SAS GRID COMPUTING ON RED HAT ENTERPRISE LINUX
FASTER, BETTER RESULTS. SCALE OUT. REDUCE COSTS.

Enterprises employ business intelligence and analytics solutions to help improve performance and make better decisions faster. SAS is a leader in business analytics, helping companies unearth insights buried in information. In the midst of an economic downturn and increased global competition, companies need the ability to scale business processes and accelerate decisions even further to effectively compete.

This requirement means supporting more users, more compute-intensive queries, and more data. Now more than ever, users need to work on complete data sets and run multiple workloads simultaneously. In a global business environment, nightly batch processing windows to run large jobs no longer exist. Knowledge workers need answers 24x7, and the faster data can be loaded and analyzed, the more it can be leveraged to create a competitive advantage.

Unfortunately, in a time of increasing demand, shrinking budgets require IT to do more with less. IT must now squeeze more utilization and performance out of compute resources, increase application availability to serve around-the-clock workloads, and expand resources quickly and inexpensively to meet the changing needs of the business.

To help enterprises leverage the power of grid in SAS environments, SAS offers a comprehensive grid solution. SAS Grid Computing combines the proven computing middleware from Platform Computing with the power of SAS Business Analytics, providing the ability to scale business processes and accelerate decisions. To help reduce costs, increase performance, and provide more flexibility, companies utilizing SAS Grid Computing run the grid nodes on Red Hat® Enterprise Linux®.

WHY SAS GRID COMPUTING?
Grid computing dramatically improves throughput, providing the ability to run analytics any time, so users are not constrained by scheduled runs. Users submit jobs to a shared pool of resources to balance the workload. Compute-intensive programs can be run in pieces across grid environments, optimizing hardware capacity, improving performance, and increasing flexibility. Reduced data processing times mean larger volumes of data can be integrated, cleansed, and analyzed in less time, producing faster, more in-depth answers.

Grid computing helps reduce operating expenses by sharing resources across many users and departments, and the grid consists of inexpensive commodity servers. Rather than overprovisioning a single, large departmental server to handle peak loads, multiple departments can leverage the increased throughput, performance, and load-balanced utilization of the many smaller servers in the grid environment.

Fine-grained scalability is one of the inherent features of a grid computing model. Increasing capacity is simply a matter of adding another server, CPU, memory, disk, etc., to the pool of resources, enabling companies to scale-out in small, budget-friendly increments. Because there are so many redundant components, fault tolerance and application availability are increased. Individual systems can be taken offline for maintenance or even fail without disrupting the business.

SAS GRID MANAGER
SAS Grid Computing enables many SAS solutions to automatically leverage a centrally managed grid infrastructure and provides high availability for critical services and business processes. With SAS Grid Computing, it’s easier, more efficient, and more cost-effective to run compute-intensive applications and support growing numbers of users and data across hardware resources.

“Red Hat Enterprise Linux was the obvious choice for our internal R&D grid because it is the most popular SAS Grid Computing platform across industries.”

– Cheryl Doninger
R&D director for Grid Computing and Open Grid Forum Advisory Committee member, SAS

www.redhat.com
• **Centralize management and reduce complexity.** IT staff can easily monitor and manage hosts, jobs, users, and queues in the grid environment from a central point.

• **Increase availability.** Scheduling, load balancing, and high-availability capabilities automate management and optimize application processing.

• **Accelerate results.** Multiprocessing capabilities allow workflows to be processed in parallel on multiple servers. Analytical and data integration processes yield faster results, accelerating decision making across the enterprise.

• **Grow incrementally.** SAS Grid Computing allows organizations to fully utilize all available computing resources. The environment can be cost-effectively scaled out using commodity hardware.

**PERFORMANCE IS KEY TO ANALYTICS**

SAS Grid Manager jobs are extremely compute and I/O intensive. An intelligent operating system is a critical component for these large workloads, especially on the new generation of multi-core CPUs. The ability to take advantage of performance enhancements in processors such as the Intel Xeon processor 5500 series and tune I/O and virtual memory makes Red Hat Enterprise Linux an ideal platform for SAS Grid Computing.

• **Game changing performance.** Intel and Red Hat have a long history of collaboration that includes aligning technology roadmaps and making complementary open source contributions. Both companies share a commitment to extending the x86 platform to its full potential, and the stunning results of Red Hat Enterprise Linux on the Intel Xeon processor 5500 series shows just how powerful that combination is.

• **Intelligent performance.** Red Hat Enterprise Linux takes advantage of the Intel Xeon processor 5500 series to adapt throughput to the workload, delivering a greater than 2x performance compared to predecessors1.

• **Automated energy efficiency.** Red Hat Enterprise Linux uses Intel Intelligent Power Technology to support low-latency changes among power states and to lower power consumption during low utilization times.

• **Virtualization performance and efficiency.** Taking advantage of next-generation Intel Virtualization Technology (Intel VT), Red Hat Enterprise Linux delivers high consolidation ratios and virtualized performance. Tuning for optimum I/O throughput. 30–50 MB/s per core throughput across the grid is key to SAS performance. The following features in Red Hat Enterprise Linux enable I/O to be tuned to match CPU performance and to handle different types of workloads.

  • Enhanced I/O schedulers enable I/O performance to be optimized on a per-device basis according to the requirements of the application.

  • Support for 10 gigabit Ethernet, iSCSI, and Fibre Channel over Ethernet to utilize the latest storage technologies.

  • MPIO allows multiple connections from servers to storage to increase availability and throughput. In active/active mode all paths are used for the I/O, which is spread in a round-robin fashion, increasing the total I/O performance.

**ACCESS YOUR DATA**

SAS Grid Computing requires a shared or clustered file system for SAS data. Red Hat Enterprise Linux offers a choice of high-performance file systems, including shared NFS. In addition, the Logical Volume Manager (LVM) can be used to create and manage pools of virtual storage devices.

“The Decision Science team at Barclay needed a reliable, cost-effective platform with the ability to scale to our needs. We chose Red Hat Enterprise Linux and SAS Grid Manager because the combination satisfied all of those requirements and more. As a result we’ve been able to markedly improve the analytic services we provide to our internal customers over our previous UNIX implementation.”

– Pablo Noguera

director Decision Science, Barclaycard.

1 press.redhat.com/2009/04/16/performance-benchmarks/
SECURE YOUR SYSTEMS

All Red Hat Enterprise Linux system services are provided with targeted policies, resulting in the highest level of out-of-the-box security in the industry. A reference policy and support for local policy modules allow private policies to be created. The Red Hat Enterprise Linux audit subsystem lets you track activities and modifications to the entire system, including file system operations, password changes, account changes, authentication services, and configuration changes. This allows Red Hat Enterprise Linux to meet US Government certifications, including CAPP/LSPP and NISPOM and assist organizations to meet regulatory requirements, including Sarbanes Oxley and HIPPA.

AUTOMATE SYSTEMS MANAGEMENT

The right management tools can help maximize the performance and efficiency of your IT systems and staff. Built on open standards, Red Hat Network Satellite is a systems management platform that makes Linux deployable, scalable, manageable, and consistent. Red Hat Network Satellite provides the tools to efficiently manage many servers as easily as one, lowering deployment and management costs.

Automated software maintenance and patching facility. The Red Hat Network Satellite server connects with Red Hat over the public Internet to download new content and updates. Red Hat Network Satellite:

- Includes an embedded database to store packages, profiles, and system information
- Instantly updates systems for security fixes or to provide packages or applications needed immediately
- Enables the creation of scripts to automate functions or integrate with existing management applications
- Enables the distribution of custom or third-party applications and updates
- Enables the creation of staged environments (development, test, production) to select, manage, and test content in a structured manner

Provisioning. With Red Hat Network Provisioning Module, it is easy to increase the compute power of your SAS grid. Red Hat Network Provisioning Module enables you to deploy, configure, manage, update, and then re-deploy systems, all from a single GUI console. Functionality includes:

- Bare metal provisioning
- Existing state provisioning
- Virtual guest provisioning
- Multi-state rollback (includes snapshot-based recovery)
- Configuration management
- RPM-based application provisioning
- Kickstart configuration writer

Node and cluster monitoring tools. The Red Hat Network Monitoring Module allows you to track the performance of systems and receive alerts regarding system performance, allowing you to react before problems arise. In addition, this module allows administrators to:

- Monitor systems using dozens of pre-built probes for Red Hat Enterprise Linux and popular applications from Oracle, MySQL, Apache, and BEA
- Create custom probes for applications not included in the pre-built probe set
- Configure warning and critical thresholds for each probe
• Receive email or pager alerts when thresholds are reached
• Group probes together for fast deployment
• View graphs of probe performance over time

YOU CAN COUNT ON RED HAT

Red Hat ranked first among software companies in the 2009 CIO Insight Vendor Value study. This is the fifth time in six years that Red Hat was named most valuable software vendor, topping the list for meeting vendor expectations for customer value, flexibility, and responsiveness. The Red Hat subscription model provides everything you need in one all-inclusive price. Multiple support services offer unlimited incidents and coverage up to 24x7 with one-hour response time. And, Red Hat’s robust certified ecosystem includes thousands of software and hardware vendors actively engaged in supporting and certifying Red Hat Enterprise Linux, often before other operating system platforms.

NEXT STEPS

→ redhat.com/sas
Ready to learn more about SAS and Red Hat? This site provides the latest information regarding Red Hat Enterprise Linux for SAS solutions, industry events, supporting documentation, and more.

→ redhat.com
Red Hat, the world’s leading provider of open source solutions and a component of the S&P 500, is headquartered in Raleigh, NC with over 65 offices spanning the globe. CIOs ranked Red Hat as one of the top vendors delivering value in enterprise software for six consecutive years in the CIO Insight Magazine Vendor Value survey. Red Hat provides high-quality, affordable technology with its operating system platform, Red Hat Enterprise Linux, together with virtualization, applications, management, and Services Oriented Architecture (SOA) solutions, including Red Hat Enterprise Virtualization and JBoss Enterprise Middleware. Red Hat also offers support, training, and consulting services to its customers worldwide.

→ sas.com/grid
Find out more about SAS and benefits of SAS Grid Computing.

Copyright © 2010 Red Hat, Inc. Red Hat, Red Hat Enterprise Linux, the Shadowman logo, JBoss, MetaMatrix, and RHCE are trademarks of Red Hat, Inc., registered in the U.S. and other countries. Linux® is the registered trademark of Linus Torvalds in the U.S. and other countries. SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries.