



# Red Hat Enterprise Linux 7 Guide de gestion de l'alimentation

---

Gestion de l'alimentation sur Red Hat Enterprise Linux 7

Red Hat Inc.  
Jacquelynn East  
Jack Reed

Marie Doleželová  
Don Domingo

Jana Heves  
Rüdiger Landmann



## Gestion de l'alimentation sur Red Hat Enterprise Linux 7

Marie Doleželová  
Red Hat Customer Content Services  
mdolezel@redhat.com

Jana Heves  
Red Hat Customer Content Services

Jacquelynn East  
Red Hat Customer Content Services

Don Domingo  
Red Hat Customer Content Services

Rüdiger Landmann  
Red Hat Customer Content Services

Jack Reed  
Red Hat Customer Content Services

Red Hat Inc.

## Notice légale

Copyright © 2016 Red Hat Inc..

This document is licensed by Red Hat under the [Creative Commons Attribution-ShareAlike 3.0 Unported License](https://creativecommons.org/licenses/by-sa/3.0/). 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.

## Résumé

Ce document explique comment gérer l'alimentation de systèmes Red Hat Enterprise Linux 7 de manière efficace. Les sections suivantes traitent de différentes techniques qui réduisent la consommation d'énergie (pour serveurs et ordinateurs portables), ainsi que la manière dont chaque technique affecte la performance générale de votre système.

---

## Table des matières

<b>Chapitre 1. Vue d'ensemble</b> .....	<b>2</b>
1.1. Importance de la gestion de l'alimentation	2
1.2. Bases de la gestion de l'alimentation	3
<b>Chapitre 2. Audit et analyse de la gestion de l'alimentation</b> .....	<b>5</b>
2.1. Aperçu général - Auditing et Analyse	5
2.2. PowerTOP	5
2.3. Diskdevstat et netdevstat	8
2.4. Battery Life Tool Kit	11
2.5. UPower	12
2.6. Gestionnaire d'alimentation GNOME	13
2.7. Autres outils d'auditing	13
<b>Chapitre 3. Infrastructure principale et mécanismes</b> .....	<b>15</b>
3.1. États inactifs du CPU	15
3.2. Utilisation des gouverneurs CPUfreq	15
3.3. Moniteurs CPU	18
3.4. Stratégies d'économie d'énergie de processeur	19
3.5. Suspension et reprise	19
3.6. Gestion de l'alimentation des périphériques en cours d'exécution	20
3.7. Gestion de l'alimentation en état actif	20
3.8. Aggressive Link Power Management	21
3.9. Optimisation d'accès au disque Relatime	22
3.10. Limitation de la consommation	23
3.11. Gestion améliorée de l'alimentation des graphismes	24
3.12. RFKill	24
<b>Chapitre 4. Cas d'utilisation</b> .....	<b>26</b>
4.1. Exemple — Serveur	26
4.2. Exemple — Ordinateur portable	27
<b>Annexe A. Astuces pour les développeurs</b> .....	<b>29</b>
A.1. Utilisation de threads	29
A.2. Réveils	30
A.3. Fsync	31
<b>Annexe B. Historique des versions</b> .....	<b>33</b>

## Chapitre 1. Vue d'ensemble

La gestion de l'alimentation fut l'un des points principaux sur lequel nous nous sommes concentrés lors des améliorations apportées à Red Hat Enterprise Linux 7. Limiter la consommation d'électricité des systèmes informatiques est l'un des aspects les plus importants de l'*IT verte* (informatique respectueuse de l'environnement), un ensemble de considérations qui englobe aussi l'utilisation de matériaux recyclables, l'impact sur l'environnement de la production de matériel, et la prise de conscience de l'environnement dans la conception et le déploiement des systèmes. Dans ce document, nous vous offrons des conseils et informations concernant la gestion de l'alimentation de vos systèmes exécutant Red Hat Enterprise Linux 7.

### 1.1. Importance de la gestion de l'alimentation

Au cœur de la gestion de l'alimentation figure la compréhension de l'optimisation efficace de la consommation d'énergie pour chaque composant d'un système. Ceci englobe l'étude des différentes tâches effectuées par votre système et la configuration de chaque composant pour vous assurer que sa performance est adéquate pour la tâche à entreprendre.

Les principales motivations pour la gestion de l'alimentation sont :

- » la réduction de la consommation d'énergie en général dans le but de réduire les coûts

Une gestion de l'alimentation correcte résulte en :

- » une réduction de chaleur pour les serveurs et les centres de calcul
- » des coûts secondaires réduits, y compris le refroidissement, l'espace, les câbles, les générateurs, et les *systèmes d'alimentation sans coupure* (UPS)
- » une augmentation de l'autonomie de la batterie pour les ordinateurs portables
- » une baisse de la production de dioxyde de carbone
- » conformité avec la réglementation du gouvernement ou avec les exigences légales concernant l'IT verte, par exemple Energy Star
- » conformité avec la ligne de conduite de la compagnie sur les nouveaux systèmes

Par définition, la baisse de consommation d'énergie pour un composant spécifique (ou du système en tant que tout) conduira à une baisse de chaleur, et évidemment, à une baisse de performance. Ainsi, vous devriez étudier et tester minutieusement les baisses de performance résultant des configurations que vous avez effectuées, tout particulièrement pour les systèmes menant à bien des missions critiques.

En étudiant les différentes tâches menées par votre système, et en configurant chaque composant afin de vous assurer que leur performance est suffisante pour la tâche, vous pouvez économiser de l'énergie, générer moins de chaleur, et optimiser l'autonomie de la batterie pour les ordinateurs portables. La plupart des principes pour l'analyse et le paramétrage d'un système par rapport à sa consommation d'électricité sont similaires à ceux du paramétrage de la performance. D'une certaine manière, la gestion de l'alimentation et le paramétrage de la performance sont des approches à la configuration d'un système opposées, car habituellement les systèmes sont optimisés soit vers la performance, soit vers une meilleure gestion de l'alimentation. Ce manuel décrit les outils fournis par Red Hat, ainsi que les techniques que nous avons développées afin de vous assister dans ce processus.

Red Hat Enterprise Linux 7 contient déjà de nombreuses fonctionnalités de gestion de l'alimentation qui sont activées par défaut. Celles-ci ont été choisies minutieusement afin de ne pas avoir d'impact sur la performance d'une utilisation typique d'un serveur ou d'un bureau. Cependant, dans certains cas très spécifiques où le débit maximal, une latence minimale, ou quand la performance CPU la plus importante est un impératif, une révision de ces valeurs par défaut est requise.

Posez-vous quelques questions afin de décider si vous devriez optimiser vos machines à l'aide des techniques décrites dans ce document :

**Q : Est-ce que je dois vraiment optimiser ?**

**R :** L'importance de l'optimisation de l'alimentation dépend de si votre compagnie donne des indications à suivre ou de s'il existe une réglementation à observer.

**Q : À quel point devrais-je optimiser ?**

**R :** Certaines des techniques que nous vous présentons ne requièrent pas que vous procédiez à la totalité du processus d'audit et d'analyse détaillée de votre machine, au contraire, elles vous fourniront un ensemble d'optimisations générales qui améliorent l'utilisation de l'alimentation. Celles-ci, bien entendu, ne seront pas aussi bonnes qu'avec un système audité et optimisé manuellement, mais elles fourniront quand même un bon compromis.

**Q : Est-ce que l'optimisation va réduire la performance du système à un niveau inacceptable ?**

**R :** La plupart des techniques décrites dans ce document auront un impact qui se ressentira sur la performance de votre système. Si vous choisissez d'implémenter une gestion de l'alimentation autre que celle installée par défaut dans Red Hat Enterprise Linux 7, vous devriez alors surveiller la performance du système après l'optimisation, puis décidez si la perte de performance est acceptable ou non.

**Q : Le temps pris et les ressources utilisées pour optimiser le système sont-ils plus importants que les gains à obtenir ?**

**R :** L'optimisation manuelle d'un système unique en suivant la totalité du processus n'en vaut pas la peine car le temps et le coût pris sont bien plus importants que les bénéfices à obtenir, et ce sur la durée entière du cycle de vie de la machine en question. Cependant, sur 10000 systèmes informatiques dans vos bureaux utilisant la même configuration et le même paramétrage, la création d'un paramétrage optimisé et s'appliquant à ces 10000 machines serait très certainement une bonne idée.

les sections suivantes vous expliqueront en quoi la performance d'un matériel optimal peut bénéficier à votre système en termes de consommation d'énergie.

## 1.2. Bases de la gestion de l'alimentation

La gestion d'une alimentation effective est construite sur les principes suivants :

### Un CPU inactif ne devrait se réveiller que lorsque nécessaire

Depuis Red Hat Enterprise Linux 6, le noyau exécute *tickless* qui remplace les anciennes interruptions de l'horloge périodique par des interruptions à la demande. Ainsi, les CPU inactifs peuvent rester inactifs jusqu'à ce qu'une nouvelle tâche soit placée en file d'attente pour traitement, et les CPU se trouvant dans des états de basse alimentation peuvent rester dans ces états plus longtemps. Cependant, les avantages de ce système peuvent être compromis si votre système possède des applications qui créent des événements de minuterie inutiles, comme par exemple, des événements d'interrogation de l'hôte, comme vérifier des changements de volume ou des mouvements du pointeur.

Red Hat Enterprise Linux 7 inclut des outils avec lesquels vous pourrez identifier et auditer des applications sur la base de l'utilisation de leur CPU. Reportez-vous à [Chapitre 2, Audit et analyse de la gestion de l'alimentation](#) pour plus de détails.

## **Le matériel et les périphériques inutilisés devraient être complètement désactivés**

Ceci est particulièrement vrai pour les périphériques possédant des pièces amovibles (par exemple des disques durs). En outre, certaines applications peuvent laisser un périphérique inutilisé mais activé « ouvert ». Lorsque cela se produit, le noyau suppose que le périphérique est en cours d'utilisation, ce qui peut empêcher le périphérique de se mettre en état d'économie d'énergie.

## **Une faible activité devrait signifier une faible puissance**

Dans de nombreux cas, ceci dépend d'un matériel moderne et d'une configuration BIOS correcte. Les composants de système plus anciens ne prennent pas en charge certaines des nouvelles fonctionnalités prises en charge dans Red Hat Enterprise Linux 7. Assurez-vous de bien utiliser le microprogramme le plus récent dans vos systèmes et que les fonctionnalités de gestion de l'alimentation dans les sections de gestion de l'alimentation ou de configuration de périphériques soient activées. Ci-dessous figurent quelques-unes des fonctionnalités à rechercher :

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

Si votre matériel prend ces fonctionnalités en charge et qu'elles sont activées dans le BIOS, Red Hat Enterprise Linux 7 les utilisera par défaut.

## **Les différentes formes des états de CPU et leurs effets**

Les CPUs modernes ainsi que l'ACPI (de l'anglais, Advanced Configuration and Power Interface) offrent différents états d'alimentation. Les trois différents états sont :

- » Veille (C-states)
- » Fréquence (P-states)
- » Sortie de chaleur (« T-states » ou « états thermaux »)

Un CPU exécuté sur l'état de veille le plus bas possible consomme un minimum de watts, mais prend aussi considérablement plus de temps à sortir de cet état lorsque nécessaire. Dans de rares cas, ceci peut faire que le CPU ne doive se réveiller immédiatement après être allé en veille. Cette situation résulte en un CPU constamment occupé et donc en une plus grande perte des économies potentielles d'énergie que si un autre état avait été utilisé.

## **Une machine éteinte consomme moins d'énergie**

Aussi évident que cela puisse paraître, l'une des meilleures manières d'économiser de l'énergie est d'éteindre les systèmes. Par exemple, votre compagnie peut développer une culture d'entreprise se concentrant sur une informatique « verte », avec comme conseil d'éteindre les machines lors de la pause déjeuner ou après le travail. Vous pourriez aussi consolider plusieurs serveurs physiques en un grand serveur et les virtualiser en utilisant la technologie de virtualisation fournie avec Red Hat Enterprise Linux 7.



## Chapitre 2. Audit et analyse de la gestion de l'alimentation

### 2.1. Aperçu général - Auditing et Analyse

L'audit, l'analyse, et un paramétrage détaillé et manuel d'un système unique est plutôt l'exception, car le temps et le coût que cela prend dépassent les bénéfices. Cependant, effectuer ces tâches une seule fois pour un grand nombre de systèmes pratiquement identiques, pour lesquels vous pourrez réutiliser les mêmes paramètres pour tous les systèmes peut être très utile. Par exemple, prenez le déploiement de milliers de systèmes de bureau ou un cluster HPC où les machines seraient quasiment identiques. Une autre raison de réaliser un audit et une analyse serait de fournir une base pour une comparaison dans laquelle vous identifieriez les régressions ou changements dans le comportement du système dans le futur. Les résultats de cette analyse peuvent se révéler très utiles lorsque des mises à jour du matériel, du BIOS, ou de logiciels se produisent de manière régulière et que vous voulez éviter toute surprise quant à la consommation d'électricité. En général, un audit et une analyse minutieuse vous offriront une bien meilleure idée de ce qu'il se passe sur un système en particulier.

Auditer et analyser un système sur sa consommation d'énergie est relativement difficile, même avec les systèmes les plus modernes. La plupart des systèmes n'offrent pas les moyens nécessaires de mesurer l'énergie utilisée via logiciel. Il existe pourtant des exceptions : la console de gestion ILO des systèmes de serveurs de Hewlett Packard offre un module de gestion de l'alimentation accessible depuis le web. IBM offre une solution similaire dans leur module de gestion de l'alimentation BladeCenter. Sur certains systèmes Dell, l'assistant « IT Assistant » offre aussi des possibilités de surveillance de l'alimentation. Il est possible que d'autres fournisseurs puissent fournir des solutions similaires pour leurs plateformes serveur, mais comme on peut l'observer, il n'existe pas de solution uniforme prise en charge par tous les fournisseurs.

Les mesures directes de consommation d'électricité sont souvent nécessaires afin de maximiser les économies autant que possible. Heureusement, il existe d'autres moyens pour mesurer si les changements sont en cours ou comment se comporte le système. Ce chapitre décrit les outils y étant nécessaires.

### 2.2. PowerTOP

Le lancement du nouveau noyau sans tic (de l'anglais, « tickless kernel ») dans Red Hat Enterprise Linux 7 permet au CPU d'entrer en état inactif plus fréquemment, en réduisant la consommation d'électricité et en améliorant la gestion de l'alimentation. L'outil **PowerTOP** identifie les composants spécifiques des applications du noyau et de l'espace utilisateur qui réveillent fréquemment le CPU. **PowerTOP** était utilisé dans le développement afin de procéder à des audits ayant mené des applications à être ajustées dans cette version, réduisant ainsi le nombre de mises en éveil du CPU non-nécessaires par un facteur 10.

Red Hat Enterprise Linux 7 est livré avec la version 2.x de **powertop**. Cette version est une réécriture complète de la base de code 1.x. Il dispose d'une interface utilisateur plus claire basée sur les onglets et utilise intensivement l'infrastructure du noyau "PERF" pour donner des données plus précises. Le comportement de puissance des dispositifs de systèmes est suivi et affiché en évidence, ainsi les problèmes peuvent être repérés rapidement. Plus expérimentalement, la CodeBase 2.x comprend un moteur d'estimation de puissance qui peut indiquer la quantité de puissance et les processus individuels consommant. Reportez-vous à [Figure 2.1, « Opération de PowerTOP »](#).

Pour installer **PowerTOP**, exécuter en tant qu'utilisateur **root**, la commande suivante :

```
yum install powertop
```

Pour exécuter **PowerTOP**, utiliser, en tant qu'utilisateur **root**, la commande suivante :

```
powertop
```

**PowerTOP** peut fournir une estimation de l'utilisation totale d'alimentation du système et afficher l'utilisation de chaque processus, périphérique, noyau, minuterie, ou gestionnaire d'interruptions. Les portables doivent être alimentés par une pile pendant cette tâche. Pour calibrer le moteur d'estimation de l'alimentation, exécuter, en tant qu'utilisateur **root**, la commande suivante :

```
powertop --calibrate
```

La calibration prend du temps. Le processus effectue différents tests, et passe par divers niveaux de luminosité, allume et éteint les systèmes. Ne pas interrompre le processus et n'interagissez pas avec la machine pendant la calibration. Lorsque le processus est terminé, **PowerTOP** commence normalement. Laissez-le fonctionner pendant environ une heure pour collecter des données. Lorsque suffisamment de données sont collectées, les chiffres d'estimation de puissance seront affichés dans la première colonne.

Si vous exécutez la commande sur un portable, il devra être alimenté par une batterie pour que toutes les données disponibles puissent se présenter.

Pendant qu'il fonctionne, **PowerTOP** rassemble les statistiques du système. Dans l'onglet **Overview** (Aperçu général), vous pouvez afficher une liste des composants qui envoient des signaux au CPU le plus fréquemment ou qui consomment le plus d'énergie (reportez-vous à [Figure 2.1, « Opération de PowerTOP »](#)). Les colonnes adjacentes affichent des estimations en matière d'alimentation, la façon dont la ressource est utilisée, les mises en éveil par seconde, la classification du composant, telle que processus, périphérique, ou minuterie, ainsi qu'une description du composant. Les réveils par seconde indiquent l'efficacité des services, périphériques ou pilotes du noyau. Moins de mises en éveil signifie une moindre consommation en alimentation. Les composants sont classés par quantité de consommation d'énergie pouvant être optimisée.

Régler les composants d'un pilote requiert le plus souvent des changements dans le noyau qui vont au-delà de l'étendue de ce document. Cependant, les processus de l'espace utilisateur qui envoient des signaux de mises en éveil sont plus facilement gérés. Premièrement, veuillez identifier si ce service ou cette application doit être exécuté(e) sur ce système. Si ce n'est pas le cas, procédez à sa désactivation. Pour fermer un service de manière permanente, veuillez exécuter :

```
systemctl disable servicename.service
```

Pour obtenir plus de détails sur le processus, exécutez, en tant qu'utilisateur **root**, les commandes suivantes :

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

Si le suivi à l'air de se répéter, vous avez probablement trouvé une attente active (de l'anglais, busy loop). Corriger ceci nécessiterait une modification du code du composant, ce qui, une fois de plus, se trouve au-delà de la portée de ce document.

Comme on le voit dans [Figure 2.1, « Opération de PowerTOP »](#), la consommation d'énergie totale et la durée de vie de la batterie restante s'affichent, si nécessaire. Ci-dessous, il s'agit d'un court résumé comportant le total des signaux de mise en éveil, par seconde, les opérations GPU par seconde, et des opérations de fichiers virtuels par seconde. Sur le reste de l'écran, il y a une liste des processus, des interruptions, des périphériques et autres ressources triés en fonction de leur utilisation. Si elle correctement calibrée, une estimation de la consommation d'énergie pour chaque élément répertorié dans la première colonne s'affiche également.

Utiliser les clés **Tab** et **Shift+Tab** pour vous déplacer à travers les onglets. Sur l'onglet **Idle stats**, l'utilisation des C-states s'affiche pour tous les processeurs et coeurs. Dans l'onglet **Frequency stats**, l'utilisation des P-states, y compris en mode Turbo (si applicable) s'affiche pour tous les processeurs et coeurs. Plus le CPU reste longtemps dans les états C- ou P- les plus élevés, le mieux (**C4** étant plus élevé

que **C3**). C'est une bonne indication pour savoir si l'utilisation du CPU a été optimisée correctement. La résidence doit correspondre à 90% ou plus, and un état C- ou P- quand le système est inactif.

L'onglet **Device Stats** donne des informations similaires à l'onglet **Overview** mais uniquement pour les périphériques.

L'onglet **Tunables** contient des suggestions qui servent à optimiser le système pour une consommation d'énergie minimisée. Les clés **up** et **down** sont là pour passer d'une suggestion à l'autre, et la clé **enter** pour faire basculer oui ou non la suggestion.

```

PowerTOP 2.3 Overview Idle stats Frequency stats Device stats Tunables
The battery reports a discharge rate of 16.7 W
The estimated remaining time is 1 hours, 25 minutes

Summary: 386.1 wakeups/second, 60.2 GPU ops/seconds, 0.0 VFS ops/sec and 42.9% CPU use

Power est.      Usage      Events/s  Category  Description
 3.79 W         2642 rpm          Device    Laptop fan
 3.39 W          53.3%           Device    Display backlight
 2.63 W         172.9 ms/s      0.00      Timer     process_timeout
 2.24 W         142.2 ms/s      17.8      Interrupt [9] acpi
 665 mW         43.6 ms/s       27.5      Process   /usr/lib64/firefox/firefox
 237 mW         10.7 ms/s       56.4      Process   /usr/lib64/seamonkey/seamonkey
 144 mW          5.7 ms/s       77.2      Interrupt PS/2 Touchpad / Keyboard / Mouse
 119 mW          7.8 ms/s       11.9      Process   /usr/bin/Xorg :0 -background none -verbose -auth /var/run/gdm
 91.3 mW         3.7 pkts/s      Device    Network interface: wlan0 (iwlfwifi)
 84.3 mW         5.5 ms/s       45.9      Timer     tick sched timer
 77.3 mW         3.3 ms/s       10.1      Process   gkrellm --geometry +1608+70
 72.9 mW         4.8 ms/s       20.6      Process   /usr/lib/polkit-1/polkitd --no-debug
 58.9 mW         3.9 ms/s       15.0      Process   /usr/lib64/seamonkey/plugin-container /usr/lib64/flash-plugin
 51.4 mW         3.4 ms/s       0.00      Interrupt [1] timer(softirq)
 42.3 mW         2.6 ms/s       13.0      Process   xfce4-screenshooter
 37.2 mW         2.4 ms/s       58.1      Timer     hrtimer_wakeup
 33.0 mW         2.2 ms/s        6.3      Interrupt [7] sched(softirq)
 31.5 mW         60.9 us/s        7.3      kWork     iwl_bg_run_time_calib_work
 29.8 mW         2.0 ms/s       41.2      kWork     od_dbs_timer
 28.9 mW         1.6 ms/s        1.7      Process   xfce4-panel
 25.2 mW         0.9 ms/s        8.6      Process   xfwm4
 21.3 mW         1.4 ms/s        0.00      Timer     delayed_work_timer_fn
 16.3 mW         1.1 ms/s        0.00      Process   /bin/dbus-daemon --system --address=systemd: --nofork --nopid
 13.1 mW         0.9 ms/s        0.5      Process   crond
 12.4 mW         0.8 ms/s        0.00      Interrupt [0] timer/1
 12.2 mW         0.8 ms/s        4.3      Interrupt [6] tasklet(softirq)
 12.1 mW         0.8 ms/s        0.05      kWork     disk_events_workfn
 12.0 mW         0.8 ms/s        0.00      Interrupt [0] timer/0
 10.0 mW         659.2 us/s       0.4      kWork     kcryptd_crypt
 10.0 mW         658.2 us/s       2.1      Process   /usr/sbin/NetworkManager --no-daemon
 8.04 mW         528.0 us/s       0.05      Process   powertop
 5.76 mW         347.4 us/s       1.6      Process   xchat
 5.59 mW         366.9 us/s       0.00      Interrupt [9] RCU(softirq)
 4.75 mW         311.5 us/s       0.00      Process   /usr/sbin/crond -n
    
```

Figure 2.1. Opération de PowerTOP

Vous pouvez également générer des rapports HTML en exécutant **PowerTOP** avec l'option **--html**. Remplacer le paramètre *htmlfile.html* par le nom désiré dans le fichier de sortie :

```
powertop --html=htmlfile.html
```

Par défaut, **PowerTOP** prend des mesures toutes les 20 secondes, mais vous pouvez changer cela par l'option **--time** :

```
powertop --html=htmlfile.html --time=seconds
```

Pour obtenir des informations supplémentaires sur **PowerTOP**, voir la page d'accueil de [PowerTOP](#).

**PowerTOP** peut également être utilisé en conjonction à l'utilitaire **turbostat**. Il s'agit d'un outil qui présente des informations comme la topologie du processeur, la fréquence, les statistiques d'état d'alimentation, la température, et le débit sur les processeurs Intel 64. Pour plus d'informations sur l'utilitaire **turbostat**, voir la page man **turbostat(8)** ou lire le guide [Performance Tuning Guide](#).

## 2.3. Diskdevstat et netdevstat

**Diskdevstat** et **netdevstat** sont des outils **SystemTap** qui collectent des informations détaillées sur les activités disque et réseau de toutes les applications exécutées sur un système. Ces outils ont été inspirés par **PowerTOP**, qui montre le nombre de réveils CPU par seconde de chaque application (reportez-vous à [Section 2.2 « PowerTOP »](#)). Les statistiques collectées par ces outils vous permettent d'identifier les applications qui utilisent de l'énergie inutilement avec beaucoup de petites opérations d'E/S, au lieu de peu d'opérations de grande taille. Les autres outils de surveillance ne mesurant que la vitesse de transfert n'aident pas à identifier ce type d'utilisation.

Installer ces outils avec la commande **SystemTap** en tant qu'utilisateur **root** :

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

Exécutez ces outils avec la commande :

```
diskdevstat
```

ou avec la commande :

```
netdevstat
```

Ces deux commandes peuvent inclure jusqu'à trois paramètres, comme suit :

```
diskdevstat update_interval total_duration display_histogram
```

```
netdevstat update_interval total_duration display_histogram
```

### *update\_interval*

Le temps en secondes entre les mises à jour de l'affichage. Par défaut : **5**

### *total\_duration*

Le temps en secondes pour la totalité de l'exécution. Par défaut : **86400** (1 jour)

### *display\_histogram*

Indiquez s'il faut inclure toutes les données collectées dans l'histogramme à la fin de l'exécution.

La sortie ressemble à celle de **PowerTOP**. Ci-dessous figure un exemple de sortie d'une plus longue exécution de **diskdevstat** :

```
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
5494  0  sda1     0    0.000    0.000    0.000    758    0.000
0.012    0.000 0logwatch
5520  0  sda1     0    0.000    0.000    0.000    140    0.000
0.009    0.000 perl
5549  0  sda1     0    0.000    0.000    0.000    140    0.000
0.009    0.000 perl
5585  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					
5429	0	sda1	0	0.000	0.000	0.000	62
0.009	0.000	crond					0.009
5379	0	sda1	0	0.000	0.000	0.000	62
0.008	0.000	crond					0.008
5473	0	sda1	0	0.000	0.000	0.000	62
0.008	0.000	crond					0.008
5415	0	sda1	0	0.000	0.000	0.000	62
0.008	0.000	crond					0.008
5433	0	sda1	0	0.000	0.000	0.000	62
0.008	0.000	crond					0.008
5425	0	sda1	0	0.000	0.000	0.000	62
0.007	0.000	crond					0.007
5375	0	sda1	0	0.000	0.000	0.000	62
0.008	0.000	crond					0.008
5477	0	sda1	0	0.000	0.000	0.000	62
0.007	0.000	crond					0.007
5469	0	sda1	0	0.000	0.000	0.000	62
0.007	0.000	crond					0.007
5419	0	sda1	0	0.000	0.000	0.000	62
0.008	0.000	crond					0.008
5481	0	sda1	0	0.000	0.000	0.000	61
0.001	0.000	crond					0.001
5355	0	sda1	0	0.000	0.000	0.000	37
0.014	0.001	laptop_mode					0.014
2153	0	sda1	26	0.003	3600.029	1290.730	0
0.000	0.000	rsyslogd					0.000
5575	0	sda1	0	0.000	0.000	0.000	16
0.000	0.000	cat					0.000
5581	0	sda1	0	0.000	0.000	0.000	12
0.002	0.000	perl					0.002
5582	0	sda1	0	0.000	0.000	0.000	12
0.002	0.000	perl					0.002
5579	0	sda1	0	0.000	0.000	0.000	12
0.001	0.000	perl					0.001
5580	0	sda1	0	0.000	0.000	0.000	12
0.001	0.000	perl					0.001
5354	0	sda1	0	0.000	0.000	0.000	12
0.170	0.014	s h					0.170
5584	0	sda1	0	0.000	0.000	0.000	12
0.002	0.000	perl					0.002
5548	0	sda1	0	0.000	0.000	0.000	12
0.014	0.001	perl					0.014
5577	0	sda1	0	0.000	0.000	0.000	12
0.003	0.000	perl					0.003
5519	0	sda1	0	0.000	0.000	0.000	12
0.005	0.000	perl					0.005
5578	0	sda1	0	0.000	0.000	0.000	12
0.001	0.000	perl					0.001
5583	0	sda1	0	0.000	0.000	0.000	12
0.001	0.000	perl					0.001
5547	0	sda1	0	0.000	0.000	0.000	11
0.002	0.000	perl					0.002
5576	0	sda1	0	0.000	0.000	0.000	11
0.001	0.000	perl					0.001

5518	0	sda1	0	0.000	0.000	0.000	11	0.000
0.001	0.000	perl						
5354	0	sda1	0	0.000	0.000	0.000	10	0.053
0.053	0.005	lm_lid.sh						

Les colonnes sont :

**PID**

l'ID du processus de l'application

**UID**

l'ID utilisateur sous lequel l'application est exécutée

**DEV**

le périphérique sur lequel les E/S ont eu lieu

**WRITE\_CNT**

le nombre total des opérations d'écriture

**WRITE\_MIN**

le temps le moins long pris pour deux écritures consécutives (en secondes)

**WRITE\_MAX**

le temps le plus long pris pour deux écritures consécutives (en secondes)

**WRITE\_AVG**

le temps moyen pris pour deux écritures consécutives (en secondes)

**READ\_CNT**

le nombre total des opérations de lecture

**READ\_MIN**

le temps le moins long pris pour deux lectures consécutives (en secondes)

**READ\_MAX**

le temps le plus long pris pour deux lectures consécutives (en secondes)

**READ\_AVG**

le temps moyen pris pour deux lectures consécutives (en secondes)

**COMMANDE**

le nom du processus

Dans cet exemple, trois applications se démarquent de manière évidente :

PID	UID	DEV	WRITE_CNT	WRITE_MIN	WRITE_MAX	WRITE_AVG	READ_CNT	READ_MIN
READ_MAX	READ_AVG	COMMANDE						
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						

Ces trois applications ont un **WRITE\_CNT** plus important que **0**, ce qui veut dire qu'elles ont effectué une sorte d'écriture pendant la mesure. De celles-ci, **plasma** fut la pire et de loin : cette application a effectuée la plupart des opérations d'écriture, et bien sûr le temps moyen entre les écritures était le plus bas. **Plasma** serait donc l'application idéale à étudier si vous vous sentiez concerné par les applications inefficaces au niveau de la consommation d'énergie.

Utilisez les commandes **strace** et **ltrace** pour examiner les applications de plus près en retraçant tous les appels système de l'ID du processus. Dans l'exemple présent, vous pourriez exécuter :

```
strace -p 2789
```

Dans cet exemple, la sortie de **strace** contient un schéma se répétant toutes les 45 secondes qui ouvre le fichier cache de l'icône KDE de l'utilisateur pour une écriture, suivi par la fermeture immédiate du fichier. Ceci a nécessairement conduit à une écriture physique sur le disque dur puisque les métadonnées du fichier (plus spécifiquement l'heure de la modification) avaient changées. Le correctif final servait à prévenir ces appels inutiles lorsqu'il n'y avait pas eu de mises à jour des icônes.

## 2.4. Battery Life Tool Kit

Red Hat Enterprise Linux 7 présente le **Battery Life Tool Kit** (en français, kit de ressources de l'autonomie de la batterie, ou BLTK), une suite de tests qui simule et analyse l'autonomie et la performance de la batterie. BLTK achève ceci en procédant à une série de tâches simulant des groupes d'utilisateurs spécifiques, tels que des développeurs, ou employés de bureau et en rapportant les résultats. Principalement développé pour tester la performance des ordinateurs portables, BLTK peut aussi être utilisé pour créer des rapports sur la performance des ordinateurs de bureau lorsque lancé avec **-a**.

BLTK vous permet de générer des charges de travail reproductibles comparables à une utilisation réelle d'une machine. Par exemple, la charge de travail **office** écrit un texte, le corrige, et refera la même chose sur une feuille de calcul. L'exécution de BLTK, combinée avec **PowerTOP**, ou tout autre outil d'audit ou d'analyse, vous permettra de vérifier si l'optimisation que vous avez effectuée a un effet quand la machine est active par rapport à quand elle n'est pas active. Comme vous pouvez exécuter la même charge de travail de multiples fois avec différents paramètres, vous pourrez comparer les résultats.

Installez BLTK par la commande :

```
yum install bltk
```

Exécutez BLTK par la commande :

```
bltk workload options
```

Par exemple, pour exécuter la charge de travail **inactive** pendant 120 secondes :

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

Les charges de travail disponibles par défaut sont :

**-I, --idle**

Le système est inactif, pour utiliser comme ligne de base pour comparer avec d'autres charges de

systeme est inactif, pour utiliser comme ligne de base pour comparer avec d'autres charges de travail

**-R, --reader**

simule la lecture de documents (par défaut avec **Firefox**)

**-P, --player**

simule la lecture de fichiers multimédia depuis un lecteur CD ou DVD (par défaut avec **mplayer**)

**-O, --office**

simule la modification de documents avec la suite bureautique **OpenOffice.org**

D'autres options vous permettent de spécifier :

**-a, --ac-ignore**

ignore si l'alimentation secteur est disponible (nécessaire pour l'utilisation d'ordinateurs de bureau)

**-T nombre\_de\_secondes, --time nombre\_de\_secondes**

le temps (en secondes) pris pour l'exécution du test ; utilisez cette option avec la charge de travail **inactive**

**-F nom\_de\_fichier, --file nom\_de\_fichier**

spécifie un fichier devant être utilisé par une charge de travail en particulier, par exemple, un fichier pour la charge de travail **player** pour lire au lieu d'accéder au lecteur CD ou DVD.

**-W application, --prog application**

spécifie une application devant être utilisée par une charge de travail, par exemple, un navigateur autre que **Firefox** pour la charge de travail **reader**

BLTK prend en charge un grand nombre d'options plus spécialisées. Pour obtenir plus de détails, référez-vous à la page man **bltk**.

bltk enregistre les résultats qu'il génère dans un répertoire spécifié dans le fichier de configuration **/etc/bltk.conf** — par défaut, **~/ .bltk/workload.results.number/**. Par exemple, le répertoire **~/ .bltk/reader.results.002/** contient les résultats du troisième test avec la charge de travail **reader** (le premier test n'est pas numéroté). Les résultats sont éparpillés sur plusieurs fichiers texte. Pour convertir ces résultats dans un format facile à lire, exécutez :

```
bltk_report chemin_vers_le_repertoire_des_resultats
```

Les résultats apparaissent maintenant dans un fichier texte nommé **Report** dans le répertoire des résultats. Pour voir les résultats dans un émulateur de terminal, utilisez l'option **-o** :

```
bltk_report -o chemin_vers_le_repertoire_des_resultats
```

## 2.5. UPower

Dans Red Hat Enterprise Linux 6 **DeviceKit-power** assumait les fonctions de gestion de l'alimentation qui faisaient partie de **HAL** ainsi que certaines fonctions qui faisaient partie du **Gestionnaire de l'alimentation GNOME** (GNOME Power Manager) dans des versions précédentes de Red Hat Enterprise Linux (reportez-



vous à la [Section 2.6, « Gestionnaire d'alimentation GNOME »](#)). Dans Red Hat Enterprise Linux 7, **DeviceKit-power** a été renommé **UPower**. **UPower** fournit un démon, un API, et un ensemble d'outils de ligne de commande. Chaque source d'alimentation du système est représentée en tant que périphérique, qu'il s'agisse d'un périphérique physique ou non. Par exemple, une batterie d'ordinateur portable et une source d'alimentation secteur sont toutes les deux considérées comme des périphériques.

Vous pouvez accéder aux outils de ligne de commande avec la commande **upower** et les options suivantes :

**--enumerate, -e**

affiche un chemin d'objet pour chaque périphérique d'alimentation sur le système, par exemple :

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

**--dump, -d**

affiche les paramètres pour tous les périphériques d'alimentation sur le système.

**--wakeups, -w**

affiche les réveils CPU sur le système.

**--monitor, -m**

surveille les modifications apportées aux périphériques d'alimentation sur le système. Par exemple, la connexion ou déconnexion de la source d'alimentation secteur, ou l'épuisement d'une batterie. Pressez sur **Ctrl+C** pour arrêter la surveillance du système.

**--monitor-detail**

surveille les modifications apportées aux périphériques d'alimentation sur le système. Par exemple, la connexion ou déconnexion de la source d'alimentation secteur, ou l'épuisement d'une batterie. L'option **--monitor-detail** présente davantage de détails que l'option **--monitor**. Pressez sur **Ctrl+C** pour arrêter la surveillance du système.

**--show-info chemin\_de\_l'objet, -i chemin\_de\_l'objet**

affiche toutes les informations disponibles sur un chemin d'objet particulier. Par exemple, pour obtenir des informations sur une batterie de votre système représentée par le chemin d'objet **/org/freedesktop/UPower/devices/battery\_BAT0**, exécutez :

```
upower -i /org/freedesktop/UPower/devices/battery_BAT0
```

## 2.6. Gestionnaire d'alimentation GNOME

**GNOME Power Manager** est un démon installé en conjonction avec l'environnement de bureau GNOME. La plus grande part de la fonctionnalité de gestion de l'alimentation que **GNOME Power Manager** fournissait dans les versions précédentes de Red Hat Enterprise Linux font partie de l'outil **DeviceKit-power** de Red Hat Enterprise Linux 6, rebaptisé **UPower** dans Red Hat Enterprise Linux 7 (voir [Section 2.5, « UPower »](#)). Cependant, **GNOME Power Manager** reste frontal pour cette fonctionnalité. À travers un applet dans la barre d'état système, **GNOME Power Manager** vous notifie des changements de statut de l'alimentation de votre système. Par exemple, un changement de la batterie à l'alimentation secteur. Il rapporte aussi le statut de la batterie et vous avertit lorsque le niveau de la batterie est bas.

## 2.7. Autres outils d'audit

## Les autres outils de monitoring

Red Hat Enterprise Linux7 offre un certain nombre d'outils avec lesquels procéder à des audits et analyses de systèmes. La plupart de ceux-ci peuvent être utilisés comme sources d'informations supplémentaires au cas où vous souhaiteriez vérifier ce que vous avez déjà découvert ou au cas où vous auriez besoin d'informations plus approfondies sur certaines parties. La majorité de ces outils sont aussi utilisés pour le paramétrage de performance. Ils incluent :

### **vmstat**

**vmstat** vous donne des informations détaillées sur les processus, la mémoire, la pagination, les E/S de bloc, les dérivements, et l'activité du CPU. Utilisez cette commande pour voir de plus près ce que le système fait et où il est actif.

### **iostat**

**iostat** est similaire à **vmstat**, mais seulement pour les E/S des périphériques de bloc. Cette commande fournit aussi une sortie et des statistiques plus verbeuses.

### **blktrace**

**blktrace** est un programme de suivi détaillé des E/S de bloc. Il décompose les informations en blocs uniques associés à des applications. Il se révèle très utile en conjonction avec **diskdevstat**.

## Chapitre 3. Infrastructure principale et mécanismes



### Important

Pour utiliser la commande **cpupower** expliquée dans ce chapitre, assurez-vous que le paquetage *kernel-tools* soit bien installé.

### 3.1. États inactifs du CPU

Les CPUs avec l'architecture x86 prennent en charge différents états dans lesquels certaines parties du CPU sont désactivées ou s'exécutent sur des paramètres de performance plus bas. Ces états, connus sous le nom de *C-states*, permettent aux systèmes d'économiser de l'énergie en désactivant partiellement les CPUs qui ne sont pas utilisés. Les C-states sont numérotés de manière ascendante, commençant par C0, et les numéros plus élevés représentent une fonctionnalité du CPU réduite et une économie d'énergie plus importante. Les C-states d'un numéro donné sont relativement similaires entre différents processeurs, même si les détails exacts des fonctionnalités spécifiques de l'état peuvent varier en fonction de la famille de processeurs. Les C-states de 0 à 3 sont définis comme suit :

#### C0

l'état d'opération, ou d'exécution. Dans cet état, le CPU fonctionne et n'est pas inactif.

#### C1, Arrêter

un état dans lequel le processeur n'exécute pas d'instructions mais n'est habituellement pas dans un état de basse alimentation. Le CPU peut continuer ses calculs avec pratiquement aucun délai. Tous les processeurs offrant des états C-States doivent pouvoir prendre cet état en charge. Les processeurs Pentium 4 prennent en charge un état C1 amélioré nommé C1E qui est en fait un état de consommation d'électricité moindre.

#### C2, Stop-Clock

un état où l'horloge est gelée pour le processeur, mais qui reste dans son état complet pour ses registres et caches, de manière à ce que les processeurs recommencent leurs calculs dès le démarrage de l'horloge. Il s'agit d'un état optionnel.

#### C3, Mise en veille

un état où le processeur s'endort réellement et n'a pas besoin de conserver son cache à jour. Le réveil à partir de cet état prend considérablement plus longtemps que de l'état C2. Une fois de plus, cet état est optionnel.

Pour afficher les états inactifs et autres statistiques du pilote CPUidle disponibles, veuillez exécuter la commande suivante :

```
cpupower idle-info
```

Les processeurs Intel récents avec la micro-architecture « Nehalem » proposent un nouvel état C-state, C6, qui peut réduire l'alimentation en tension d'un processeur à zéro, mais réduit habituellement la consommation de l'alimentation entre 80% et 90%. Le noyau dans Red Hat Enterprise Linux 7 inclut des optimisations pour ce nouvel état C-state.

### 3.2. Utilisation des gouverneurs CPUfreq

L'une des manières les plus efficaces de réduire la consommation d'énergie et la sortie de chaleur sur votre système est d'utiliser CPUfreq. CPUfreq — aussi connu sous le nom de mise à l'échelle de la vitesse du CPU — permet d'ajuster la vitesse d'horloge du processeur spontanément. Ceci permet au système de fonctionner à une vitesse d'horloge réduite afin d'économiser de l'énergie. Les règles pour les changements de fréquences, que ce soit vers une vitesse d'horloge plus ou moins élevée, ou à quel moment les changer, sont définies par le gouverneur CPUfreq.

Le gouverneur définit les caractéristiques du CPU du système, qui, elles-mêmes, affecteront la performance du CPU. Chaque gouverneur possède son propre comportement, but, et sa propre pertinence en termes de charge de travail. Cette section décrit comment choisir et configurer un gouverneur CPUfreq, les caractéristiques de chaque gouverneur, et le type de charge de travail que chaque gouverneur peut effectuer.

### 3.2.1. Types de gouverneurs CPUfreq

Cette section liste et décrit les différents types de gouverneurs CPUfreq disponibles dans Red Hat Enterprise Linux 7.

#### **cpufreq\_performance**

Le gouverneur Performance force le CPU à utiliser la plus haute fréquence d'horloge possible. Cette fréquence est statiquement réglée, et ne changera pas. Ainsi, ce gouverneur en particulier n'offre *aucune économie d'énergie*. Il n'est utile que pour de longues heures de charges de travail intensives, et lorsque le CPU est rarement (ou n'est pas) inactif.

#### **cpufreq\_powersave**

En revanche, le gouverneur Powersave force le CPU à utiliser la plus basse fréquence d'horloge possible. Cette fréquence est réglée statiquement, et ne changera pas. Ainsi, ce gouverneur en particulier offre une économie d'énergie maximale, mais ceci au prix d'une *performance de CPU plus basse*.

Le terme « Powersave » (économie d'énergie) peut être trompeur, même si, en principe, un processeur lent avec une pleine charge consomme plus d'énergie qu'un processeur rapide non chargé. Ainsi, tandis qu'il peut être recommandable de paramétrer le processeur pour qu'il utilise le gouverneur « Powersave » pendant les moments de basse activité, toute haute charge de travail inattendue à ce moment peut, en réalité, causer au système de consommer plus d'énergie.

Le gouverneur « Powersave » est, en d'autres termes, plus un « limiteur de vitesse » qu'un « économiseur d'énergie ». Il est particulièrement utile sur les systèmes et environnements où la surchauffe est un problème.

#### **cpufreq\_ondemand**

Le gouverneur Ondemand est un gouverneur dynamique qui permet au CPU d'atteindre une fréquence d'horloge maximale lorsque la charge du système est importante, et d'atteindre une fréquence minimale lorsque le système est inactif. Tandis que ceci permet d'ajuster la consommation d'énergie en fonction de la charge du système, cette opération se produit au détriment de la *latence entre les changements de fréquences*. Ainsi, la latence peut contrebalancer les bénéfices de performance/d'énergie du gouverneur Ondemand si le système change trop souvent entre inactivité et charges de travail intenses.

Pour la plupart des systèmes, le gouverneur Ondemand peut fournir le meilleur compromis qui soit entre émissions de chaleur, consommation d'énergie, performance, et manageabilité. Lorsque le système est uniquement occupé à certains moments spécifiques de la journée, le gouverneur Ondemand changera automatiquement entre les fréquences maximales et minimales en fonction de la charge et sans la moindre autre intervention.

#### **cpufreq\_userspace**

Le gouverneur Userspace permet aux programmes de l'espace utilisateur (ou à tout autre processus exécuté en tant que root) de définir la fréquence. De tous les gouverneurs, Userspace est le plus personnalisable ; et selon sa configuration, il peut fournir le meilleur équilibre entre performance et consommation pour votre système.

### cpufreq\_conservative

Tout comme le gouverneur Ondemand, le gouverneur Conservative ajuste aussi la fréquence de l'horloge en fonction de l'utilisation (comme le gouverneur Ondemand). Cependant, tandis que le gouverneur Ondemand le fait d'une manière plus agressive (du maximum au minimum, et vice-versa), le gouverneur Conservative change de fréquence de manière plus graduelle.

Ceci signifie que plutôt que de simplement choisir entre maximum et minimum, le gouverneur Conservative s'ajustera à une fréquence d'horloge qu'il reconnaîtra appropriée en fonction de la charge. Alors que ceci peut offrir d'importantes économies d'énergie, cette opération est effectuée aux dépens d'une *latence bien plus élevée* que celle du gouverneur Ondemand.



#### Note

Vous pouvez activer un gouverneur à l'aide des tâches **cron**. Ceci vous permettra de définir automatiquement des gouverneurs particuliers à des moments spécifiques de la journée. Ainsi, vous pouvez spécifier un gouverneur de basse latence pendant les périodes d'inactivité (par exemple hors des heures de travail traditionnelles), puis retourner à un gouverneur de fréquence plus importante durant les heures de charges de travail plus importantes.

Pour des instructions sur comment activer un gouverneur spécifique, veuillez vous reporter à [Section 3.2.2, « Installation de CPUfreq »](#).

### 3.2.2. Installation de CPUfreq

Tous les pilotes CPUfreq sont construits à part entière du package kernel-tools, et sont sélectionnés automatiquement, donc pour installer CPUfreq, vous avez juste besoin de sélectionner un gouverneur.

Vous pourrez voir quels gouverneurs sont disponibles pour un CPU spécifique avec :

```
cpupower frequency-info --governors
```

Vous pourrez alors activer un de ces gouverneurs sur tous les CPU avec :

```
cpupower frequency-set --governor [governor]
```

Pour activer un gouverneur sur des coeurs spécifiques, utiliser **-c** en précisant une plage ou une liste séparée par des virgules de numéros de CPU. Par exemple, pour activer le gouverneur Userspace pour les CPUs 1-3 et 5, voici la commande :

```
cpupower -c 1-3,5 frequency-set --governor cpufreq_userspace
```

### 3.2.3. Paramétrer la politique CPUfreq et la vitesse

Une fois que vous aurez choisi le gouverneur CPUfreq approprié, vous pourrez voir la vitesse du processeur et les informations sur la stratégie avec la commande **cpupower frequency-info**, vous pourrez aussi

régler la vitesse de chaque processeur plus précisément avec des options pour **cpupower frequency-set**.

Pour **cpupower frequency-info**, les options suivantes sont disponibles :

- » **--freq** — Affiche la vitesse actuelle du CPU en fonction du cœur de CPUfreq, en KHz.
- » **--hwfreq** — Affiche la vitesse actuelle du CPU en fonction du matériel, en KHz (uniquement disponible en tant que root).
- » **--driver** — Affiche quel pilote de CPUfreq est utilisé pour paramétrer la fréquence sur ce CPU.
- » **--governors** — Affiche les gouverneurs CPUfreq disponibles dans ce noyau. Si vous souhaitez utiliser un gouverneur CPUfreq qui n'est pas listé dans ce fichier, voir [Section 3.2.2, « Installation de CPUfreq »](#) pour obtenir des instructions.
- » **--affected-cpus** — Répertoire les CPU nécessitant un logiciel de coordination de fréquence.
- » **--policy** — Affiche la portée de la stratégie CPUfreq actuelle, en KHz, ainsi que le gouverneur actuellement actif.
- » **--hwlimits** — Répertoire les fréquences disponibles pour le CPU, en KHz.

Pour **cpupower frequency-set**, les options suivantes sont disponibles :

- » **--min <freq>** et **--max <freq>** — Définit les *limites de la stratégie* du CPU, en KHz.



### Important

Lors de la définition des limites de la politique, vous devriez définir **--max** avant **--min**.

- » **--freq <freq>** — Définit une vitesse d'horloge spécifique pour le CPU, en KHz. Vous pouvez uniquement définir une vitesse située dans les limites de la stratégie du CPU (selon **--min** et **--max**).
- » **--governor <gov>** — Définit un nouveau gouverneur CPUfreq.



### Note

Si le paquetage *cpupowerutils* n'est pas déjà installé, les paramètres CPUfreq peuvent être affichés à partir des réglables se trouvant dans `/sys/devices/system/cpu/[cpuid]/cpufreq/`. Les paramètres et valeurs peuvent être modifiés en écrivant sur ces réglables. Par exemple, pour définir la vitesse d'horloge minimum du cpu0 sur 360 KHz, veuillez utiliser :

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

## 3.3. Moniteurs CPU

**cpupower** propose une sélection d'analyseurs fournissant des statistiques sur les états de veille et d'inactivité et des informations sur la fréquence et des rapports sur la topologie du processeur, Certains

analyseurs sont spécifiques aux processeurs, alors que d'autres sont compatibles avec n'importe quel processeur. Reportez-vous à la page man **cpupower-monitor** pour plus de détails sur ce que chaque analyseur mesure et avec quels systèmes ils sont compatibles.

Les options suivantes peuvent être utilisées avec la commande **cpupower monitor** :

- ✦ **-l** — répertorie tous les moniteurs disponibles sur votre système.
- ✦ **-m <monitor1>, <monitor2>** — affiche des moniteurs spécifiques. Leurs identifiants peuvent être trouvés en exécutant **-l**.
- ✦ **command** — affiche les statistiques d'inactivité et les demandes de processeur d'une commande spécifique.

### 3.4. Stratégies d'économie d'énergie de processeur

**cpupower** offre plusieurs manières de réguler les stratégies d'économie d'énergie de votre processeur.

Les options suivantes peuvent être utilisées avec la commande **cpupower set** :

#### **--perf-bias <0-15>**

Permet au logiciel sur processeurs Intel pris en charge de contribuer de manière plus active à déterminer l'équilibre entre une performance optimale et de meilleures économies d'énergie. Ceci ne remplace pas les autres stratégies d'économies d'énergie. Les valeurs assignées vont de 0 à 15, où 0 représente une performance optimale et 15 représente une efficacité énergétique optimale.

Par défaut, cette option s'applique à tous les cœurs. Pour l'appliquer à un cœur en particulier, veuillez ajouter l'option **--cpu <cpulist>**.

#### **--sched-mc <0|1|2>**

Restreint l'utilisation d'énergie par les processus systèmes aux cœurs dans un seul paquetage processeur (aussi appelé paquetage CPU) avant de puiser dans d'autres paquetages CPU. 0 définit aucune restriction, 1 utilise un seul paquetage CPU, et 2 réalise ceci en plus de favoriser les paquetages CPU semi-inactifs pour gérer les réveils des tâches.

#### **--sched-smt <0|1|2>**

Restreint l'utilisation d'énergie par les processus systèmes aux frères du fil d'un cœur de CPU avant de puiser dans d'autres cœurs. 0 définit aucune restriction, 1 utilise un seul paquetage CPU, et 2 réalise ceci en plus de favoriser les paquetages CPU semi-inactifs pour gérer les réveils des tâches.

### 3.5. Suspension et reprise

Lorsqu'un système est suspendu, le noyau appelle les pilotes pour qu'ils stockent leur état, puis il les décharge. Lorsque le système est rétabli, il recharge ces pilotes, qui tentent de reprogrammer leurs périphériques. La capacité des pilotes à accomplir cette tâche détermine si le système peut être rétabli avec succès.

Les pilotes vidéo sont particulièrement problématiques car la spécification *ACPI* (Advanced Configuration and Power Interface) ne requiert pas du microprogramme du système qu'il ne reprogramme le matériel vidéo. Ainsi, à moins que les pilotes vidéo ne soient capables de programmer le matériel depuis un état non-initialisé, ils pourront empêcher la reprise du système.

Red Hat Enterprise Linux 7 inclut une meilleure prise en charge des chipsets graphiques, ce qui permet à la suspension et à la reprise de fonctionner sur un plus grand nombre de plateformes. La prise en charge des chipsets NVIDIA a été particulièrement améliorée ; notamment pour les séries GeForce 8800.

### 3.6. Gestion de l'alimentation des périphériques en cours d'exécution

La gestion de l'alimentation des périphériques en cours d'exécution (RDPM, de l'anglais Runtime device power management) contribue à réduire la consommation d'énergie en échange d'un impact minimum visible par l'utilisateur. Si un périphérique demeure inactif pendant un moment suffisamment long, et que le support matériel RDPM existe à la fois dans le périphérique et le pilote, le périphérique se retrouve dans un état inférieur. Le recouvrement de l'état d'alimentation inférieur est assuré par un événement E/S externe, qui amène le pilote du périphérique et le noyau à retourner à un état « en cours d'exécution ». Tout cela a lieu automatiquement, car RDPM est activé par défaut.

Les utilisateurs sont autorisés à contrôler RDPM sur un périphérique en configurant l'attribut dans un fichier spécial. Les fichiers de configuration de périphériques particuliers se trouvent dans le répertoire `/sys/devices/device/power/`, où *device* remplace le chemin qui mène au répertoire d'un périphérique particulier.

Ainsi, pour configurer le RDPM d'un CPU, accéder au répertoire suivant :

```
/sys/devices/system/cpu/power/
```

Le fait de ramener un périphérique d'un état d'alimentation inférieur à l'état d'exécution en cours ajoute une latence supplémentaire à l'opération e/s suivante. La durée de ce délai supplémentaire est spécifique au périphérique. Le schéma de configuration décrit ici permet à l'administrateur système de désactiver RDPM sur une base périphérique par périphérique et à la fois d'examiner et de contrôler certains des autres paramètres. Chaque répertoire `/sys/devices/device/power` contient les fichiers de configuration suivants :

#### control

Ce fichier est utilisé pour activer ou désactiver RDPM pour un périphérique donné. Tous les périphériques ont une des deux valeurs suivantes pour l'attribut dans le fichier **control** :

##### auto

valeur par défaut pour tous les périphériques, qui peuvent être en RDPM automatique, suivant le pilote

##### on

empêche le pilote de gérer l'état d'alimentation du périphérique en cours d'exécution

#### autosuspend\_delay\_ms

Ce fichier contrôle auto-suspend, qui est la période minimale d'inactivité entre l'état inactif et la mise en attente du périphérique. Le fichier contient la valeur de délai de suspension automatique en millisecondes. Une valeur négative empêche le périphérique d'être mis en attente en cours d'exécution, ce qui a le même effet que de définir l'attribut du fichier `/sys/devices/device/power/control` à **on**. Les valeurs supérieures à 1000 secondes sont arrondies à la seconde la plus proche.

### 3.7. Gestion de l'alimentation en état actif



La *gestion de l'alimentation en état actif* (ASPM, de l'anglais Active-State Power Management) permet de réaliser des économies d'énergie dans le sous-système *interconnexion de composants périphériques* (PCI Express, ou PCIe) en définissant un état d'alimentation plus faible pour les liens PCIe lorsque les périphériques auxquels ils sont connectés ne sont pas en cours d'utilisation. ASPM contrôle l'état de l'alimentation des deux côtés du lien, et permet d'économiser de l'énergie dans le lien même lorsque le périphérique se trouvant à la fin de celui-ci est dans un état d'alimentation complète.

Lorsque ASPM est activé, la latence de périphérique augmente à cause du temps requis pour effectuer la transition du lien entre les différents états d'alimentation. ASPM possède trois politiques afin de déterminer les états d'alimentation :

#### défaut

établit les états d'alimentation du lien PCIe en fonction des valeurs par défaut spécifiées par le microprogramme du système (par exemple, BIOS). Ceci est l'état par défaut d'ASPM.

#### powersave

règle ASPM de manière à économiser de l'énergie à chaque occasion, sans tenir compte de la performance en résultant.

#### performance

désactive ASPM afin de permettre aux liens PCIe d'opérer avec une performance maximale.

La prise en charge ASPM peut être activée ou désactivée par le paramètre du noyau **pcie\_aspm**, dans lequel **pcie\_aspm=off** désactive ASPM et **pcie\_aspm=force** active ASPM, même sur les périphériques ne prenant pas en charge ASPM.

Les stratégies ASPM sont définies dans **/sys/module/pcie\_aspm/parameters/policy**, mais peuvent aussi être spécifiées lors du démarrage avec le paramètre de noyau **pcie\_aspm.policy**, dans lequel **pcie\_aspm.policy=performance** définira par exemple la stratégie de performance ASPM.



### Avertissement

Si **pcie\_aspm=force** est défini, le matériel qui ne prend pas en charge ASPM peut provoquer une cessation de réponse du système. Avant de définir **pcie\_aspm=force**, assurez-vous que tout le matériel PCIe sur le système prend en charge ASPM.

## 3.8. Aggressive Link Power Management

*ALPM* (de l'anglais, Aggressive Link Power Management) est une technique d'économie d'énergie qui aide le disque à économiser l'énergie en paramétrant un lien SATA sur une faible puissance lors d'un temps d'inactivité (lorsqu'il n'y a pas d'E/S). ALPM règle automatiquement le lien SATA sur un état actif une fois que les requêtes d'E/S sont en file d'attente sur ce lien.

Les économies d'énergie offertes par ALPM se font aux dépens de la latence de disque. Ainsi, ALPM ne devrait être utilisé que dans le cas où le système connaîtra de longues périodes d'inactivité d'E/S.

ALPM est uniquement disponible sur les contrôleurs SATA qui utilisent les interfaces *AHCI* (de l'anglais, Advanced Host Controller Interface). Pour obtenir plus d'informations sur AHCI, voir <http://www.intel.com/technology/serialata/ahci.htm>.

Lorsque disponible, ALPM est activé par défaut. ALPM possède trois niveaux :

### min\_power

Ce mode définit le lien sur son état le plus bas (SLUMBER) lorsqu'il n'y a pas d'E/S sur le disque. Ce mode est utile lorsque de longues périodes d'inactivité sont à prévoir.

### medium\_power

Ce mode définit le lien sur le second état le plus bas (PARTIAL) lorsqu'il n'y a pas d'E/S sur le disque. Ce mode est conçu afin de permettre des transitions entre états d'alimentation (par exemple lorsqu'il y a des alternances entre une utilisation intense d'E/S et une période d'inactivité d'E/S) tout en limitant au maximum l'impact sur la performance.

Le mode **medium\_power** permet au lien de faire une transition entre les états « PARTIAL » (partiels) et les états « ACTIVES » (états actifs, aussi nommés « fully-powered » en anglais) selon la charge. Remarquez qu'il n'est pas possible de faire une transition de « PARTIAL » à « SLUMBER » et vice-versa. Dans ce cas, soit l'état de l'alimentation ne peut pas faire de transition vers l'autre sans que la transition ne passe par l'état « ACTIVE » en premier.

### max\_performance

ALPM est désactivé, le lien ne peut pas entrer dans un état de faible alimentation lorsqu'il n'y a pas d'E/S sur le disque.

Pour vérifier si vos adaptateurs hôtes SATA prennent en charge ALPM, vous pouvez vérifier que le fichier `/sys/class/scsi_host/host*/link_power_management_policy` existe. Pour changer les paramètres, écrivez simplement les valeurs décrites dans cette section sur ces fichiers, ou affichez les fichiers pour vérifier le paramètre actuel.



#### Important

Définir ALPM sur **min\_power** ou **medium\_power** désactivera automatiquement la fonctionnalité « Hot Plug ».

## 3.9. Optimisation d'accès au disque Relatime

Le standard POSIX requiert que les systèmes d'exploitation effectuent la maintenance des métadonnées du système de fichiers qui enregistrent la dernière date d'accès à un fichier. Cette date de dernière modification est appelée **atime**, et la maintenir requiert une constante série d'opérations d'écriture sur le stockage. Ces écritures gardent les périphériques de stockage et leurs liens occupés et allumés. Comme peu d'applications utilisent les données **atime**, l'activité de ce périphérique de stockage gaspille de l'énergie. De plus, l'écriture sur le stockage se produirait même si le fichier n'était pas lu à partir du stockage, mais à partir du cache. Depuis quelques temps, le noyau Linux prend en charge une option **noatime** pour **mount** et n'écrit pas de données **atime** sur les systèmes de fichiers montés avec cette option. Cependant, l'arrêt de cette fonctionnalité est problématique car certaines applications reposent sur les données **atime**, et elles échoueront si ces données ne sont pas disponibles.

Le noyau utilisé dans Red Hat Enterprise Linux 7 prend en charge une solution alternative — **relatime**. **Relatime** effectue la maintenance de **atime**, mais pas à chaque fois qu'on accède à un fichier. Avec l'activation de cette option, les données **atime** sont écrites sur le disque seulement si le fichier a été modifié depuis que la dernière fois que les données **atime** ont été mises à jour (**mtime**), ou si le dernier accès au fichier a eu lieu depuis plus d'un certain temps (par défaut, un jour).

Par défaut, tous les systèmes de fichiers sont maintenant montés avec l'option **relatime** activée. Vous pouvez la supprimer pour n'importe quel système de fichiers avec l'option **norelatime**.

### 3.10. Limitation de la consommation

Red Hat Enterprise Linux 7 prend en charge les fonctionnalités de limitation de la consommation trouvées dans le matériel récent, tel que dans le *DPC* (Dynamic Power Capping, ou limitation de consommation dynamique) de HP et la technologie de gestionnaire de nœuds (Node Manager) d'Intel. La limitation de consommation permet aux administrateurs de limiter la consommation d'énergie des serveurs, mais elle permet aussi aux gérants de planifier les centres de données de manière plus efficace, car le risque de surcharger les réserves d'énergie est grandement réduit. Les gérants peuvent ainsi placer plusieurs serveurs sur le même espace physique et savoir que si la consommation est limitée, la demande d'énergie lors de lourdes charges n'excédera pas la quantité disponible.

#### Dynamic Power Capping de HP

Le Dynamic Power Capping est une fonctionnalité disponible sur les serveurs sélectionnés ProLiant et BladeSystem. Elle permet aux administrateurs systèmes de limiter la consommation de l'alimentation d'un serveur ou d'un groupe de serveurs. Cette limitation est une limite définitive que le serveur ne pourra pas dépasser, peu importe sa charge de travail. Celle-ci n'aura aucun effet jusqu'à ce que le serveur n'atteigne la limite de consommation. À ce moment, un processeur de gestion ajustera les états P (ou P-states) du CPU, et la limitation d'horloge afin de limiter l'énergie consommée.

Le Dynamic Power Capping modifie le comportement du processeur indépendamment du système d'exploitation. Cependant, *iLO2* (*integrated Lights-Out 2*), le microprogramme de HP, autorise les systèmes d'exploitation à accéder au processeur de gestion. Ainsi, les applications dans l'espace utilisateur peuvent effectuer des requêtes sur le processeur de gestion. Le noyau utilisé dans Red Hat Enterprise Linux 7 inclut un pilote pour les microprogrammes *iLO* et *iLO2* de HP, ce qui autorise les programmes à effectuer des requêtes sur les processeurs de gestion sur `/dev/hpilo/dXccbM`. Le noyau inclut aussi une extension de l'interface **hwmon sysfs** pour prendre en charge les fonctionnalités de limitation de l'alimentation (« Power capping ») et un pilote **hwmon** pour les jauges d'alimentation ACPI 4.0 qui utilisent l'interface **sysfs**. Ensemble, ces fonctionnalités permettent au système d'exploitation et aux outils de l'espace utilisateur de lire la valeur configurée pour la limitation de l'alimentation ainsi que l'utilisation actuelle de l'alimentation du système.

Pour obtenir plus de détails sur la limitation de consommation et le Dynamic Power Capping de HP, reportez-vous au document *HP Power Capping and HP Dynamic Power Capping for ProLiant Servers*, disponible sur <http://h20000.www2.hp.com/bc/docs/support/SupportManual/c01549455/c01549455.pdf>.

#### Gestionnaire de nœuds Intel

Le gestionnaire de nœuds Intel impose une limitation de la consommation aux systèmes, à l'aide du processeur des états P et T (T-states et P-states) pour limiter la performance du CPU et par conséquent la consommation de l'alimentation. En définissant une politique de gestion de l'alimentation, les administrateurs peuvent configurer les systèmes afin qu'ils utilisent moins d'énergie lorsque les charges du système sont basses, par exemple pendant la nuit ou les week-ends.

Le gestionnaire de nœuds Intel ajuste la performance du CPU à l'aide de l'*OSPM* (de l'anglais, Operating System-directed configuration and Power Management) avec l'*ACPI* (Advanced Configuration and Power Interface) standard. Lorsque le gestionnaire de nœuds Intel notifie le pilote *OSPM* des changements des T-states, le pilote procède aux changements correspondants sur le processeur des P-states. De la même manière, lorsque le gestionnaire de nœuds Intel notifie le pilote *OSPM* des changements des P-states, celui-ci effectuera les changements correspondants sur les T-states. Ces changements s'effectuent automatiquement et ne requièrent pas l'ingérence du système d'exploitation. Les administrateurs configurent et surveillent le gestionnaire de nœuds Intel avec un logiciel *DCM* (gestionnaire de centre de données Intel, de l'anglais, Intel Data Center Manager).

Pour obtenir plus de détails sur le gestionnaire de nœuds Intel, reportez-vous à *Node Manager — A Dynamic Approach To Managing Power In The Data Center*, disponible sur <http://communities.intel.com/docs/DOC-4766>

### 3.11. Gestion améliorée de l'alimentation des graphismes

Red Hat Enterprise Linux 7 permet d'économiser de l'énergie sur les graphismes et moniteurs en éliminant plusieurs sources de consommation non-nécessaires.

#### Regénération d'horloge LVDS

La *signalisation différentielle à basse tension* (de l'anglais, low-voltage differential signalling, ou LVDS) est un système transportant des signaux électroniques sur des fils de cuivre. Une application importante de ce système est la transmission d'informations de pixels sur les écrans *LCD* (de l'anglais, liquid crystal display) d'ordinateurs portables. Tous les affichages possèdent des *fréquences de rafraîchissement* — la fréquence à laquelle ils reçoivent des nouvelles données depuis un contrôleur graphique et redessinent l'image à l'écran. Habituellement, l'écran reçoit de nouvelles données soixante fois par seconde (une fréquence de 60 Hz). Lorsqu'un écran et un contrôleur graphique sont liés par LVDS, le système LVDS utilise de l'énergie à chaque cycle de rafraîchissement. En période d'inactivité, la fréquence de rafraîchissement de nombreux écrans LCD peut descendre jusqu'à 30 Hz sans que cela ne puisse particulièrement se remarquer (contrairement aux moniteurs à *tube cathodique* (ou CRT), pour lesquels la baisse de la fréquence de rafraîchissement produit un scintillement distinct). Le pilote des adaptateurs graphiques Intel, qui est intégré dans le noyau utilisé dans Red Hat Enterprise Linux 7 effectue cette *baisse d'horloge* automatiquement, et permet d'économiser environ 0.5 W lorsque l'écran est inactif.

#### Activation du rafraîchissement automatique de la mémoire

La *mémoire vive dynamique synchrone* (ou SDRAM) — comme celle utilisée pour la mémoire vidéo dans les adaptateurs graphiques — est rechargée des milliers de fois par seconde de manière à ce que les cellules de mémoire individuelles puissent retenir les données stockées en elles. À part sa fonction principale de gestion de données circulant vers et hors de la mémoire, le contrôleur de mémoire est habituellement responsable de l'initiation de ces cycles de rafraîchissement. Toutefois, SDRAM possède aussi un mode de *rafraîchissement* de basse consommation. Dans ce mode, la mémoire utilise un compteur interne pour générer ses propres cycles de rafraîchissement, permettant ainsi au système de fermer le contrôleur de mémoire sans pour mettre en danger les données présentes en mémoire. Le noyau utilisé dans Red Hat Enterprise Linux 7 peut ainsi déclencher le rafraîchissement automatique de la mémoire avec les adaptateurs graphiques d'Intel lorsqu'ils sont inactifs, ce qui permet déconomiser autour de 0,8 W.

#### Réduction de l'horloge GPU

Les GPUs (de l'anglais, graphical processing units) typiques contiennent des horloges internes gouvernant différentes parties de leur circuiterie interne. Le noyau utilisé dans Red Hat Enterprise Linux 7 peut réduire la fréquence de certaines horloges internes des GPUs d'Intel et d'ATI. La réduction du nombre de cycles effectués par les composants GPU sur un temps donné permet d'économiser de l'énergie qu'ils auraient consommé lors de cycles qu'ils n'avaient pas à effectuer. Le noyau réduit automatiquement la vitesse de ces horloges lorsque le GPU est inactif, et l'augmente lorsque l'activité de celui-ci augmente. La réduction des cycles d'horloge des GPUs peut vous économiser jusqu'à 5 W.

#### Mise hors-tension GPU

Les pilotes graphiques Intel et ATI dans Red Hat Enterprise Linux 7 peuvent détecter lorsqu'aucun moniteur n'est attaché à un adaptateur et éteint ainsi complètement le GPU. Cette fonctionnalité est particulièrement significative pour les serveurs ne possédant pas de moniteur de manière régulière.

### 3.12. RfKill

De nombreux systèmes informatiques contiennent des transmetteurs radio, notamment des périphériques Wi-Fi, Bluetooth, et 3G. Ces périphériques consomment de l'électricité qui est gaspillée quand le périphérique n'est pas en cours d'utilisation.

*RFKill* est un sous-système dans le noyau Linux qui fournit une interface au travers de laquelle les transmetteurs radio d'un système informatique peuvent être questionnés, activés, et désactivés. Lorsque les transmetteurs sont désactivés, ils peuvent être placés dans un état où le logiciel peut les réactiver (un *soft block*), ou dans un état où le logiciel ne peut pas les réactiver (un *hard block*).

Le cœur RFKill fournit l'interface de programmation (API) pour le sous-système. Les pilotes de noyau désignés pour prendre RFKill en charge utilisent cette API pour s'enregistrer auprès du noyau, et incluent des méthodes pour activer et désactiver le périphérique. De plus, le cœur RFKill offre des notifications que les applications de l'utilisateur peuvent interpréter, ainsi que des manières pour que les applications de l'utilisateur puissent questionner les états des transmetteurs.

L'interface RFKill se trouve dans `/dev/rfkill`, qui contient l'état actuel de tous les transmetteurs radio sur le système. Chaque périphérique possède son propre état RFKill enregistré dans `sysfs`. De plus, RFKill fournit des *uevents* pour chaque changement d'état dans un périphérique pour lequel RFKill est activé.

**Rfkill** est un outil de la ligne de commande qui vous permet de questionner et de modifier des périphériques du système qui possèdent RFKill activé. Pour obtenir cet outil, installez le paquet *rfkill*.

Utilisez la commande **rfkill list** pour obtenir une liste des périphériques, chacun d'entre eux aura un *numéro d'index* qui lui est associé, commençant par **0**. Vous pourrez utiliser ce numéro d'index pour dire à **rfkill** de bloquer ou débloquer un service, par exemple :

```
rfkill block 0
```

bloque le premier périphérique RFKill sur le système.

Vous pouvez aussi utiliser **rfkill** pour bloquer certaines catégories de périphériques, ou même tous les périphériques RFKill. Par exemple :

```
rfkill block wifi
```

bloque tous les périphériques Wi-Fi sur le système. Pour bloquer tous les périphériques RFKill, exécutez :

```
rfkill block all
```

Pour débloquer des périphériques, exécutez **rfkill unblock** au lieu de **rfkill block**. Pour obtenir la liste complète des catégories de périphériques que **rfkill** peut bloquer, exécutez **rfkill help**

## Chapitre 4. Cas d'utilisation

Ce chapitre décrit deux types de cas d'utilisation dans le but d'illustrer les méthodes d'analyse et de configuration décrites ailleurs dans ce guide. Le premier exemple prend en considération des serveurs typiques et le deuxième est un ordinateur portable commun.

### 4.1. Exemple — Serveur

De nos jours, un serveur standard typique est fourni avec toutes les fonctionnalités de matériel nécessaires prises en charge dans Red Hat Enterprise Linux 7. La première chose à faire est de prendre en considération les types de charges de travail pour lesquelles le serveur va être utilisé. En vous basant sur ces informations, vous pourrez ensuite décider quels composants pourront être optimisés pour réaliser des économies d'énergie.

Peu importe le type de serveur, la performance des graphismes n'est généralement pas requise. Ainsi, les économies d'énergie GPU peuvent être laissées en cours d'utilisation.

#### Serveur web

Un serveur web nécessite un réseau et une E/S de disque. Selon la vitesse de connexion extérieure, 100 Mo/s peut suffire. Si la machine sert principalement à des pages statiques, la performance du CPU peut ne pas être très importante. Ainsi, les choix de gestion de l'alimentation qui s'offrent à vous incluent :

- aucun plugin de réseau ou de disque pour **tuned**.
- ALPM en marche.
- Gouverneur **ondemand** en marche.
- carte réseau limitée à 100 Mo/s.

#### Serveur de calcul

Un serveur de calcul nécessite principalement des CPU. Les choix de gestion de l'alimentation s'offrant à vous incluent :

- des plugins de réseau ou de disque pour **tuned** selon les tâches et l'endroit où sont stockées les données ; ou pour des systèmes en mode batch, **tuned** est totalement actif.
- selon l'utilisation, le gouverneur **performance**.

#### Serveur mail

Un serveur mail nécessite principalement des E/S de disque et des CPU. Les choix de gestion de l'alimentation s'offrant à vous incluent :

- Le gouverneur **ondemand** activé, car les derniers pourcentages de la performance du CPU ne sont pas importants.
- aucun plugin de réseau ou de disque pour **tuned**.
- la vitesse du réseau ne devrait pas être limitée, car les mails sont souvent internes et peuvent ainsi bénéficier d'un lien de 1 Go/s ou de 10 Go/s.

#### Serveur de fichiers

Les prérequis d'un serveur de fichiers sont similaires à ceux d'un serveur mail, mais selon le protocole utilisé, celui-ci peut nécessiter une meilleure performance CPU. Typiquement, les serveurs basés sur Samba requièrent davantage de CPU que NFS, et NFS en requiert plus que iSCSI. Même ainsi, vous devriez être en mesure d'utiliser le gouverneur **ondemand**.

### Serveur de répertoires

Typiquement, un serveur d'annuaire requiert moins d'E/S de disque, particulièrement s'il est équipé de suffisamment de RAM. La latence du réseau est importante, les E/S de réseau le sont moins. Prenez en considération le paramétrage de la latence du réseau à une vitesse de lien moindre, mais testez soigneusement ceci pour le réseau en question.

## 4.2. Exemple — Ordinateur portable

Un autre domaine très commun dans lequel la gestion et l'économie d'énergie peut réellement faire une différence est l'ordinateur portable. Comme les ordinateurs portables, par définition, utilisent beaucoup moins d'énergie que les stations de travail ou les serveurs, la possibilité de réaliser des économies est, dans l'absolu, moindre que pour d'autres machines. Par contre, en mode batterie, toute économie peut se révéler utile pour obtenir quelques minutes de plus. Même si cette section se concentre sur les ordinateurs portables en mode batterie, vous pouvez sans aucun doute utiliser certains, ou la totalité, de ces paramètres sur une alimentation secteur.

Les économies sur des composants uniques font généralement une différence relative plus importante que celle qu'elles ont sur des stations de travail. Par exemple, une interface réseau de 1 Go/s exécutée à 100 Mo/s économisera environ 3–4 watts. Pour un serveur typique avec une consommation totale d'environ 400 watts, cette économie est d'environ 1 %. Sur un ordinateur portable consommant un total d'à peu près 40 watts, l'économie d'énergie de ce seul composant représenterait 10 % de l'énergie totale consommée.

Les optimisations spécifiques pour réaliser des économies d'énergie sur un ordinateur portable typique incluent :

- ✦ La configuration du BIOS du système afin de désactiver tout le matériel que vous n'utilisez pas. Par exemple, les ports parallèles ou sériels, lecteurs de cartes, webcams, WiFi, Bluetooth, etc.
- ✦ Atténuez l'affichage dans des environnements plus sombres, lorsque vous n'avez pas besoin d'un éclairage maximum pour lire l'écran de manière confortable. Utilisez **Système+Préférences** → **Gestion de l'alimentation** sur le bureau GNOME, **Lanceur d'application Kickoff+Ordinateur+Paramètres système+Avancé** → **Gestion de l'alimentation** sur le bureau KDE ; ou **gnome-power-manager** ou **xbacklight** dans la ligne de commande ; ou utilisez les touches de fonction sur votre ordinateur portable.

De plus, (ou alternativement) vous pouvez effectuer de nombreux petits ajustements aux différents paramètres du système :

- ✦ utilisez le gouverneur **ondemand** (par défaut activé dans Red Hat Enterprise Linux 7)
- ✦ active l'économie d'énergie AC97 (activé par défaut dans Red Hat Enterprise Linux 7) :

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

- ✦ active l'auto-suspension USB :

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

Remarquez que l'auto-suspension USB ne fonctionne pas correctement avec tous les périphériques USB.

- ✦ monte un système de fichiers à l'aide de relatime (par défaut dans Red Hat Enterprise Linux 7) :

```
mount -o remount,relatime point_de_montage
```

- ✦ réduit la luminosité de l'écran à **50** ou moins, par exemple :

```
xbacklight -set 50
```

- ✦ active DPMS pour écran inactif :

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

- ✦ désactive le Wi-Fi :

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



## Annexe A. Astuces pour les développeurs

Tout bon manuel de programmation couvrira les problèmes liés à la mémoire et à l'allocation de mémoire et à la performance de certaines fonctions spécifiques. En développant votre logiciel, soyez conscient des problèmes qui peuvent augmenter la consommation d'énergie sur les systèmes sur lesquels le logiciel est en cours d'exécution. Même si ces considérations n'affectent pas chaque ligne de code, vous pouvez l'optimiser dans les domaines où se créent des goulots d'étranglement de performance.

Les techniques les plus souvent problématiques incluent :

- ✦ l'utilisation de threads.
- ✦ les réveils CPU inutiles et l'utilisation incorrecte des réveils. Si vous devez effectuer un réveil, faites-le en une seule fois et aussi vite que possible (en anglais, « race to idle »).
- ✦ l'utilisation de `[f]sync()` sans que cela ne soit nécessaire.
- ✦ les interrogations actives inutiles, ou l'utilisation de délais d'attente réguliers et de courtes durées (réagit aux événements à la place).
- ✦ une utilisation des réveils inefficace.
- ✦ un accès disque inefficace. Utilisez des mémoires-tampons de grande taille pour éviter un accès au disque trop fréquent. N'écrivez qu'un grand bloc à la fois.
- ✦ une utilisation inefficace des minuteurs. Groupez les minuteurs sur plusieurs applications (ou même sur plusieurs systèmes) si possible.
- ✦ des E/S, ou une consommation d'énergie ou utilisation de mémoire excessive (y compris les fuites mémoires)
- ✦ l'exécution de calculs inutiles.

Les sections suivantes examinent certains de ces domaines en détails.

### A.1. Utilisation de threads

Il est communément admis que l'utilisation de threads permet aux applications de fonctionner mieux et plus vite, mais ce n'est pas toujours le cas.

#### Python

Python uses the Global Lock Interpreter <sup>[1]</sup>, so threading is profitable only for larger I/O operations. **Unladen-swallow** <sup>[2]</sup> is a faster implementation of Python with which you might be able to optimize your code.

#### Perl

Perl threads were originally created for applications running on systems without forking (such as systems with 32-bit Windows operating systems). In Perl threads, the data is copied for every single thread (Copy On Write). Data is not shared by default, because users should be able to define the level of data sharing. For data sharing the **threads::shared** module has to be included. However, data is not only then copied (Copy On Write), but the module also creates tied variables for the data, which takes even more time and is even slower. <sup>[3]</sup>

#### C

C threads share the same memory, each thread has its own stack, and the kernel does not have to create new file descriptors and allocate new memory space. C can really use the support of more CPUs for more threads. Therefore, to maximize the performance of your threads, use a low-level language like C or C++. If you use a scripting language, consider writing a C binding. Use profilers to identify poorly performing parts of your code. [4]

## A.2. Réveils

De nombreuses applications scannent les fichiers de configuration en cherchant des modifications. Dans de nombreux cas, le scan est effectué à intervalle fixe, toutes les minutes par exemple. Ceci peut poser problème, car le disque est forcé de se réveiller après les périodes d'interruption de la rotation du disque. La meilleure solution est de trouver un bon intervalle, un bon mécanisme de vérification, ou de vérifier les modifications avec **inotify** et de réagir aux événements. **Inotify** peut vérifier une variété des changements sur un fichier ou sur un répertoire.

Par exemple :

```
#include <stdio.h>
#include <stdlib.h>
#include <sys/time.h>
#include <sys/types.h>
#include <sys/inotify.h>
#include <unistd.h>

int main(int argc, char *argv[]) {
    int fd;
    int wd;
    int retval;
    struct timeval tv;

    fd = inotify_init();

    /* checking modification of a file - writing into */
    wd = inotify_add_watch(fd, "./myConfig", IN_MODIFY);
    if (wd < 0) {
        printf("inotify cannot be used\n");
        /* switch back to previous checking */
    }

    fd_set rfd;
    FD_ZERO(&rfd);
    FD_SET(fd, &rfd);
    tv.tv_sec = 5;
    tv.tv_usec = 0;
    retval = select(fd + 1, &rfd, NULL, NULL, &tv);
    if (retval == -1)
        perror("select()");
    else if (retval) {
        printf("file was modified\n");
    }
    else
        printf("timeout\n");

    return EXIT_SUCCESS;
}
```

L'avantage de cette approche est la variété des vérifications que vous pouvez effectuer.

La limitation principale est que le nombre de suivis disponibles sur un système est limité. Ce nombre peut être obtenu sur `/proc/sys/fs/inotify/max_user_watches` et même s'il peut être changé, ceci n'est pas recommandé. De plus, au cas où `inotify` échouerait, le code devra utiliser une différente méthode de vérification, ce qui devrait signifier qu'il y aura de nombreuses occurrences de `#if #define` dans le code source.

For more information on `inotify`, refer to the `inotify(7)` man page.

### A.3. Fsync

`Fsync` est connu pour être une opération utilisant beaucoup d'E/S, mais cela n'est pas totalement vrai.

`Firefox` avait pour habitude d'appeler la bibliothèque `sqlite` à chaque fois que l'utilisateur cliquait sur un lien vers une nouvelle page. `Sqlite` appelait `fsync` et à cause des paramètres du système de fichiers (principalement `ext3` avec le mode `data-ordered`), il y avait une longue latence lorsque rien ne se passait. Ceci pouvait ainsi prendre longtemps (jusqu'à 30 secondes) si un autre processus copiait un fichier de grande taille au même moment.

Cependant, dans d'autres cas où `fsync` n'était pas du tout utilisé, des problèmes sont survenus lors du passage au système de fichiers `ext4`. `Ext3` était réglé sur un mode de données ordonnées, qui vidait la mémoire toutes les quelques secondes et l'enregistrait sur un disque. Mais avec `ext4` et `laptop_mode`, l'intervalle entre les sauvegardes est devenu plus long et des données pouvaient se perdre lorsque le système était éteint de manière inattendue. Maintenant `ext4` est corrigé, mais nous devons quand même prendre en considération le design de nos applications en faisant attention, et utiliser `fsync` lorsqu'approprié.

L'exemple suivant illustrant simplement la lecture et l'écriture dans un fichier de configuration montre comment une copie de sauvegarde peut être réalisée, ou comment des données peuvent être perdues :

```
/* open and read configuration file e.g. ./myconfig */
fd = open("./myconfig", O_RDONLY);
read(fd, myconfig_buf, sizeof(myconfig_buf));
close(fd);
...
fd = open("./myconfig", O_WRONLY | O_TRUNC | O_CREAT, S_IRUSR | S_IWUSR);
write(fd, myconfig_buf, sizeof(myconfig_buf));
close(fd);
```

Une meilleure approche aurait été comme suit :

```
/* open and read configuration file e.g. ./myconfig */
fd = open("./myconfig", O_RDONLY);
read(fd, myconfig_buf, sizeof(myconfig_buf));
close(fd);
...
fd = open("./myconfig.suffix", O_WRONLY | O_TRUNC | O_CREAT, S_IRUSR |
S_IWUSR
write(fd, myconfig_buf, sizeof(myconfig_buf));
fsync(fd); /* paranoia - optional */
...
close(fd);
rename("./myconfig", "./myconfig~"); /* paranoia - optional */
rename("./myconfig.suffix", "./myconfig");
```

---

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

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

[3] [http://www.perlmonks.org/?node\\_id=288022](http://www.perlmonks.org/?node_id=288022)

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

## Annexe B. Historique des versions

<b>Version 2.1-3.1</b>	<b>Thu Jun 22 2017</b>	<b>Terry Chuang</b>
Fichiers de traduction synchronisés avec les sources XML 2.1-3		
<b>Version 2.1-3</b>	<b>Mon Oct 24 2016</b>	<b>Marie Doleželová</b>
Version pour la distribution GA 7.3.		
<b>Version 2.0-1</b>	<b>Wed 11 Nov 2015</b>	<b>Jana Heves</b>
Distribution 7.2 GA.		
<b>Version 1-3</b>	<b>Fri 19 Jun 2015</b>	<b>Jacquelynn Est</b>
Correction de nom de package erroné dans Infrastructure principale et mécanismes.		
<b>Version 1-2</b>	<b>Wed 18 Feb 2015</b>	<b>Jacquelynn Est</b>
Version 7.1 GA.		
<b>Version 1-1</b>	<b>Thu Dec 4 2014</b>	<b>Jacquelynn Est</b>
Version 7.1 Beta		
<b>Version 0.9-9</b>	<b>Tue Jun 9 2014</b>	<b>Yoana Ruseva</b>
Version pour la distribution GA 7.0		
<b>Version 0.9-1</b>	<b>Fri May 9 2014</b>	<b>Yoana Ruseva</b>
Reconstruit pour les changements de style.		
<b>Version 0.9-0</b>	<b>Wed May 7 2014</b>	<b>Yoana Ruseva</b>
Sortie de Red Hat Enterprise Linux 7.0 à but de révision.		
<b>Version 0.1-1</b>	<b>Thu Jan 17 2013</b>	<b>Jack Reed</b>
Branché à partir de la version Red Hat Enterprise Linux 6 de ce document		