



redhat.[®]

RH442

RED HAT PERFORMANCE TUNING

Red Hat Performance Tuning
Red Hat Enterprise Linux 6.2
Release en-2-20111219

RED HAT PERFORMANCE TUNING

Red Hat Enterprise Linux 6.2 RH442

Red Hat Performance Tuning

Edition 2

Author	Forrest Taylor
Author	Wander Boessenkool
Author	Bowe Strickland
Author	Scott McBrien
Author	George Hacker
Editor	Steven Bonneville

Copyright © 2011 Red Hat, Inc.

The contents of this course and all its modules and related materials, including handouts to audience members, are Copyright © 2011 Red Hat, Inc.

No part of this publication may be stored in a retrieval system, transmitted or reproduced in any way, including, but not limited to, photocopy, photograph, magnetic, electronic or other record, without the prior written permission of Red Hat, Inc.

This instructional program, including all material provided herein, is supplied without any guarantees from Red Hat, Inc. Red Hat, Inc. assumes no liability for damages or legal action arising from the use or misuse of contents or details contained herein.

If you believe Red Hat training materials are being used, copied, or otherwise improperly distributed please e-mail training@redhat.com or phone toll-free (USA) +1 (866) 626-2994 or +1 (919) 754-3700.

Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, Hibernate, 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 registered trademark of Silicon Graphics International Corp. or its subsidiaries in the United States and/or other countries.

All other trademarks are the property of their respective owners.

Document Conventions	vii
Notes and Warnings	vii
Introduction	ix
Welcome to class!	ix
About Red Hat Enterprise Linux	ix
Additional Red Hat Enterprise Linux Software	x
Contacting Red Hat Technical Support	xii
About This Course	xv
Red Hat Performance Tuning	xv
Structure of the Course	xv
Orientation to the Classroom Network	xvi
Internationalization	xvii
Language Support	xvii
System-wide Default Language	xvii
Per-user Language Selection	xvii
Input Methods	xviii
Language Codes Reference	xviii
1. Introduction to Performance Tuning	1
Performance Tuning Is	2
Performance Level Agreements	3
Disable Unused Services	4
Monitoring vs. Profiling	5
2. Collecting, Graphing and Interpreting Data	9
Units and Unit Conversions	10
Profiling Tools	14
Using awk to Format Data	20
Plotting Data	24
Unit Test	30
3. General Tuning	33
Queueing Theory	34
Asymptotic complexity	37
Displaying and Configuring Module Parameters	40
Installing and enabling tuned	43
Configuring tuned	45
File systems: Fragmentation and RAID layouts	47
Unit Test	51
4. Limiting Resource Usage	55
Limiting Resource Usage	56
Unit Test	61
5. Hardware Profiling	65
Generating a Hardware Profile	66
Determining SMBIOS/DMI Information	76
Generating a Profile of the Whole System	80
Assessing NUMA Topologies	83
Profiling Storage	87
I/O Scheduling	91

Unit Test	95
6. Software Profiling	99
CPU Scheduling	100
strace and ltrace Usage	106
Use valgrind to Profile Cache Usage	110
7. Systemtap	123
Configure and Use systemtap	124
Unit Test	131
8. Mail Server Tuning	135
Rotational Delay and Disk Elevators	136
Further Mail Server Tuning Tips	143
Unit Test	144
9. Large Memory Workload Tuning	147
Memory Management	148
Finding Memory Leaks	156
Tuning Swap	158
Memory Reclamation	162
Non-Uniform Memory Access (NUMA)	168
Unit Test	170
10. Tuning for a CPU Intensive Workload	173
Limiting CPU Access with cgroups	174
Balancing Interrupts	176
Pin Processes to a Specific CPU with cgroups	178
Realtime Scheduling	180
Unit Test	181
11. File Server Tuning	185
File System Journaling	186
Selecting a tuned Profile for a File Server Workload	189
Network Performance Tuning	191
BDP and Window Scaling	194
Bonding and Link Aggregation	198
Jumbo Frames	201
Unit Test	203
12. Database Server Tuning	207
Database Server tuned Profile	208
Tune the Network for Latency	210
Tune SysV IPCS	212
Huge Pages	215
Tune overcommit and swappiness	218
Unit Test	220
13. Power Usage Tuning	223
Power Saving Strategies	224
Power Saving tuned Profiles	227
Profile Power Usage with powertop	230
Unit Test	233
14. Virtualization Tuning	237

Virtualization tuned Profile	238
CPU Pinning	240
Kernel Samepage Merging (KSM)	244
Limit Virtualization Guests Using cgroups	247
Virtual Machine Storage	250
Unit Test	252
A. Solutions	255
Collecting, Graphing and Interpreting Data	255
General Tuning	262
Limiting Resource Usage	266
Hardware Profiling	268
Software Profiling	272
Systemtap	275
Mail Server Tuning	278
Large Memory Workload Tuning	281
Tuning for a CPU Intensive Workload	285
File Server Tuning	290
Database Server Tuning	298
Power Usage Tuning	307
Virtualization Tuning	313

Document Conventions

Notes and Warnings



Note

"Notes" are tips, shortcuts or alternative approaches to the task at hand. Ignoring a note should have no negative consequences, but you might miss out on a trick that makes your life easier.



Comparison

"Comparisons" look at similarities and differences between the technology or topic being discussed and similar technologies or topics in other operating systems or environments.



References

"References" describe where to find external documentation relevant to a subject.



Important

"Important" boxes detail things that are easily missed: configuration changes that only apply to the current session, or services that need restarting before an update will apply. Ignoring a box labeled "Important" will not cause data loss, but may cause irritation and frustration.



Warning

"Warnings" should not be ignored. Ignoring warnings will most likely cause data loss.

Introduction

Welcome to class!

Thank you for attending this Red Hat training class. Please let us know if you have any special needs while at our training facility.

Please ask the instructor if you have any questions about the facility, such as operating hours of the facility and when you will have access to the classroom, locations of restrooms and break rooms, availability of telephones and network connectivity, and information about the local area.

As a courtesy to other students, please place your pager or cell phone's ringer on vibrate or mute, or turn off your devices during class. We ask that you only make calls during break periods.

If you have a personal emergency and are unable to attend or complete the class, please let us know. Thank you!

About Red Hat Enterprise Linux

This course is taught using Red Hat Enterprise Linux, an enterprise-targeted Linux distribution focused on mature open source software designed specifically for organizations using Linux in production settings.

Red Hat Enterprise Linux is sold on a subscription basis, where the subscription gives you continuous access to all supported versions of the operating system in binary and source form, not just the latest one, including all updates and bug fixes. Extensive support services are included which vary depending on the subscription level selected; see <http://www.redhat.com/subscriptions/> for details. Various Service Level Agreements are available that may provide up to 24x7 coverage with a guaranteed one hour response time for Severity 1 issues. Support will be available for up to seven years after a particular major release (ten years with the optional "Extended Lifecycle Support" Add-On).

Red Hat Enterprise Linux is released on a multi-year cycle between major releases. Minor updates to major releases are released periodically during the lifecycle of the product. Systems certified on one minor update of a major release continue to be certified for future minor updates of the major release. A core set of shared libraries have APIs and ABIs which will be preserved between major releases. Many other shared libraries are provided, which have APIs and ABIs which are guaranteed within a major release (for all minor updates) but which are not guaranteed to be stable across major releases.

Red Hat Enterprise Linux is based on code developed by the open source community, which is often first packaged through the Red Hat sponsored, freely-available Fedora distribution (<http://fedoraproject.org/>). Red Hat then adds performance enhancements, intensive testing, and certification on products produced by top independent software and hardware vendors. Red Hat Enterprise Linux provides a high degree of standardization through its support for four processor architectures (32-bit Intel x86-compatible, AMD64/Intel 64 (x86-64), IBM POWER, and IBM mainframe on System z). Furthermore, we support the 4000+ ISV certifications on Red Hat Enterprise Linux whether the RHEL operating system those applications are using

is running on “bare metal”, in a virtual machine, as a software appliance, or in the cloud using technologies such as Amazon EC2.

Currently, the Red Hat Enterprise Linux product family includes:

- *Red Hat Enterprise Linux for Servers*: the datacenter platform for mission-critical servers running Red Hat Enterprise Linux. This product includes support for the largest x86-64 and x86-compatible servers and the highest levels of technical support, deployable on bare metal, as a guest on the major hypervisors, or in the cloud. Subscriptions are available with flexible guest entitlements of one, four, or unlimited guests per physical host. Pricing is based on the basis of the number of socket-pairs populated on the system motherboard, the number of guests supported, the level of support desired, and the length of subscription desired.

Red Hat Enterprise Linux for IBM POWER and *Red Hat Enterprise Linux for IBM System z* are similar variants intended for those system architectures.

- *Red Hat Enterprise Linux Desktop*: built for the administrator and end-user, Red Hat Enterprise Linux Desktop provides an attractive and highly productive environment for knowledge workers on desktops and laptops. Client installations can be finely tailored and locked down for simplicity and security for any workstation task.

The basic *Desktop* variant is designed for task workers who have a limited amount of administrative control over the system, who primarily use productivity applications like Firefox Evolution/Thunderbird, OpenOffice.org, and Planner/TaskJuggler. The more sophisticated *Workstation* variant is designed for advanced Linux users who need a stand-alone development environment, and who are expected to have local super-user privileges or selected super-user privileges.

In addition, other variants exist such as *Red Hat Enterprise Linux for HPC Head Node* and *Red Hat Enterprise Linux for HPC Compute Node* (targeted at high-performance computing clusters), and *Red Hat Enterprise Linux for SAP Business Applications*. For more information please visit <http://www.redhat.com/>.

Additional Red Hat Enterprise Linux Software

Two additional software update channels are provided with Red Hat Enterprise Linux beyond the core software packages shipped:

- *Supplementary*: the “Supplementary” channel provides selected closed source packages, built for Red Hat Enterprise Linux as a convenience to the customer. These include things like Adobe Flash or proprietary Java JVMs.
- *Optional*: the “Optional” channel provides selected open source packages, as a convenience only. They are generally included in another Red Hat Enterprise Linux variant as a fully-supported package, or are a build requirement for the distribution. These packages are only available through a Red Hat Network child channel.



Important

Supplementary and *Optional* packages are provided with limited support, as a customer convenience only.

Red Hat also offers a portfolio of fully-supported *Add-Ons for Red Hat Enterprise Linux* which extend the features of your Red Hat Enterprise Linux subscription. These add-ons allow you to add capabilities and tailor your computing environment to your particular needs. These Add-Ons include support for high availability application clustering, cluster file systems and very large file systems, enhanced system management with Red Hat Network, extended update support for minor releases, and more.



Note

Please visit <http://www.redhat.com/rhel/add-ons/> for more information about available *Add-Ons for Red Hat Enterprise Linux*.

For information about other products which are provided by Red Hat, such as Red Hat Enterprise Virtualization, JBoss Enterprise Middleware, Red Hat Enterprise MRG, and various custom consulting and engineering services, <http://www.redhat.com/products/> also has useful information.

The Fedora Project also provides additional packages for Red Hat Enterprise Linux through *EPEL* (*Extra Packages for Enterprise Linux*). EPEL is a volunteer-based community effort to create a repository of high-quality add-on packages which can be used with Red Hat Enterprise Linux and compatible derivatives. It accepts legally-unencumbered free and open source software which does not conflict with software in Red Hat Enterprise Linux or Red Hat add-on products. EPEL packages are built for a particular major release of Red Hat Enterprise Linux and will be updated by EPEL for the standard support lifetime of that major release.

Red Hat does not provide commercial support or service level agreements for EPEL packages. While not supported by Red Hat, EPEL provides a useful way to reduce support costs for unsupported packages which your enterprise wishes to use with Red Hat Enterprise Linux. EPEL allows you to distribute support work you would need to do by yourself across other organizations which share your desire to use this open source software with RHEL. The software packages themselves go through the same review process as Fedora packages, meaning that experienced Linux developers have examined the packages for issues. As EPEL does not replace or conflict with software packages shipped in RHEL, you can use EPEL with confidence that it will not cause problems with your normal software packages.

For developers who wish to see their open source software become part of Red Hat Enterprise Linux, often a first stage is to sponsor it in EPEL so that RHEL users have the opportunity to use it, and so experience is gained with managing the package for a Red Hat distribution.

Visit <http://fedoraproject.org/wiki/EPEL> for more information about EPEL.



Important

EPEL is supported by the community-managed Fedora Project and not by Red Hat Support.

Contacting Red Hat Technical Support

One of the benefits of your subscription to Red Hat Enterprise Linux is access to technical support through Red Hat's customer portal at <http://access.redhat.com/>. If you do not have a Red Hat account on the customer portal or are not able to log in, you can go to <https://access.redhat.com/support/faq/LoginAssistance.html> or contact Customer Service for assistance.

You may be able to resolve your problem without formal technical support by searching Knowledgebase (<https://access.redhat.com/kb/knowledgebase/>). Otherwise, depending on your support level, Red Hat Support may be contacted through a web form or by phone. Phone numbers and business hours for different regions vary; see <https://access.redhat.com/support/contact/technicalSupport.html> for current information. Information about the support process is available at https://access.redhat.com/support/policy/support_process.html.

Some tips on preparing your bug report to most effectively engage Red Hat Support:

- *Define the problem.* Make certain that you can articulate the problem and its symptoms before you contact Red Hat. Be as specific as possible, and detail the steps you can use (if any) to reproduce the problem.
- *Gather background information.* What version of our software are you running? Are you using the latest update? What steps led to the failure? Can the problem be recreated and what steps are required? Have any recent changes been made that could have triggered the issue? Were messages or other diagnostic messages issued? What exactly were they (exact wording may be critical)?
- *Gather relevant diagnostic information.* Be ready to provide as much relevant information as possible; logs, core dumps, traces, the output of **sosreport**, etc. Technical Support can assist you in determining what is relevant.
- *Determine the Severity Level of your issue.* Red Hat uses a four-level scale to indicate the criticality of issues; criteria may be found at https://access.redhat.com/support/policy/GSS_severity.html.



Warning

Bugzilla is not a support tool! For support issues affecting Red Hat Enterprise Linux, customers should file their bugs through the support channels discussed above in order to ensure that Red Hat is fully aware of your issue and can respond under the terms of your Service Level Agreement. Customers should *not* file bugs directly in the <http://bugzilla.redhat.com/> web interface.

For Red Hat Enterprise Linux, Bugzilla is used by engineering to track issues and changes, and to communicate on a technical level with Engineering partners and other external parties. Anyone, even non-customers, can file issues against Bugzilla, and Red Hat does monitor them and review them for inclusion in errata.

However, Red Hat does not guarantee any SLA for bugs filed directly in Bugzilla (bypassing normal support channels). A review might happen immediately, or after a time span of any length. Issues coming through Support are always prioritized above issues of similar impact and severity filed against Bugzilla. Also, workarounds and hotfixes if possible and appropriate may be provided to customers by Support even before a permanent fix is issued through Red Hat Network.

Red Hat considers issues directly entered into Bugzilla important feedback, and it allows us to provide efficient interaction with the open source development community and as much transparency as possible to customers as issues are processed. Nevertheless, for customers encountering production issues in Red Hat Enterprise Linux, Bugzilla is not the right channel.

About This Course

Red Hat Performance Tuning

The Red Hat Performance Tuning course strives to teach real-world skills with performance tuning. This course has been enhanced with case studies that focus on specific scenarios encountered by system administrators the world over. Queuing theory and other concepts are taught to help analyze and tune systems for best performance for a given purpose.

Objectives

- Implement open source tools to profile memory, CPU, network and other subsystems
- Tune systems for a certain workload (e.g., HPC, Large Memory, Database Server, File Server, etc)
- Tune systems based on monitoring and tuning analysis
- Tune virtual machines (host and guest)
- Understand hardware designs and limitations
- Understand queuing theory, BDP and other tuning concepts

Audience and Prerequisites

- Students who are senior Linux system administrators with *at least* five years of full-time Linux experience, preferably using Red Hat Enterprise Linux
- Students should enter the class with current RHCE credentials.

Structure of the Course

Red Hat training courses are interactive, hands-on, performance-based, real world classes meant to engage your mind and give you an opportunity to use real systems to develop real skills. We encourage students to participate in class and ask questions in order to get the most out of their training sessions.

This course is divided up into a number of *Units* organized around a particular topic area. Each Unit is divided up into multiple *Sections* which focus on a specific skill or task. The unit will start with an introduction to the material, then move on to the first section.

In each section, there will be a *presentation* led by the instructor. During the presentation, it may be a good idea to take notes in your student workbook (this book), and the instructor may remind you to do so. The presentation is followed by a short activity or *assessment* to give you the opportunity to practice with the material or review procedures. After a review of the assessment, the instructor will move on to the next section. At the end of the unit, there will normally be a hands-on lab exercise of some sort (a "unit test") which will give you an opportunity to learn by doing and review your understanding of the unit's content. Please feel free to ask questions in

class, or asking the instructor for advice and help during the end-of-unit exercise. We want the classroom environment to be a "low risk" place where you feel comfortable asking questions and learning from things that work and things that do not at first.

Orientation to the Classroom Network

Two subnets may be used in this course. The primary classroom network is 192.168.0.0/24, and belongs to hosts in the DNS domain "example.com". This network will be used for most classroom activities. Some courses use a second subnet, 192.168.1.0/24, belonging to hosts in the DNS domain "remote.test". This network can be reached from hosts in example.com, and is used in lab exercises which require testing services or security settings from machines (theoretically) outside your administrative control.

Students are each assigned a physical machine (desktopX.example.com on 192.168.0.X) which may host two or more virtual machines for lab activities, serverX.example.com and hostX.example.com.

In some courses, students may also use a non-root account on a test machine in the remote.test domain, remoteX.example.com (192.168.1.X) to test access to network services on their example.com machines in lab activities.

The instructor controls a number of machines which students may see as well. The machine instructor.example.com (also known as instructor.remote.test) is the classroom utility server, providing default routing services, DHCP, DNS name service, one or more YUM repositories of software used by the class, and other network services. It is also connected to the classroom video projector to allow the instructor to display slides and demonstrations. It provides a virtual machine for the instructor, demo.example.com, which the instructor will use for in-class demonstrations.

Machine name	IP addresses	Role
desktopX.example.com	192.168.0.X	Physical student workstation
serverX.example.com	192.168.0.(X+100)	Main student virtual machine
hostX.example.com	192.168.0.(X+200)	Secondary student virtual machine
remoteX.remote.test	192.168.1.X	Student test machine in remote.test domain (shared)
instructor.example.com	192.168.0.254	Physical instructor machine and utility server
instructor.remote.test	192.168.1.254	Identity of instructor.example.com on remote.test network
demo.example.com	192.168.0.250	Instructor virtual demonstration machine

Classroom Machines

Internationalization

Language Support

Red Hat Enterprise Linux 6 officially supports twenty-two languages: English, Assamese, Bengali, Chinese (Simplified), Chinese (Traditional), French, German, Gujarati, Hindi, Italian, Japanese, Kannada, Korean, Malayalam, Marathi, Oriya, Portuguese (Brazilian), Punjabi, Russian, Spanish, Tamil, and Telugu. Support for Maithili, Nepalese, and Sinhala are provided as Technology Previews.

System-wide Default Language

The operating system's default language is normally set to US English (en_US.UTF-8), but this can be changed during or after installation.

To use other languages, you may need to install additional package groups to provide the appropriate fonts, translations, dictionaries, and so forth. By convention, these package groups are always named **language-support**. These package groups can be selected during installation, or after installation with PackageKit (System > Administration > Add/Remove Software) or **yum**.

A system's default language can be changed with **system-config-language** (System > Administration > Language), which affects the **/etc/sysconfig/i18n** file.

Per-user Language Selection

Users may prefer to use a different language for their own desktop environment or interactive shells than is set as the system default. This is indicated to the system through the **LANG** environment variable.

This may be set automatically for the GNOME desktop environment by selecting a language from the graphical login screen by clicking on the **Language** item at the bottom left corner of the graphical login screen immediately prior to login. The user will be prompted about whether the language selected should be used just for this one login session or as a default for the user from now on. The setting is saved in the user's **~/.dmrc** file by GDM.

If a user wants to make their shell environment use the same **LANG** setting as their graphical environment even when they login through a text console or over **ssh**, they can set code similar to the following in their **~/.bashrc** file. This code will set their preferred language if one is saved in **~/.dmrc** or will use the system default if one is not:

```
i=$(grep 'Language=' ${HOME}/.dmrc | sed 's/Language=//')
if [ "$i" != "" ]; then
    export LANG=$i
fi
```

Languages with non-ASCII characters may have problems displaying in some environments. Kanji characters, for example, may not display as expected on a virtual console. Individual commands can be made to use another language by setting **LANG** on the command-line:

```
[user@host ~]$ LANG=fr_FR.UTF-8 date
lun. oct. 24 10:37:53 CDT 2011
```

Subsequent commands will revert to using the system's default language for output. The **locale** command can be used to check the current value of **LANG** and other related environment variables.

Input Methods

IBus (Intelligent Input Bus) can be used to input text in various languages under X if the appropriate language support packages are installed. You can enable IBus with the **im-chooser** command (**System > Preferences > Input Method**).

Language Codes Reference

Language	\$LANG value	Language package group
English (US)	en_US.UTF-8	(<i>default</i>)
Assamese	as_IN.UTF-8	assamese-support
Bengali	bn_IN.UTF-8	bengali-support
Chinese (Simplified)	zh_CN.UTF-8	chinese-support
Chinese (Traditional)	zh_TW.UTF-8	chinese-support
French	fr_FR.UTF-8	french-support
German	de_DE.UTF-8	german-support
Gujarati	gu_IN.UTF-8	gujarati-support
Hindi	hi_IN.UTF-8	hindi-support
Italian	it_IT.UTF-8	italian-support
Japanese	ja_JP.UTF-8	japanese-support
Kannada	kn_IN.UTF-8	kannada-support
Korean	ko_KR.UTF-8	korean-support
Malayalam	ml_IN.UTF-8	malayalam-support
Marathi	mr_IN.UTF-8	marathi-support
Oriya	or_IN.UTF-8	oriya-support
Portuguese (Brazilian)	pt_BR.UTF-8	brazilian-support
Punjabi	pa_IN.UTF-8	punjabi-support
Russian	ru_RU.UTF-8	russian-support

Language	\$LANG value	Language package group
Spanish	es_ES.UTF-8	spanish-support
Tamil	ta_IN.UTF-8	tamil-support
Telugu	te_IN.UTF-8	telugu-support
<i>Technology Previews</i>		
Maithili	mai_IN.UTF-8	maithili-support
Nepali	ne_NP.UTF-8	nepali-support
Sinhala	si_LK.UTF-8	sinhala-support

Language Codes



UNIT 1

INTRODUCTION TO PERFORMANCE TUNING

Introduction

Unit Details	
Unit Goal	Have a high-level understanding of performance tuning and its goals
Performance Objectives	<ul style="list-style-type: none">• Understand the goals of performance tuning• Prepare performance level agreements• Disable unused services• Understand how system monitoring and profiling benefit performance tuning
Unit Sections	<ul style="list-style-type: none">• Performance Tuning Is ...• Performance Level Agreements• Disable Unused Services• Monitoring vs. Profiling

Performance Tuning Is ...

Performance tuning has sometimes been called a "black art" due to the sometimes arcane knowledge required to effectively tune a system. It requires an understanding of both the hardware and the software components of a system as well as the many interactions between the systems. Performance tuning is the process of tailoring the configuration of a system, or systems, for a particular application so that the application gets the best possible response time or throughput.

Tuning goals must be established. Is the system used by interactive users or is it used to perform batch processing of reports? Often there are tradeoffs, for example a system could be tuned for low latency (fast response time) at the expense of high throughput. Real-time applications could require predictable latency, so the system would be tuned to minimize or eliminate unpredictable latency.

Human factors must also be considered. Management may ask for improvements to perceived performance versus actual system efficiency. The microprocessor that is at the heart of your computer system does not care whether it is running at five percent utilization or ninety-five percent utilization, but you might have a definite opinion about which situation is preferable. Usually, performance tuning is undertaken to remove blockages that slow data transfer. Distribution of workload by equipment or time of day can sometimes be used to recover scarce system resources during peak times.

Performance management is the process of making sure that adequate computing resources are available to accomplish the business needs of all users. As part of the process of performance management you may have to manage or set expectations about what is considered acceptable performance for a system. Users may have unrealistic performance expectations because they do not compensate for the physical characteristics of hardware or other factors.