RED HAT ENTERPRISE LINUX 5.3 TECHNICAL OVERVIEW

CONTENTS

PURPOSE

In January of 2009, Red Hat delivered the latest release in the Red Hat Enterprise Linux 5 family. In keeping with Red Hat's goals of stable, scalable, and supportable high-performance open source software, Red Hat Enterprise Linux 5.3 delivers enhancements to the core operating system, the Linux developer environment, and the integrated infrastructure. Although there are nearly 3,400 separate changes comprising Red Hat Enterprise Linux 5.3, this paper will concentrate on the new platform features that are relevant for developers, system administrators, and IT architects. New tools for performance and debugging make Red Hat Enterprise Linux 5.3 easier to use for Linux developers and maintains the ABI/API stability that they expect from a Red Hat subscription. For Java developers, the first certified Open Source JDK is included with this release. By integrating OpenJDK with the operating system, Red Hat becomes one of the few companies that has a complete enterprise Java environment.

Red Hat works very closely with the IHV community and it is reflected in the expanded hardware support in Red Hat Enterprise Linux 5.3. Full support for Intel's next-generation Nehalem processors and AMD's next chipset are included in this release. By working collaboratively, the CPU advancements in virtualization will be immediately available to those Red Hat subscribers when they acquire the latest hardware. In addition, Red Hat provides several new options for network storage and file systems to meet the diverse needs of Red Hat customers. Virtualization performance and management shows the results of continuous improvement as Red Hat sets new levels of scalability in both physical and virtual operations.

Before describing new features, it is important to keep in mind that the key values of any update Red Hat Enterprise Linux release are the hundreds of fixes that won't get detailed mention in this whitepaper. These fixes comprise the majority of the maintenance work that makes Red Hat Enterprise Linux the leading enterprise distribution. They address widely diverse bugs discovered during Red Hat's ongoing proactive testing initiatives, as well as through customer reports. Of equal importance is Red Hat's ongoing vigilance to plug all known security vulnerabilities. Red Hat Enterprise Linux update releases aggregate the full collection of bug fixes and security hardening into a fully integrated and tested release distribution. This work is the cornerstone of Red Hat Enterprise Linux. The remainder of this whitepaper focuses on the new features which debut in Red Hat Enterprise Linux 5.3.

CAPABILITIES

This paper is organized into functional areas. Note that some features span multiple functional areas, including:

- Virtualization: describing scalability and performance enhancements

- Development / deployment tools

- Hardware enablement: including performance enhancements and subsections on platform features and device drivers

- Filesystem and storage: enhancements above the basic block driver layer

- System services: including new security features and performance tuning

- Cluster: storage and control capabilities

Changes in each of these areas are described in the upcoming sections. Please keep in mind that the paper contains only high-level summary information and is not comprehensive.

VIRTUALIZATION

One of the most compelling features that debuted in Red Hat Enterprise Linux 5 was fully integrated support for virtualization. A huge amount of work went into the integration, all aimed at providing customers with a simple and consistent configuration and operation experience. Virtualization remains a substantial emphasis for Red Hat's development team, continuing the improvements in Red Hat Enterprise Linux 5.3.

Since the initial delivery of Red Hat Enterprise Linux 5 in March 2007, the growing scalability of x86-64-based hardware has provided the motivation to support increasingly large virtualization platforms. In addition to enhancements to the core hypervisor (hosting) layer, Red Hat Enterprise Linux has also received improved guest capabilities. These enhancements provide benefits for customers who wish to deploy a few large guest instances, and also those who wish to deploy numerous smaller guests. Both deployment styles are valid, as customers are increasingly using virtualization to lower total cost of ownership (TCO) by increasing system management flexibility, for example by enabling migration of workloads based on growth needs, and for high availability, such as via guest instance failover and migration for planned maintenance. Examples of the scalability enhancements in Red Hat Enterprise Linux 5.3,[1] include:

- Support for up to physical 126 CPUs and 32 CPUs per virtual server

- Support for up to 1TB of memory per server and 80GB per virtual server

- Support for more than 16 disks per guest

- Support for more than four network adapters per guest

---

1   System scalability limits describe the practical software support limits. Certification of hardware configurations is dependent on Red Hat's requirements of passing the certification test suite and availability of systems for Red Hat's development and test teams to have direct access to for testing purposes.

Improved paravirtualization has been an important factor in driving the demand for increased scalability. In traditional fully virtualized environments, application workloads that had high levels of network and disk IO could incur up to 30 percent performance overhead compared to bare-metal deployments. Paravirtualization provides device drivers that cleanly plug into older Red Hat Enterprise Linux releases, which operate as virtualized guests in Red Hat Enterprise Linux 5. These paravirtualization device drivers are able to utilize enhanced hardware capabilities to bypass the majority of the virtualization overhead, resulting in minimal performance degradation. This allows IO-intensive applications to become candidates for virtualization. Red Hat Enterprise Linux 5.3 includes enhancements to the previously existing paravirtualization drivers as well as optimizations such as utilizing large 2MB page tables. These paravirtualization enhancements enable deployment of applications such as database and messaging workloads in virtualized environments, thereby driving the need for increasingly larger hardware configurations.

Another new virtualization feature in Red Hat Enterprise Linux 5.3 is **libvirt-cim**. Since the initial release of Red Hat Enterprise Linux 5, Red Hat has provided an abstraction layer called libvirt as the system management interface to virtualization. Libvirt is designed to hide the differences and changes in low-level virtualization implementations and changing interfaces from the system management tools. The libvirt library has proven popular and has been increasingly adopted and utilized by a variety of system management tool vendors. This benefits customers, who can choose the tools with which they are familiar and which best meet their use cases. The vibrant community that has rallied around libvirt has truly been a win-win. Meanwhile, numerous commercial system management frameworks utilize an architecture called Common Information Model (CIM) as the interface to interact with managed services. Combining the features of both management standards, Red Hat Enterprise Linux 5.3 introduces libvirt-cim, widening the set of virtualization configuration and operational management capabilities to include CIM compliant interfaces.

DEVELOPMENT AND DEPLOYMENT TOOLS

SYSTEMTAP

A compelling new feature in Red Hat Enterprise Linux 5.3 is called SystemTap. This utility provides groundbreaking new capabilities for application performance monitoring and problem diagnosis. What makes SystemTap special is the fact that it enables customers to dynamically monitor applications live in production deployments. For example, SystemTap allows for the creation of runtime trace points where performance-related statistics can be gathered. As useful as raw performance statistics are, SystemTap takes it a step further by providing a scripting language which allows for dynamic insertion of runtime code. This allows for sophisticated instrumentation. For example, it is possible to trigger the execution of debugging code when the value of a certain parameter exceeds a threshold value. Previously, it was necessary to build a version of the application with the debug code and deploy it into production environments in order to diagnose issues. In practice, that approach is problematic because complex production environments consist of layers of software like kernel, operating system libraries, middleware, and business logic, which are often provided by a variety of vendors. The prospect of instrumenting and delivering new debug versions into production environments to diagnose issues was extremely time-consuming at best.

SystemTap overcomes these obstacles by enabling dynamic addition of diagnostic code at runtime. This obviates the need for custom debug versions of software. Instead, all that is needed is the creation of small debug modules, referred to as SystemTap scripts, which get dynamically loaded onto the running system to produce the desired diagnostic information. SystemTap capabilities span from low-level kernel, to runtime system libraries and up through the application stack. For example, it is possible to diagnose whether performance bottlenecks are in the kernel, glibc, middleware, or application layer. By using SystemTap the respective layer was quickly determined and from there, the SystemTap capabilities were used to drill down to the root cause.

To help educate Red Hat Enterprise Linux customers on the benefits of SystemTap, a Beginner's Guide is included, as is a large set of example SystemTap scripts. These and other documentation are available on the SystemTap documentation page: **http://sourceware.org/systemtap/documentation.html**.

OPENJDK(TM)

Red Hat and its customers understand the power of open source collaborative development. That is why the introduction of the industry's first product version of an open source Java™ runtime is exciting news. Red Hat engineers have worked closely with Sun and the community to produce this release, which is a full implementation of the Sun Java SE 1.6 specification and was tested against the Java SE 1.6 Technical Compatibility Kit with zero failures. The existence of a truly open community for Java™ runtime development opens up huge potential for optimization of Java™ on the Linux platform. Beyond the base platform, Java™ is highly instrumental to Red Hat's JBoss Enterprise Middleware, as well as the high-performance Red Hat Enterprise MRG Messaging offering. Watch for future optimizations of Red Hat's complete Java™ stack.

GCC 4.3

Red Hat's tools team is a driving force behind rapid innovation in the gcc compiler and associated library and diagnostic tools suite, and gcc is the compiler included in Red Hat Enterprise Linux. The latest release of the compiler, **gcc 4.3**, is provided as a technology preview in Red Hat Enterprise Linux 5.3 for use in development environments. This release provides many advances over the previous compiler, gcc 4.1, including:

- Experimental C++0x support in C++ as well as in libstdc++ options are available for compliance checking

- Interprocedural optimization and inlining performed on static single assignment

- Improved math optimizations via inclusion of the MPFR library

- Vectorizer enabled by default at -O3

- Warray-bounds added to enable bounds checking where possible

- GCJ now uses the eclipse compiler, enabling all Java 1.5 features

- Many libstdc++ improvements and features including TR1 support and parallel support

- Support for the OpenMP 3.0 standard

- Many other smaller optimization and target tuning performance enhancements

Complete details of the enhancements in GCC 4.3 and GCC 4.2 can be found at **http://gcc.gnu.org/gcc-4.3/changes.html** and **http://gcc.gnu.org/gcc-4.2/changes.html**.

HARDWARE ENABLEMENT

HARDWARE PLATFORM FEATURES AND PERFORMANCE ENHANCEMENTS

Historically, the challenge of providing support for hundreds of new hardware devices could be limited to basic system bring-up tasks with some performance optimization. While this entails a substantial effort to this day with the close cooperation of Red Hat partners, including chip vendors, OEMs, and peripheral providers, exploiting modern, complex system hardware capabilities often necessitates additional, sophisticated software. To summarize just a few of the many hardware-related features in Red Hat Enterprise Linux 5.3:

- **Intel® Tylersburg / Nehalem platform:** Intel's newest generation of motherboards is called the Intel® Tylersburg / Nehalem platform. Through an intensive development initiative between Intel and Red Hat engineers, the upstream community and OEM partners, Red Hat Enterprise Linux 5.3 contains performance optimizations for the new platform. Initial benchmarks have demonstrated dramatic performance improvements over the previous generation of processors.

- **Power management:** operational cost savings and physical space considerations drive the need for higher efficiency and low power consumption. Red Hat Enterprise Linux 5.3 includes low-level kernel optimizations aimed at operating equipment efficiently in the lowest possible power states. This includes exploiting the new deep **C states** available on the Intel® Tylersburg / Nehalem platform. Additionally, the inclusion of **ACPI T-state** support facilitates effective processor throttling.

- **I/O Memory Management Unit (IOMMU):** Certain **AMD** processors and supporting chipsets have an IOMMU, which is a service processor that optimizes the passing of data between system IO devices and main memory. New optimizations are provided in Red Hat Enterprise Linux 5.3 to support secure DMA transfers. This is particularly useful in virtualized deployments to provide hardware-level I/O isolation between guests. The kernel boot option for enabling the driver is 'iommu="amd"' and device isolation is enabled by 'amd_iommu="isolation."

DEVICE DRIVERS

- Every Red Hat Enterprise Linux update release includes a large number of device driver updates that enable new hardware, optimize performance, and support new features. Updated drivers are provided in Red Hat Enterprise Linux 5.3 for disk, network, both wired and wireless, graphics, and a variety of other peripherals. The release notes itemize driver enhancements and provide full details. But here are a few categories of driver types which are new to Red Hat Enterprise Linux 5.3 and warrant mention:

- dmraid: Historically, RAID storage adapters tended to be high-end add-on devices, usually utilized exclusively in high-end server configurations. Recently, Intel began shipping hardware-based RAID features, called dmraid, on some newer motherboards. Red Hat Enterprise Linux 5.3 includes the device drivers for dmraid as well as a configuration management interface in the anaconda installer (including kickstart). Support was provided for RAID levels 0, 1 in prior Red Hat Enterprise Linux updates, while support for RAID level 5 has been added in Red Hat Enterprise Linux 5.3.

- iSCSI target support: is now fully supported, allowing customers the option of efficiently accessing iSCSI storage servers deployed in their environment.

- FCoE (Fibre Channel over Ethernet): is a new alternative to approaches such as NFS and iSCSI for network based storage. FCoE allows a convergence of storage and networking traffic on a single adapter. It does this while maintaining a high degree of compatibility with Fibre Channel, including the guarantees for reliable performance that the Fibre Channel protocol provides. Red Hat Enterprise Linux 5.3 provides full support for FCoE on three specialized hardware implementations. These include Cisco fnic driver, the Emulex lpfc driver, and the QLogic qla2xx driver. Support for FCoE over a standard Ethernet card is available as a technology preview.[2]

FILESYSTEM AND STORAGE

STORAGE

In addition to the above-mentioned device driver enhancements, there are new capabilities in the upper layers of Red Hat Enterprise Linux's storage stack, such as at the volume management layer. Key highlights include:

- **LVM cluster 2-way mirroring** is now fully supported in the Red Hat Enterprise Linux Advanced Platform edition. Cluster mirroring is a particularly useful feature for virtualized environments, which typically require consistent cluster storage when using virtual server live migration.

- Speedups in system boot and startup time in LVM deployments by more effective usage of the **lvmcache**.

- Block device encryption support, **dm-crypt**, and system integration have been enhanced in the anaconda installer enabling encrypted support for the root filesystem as well as swap partition. Encrypted filesystems are well-suited for laptop usage, but are additionally finding increased use in server environments where they help mitigate concerns for sensitive data falling into the wrong hands with the disposal of old hardware.

FILESYSTEM

A variety of filesystem enhancements appear in Red Hat Enterprise Linux 5.3, ranging from new filesystem types to improved interoperability. Highlights of the filesystem enhancements include:

- **GFS2** is fully supported in Red Hat Enterprise Linux 5.3 Advanced Platform edition, which was previously available in Red Hat Enterprise Linux 5 as a technology preview. GFS2 is a second-generation parallel, fully coherent cluster filesystem that delivers enhanced performance and scalability beyond the initial GFS1 implementation. Cluster filesystems are becoming increasingly important in virtualized environments as they form a consistent storage pool over which live migration and failover of virtualized guests can be performed. They also form a highly available platform for many critical datacenter applications like web or file servers.

---

2   Technology preview, tech preview. The majority of new features provided in Red Hat Enterprise Linux releases are fully supported, however it is also common for releases to include some new features which are not sufficiently mature to be ready for fully supported production deployment. To provide customers with early access to these promising new features, and to enable them to participate in validation testing, these features are designated as tech preview.

- The default filesystem in Red Hat Enterprise Linux 5 is ext3. An incremental evolution to ext3 debuts in Red Hat Enterprise Linux 5.3 and is unsurprisingly called **ext4**. The primary objective of ext4 is to provide scalability enhancements, these are particularly beneficial as the size of disks has grown substantially over the years. Ext4 can also create new filesystems (mkfs) and perform consistency checks and repairs after an unclean shutdown much more rapidly than ext3. It is provided as technology preview status.

- **eCryptfs** fixes and integration are also available in Red Hat Enterprise Linux 5.3. The above storage subsection referenced dmraid, which provides encryption at the LVM block device layer. An alternative encryption approach is to perform it at the filesystem layer. There are tradeoffs in terms of flexibility and ease of storage management which determine whether block-layer or filesystem-layer encryption is most effective, depending on the use case. While eCryptfs first appeared in Red Hat Enterprise Linux 5.2, Red Hat Enterprise Linux 5.3 includes numerous fixes and full integration with the anaconda installer, enabling system administrators to easily configure encrypted filesystems. eCryptfs is provided as tech preview status.

SYSTEM SERVICES

- Microsoft Windows®[3] interoperability enhancements:

  - **CIFS** is a network-based filesystem-based protocol that enables Red Hat Enterprise Linux systems to be either client or server to storage among heterogeneous Linux/Windows environments. New to Red Hat Enterprise Linux 5.3 are security provisions enabling usage of kerberos credentials.

  - **Samba** is a suite of interoperability tools allowing integration of user account data, filesystem sharing and active directory integration. Red Hat Enterprise Linux 5.3 updates to the most recent samba suite bringing enhanced integration with Windows Vista.

- The Linux kernel strives to do dynamic, adaptive tuning without necessitating manual performance tuning by the system administrator. While the number of tuning parameters which adaptively self-tune is growing, there remain a number of static tuning parameters which, depending on hardware configuration and application workload, can be modified to yield performance benefits. All kernel tuning parameters have a single default value and, in practice, this one-size-fits-all approach can be suboptimal, necessitating manual tuning. Through experience working with customers to optimize large production environments, a number of recognized patterns where similar tuning settings are beneficial have been identified. Consequently, Red Hat Enterprise Linux 5.3 includes a new utility, **ktune**, a script which adjusts kernel tuning parameters to a profile suited for large memory systems, larger than 64 GB, that have high disk and network IO rates. Customers with large systems may wish to experiment with ktune and observe whether it can provide performance benefits. ktune is delivered as a separate, optional script because, although it was considered too invasive to modify default kernel parameters in a update release, Red Hat wanted customers to benefit from its performance tuning experience. Even if you do not elect to run the ktune script on your server as-is, it may be helpful to look at the parameters it alters; some adjustment of individual settings may be applicable in your environment. The ktune script is provided as a technology preview and Red Hat invites customer feedback.

3   Windows is a registered trademark of Microsoft Corporation in the United States and other countries.

- The **CUPS** printing subsystem now includes kerberos authentication.

- The **RPM** software package manager utility has received a substantial number of performance enhancements and fixes.

- The **audit** subsystem has been enhanced to include remote logging functionality.

- The **NetworkManager** daemon now provides support for sharing multiple active connections – for example to bridge other laptops through a common bluetooth phone. Mobile broadband support is provided for a select set of devices, and static IP addresses are now configurable.
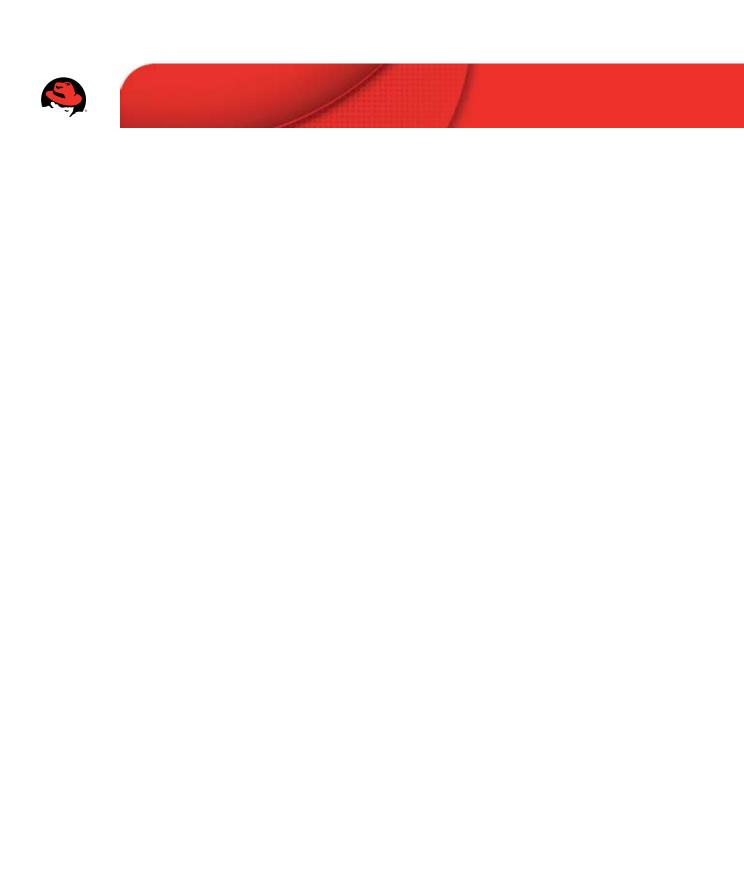
CLUSTER

The clustering capabilities provided in Red Hat Enterprise Linux 5, included in the Advanced Platform edition, are gaining increased usage as a high-availability augmentation to the virtualization capabilities. Examples include automated failover of virtualized guests.

- Cluster volume management and filesystems are a foundational underpinning of virtualized guest migration, and provide a uniform, consistent view of storage regardless of which computer the virtualized guest is running on. Red Hat Enterprise Linux 5.3 enhancements for cluster storage include:

  - The LVM volume manager now supports **cluster volume 2-way mirroring**.

  - The **GFS2** parallel cluster filesystem is now fully supported for use in production deployments.

- A fencing agent has been provided for the VMware ESX and VirtualCenter server as a technology preview. This is a control module which ensures that a failed VMware instance is fully terminated before initiating restart, thus preventing concurrently running multiple instances. This capability enables Red Hat Enterprise Linux virtual guests to be clustered on a VMware server.

CONCLUSION

Red Hat Enterprise Linux 5.3 provides many new features, bug fixes, and security improvements. As Red Hat's flagship product, it provides enterprise readiness and compatibility across updates. The preservation of compatibility is one of the principal benefits that differentiate Red Hat Enterprise Linux releases from general upstream community development, which in many cases tends to be revolutionary rather than evolutionary. While the rapid pace of upstream advancement yields plentiful innovation from which customers and developers eventually benefit, invasive, disruptive, or incompatible features are normally deferred to major releases and not included in update Red Hat Enterprise Linux releases. Through strict vigilance to compatibility and stability, customers can deploy Red Hat Enterprise Linux 5.3 with confidence without any need to recertify their application and middleware stacks.

www.redhat.com
#923069_0109