limitar red de cable a 100 Mbit/s (parte del perfil de **laptop-battery-powersave**):

```
ethtool -s eth0 advertise 0x0F
```

APÉNDICE A. CONSEJOS PARA DESARROLLADORES

Cada texto de programación cubre problemas sobre asignación de memoria y rendimiento de funciones específicas. Cuando usted desarrolle su software, tenga en cuenta los aspectos que pueden aumentar el consumo de energía en los sistemas en los cuales se ejecuta el software. Aunque estas consideraciones no afectan cada línea de código, puede optimizar su código en áreas en las cuales se presentan cuellos de botella para rendimiento.

Algunas técnicas que suelen ser problemáticas incluyen:

- el uso de hilos.
- los despertadores de CPU innecesarios y el uso ineficiente de despertadores. Si debe despertar, haga todo al mismo tiempo (carrera a ralentí) y tan pronto como sea posible.
- el uso innecesario de `[f]sync()`.
- el polling activo innecesario o uso de plazos cortos o regulares. (Reaccionar a eventos en su lugar)
- no usar eficientemente despertadores.
- acceso ineficiente de disco. Usar grandes buffers para evitar el acceso frecuente a disco. Escribir un gran bloque a la vez.
- el uso ineficiente de temporizadores. Temporizadores de grupo a través de aplicaciones (o incluso a través de sistemas) si es posible.
- excesiva E/S. consumo de energía o uso de memoria (incluyendo filtraciones de memoria)
- realización innecesaria de computos.

Las siguientes secciones examinan algunas de estas áreas de una forma más detallada.

A.1. USO DE HILOS

Se cree que el uso de hilos hace que las aplicaciones tengan un rendimiento mejor y más ágil, pero no es cierto en todos los casos.

Python

Python usa el Global Lock Interpreter^[1], por lo tanto, el uso de hilos es solo útil para operaciones de E/S mayores. **Unladen-swallow**^[2] es una aplicación de Python más rápida con la cual puede optimizar su código.

Perl

Los hilos de Perl fueron creados originalmente para aplicaciones que se ejecutan en sistemas sin bifurcaciones (como los sistemas con sistemas operativos de Windows de 32 bits). En hilos de Perl, los datos se copian para cada hilo (Copia en Escritura). Los datos no son compartidos de forma predeterminada, puesto que los usuarios deben definir el nivel de datos compartidos. Para compartir datos, el módulo **threads::shared** tiene que estar incluido. Sin embargo, los datos no son solamente copiados (Copia en Escritura), sino que el módulo también crea variables vinculadas para los datos, lo cual toma aún más tiempo y es aún mucho más lenta.^[3]

C

Los hilos C comparten la misma memoria, cada hilo tiene su propia pila y el kernel no tiene que crear nuevos descriptores de archivo ni asignar nuevo espacio de memoria. C puede usar el soporte de más CPU para más hilos. Por lo tanto, para maximizar el rendimiento de sus hilos, utilice un lenguaje de bajo nivel como C o C++. Si utiliza un lenguaje de scripts, considere escribir un vínculo C. Use perfiladores para identificar las partes de su código que se ejecutan pobremente. [4]

A.2. DESPERTADORES

Muchas aplicaciones pueden examinar archivos de configuración para cambios. En algunos casos, el examen se realiza en un intervalo fijo, por ejemplo, cada minuto. Esto puede ser un problema, puesto que fuerza al disco a despertar de los spindowns. La mejor solución es hallar un buen intervalo, un buen mecanismo de verificación o revisar los cambios con **inotify** y reaccionar a eventos. **Inotify** puede verificar la variedad de cambios en un archivo o un directorio.

Por ejemplo:

```
int fd;
fd = inotify_init();
int wd;
/* checking modification of a file - writing into */
wd = inotify_add_watch(fd, "./myConfig", IN_MODIFY);
if (wd < 0) {
    inotify_cant_be_used();
    switching_back_to_previous_checking();
}
...
fd_set rdfs;
struct timeval tv;
int retval;
FD_ZERO(&rdfs);
FD_SET(0, &rdfs);

tv.tv_sec = 5;
value = select(1, &rdfs, NULL, NULL, &tv);
if (value == -1)
    perror(select);
else {
    do_some_stuff();
}
...
```

La ventaja de este enfoque es la variedad de verificaciones que se pueden realizar.

La principal limitante es que solo un número limitado de relojes está disponible en el sistema. El número puede obtenerse en `/proc/sys/fs/inotify/max_user_watches` y aunque se puede cambiar, no es recomendable. Además, si **inotify** falla, el código tiene que resguardarse a un método de verificación diferente, lo cual suele significar muchas apariciones de `#if #define` en el código fuente.

Para obtener mayor información sobre **inotify**, consulte la página `man inotify`.

A.3. FSYNC

Fsync se conoce como una operación de E/S costosa, pero esto no es totalmente cierto. Por ejemplo, consulte el artículo de Theodore Ts *No le tema a fsync!* [5] y la discusión que lo acompaña.

Firefox llamaba a la biblioteca **sqlite** cada vez que el usuario hacía clic en un enlace para ir a una nueva página. **Sqlite** llamaba a **fsync** y debido a las configuraciones del sistema de archivos (principalmente ext3 con datos en modo ordenado), había más latencia cuando nada sucedía. Esto podía tardar un largo tiempo (más de 30 segundos) si otro proceso estaba copiando un archivo grande al mismo tiempo.

Sin embargo, en otros casos, donde **fsync** wasn't no se utilizaba, se presentaban problemas con el cambio al sistema de archivos ext4. Ext3 se establecía de modo datos-ordenados, lo cual borraba la memoria cada pocos segundos y le ahorraba al disco. Sin embargo, con ext4 y modo_portátil, el intervalo entre almacenamiento utilizaba más tiempo y los datos podían perderse cuando el sistema inesperadamente era apagado. Ahora ext4 está corregido, pero debe aún considerarse cuidadosamente el diseño de sus aplicaciones y usar **fsync** como es apropiado.

El siguiente ejemplo de lectura y escritura en un archivo de configuración muestra cómo se puede hacer una copia de seguridad de un archivo y cómo se pueden perder datos:

```
/* open and read configuration file e.g. ~/.kde/myconfig */
fd = open("./kde/myconfig", O_WRONLY|O_TRUNC|O_CREAT);
read(myconfig);
...
write(fd, bufferOfNewData, sizeof(bufferOfNewData));
close(fd);
```

Un mejor enfoque sería:

```
open("./kde/myconfig", O_WRONLY|O_TRUNC|O_CREAT);
read(myconfig);
...
fd = open("./kde/myconfig.suffix", O_WRONLY|O_TRUNC|O_CREAT);
write(fd, bufferOfNewData, sizeof(bufferOfNewData));
fsync; /* paranoia - optional */
...
close(fd);
rename("./kde/myconfig", "./kde/myconfig~"); /* paranoia - optional */
rename("./kde/myconfig.suffix", "./kde/myconfig");
```

[1] <http://docs.python.org/c-api/init.html#thread-state-and-the-global-interpretor-lock>

[2] <http://code.google.com/p/unladen-swallow/>

[3] http://www.perlmonks.org/?node_id=288022

[4] <http://people.redhat.com/drepper/lt2009.pdf>

[5] <http://thunk.org/tytso/blog/2009/03/15/dont-fear-the-fsync/>

APÉNDICE B. HISTORIAL DE REVISIONES

Revisión 1.0-10.400 Rebuild with publican 4.0.0	2013-10-31	Rüdiger Landmann
Revisión 1.0-10 Rebuild for Publican 3.0	2012-07-18	Anthony Towns
Revisión 1.0-2 Corrección de errores tipográficos menores	Fri Oct 22 2010	Rüdiger Landmann
Revisión 1.0-1 remoción de la etiqueta "borrador"	Thu Oct 7 2010	Rüdiger Landmann
Revisión 1.0-0 Lanzamiento de GA	Thu Oct 7 2010	Rüdiger Landmann