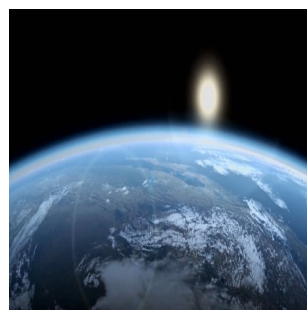


HPC on Oracle Cloud Infrastructure Delivers Performance and Stability

Disclaimer: This document is for informational purposes. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described in this document remains at the sole discretion of Oracle.

High Performance Computing (HPC) delivers value for the money invested, revenues increase by more than \$500 for every \$1 invested in HPC. Rising capital costs make large outlays for HPC environments and datacenters challenging. With Oracle Cloud's HPC offering, enterprises running HPC workloads can move HPC workloads into the cloud without sacrificing performance. Built on Bare Metal servers that can be provisioned immediately, Oracle offers the same performance and stability that traditional on-premises HPC center's, but it requires no capital expenditure. Combined with Oracle's enterprise experience, best-in-class security and low data migration costs, Oracle Cloud is the ideal place to run HPC workloads.



"With Oracle Cloud our Deep Space Trajectory Explorer can evaluate 50x more trajectory options, which gives us better information and decisions."
– Nick Netreba
Director of Mission Services, a.i. solutions

SCALABLE, ON-DEMAND RESOURCES FOR HIGH PERFORMANCE COMPUTING

Performance and control with Oracle Cloud Infrastructure

Organizations continue to move business applications and databases to the cloud to reduce the cost of purchasing, updating, and maintaining on-premises hardware and software. However, most high performance computing (HPC) workloads remain on premises—mainly because these jobs have specialized needs that traditional cloud offerings can't handle. HPC applications consistently consume 100 percent of available compute power to achieve a specific outcome or result. They require dedicated network performance, fast storage, high compute capabilities, and significant amounts of memory—resources that are in short supply in the virtualized infrastructure that constitutes today's commodity clouds.

And yet, running HPC jobs in the cloud makes a lot of sense—assuming the cloud service has been architected in an optimum way for high performance business or technical workloads. These jobs require immense amounts of resources during peak periods—such as when running a simulation, creating a model, or rendering an animated movie—and require little or no resources during the idle time between jobs.

Oracle Cloud Infrastructure offers exceptional performance, security, and control for today's most demanding HPC workloads. From bare metal offerings with the newest and fastest CPUs and GPUs, to extreme low-latency flash storage, and lightning fast RDMA networks, Oracle's cloud-based infrastructure enables customers to solve complex problems fast.

A BETTER SOLUTION FOR HPC WORKLOADS

Oracle's infrastructure as a service (IaaS) offering brings you consistently leading-edge compute, storage, networking, and software technologies—all a fraction of the cost of building high-end infrastructure on premises.

Oracle HPC customers can "lift and shift" their on-premises applications easily to Oracle Cloud Infrastructure, gaining a performance boost while saving an expensive on-premises hardware refresh. For building new HPC applications, Oracle supports key technologies such as schedulers, high performance file storage, and GPUs. Key HPC independent software vendors (ISVs) such as Altair, IBM, ANSYS, Cadence, and Siemens are also available for Oracle Cloud Infrastructure.

Challenges Facing HPC Customers:

- Overspending or under-utilization: It's difficult to correctly size infrastructure for HPC workloads
- Massive upfront capital cost outlays for on-premises infrastructure: HPC deployments are larger than typical enterprise environments, and may present substantial risk
- Slower time to results in the cloud: Most commodity cloud offerings only use virtual machines that reduce performance
- Software frameworks for rendering, scientific, manufacturing, design, AI, machine learning, and more

From cancer research to drug testing, autonomous driving to electronic design automation, this infrastructure meets the exacting needs of data scientists, engineers, designers, and researchers.

Advantages of OCI

- Faster time to value. Building your on-premises HPC Cluster isn't just expensive, it takes a long time.
- Put control directly in the hands of teams that consume HPC infrastructure
- Only pay for the resources that you consume, as you consume them
- Gain the ability to increase capacity to meet spikes in demand
- Focus on your strategic priorities and HPC outcomes, not on managing infrastructure or constantly fighting for capital budget for new hardware.

"Oracle's mission is to build a best-in-class HPC infrastructure with uncompromising performance based on top-of-the-line CPU and GPU offerings. The OCI offering includes bare metal instances that are far superior to the virtualized environments found in commodity clouds."

- Aaron Spinks, Vice President, Oracle Cloud Infrastructure

EXCEPTIONAL SCALABILITY AND PERFORMANCE

Oracle Cloud Infrastructure provides extremely fast cluster networking, high-powered bare metal CPU instances, and ultra-low-latency NVMe flash storage to help you build powerful HPC clusters. Oracle Cloud Infrastructure now combines its proven HPC instance with a low-latency network that can span up to 20,736 cores. The core HPC building block ("shape" in Oracle Cloud parlance) has 36 cores from two 3.7GHz Intel® Xeon® Gold 6154 Processors, 384 GB RAM, and 6.4 TB NVME local storage.

"By partnering with Intel to deploy Intel Xeon Gold 6154 Processors, OCI has been able to leverage the large performance gains from this generation of Intel Xeon Scalable processors and to accelerate the performance of HPC workloads like scientific simulation with the new AVX-512 instructions," said Jeff Wittich, Senior Director of Cloud Service Providers at Intel.

Oracle uses Mellanox technology to enable RDMA through RDMA over Converged Ethernet (RoCE) v2, an extremely low-latency protocol delivered over (fast) standard ethernet. Many highly parallel HPC workloads are built around a message-passing interface (MPI) that demands lower latency than can be offered in a typical cloud offering. Compared to virtual machines (VMs), bare metal compute instances yield more consistent results that are comparable to an on-premises compute infrastructure. With Oracle Cloud Infrastructure, you don't need to settle for a virtualized HPC network.

"Now is the time for cloud infrastructure architects to look at RDMA based interconnect to increase their compute, memory and storage performance," said Motti Beck, Senior Director Of Enterprise Market Development at Mellanox. "Deploying Mellanox's RoCE v2 solutions in OCI enables Oracle to achieve their SLAs at a lower cost while fulfilling challenging Cloud performance and efficiency metrics."

Scale

What is HPC?

HPC aggregates multiple computers and storage devices into a cohesive fabric, architected to solve large problems in science, engineering, and business. In an HPC environment, individual nodes work together to solve a problem larger than any one computer can easily solve. The popularity of this type of infrastructure is growing fast: The global HPC market is expected to increase from \$32 billion in 2017 to \$45 billion by 2022.³

Zenotech

Zenotech Ltd. is a Bristol, U.K.-based company that provides cloud solutions for computational fluid dynamics (CFD). Aerospace companies, automotive manufacturers, civil engineering firms, and many types of scientific organizations use Zenotech's Elastic Private Interactive Cloud (EPIC) and novel zCFD solver to run performance-intensive CFD workloads

"What we really like about the bare metal offering from Oracle is that there is very little technology between us and the hardware."

-David Standingford Director and Cofounder, Zenotech

³ <https://www.reportbuyer.com/product/3607622/high-performance-computing-market-by-component-and-services-deployment-type-organization-size-server-price-band-application-area-and-region-global-forecast-to-2022.html>

When virtualized in your data center, CPU-intensive tasks that require little system interaction, normally, experience very little impact or CPU overhead. However, virtualized environments in the cloud include monitoring and management layers, which can add significant overhead.

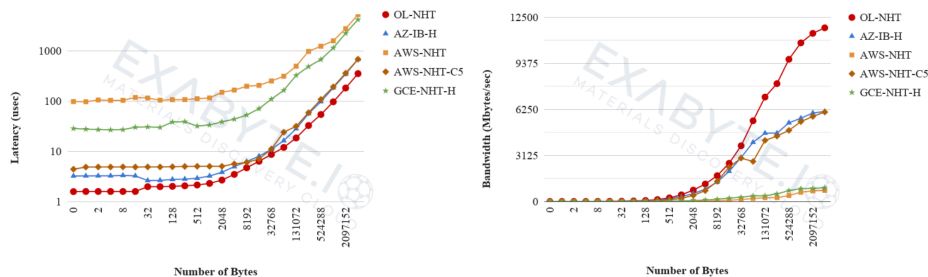
Virtualization overhead is not synchronized across an entire cluster, which creates problems for MPI jobs, which effectively have to wait for the slowest node in a cluster to return data before advancing to the next simulation iteration. You're only as fast as your slowest node, noisiest neighbor, or overburdened network.

With Oracle Cloud Infrastructure's bare metal environment, no hypervisor or monitoring software runs on your compute instance. With limited overhead, HPC applications scale across multiple nodes just as well as it would in your data center.

RDMA Network

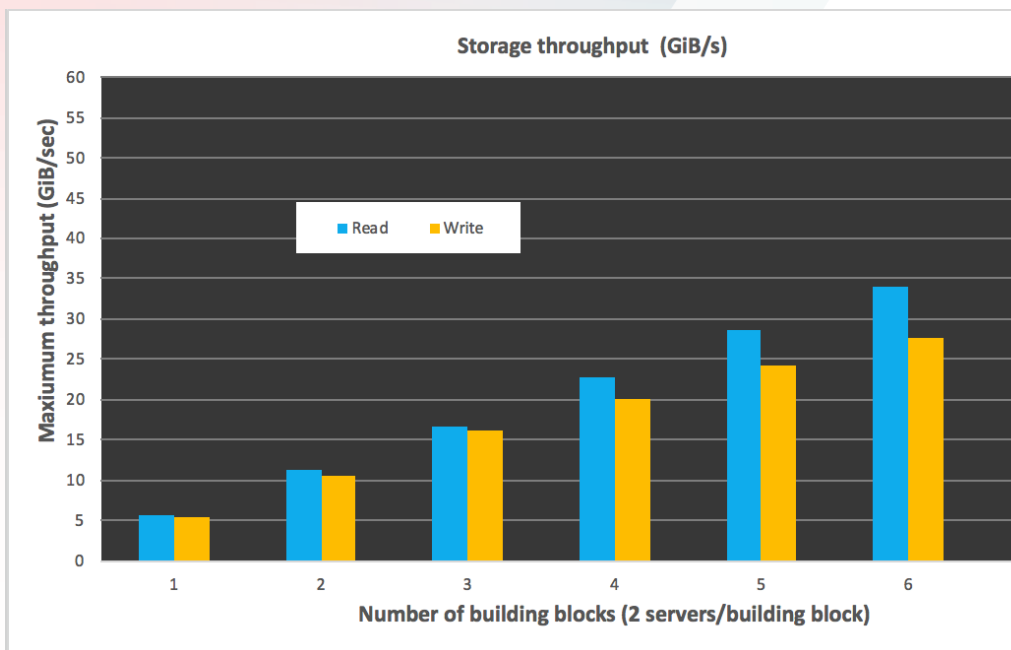
Ultra-low node-to-node latency and high throughput is required on HPC systems. Oracle Cloud Infrastructure's Cluster Networking leads the cloud in both categories. In order to run HPC applications like Computational Fluid Dynamics (CFD) and automotive crash simulations at scale, latency needs to be below 2us and throughput must be high. OCI provides this with Cluster Networking. Because only the user's HPC traffic runs on Cluster Networking latency and throughput ... because HPC traffic does not compete with the server's external network traffic, latency and throughput remain high ... When you have applications running on thousands of cores, low node-to-node latency isn't enough. The ability to scale models down to a very small size is more important. In computational fluid dynamics (CFD), users typically want to know the smallest amount of work they can do on a node before they hit a network bottleneck that limits the scalability of their cluster. This is the network efficiency of an HPC cluster or, in other words, getting the most "bang for your buck"! go see exabyte.io's blog

<https://docs.exabyte.io/benchmarks/2018-11-12-comparison/>



File Servers

File Servers using Oracle Cloud Infrastructure's bare metal compute and networking to enable performance that hasn't previously been seen in the cloud. The following graph shows the throughput different-size file servers. On as small as two nodes, Oracle Cloud Infrastructure provides over 5 GB/s throughput, easily achieving over 50 GB/s



If you are running Computational Fluid Dynamics, Electronic Design Analysis, Financial Risk Modeling, or Crash simulation you will get superior HPC performance on Oracle Cloud.

“Oracle offers bare metal and virtual compute instances as well as NVMe versions with up to 25TB of NVMe storage for performance unlike anything else seen in the cloud.”

– Storage Review Enterprise Lab, 2018¹

EXCEPTIONAL END TO END HPC EXPERIENCE

While Oracle does offer the best HPC performance on the cloud our HPC offering does not stop there. Our customers have an impressive end-to-end experience on Oracle Cloud. Oracle offers the best internal network, based on our baremetal compute, the lowest cost and highest performing block storage, which for HPC leads to power file servers. GPUs and strong ISV relationships enable virtual desktoping and engineering environments, best-in-class security and industry first performance level SLA's provide an end-to-end HPC experience that matches your on-premises experience with the scale and flexibility of the cloud.

Network and Data Transfer

OCI offers flat networking and ‘off-box virtualization’ which means that the network is completely isolated and just as importantly for HPC it means that servers can achieve full line rate between servers. All bare-metal servers come with at least 2x 25Gbps Network Interface Cards (NICs) and the HPC shapes offer 125 Gbps (1x 25 1x 100).

In addition to data transfer services for moving large amounts of data, Oracle offers no-cost ingestion for all data and the first 10 TB of data egress for free. Typically HPC customers opt for a dedicated ‘Fast Connect’ through their ISP. Oracle charges per port for the Fast Connect which means that you pay 1 time for a port and then there is no additional charge for data

¹ storagereview.com/oracle_cloud_infrastructure_compute_bare_metal_instances_review

egress, this can mean tens or hundreds of thousands of dollars in savings and of course it removes concern about data transfer.

Storage

Like most cloud providers, Oracle offers a range of basic storage options including various flavors of object, file, and block storage. But Oracle also has offerings that more specifically map to the needs of high performance computing workloads.

- Local NVMe SSD — high speed local flash storage ideal for large databases, high performance computing HPC (local scratch and logs), and big data workloads, like Apache Spark and Hadoop.
- Block Volumes -- networked block storage services offering up to 60 IOPS per GB, up to a maximum of 25,000 IOPS per volume, backed by Oracle's highest-performance SLA.
- Parallel File Systems – HPC systems often require larger data sets and higher performance than standard enterprise file servers can provide. Parallel file systems built with Oracle's high performance bare metal instances can easily achieve over 60 GB/s throughput at a price of less than five cents per GB per month.
- The Fastest File servers in the cloud – Oracle and IBM have partnered to provide high-performance file servers at scale on Oracle Cloud Infrastructure. Using IBM Spectrum Scale, you can now build a high-performance computing (HPC) file server on Oracle Cloud Infrastructure Block Storage for your HPC or Big Data applications. This file server deploys in minutes, scales easily, costs a few cents per gigabyte per month, and provides performance that's consistent with on-premises HPC file servers.

Security

OCI features “core-to-edge” security that includes customer isolation, data protection, protection against internet threats, and automated threat remediation. Within this advanced infrastructure, Oracle isolates compute and network resources, ensuring that no customer can see any other customer's data or traffic. Customer code, data, and resources reside on separate machines from the management machines. This isolation is designed to prevent an attacker from moving laterally inside the cloud to steal or manipulate data – even Oracle, as the cloud provider, cannot see customer data.

- Network Isolation: Each customer's network activity travels through a private Layer 3 software-defined overlay network, encapsulating traffic as soon as it enters the edge of Oracle Cloud. Oracle moves the virtualization layer into this physical network—sometimes called “off-box virtualization.” Essentially, each customer receives an isolated virtual cloud network.
- Compute Isolation: Bare metal compute instances are completely dedicated to the subscribing customer and isolated from all other tenants—with no shared processor or memory resources.
- Isolation between vendor and customer: OCI customers can deploy workloads on compute and storage hosts on which no Oracle software of any kind runs, providing an additional level of isolation from service provider activities. Oracle has no access to the contents of the server and cannot see customer data.

The Best Guarantees in the Industry

Oracle is the only major cloud provider to back its service with a performance service level agreement (SLA). If Oracle cannot provide at least 90 percent of published performance, at least 99 percent of the time, for local storage IOPS, block storage IOPS, or networking bandwidth, customers can claim up to 25 percent credit on affected services.

In addition to performance SLAs, Oracle also offers management and availability SLAs. Availability SLAs cover compute, block volumes, object storage and FastConnect. Manageability SLAs cover compute, block volumes, and database management systems.

“The bottom line is that forward-thinking businesses that can derive measurable value from high-performance cloud should definitely be evaluating what Oracle has going on.”

– Storage Review²

A NEW PHILOSOPHY FOR IAAS

Oracle developed its IaaS offering based the strategy of taking the best attributes of on-premises computing environments and making them available in the Cloud. For business applications, that was Exadata. For High Performance applications, it's the HPC “shapes” with bare-metal CPUs, and flat networking. Thus every element of Oracle Cloud Infrastructure, from its best-in-class server hardware to its fully switched networks to Oracle's service level agreements (SLAs) has been chosen and executed with this objective in mind.

Oracle's flexible pricing works well for the iterative nature of HPC applications, since you can spin up plenty of capacity for big jobs, then spin it back down when you're done. Or spin it back up with a completely different configuration to meet the needs of a different job.

A NEW GENERATION OF CLOUD

Oracle Cloud Infrastructure (OCI) is qualitatively different from the first-generation clouds popularized by Amazon, Microsoft, and other cloud vendors. The infrastructure has been explicitly designed to provide the performance predictability, core-to-edge security, and governance necessary for enterprise workloads. As a result, these workloads require fewer compute servers and block storage volumes, and hence lower cost to deliver the same or better performance than competitors.

Oracle Cloud Infrastructure is designed around highly scalable, high bandwidth, low-latency networks, deployed on an isolated fabric, and accessible through a common set of APIs. Hardware options include everything from small virtual machines (VMs) to large bare metal clusters, powered by engineered systems such as Oracle Exadata.

Whether you use small VMs or large bare metal clusters, your cloud infrastructure is accessible through the same APIs and the same console, and can access the same low-latency, high-performance databases and storage devices—including file, block, and object storage options.

“Oracle Bare Metal Cloud Services was twice as fast as any other instance we have running, with one result nearly 10 times faster.”

– Gareth Williams, CEO and Founder, YellowDog Ltd

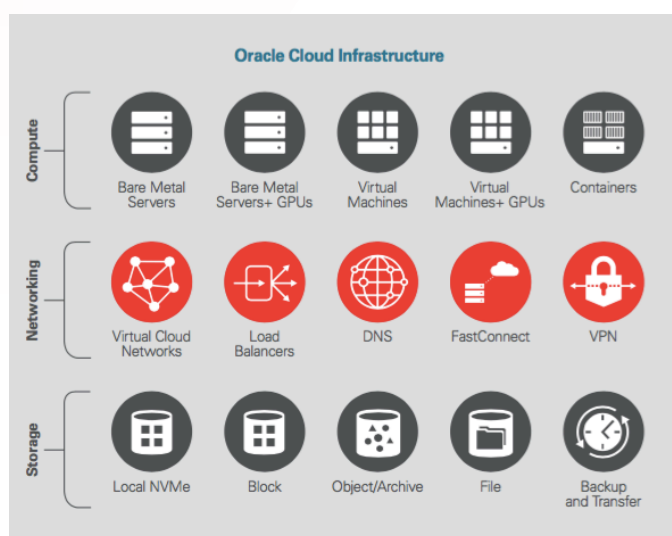
WHAT MAKES OCI UNIQUE?

- Oracle doesn't oversubscribe its compute, network, or storage resources, ensuring unbeatable performance and consistency across enterprise, cloud native, and HPC workloads.

² Storage Review Enterprise Lab, 2018
(storagereview.com/oracle_cloud_infrastructure_compute_bare_metal_instances_review)

- Oracle bare metal servers are some of the fastest in the industry, opening the door for enterprises to run HPC simulations requiring fast and scalable local storage in the cloud
- Each bare metal server is isolated from all other tenants, giving you complete control over the environment.

By contrast, the most popular offerings from Amazon Web Services (AWS) and Microsoft Azure are focused on rapid, flexible deployments of non mission-critical workloads. These first-generation cloud offerings simply can't handle performance-intensive workloads because they are virtualized and oversubscribed. Furthermore, Oracle IaaS pricing is straightforward and predictable. While AWS has hundreds of SKUs for their most basic services, and has the potential for unexpected charges, Oracle has no hidden costs associated with using its cloud, such as customers often encounter with Amazon's inter-region and inter-zone data transfer costs.



CASE IN POINT: FLUID DYNAMICS FOR SCIENCE, ENGINEERING, AND RESEARCH

Organization: Altair provides software and cloud solutions for product development, high performance computing (HPC), and data intelligence. Its software applications optimize design performance across multiple disciplines encompassing structures, motion, fluids, thermal management, electromagnetics, system modeling and embedded systems.

The Challenge: Altair's customers run complex physics and flow simulations to study the aerodynamic properties of buildings, vehicles, and environmental problems. For example, Altair's nanoFluidX application predicts the flow in complex geometries with complex motion, such as oiling in powertrain systems with rotating gears and shafts. High-end GPUs provide the ideal computing platform to handle these strenuous workloads, but they aren't always available for Altair's customers.

The Solution: Thanks to a partnership between Oracle and Altair, customers can run Altair's computational fluid dynamics solvers in a scalable, high-performance cloud environment. Altair's advanced CFD solvers yield rapid simulation results for even the most complex cases.

Oracle offers cutting-edge GPU instances and accompanying storage capabilities that preclude the need to move large amounts of data.

The Results: With this on-demand solution from Oracle and Altair, engineers don't have to wait in job queues or endure long HPC hardware procurement cycles. Altair provides their product as a service on Oracle Cloud Infrastructure, making it easy for users to submit and monitor jobs and visualize results in the cloud. Oracle's state-of-the-art NVIDIA GPUs provide a 4x speed increase over CPU-based solutions. Customers obtain exceptional computational performance immediately, without the cost or build-out time required on-premises.

"We were looking for the best price/performance and we found it with Oracle Cloud Infrastructure. Oracle is the first cloud provider to provide bare metal computing resources that match on-premises performance."

- Sam Mahalingam, Chief Technical Officer for Enterprise Solutions, Altair

CASE IN POINT: SPECIAL EFFECTS, ANIMATION, AND 3D RENDERING

Organization: Snowball Studios produces top quality 3D animation and visual effects for clients in the film, television, computer games, advertising, and mobile industries. Positioned at the heart of the Israeli hi-tech scene, and with offices in London and Toronto, Snowball's digital creations are aired all over the world.

The Challenge: Large brands such as Disney, Mattel, and many other industry leaders depend on snowball to run commercial animation projects that involve rendering thousands of frames. These projects place huge demands on the infrastructure during peak periods—such as when creating a new commercial or rendering a full-length film. It often takes several hours of rendering time to create each individual frame. Acquiring and maintaining the necessary hardware and software to run these resource-intensive jobs was a mounting challenge.

The Solution: Snowball relies on Oracle to supply advanced compute, storage, networking, and software technologies on demand. Oracle Cloud Infrastructure can be scaled instantly to handle complex processing tasks, then shut down between projects.

The Results: Snowball has grown to be the leading 3D Animation and Visual Effects facility in its region, combining the best international creative talent with cutting-edge HPC technology. The company does not need to acquire, manage, update, and maintain a massive technology infrastructure to do this work efficiently.

"Oracle uses superior servers with more memory, stronger processors, and more cores. When we considered what we would get for the same amount of money from other cloud providers, we saw that we were getting more from Oracle."

- Yoni Cohen, founder and chief creative officer, Snowball Studios

GET STARTED TODAY

HPC entails compute-intensive, complex problems in manufacturing, media and entertainment, seismic research, artificial intelligence, fluid dynamics, and many other computationally

intensive domains. These applications often require high network performance, fast storage, high compute capabilities, and significant memory.

You can now run any HPC workload on Oracle Cloud Infrastructure with the same robust, scalable, and predictable performance that you would expect from a high-end on-premises HPC infrastructure. It's easy to get started: just add your data—no configuration is required.

Why Move HPC Workloads to OCI?

- Pay for compute-intensive workloads as predictable operational expenses, and avoid massive capital expenditures
- Gain performance that matches and even exceeds on-premises deployments
- Scale infrastructure up and down quickly, and pay only for the resources that you use
- Complete simulations, renderings, and AI training fast with the highest performing IaaS among public clouds
- Focus on your simulations and renderings—not on keeping up with the latest hardware and software

Visit cloud.oracle.com/iaas/hpc to test-drive an HPC cluster for yourself or to sign up for Oracle's free HPC benchmarking service.

Copyright © 2020, Oracle and/or its affiliates. All rights reserved. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Oracle and Java are registered trademarks of Oracle and/or its affiliates. Other names may be trademarks of their respective owners.

Intel and Intel Xeon are trademarks or registered trademarks of Intel Corporation. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. AMD, Opteron, the AMD logo, and the AMD Opteron logo are trademarks or registered trademarks of Advanced Micro Devices. UNIX is a registered trademark of The Open Group.