



Enhance Data Center Utilization by **95%** with **Enterprise Scale Workload Management**

Today's global business environment is rife with challenges as companies strive to meet increasingly stiff requirements from customers, partners and regulatory agencies while operating under constrained economic conditions.

As complexity grows, IT organizations need more efficient ways to manage the application and server infrastructure for these data-intensive workloads. Without automation to schedule and manage workloads consisting of hundreds of thousands of tasks in a single day, users schedule tasks or jobs at random, which means that jobs aren't often handled according to business priority or the best available IT resources at that moment. Problems can arise with customers and revenue streams if a high-priority project misses a deadline because the job either couldn't be scheduled fast enough or wasn't assigned the appropriate resources.

An unmanaged workload environment can also cause discord among employees and departments competing for resources. For the IT manager, this situation is also time-intensive, causing needless hours shifting workloads around and

InfoWorld
Custom Solutions Group

troubleshooting performance problems.

Then there's the issue of IT value. The typical enterprise data center has an average utilization rate as low as 4 percent — which means that a significant amount of server resources are sitting idle when they could be used to help meet critical project needs and on-time delivery for customers. Expensive testing and simulation applications may also sit idle if hardware resources are tied up, another potential drain on an organization if a license must be renewed to accommodate delayed schedules.

For organizations with intensive compute and data processing needs, managing and scheduling workloads manually is ultimately a money-losing proposition. At a time when many IT departments are required to do more with less, this becomes an issue of competitiveness: lean is mean.

A better way to handle these jobs to meet business objectives and maximize the infrastructure is through workload management software, which offers advanced scheduling tools ideal for cluster computing environments. This technology makes it possible to do a number of critical tasks, including:

- Optimally manage execution of hundreds of jobs from multiple applications on multiple machines or processors, using business rules;
- Ensure that all processors are always busy, which speeds up work for internal customers and optimizes resources;
- Increase end user productivity by relieving them of the need to track their jobs and the servers they need to use, and;
- Efficiently troubleshoot jobs because IT has a single console to view status of all jobs across the organization. As well, IT no longer needs to manage users' interactions with the machines.

Workload management technology usually does a much better job of keeping machines busy than users do on their own, resulting in greater hardware utilization and productivity. For the fiscally-aware CIO, the organization can effectively run more jobs on the existing infrastructure, eliminating unnecessary capital investments.

Many large, data-intensive organizations have some solution for workload management today, yet requirements are constantly changing. Organizations using the popular Apache Hadoop distributed data processing software must schedule those jobs separately due to integration conflicts. Increasingly,

enterprises are using cloud services to handle at least part of their processing and application access needs. Today's tools must be flexible to support these requirements and also scalable to support the largest of environments.

ORACLE GRID ENGINE

The Texas Advanced Computing Center at the University of Texas is an example of an extremely high-demand, high-CPU environment. The research organization uses workload management software from Oracle to manage its 63,000-core cluster, which supports simulations and calculations for projects across many fields. One group, for instance, is using Oracle Grid Engine to do real-time face recognition in streaming video feeds.

The Oracle solution allows TACC to manage roughly 4,000 machines and 1,000 projects throughout the United States. The TACC cluster has been able to maintain 99 percent hardware utilization rates through workload management.

Smaller organizations outside of research and manufacturing can also take advantage of this technology. Australia-based Rising Sun Pictures uses Grid Engine to orchestrate its video rendering process to create digital effects for blockbuster films such as the Harry Potter series. Grid Engine ensures that the specific steps in the rendering process are performed as efficiently as possible, which is critical for a company that must meet tight deadlines for its Hollywood clients.

WHAT'S NEW: INTEGRATION, CLOUD SUPPORT AND ADVANCED SCHEDULING

As enterprise requirements change, so must workload management technology. Oracle Grid Engine 6.2 helps organizations achieve the same scalability and reliability of the past, but has added integration with Apache Hadoop, topology-aware scheduling and support for cloud computing provisioning, among other new features.

At a high-level, here's how the software works: When users submit their work to Oracle Grid Engine as jobs, the software monitors the current state of all resources in the cluster and assigns jobs to the most suitable resources. Using the software, administrators can accurately model computing resources and translate business rules into policies that govern the use of those resources. Grid Engine is flexible enough to allow customization, such as enabling IT to assign network bandwidth as a computing resource. In this example, when a job requests a given bandwidth, it is scheduled only on machines that can provide that bandwidth.

Here's more detail on some of the specific features of Grid Engine 6.2:

Apache Hadoop integration

For Web and life sciences companies, financial services firms and others, the Apache Hadoop open-source application suite has grown in importance because of its strength in analyzing unstructured and complex data sets from both new and legacy systems. Typically, organizations running Apache Hadoop for large-scale data processing had to run those applications on a dedicated cluster. Now, Oracle Grid Engine allows organizations to share infrastructure resources between Hadoop and other data center applications, saving on hardware as well as the time and hassle of managing dual environments. Oracle Grid Engine's scheduling policies and accounting database also shave costs from managing Hadoop applications.

Support for the cloud

As high-tech vendors offer a growing array of cloud-based services, enterprises are taking notice and doing pilot projects. Oracle Grid Engine 6.2 allows IT to incorporate cloud resources into the scheduling tool and automatically access them as needed when application workloads surge. The ability to use on-demand cloud resources for large-scale data processing reduces the need to rely solely upon data center capacity, which in turn helps lower data center operating costs. Users benefit from faster response times but won't know (or need to know) that cloud-based servers are helping do the job.

Advanced scheduling

Many applications run optimally when tied to specific CPU sockets and/or cores. Oracle Grid Engine 6.2 allows users to specify these topology preferences when submitting jobs, which can maximize application performance and cut execution times in some cases by more than 50 percent. Administrators can also configure a cluster so that the job can request exclusive, non-shared access to a hardware resource, including processors, memory or built-in accelerator boards.

Among other new features, Oracle Grid Engine 6.2 helps organizations be more energy-efficient by automatically managing the power settings of idle systems during off-peak hours, thereby reducing power usage. Improved preemption helps achieve higher levels of data center utilization and, for large parametric jobs, the Oracle Grid Engine scheduler can now share resources more fairly among users for better efficiency and faster completion of jobs.

HOW COMPANIES ARE BENEFITING FROM WORKLOAD MANAGEMENT

Organizations primarily adopt workload management software to increase value from their data centers and ease the pain of managing a million jobs simultaneously. The technology is particularly useful in environments where companies are producing complex products such as electronics and pharmaceuticals, which require heavy computation, analysis, modeling and simulation across many machines. Managing the complexity of multiple applications and departments competing for the same set of resources is the sweet spot for workload management. Here are a few industry examples of how workload management software makes a difference:

Genetics research firm scales with demand: Complete Genomics is using Oracle Grid Engine to manage the calculations needed to sequence the human genome, work that requires enormous amounts of computing power and storage. Using Oracle Grid Engine, Complete Genomics will be able to run cost-effective comparisons of genomes from thousands of individuals, which it hopes will transform the way disease is studied, diagnosed and treated. Workload management will be critical in helping the organization grow its data center capacity by an order of magnitude so that it can process much larger data sets to further genetic research and drug development across the industry.

Faster time to results for software development: Mentor Graphics, a leading provider of electronic design automation software, uses Oracle Grid Engine to manage its regression tests. To test its software, Mentor Graphics submits the tests as thousands of jobs to be run on the cluster. Oracle Grid Engine makes sure that every machine is busy running tests. When a machine completes a test run, Oracle Grid Engine assigns it another, until all tests are completed. As a result, Mentor Graphics has been able to speed its testing time by 500 percent, which means it can deliver new versions of its software to customers faster. The company is also using Oracle Grid Engine to manage testing software licenses, ensuring that it does not use more licenses than are available during jobs.

Supporting financial calculations with high reliability: Global investment firm D.E. Shaw uses Oracle Grid Engine to manage various financial calculations for its customers, including risk determination and market prediction. The company's core business runs through the Oracle Grid Engine cluster and supports a 99 percent availability SLA. The workload management software has also enabled D.E. Shaw to

integrate many of its custom-developed financial applications into the cluster with little or no modifications.

Saving money and time for IT: Oracle Grid Engine consistently helps organizations increase data center utilization from up to 95 percent through more consistent optimization of hardware resources. Software utilization, as well, is critical in research and manufacturing environments, as third-party licenses for testing and modeling may cost 10 times more than the hardware used to run it. With workload management, software resources can be used when needed instead of sitting idle while waiting for an available machine, which helps maximize the value of licenses. But beyond the savings on infrastructure, IT benefits greatly from Oracle Grid Engine. Administrators have a single console to view job status and reassign workloads if needed, while automated scheduling and resource management allows IT to spend time on more strategic activities for the business.

Enabling processes for operational decisions:

Increasingly, Oracle Grid Engine is being used in non-conventional ways to help companies better manage fluctuating demand. For instance, an airline used Oracle Grid Engine to help run calculations for passenger configurations, modeling seats as computing resources and passengers as workload. While not typical, this shows Oracle Grid Engine's flexibility in helping companies and organizations save time and money, increase customer satisfaction and efficiency, and gain a competitive edge.

USING ORACLE GRID ENGINE

For IT managers, interacting with Oracle Grid Engine is straightforward. First, they need to determine and model what resources, including hardware, software or bandwidth, to manage with Oracle Grid Engine. The next step is to deploy Oracle Grid Engine across all of the machines and nodes it will manage. Since Oracle Grid Engine has a zero-touch interface to the application workloads it runs, there's no need to make any modifications to get started. In most cases, the applications aren't even aware that the workload management software is running.

Oracle Grid Engine also gives administrators the power to model business objectives as policies and service level metrics. This ensures that computing resources are assigned and managed according to business priorities. Policies range

from something as simple as honoring project deadlines to more complex rules governing how departments, projects and individual users share a cluster's resources. These policies can also govern decisions such as when it's appropriate to apply power management to idle resources, when it's time to lease more capacity from a cloud and when it's time to shift resource assignments among clusters.

End users can still specify requirements for their jobs in a managed environment. They can describe their jobs, related resource needs and scheduling requirements, and then Oracle Grid Engine takes it from there, honoring the requirements specified and doing its best to honor user requests.

Once jobs are in progress, administrators have a bird's eye view of all of the jobs running across the enterprise. They can see, for instance, that 97 percent of the cores are busy and 3 percent are idle. If there is a backlog of jobs, an administrator can investigate why the idle machines are not working. Fortunately, the software takes care of some of the troubleshooting automatically; if a server goes down Oracle Grid Engine can quickly switch the job to another server without any noticeable impact. Handling server repairs in batch mode saves managers time. Oracle Grid Engine can streamline repair and maintenance of machines by allowing machines that are down to sit idle for repair later with other failed machines.

Ultimately, through proactive management and automated scheduling of computing resources, data center managers can look at historical workload trends and make business or technology adjustments that may prevent common problems from recurring. Without advance workload management technology, however, understanding the relationship between workloads, resources and business needs in a large environment is nearly impossible.

With the volume of high priority workloads growing daily for data center and IT managers across many industries, there's no room for errors or slow response times. Internal and external customers shouldn't have to wait for resources, products and services because your IT infrastructure can't manage jobs effectively. Using Oracle Grid Engine, however, you can not only meet current needs more productively, but also expand your business and save money on infrastructure at the same time. ||