ORACLE
SUPERCLUSTER

A Technical Overview of Oracle SuperCluster

ORACLE WHITE PAPER | APRIL 2016

ORACLE

Table of Contents

Introduction

Oracle SuperCluster is Oracle's fastest, most secure, and most scalable engineered system. It is a complete engineered system for optimally running databases and applications on a single integrated system. Oracle SuperCluster is ideal for massive consolidation and private clouds. It is designed, tested, and integrated to run business-critical enterprise applications and rapidly deploy cloud services while delivering the highest levels of security, extreme efficiency, cost savings, and performance. To meet the security needs of the modern data center, Oracle SuperCluster employs a comprehensive defense-in-depth security strategy that spans the database, application, server, storage, and networking components. As a result of its architectural design, Oracle SuperCluster models are well suited for multitier enterprise applications with web, database, and application components. This versatility, along with powerful, bundled virtualization capabilities, makes them ideal systems on which to consolidate large numbers of applications, databases, and middleware workloads, or to deploy complex, multiuser development, test, and deployment environments. They combine highly available and scalable technologies, such as optional Oracle Database 12*c* and Oracle Database 11*g* with Oracle Real Application Clusters (Oracle RAC) and optional Oracle Solaris Cluster software. The architectures enable a high degree of isolation between concurrently deployed applications, which may have varied security, reliability, and performance requirements.

Oracle SuperCluster provides optimal solutions for all database workloads, ranging from scan-intensive data warehouse applications to highly concurrent online transaction processing (OLTP) applications. With their combination of Oracle Exadata Storage Server, Oracle Database software, and the latest hardware components, Oracle SuperCluster M7 and Oracle SuperCluster M6-32 deliver extreme performance in a highly available, highly secure environment. Each Oracle Exadata Storage Server uses Exadata Smart Flash Logging, a feature of Oracle Exadata that both improves user-transaction response times and increases overall database throughput for I/O-intensive workloads by accelerating performance-critical database algorithms.

For example, it provides more processing power with up to 32 SPARC M6 processors from Oracle, more memory with up to 1 TB per physical SPARC M6 processor, more built-in I/O with 64 PCIe Gen 3 slots, and, finally, three levels of virtualization: physical domains (or PDoms), logical domains (LDoms), and Oracle Solaris Zones.

Customers can integrate Oracle SuperCluster systems with Oracle Exadata or Oracle Exalogic machines by using the available InfiniBand expansion ports and optional data center switches. The InfiniBand technology used by Oracle SuperCluster systems offer high-bandwidth, low latency, hardware-level reliability and security. For application environments that follow Oracle's best practices for highly scalable, fault-tolerant systems, no application architecture or design changes are required to benefit from Oracle SuperCluster systems. Deployments also can connect many Oracle SuperCluster M7 and Oracle SuperCluster M6-32 systems or a combination of Oracle SuperCluster systems and Oracle Exadata to develop a single, large-scale environment sharing the same InfiniBand fabric. Customers can integrate Oracle SuperCluster systems with their current data center infrastructure using the available 10 GbE ports in each of Oracle’s SPARC servers within Oracle SuperCluster.

Oracle SuperCluster Components

Oracle SuperCluster is comprised of many Oracle products that are preconfigured, pretuned, and pretested by Oracle experts, eliminating weeks or months of effort typically required to design, integrate, and deploy a high-performance and highly available system. Extensive end-to-end testing ensures all components work seamlessly together and there are no performance bottlenecks or single points of failure that can affect the system. The main differences between Oracle SuperCluster M7 and Oracle SuperCluster M6-32 is the SPARC server.

Oracle's SPARC M7 Component

Oracle SuperCluster M7 is a complete, preconfigured, cluster-capable solution utilizing the following components:

» **Oracle's SPARC M7 Servers**—Oracle's new SPARC M7 processor–based servers take Oracle's server technology to new levels by offering the world's first implementation of Oracle's Software in Silicon technology to build clouds with the most secure platforms in the world. Offering both database and application security and acceleration, these servers offer at their core Software in Silicon features like Silicon Secured Memory and In-Memory Query Acceleration, along with data compression and decompression and encryption.

Each SPARC M7 chassis has the following:

- » **Utilization of Oracle virtualization technology to electrically isolate the SPARC M7 chassis into two compute nodes (physical domains)**—Each compute node is equivalent to a physical server and can be configured with one, two, or four SPARC M7 processors.
- » **Minimum of 512 GB and a maximum of 2 TB of memory per SPARC M7 compute node**—512 GB (32 GB dual inline memory modules [DIMMs]) of memory is provided per SPARC M7 processor.
- » **Minimum of two compute nodes and maximum of four compute nodes per rack**—Each SPARC M7 chassis can contain one or two compute nodes, and a rack can contain up to two SPARC M7 chassis for a maximum of four compute nodes. A minimum of two compute nodes is required for application or database high availability.
- » **Oracle's Sun PCIe dual-port QDR InfiniBand host channel adapter (HCA)**—This is a low-latency 40 Gb/sec InfiniBand HCA. There is one InfiniBand HCA configured per processor.
- » **10 Gigabit Ethernet adapters**—One 10 GbE adapter is configured per processor.

Oracle's SPARC M6-32 Component

Oracle SuperCluster M6-32 is a complete, preconfigured, cluster-capable solution utilizing the following components:

» **SPARC M6-32 servers**—Oracle's SPARC M6-32 server offers a large memory capacity and a highly integrated design that supports virtualization and consolidation of mission-critical applications. Oracle SuperCluster M6-32 can be flexibly configured to meet the most demanding data center workloads.

Each SPARC M6-32 server has the following:

- » **Utilization of Oracle virtualization technology to electrically isolate the SPARC M6-32 rack into two or four compute nodes (physical domains)**—Each compute node can be configured with 4, 8, 12, or 16 SPARC M6 processors based on specific compute node configurations.
- » **Minimum of 4 TB and a maximum of 32 TB of memory per SPARC M6-32 rack**—512 GB (16 GB DIMMs) or 1 TB (32 GB DIMMs) of memory is provided per SPARC M6 processor.
- » **Minimum of 8 and a maximum of 32 SPARC M6 processors per SPARC M6-32 rack**—Each processor comes with 12 cores and 8 threads per core.

- **Sixteen to 32 disk drives per rack**—There are eight 1.2 TB 10,000 RPM SAS disk drives per I/O unit (IOU).
- **Oracle's Sun PCIe dual-port QDR InfiniBand host channel adapter**—This is a low-latency 40 Gb/sec InfiniBand HCA. There are four InfiniBand HCAs in each IOU.
- **Eight to 16 base I/O cards**—These cards provide SAS controllers for the disks in an IOU. There are four in each IOU, and each card also provides two 10 GbE ports. The 10 GbE ports are for client access to the Oracle SuperCluster M6-32.

Common Oracle SuperCluster Components

Both Oracle SuperCluster M7 and Oracle SuperCluster M6-32 offer these elements:

- **Oracle Exadata Storage Server**—These servers provide three critical technologies to improve database performance: the smart scale-out storage and the Exadata Smart Flash Cache and Exadata Hybrid Columnar Compression features of Oracle Exadata. There are a minimum of three Oracle Exadata Storage Server systems in an Oracle SuperCluster M7 base configuration and a minimum of nine in the Oracle SuperCluster M6-32 base configuration. There are two storage technology options for Oracle Exadata Storage Server. The first is using flash-based storage and the second is using high-capacity disks storage. Oracle SuperCluster can scale storage capacity by configuring additional Oracle Exadata Storage Server systems.
- **Oracle ZFS Storage ZS3-ES appliance**—Providing 160 TB of raw disk capacity, this appliance uses the flash-enabled technology of Hybrid Storage Pool, a feature of Oracle ZFS Storage Appliance, to improve application response times. Its performance scalability for file-based I/O and ease of management make it a good fit for managing shared application data files within Oracle SuperCluster.
- **Oracle's Sun Datacenter InfiniBand Switch 36**—This switch provides a high-throughput, low-latency, and scalable fabric suitable for fabric consolidation of interprocess communication, network, and storage. InfiniBand delivers up to 63 percent higher transactions per second for Oracle Real Application Clusters (Oracle RAC) over GbE networks.
- **Ethernet management switch**—This switch provides network management connectivity to the management ports on all servers and switches used in Oracle SuperCluster**.**
- **Oracle Solaris operating system**—Oracle Solaris 11 is a secure, integrated, and open platform engineered for large-scale enterprise cloud environments, allowing users to deploy enterprise mission-critical applications safely and securely with no compromise.
- **Integrated virtualization**—Enhanced security, increased utilization, and improved reliability are delivered through Oracle Solaris Zones and through Oracle VM Server for SPARC (previously known as Oracle's Sun Logical Domains).
- **Oracle Enterprise Manager Ops Center**—Oracle Enterprise Manager Ops Center delivers a converged hardware management solution that integrates management across the infrastructure stack to help IT managers deploy and manage Oracle SuperCluster more efficiently.
- **Oracle's Exalogic Elastic Cloud Software (optional)**—Oracle Exalogic provides extreme performance for Java applications, Oracle Applications, and all other enterprise applications, and it reduces application implementation costs and ongoing costs compared to traditional enterprise application platforms and private clouds assembled from separately sourced components. Oracle SuperCluster configurations contain a specific combination of servers, storage, network, and software elements to ensure the highest quality for integration and testing during production. The systems can expand the amount of storage.

Oracle SuperCluster Component Details

All components selected and integrated into the Oracle SuperCluster platform provide unique and differentiated features. These features are integrated and tested together to provide a feature-rich solution for the most demanding data center requirements.

SPARC M7 Servers

The SPARC M7 servers are designed for modern cloud infrastructures. They are ideal for database and commercial business applications requiring operational efficiency, reliability, and scalability for large mission-critical computing environments.

Oracle's new SPARC M7 processor–based servers take Oracle's server technology to new levels by offering the world's first implementation of Oracle's Software in Silicon technology to build clouds with the most secure platforms in the world. Offering both database and application security and acceleration, these servers offer Silicon Secured Memory, In-Memory Query Acceleration, data compression and decompression, and encryption at their core.

With its new Software in Silicon capabilities coupled with an innovative cache and memory hierarchy, Oracle's SPARC M7 processor delivers dramatically higher processing speed and revolutionary protection against malware and software errors.

The per-thread performance is improved with the entirely new on-chip L2 and L3 cache design and increased processor frequency. The 64 MB L3 cache is partitioned and fully shared, and hot cache lines are migrated to the closest partition to minimize latency and maximize performance. The architecture of the core clusters and partitioned cache is ideal for server virtualization and pluggable databases. System administration and performance tuning are easier, because the design minimizes interaction between logical domains or between databases. The processor can dynamically trade per-thread performance for throughput by running up to 256 threads, or it can run fewer higher-performance threads by devoting more resources to each thread. This flexibility allows the system to balance overall throughput versus per-thread performance for optimal results.

The SPARC M7 processor incorporates eight on-chip accelerators to offload in-memory database query processing and perform real-time data decompression, while crypto instruction accelerators are integrated directly into each processor core. Together, the Software in Silicon features deliver significant performance advantages, including the following:

- Silicon Secured Memory provides real-time data integrity checking to guard against pointer-related software errors and malware, replacing very costly software instrumentation with low-overhead hardware monitoring. Silicon Secured Memory enables applications to identify erroneous or unauthorized memory access, diagnose the cause, and take appropriate recovery actions.
- In-Memory Query Acceleration provided by the accelerators delivers performance that is up to 10 times faster compared to other processors.
- The In-Line Decompression feature enables storing up to three times more data in the same memory footprint, without a performance penalty.
- Accelerated cryptography helps eliminate the performance and cost barriers typically associated with secure computing—which is increasingly essential for modern business operation. These accelerators enable high-speed encryption for more than a dozen industry-standard ciphers, eliminating the performance and cost barriers typically associated with secure computing.

Each of the eight in-silicon data analytics accelerators (DAXs) or Software in Silicon coprocessors included on the SPARC M7 die provides the following:

» A data/message pipe for very fast local data movement to offload the CPU and provide secure remote cluster messaging

» A query pipe for in-memory columnar acceleration that scans data vectors and applies predicates

Together, the coprocessors provide very low-overhead interprocess communication and very fast atomic operations. For example, DAXs located on different processors can exchange messages and access remote memory locations, exchanging locks without CPU involvement. Utilizing this functionality requires Oracle Database 12*c* with the Oracle Database In-Memory option and Oracle Solaris 11.3 or later.

CPU, Memory, and I/O Unit Chassis

The SPARC M7 server is contained within one CPU, memory, and I/O unit (CMIOU) chassis—the enclosure that houses the processor/memory boards, service processors (SPs), and connectors for the interconnect assemblies. The enclosure also includes power supplies and cooling fans. Figure 1 shows the front and rear view of the CMIOU chassis and its key components.

The SPARC M7 chassis used with Oracle SuperCluster M7 supports up to total of eight processors. However, the Oracle SuperCluster M7 chassis consists of two electrically isolated hardware partitions. This is done using Oracle's Physical Domains, to create two isolated compute nodes within each chassis. Therefore the maximum number of processors per compute node is four, and there are two compute nodes within a single chassis. All SPARC M7 server hardware required to operate both compute nodes is contained in a single CMIOU chassis.

The front of the chassis features hot-swappable fan modules and power supply units (PSUs) as well as interconnect assemblies that connect the CMIOU boards together into a system. Front-facing features include the following:

» Eight hot-swappable fan modules

» Six hot-swappable N+N redundant PSUs, 3,000 W at 200 VAC to 240 VAC

The rear of the CMIOU chassis includes the following:

» Up to eight CMIOU boards

» Two redundant SPs

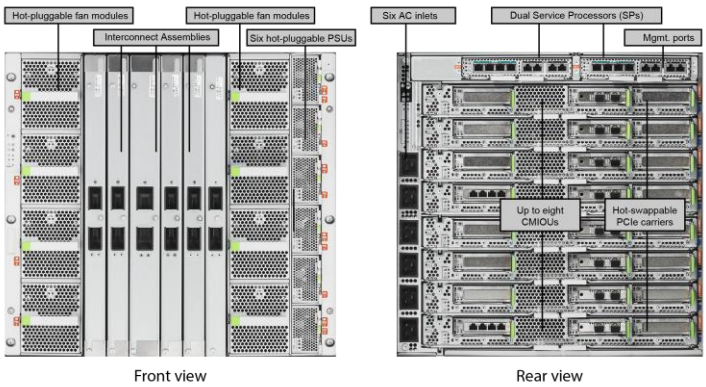

Figure 1. SPARC M7 chassis

CPU, Memory, and I/O Unit Board

Each CMIOU board assembly contains one SPARC M7 processor on a mezzanine board plus associated memory and I/O. All 16 memory DIMM slots are on the board. An I/O controller ASIC provides dedicated root complexes for three PCIe 3.0 (x16) slots. PCIe hot-pluggable carriers are included with the board. When inserted into the CMIOU chassis, the CMIOU board connects with the interconnect assemblies that provide the connectivity between the CMIOU boards and SPs.

The Oracle SuperCluster M7 chassis uses CMIOU boards with the following configuration:

- One SPARC M7 processor with 32 cores @ 4.13 GHz
- 512 GB of memory (16 x 32 GB DDR4 DIMMs)
- One InfiniBand host channel adapter
- One 10 GbE adapter
- One four-port Ethernet adapter (only for the first CMIOU per compute node [physical domain])

Interconnect Assembly

The systems interconnect for the SPARC M7 server is implemented with the interconnect assemblies. The SPARC M7 server uses interconnect assemblies that are contained within the 10U space of the CMIOU chassis. Five interconnect assemblies provide the coherency link (CL) connections (glueless systems interconnect) between the CMIOU boards. One interconnect assembly is used for communication between the CMIOU boards and the dual SPs at the top of the CMIOU chassis.

Service Processor, Service Processor Proxy, and Service Processor Module

The SPARC M7 server features redundant hot-pluggable service processors (SPs). There are two SPs located in the CMIOU chassis. Each SP has one serial (RJ45) and one 1000BASE-T management port (RJ45, autonegotiates to 10/100/1000 Mb/sec). The SP communicates with the CMIOU boards via the SP interconnect assembly.

Each SP includes two service processor modules (SPMs) in order to always have redundant SP functionality and failover capability. The SPM is the component that runs the Oracle Integrated Lights Out Manager (Oracle ILOM) software and provides the SP functionality for the server system.

SPARC M6-32 Servers

The SPARC M6-32 server provides high throughput and computing density along with built-in virtualization and extreme scalability and is a highly efficient platform for deploying large-scale, mission-critical applications.

Figure 2. SPARC M6-32 server

Architected to reduce planned and unplanned downtime, the SPARC M6-32 server includes advanced reliability, availability, and serviceability (RAS) capabilities to avoid outages and reduce recovery time. These design elements are vital for a mission-critical system such as Oracle SuperCluster. Design features that boost the reliability of the SPARC M6-32 server include:

- **Advanced CPU integration**—The SPARC M6 processor is a 12-core processor, with each core featuring eight threads. Each processor provides 48 MB of L3 cache, which is critical in reducing the time required to read and process data.
- **Extended-ECC memory**—The memory system has redundant components built on to each memory DIMM, which allow a DIMM to continue to operate even with partial failure. Utilizing 16 GB DIMMs or 32 GB DIMMs, each SPARC M6-32 server in Oracle SuperCluster M6-32 provides up to 16 TB or 32 TB of memory.
- **Memory lane sparing**—Each of the two built-in memory controllers in the SPARC M6 processor has memory lane sparing. In the event that a particular lane fails, the memory controller can continue without interruption to access the main memory.
- **Processor lane sparing**—The interconnect on all SPARC M6 processor/memory boards has lane sparing. This means there are no lost packets of data during cache coherency or for remote memory access.
- **Fault-resilient power options and hot-swappable components**—Systems feature redundant, hot-swappable service processors, power supply units, and fan units. Redundant storage can be created using hot-swappable disk drives with disk-mirroring software.
- **Hardware redundancy**— SPARC M6-32 servers feature redundant system interconnect boards, redundant clock boards, redundant power, redundant fans, and redundant data paths.

The 10 GbE interfaces on the base I/O cards are used by clients for access to the Oracle SuperCluster M6-32 compute nodes. In addition to the 10 GbE interfaces, also included are Quad Data Rate (QDR) InfiniBand cards. These are used for internal communication and access to the instances of Oracle Exadata Storage Server and the Oracle ZFS Storage ZS3-ES appliance.

All CPUs and memory are contained on boards called CPU/memory units (CMUs). All I/O components are placed in an I/O unit (IOU). There is one IOU for every four CMUs. A grouping of four CMUs and an IOU is called a domain

configuration unit (DCU). In the SPARC M6-32 server, four DCUs are possible, and each DCU can have either two or four CMUs.

Figure 3. DCU assignments

The DCUs are building blocks that are used to configure the physical domains (PDoms). In Oracle SuperCluster M6-32, there are either two or four PDoms, with two-PDom configurations supporting either one or two DCUs each. There are two types of configurations, basic and extended. Base configurations always have one DCU per PDom, and extended configurations always have two DCUs per PDom.

Because each DCU has its own IOU, there are enough PCIe slots to provide I/O connectivity for the standard I/O cards included. Each IOU has four QDR InfiniBand cards, four base I/O cards (also called Express Module SAS [EMS] cards) with two 10 GbE ports each, eight 1.2 TB 10,000 RPM SAS disk drives, and one 1 GbE card for domain management. Figure 4 illustrates the IOU layout.

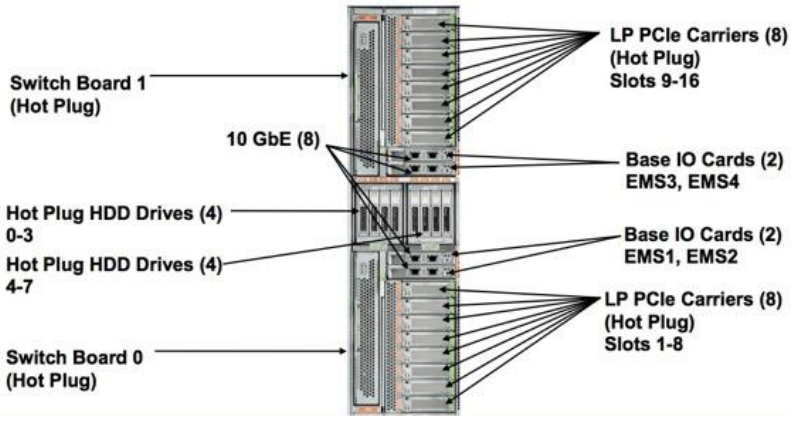

Figure 4. IOU layout

The 10 GbE ports on the EMS cards are the primary network ports for client access to Oracle SuperCluster M6-32.

Platform management is done via the service processor, or Oracle Integrated Lights Out Manager (Oracle ILOM) 3.0. Oracle ILOM provides a command-line interface (CLI), a web-based graphical user interface (GUI), and Intelligent Platform Management Interface (IPMI) functionality to aid out-of-band monitoring and administration. The management software, Oracle Enterprise Manager Ops Center, communicates with Oracle ILOM to manage and monitor Oracle SuperCluster M6-32.

Oracle Exadata Storage Server

Oracle Exadata Storage Server runs Oracle's Exadata Storage Server Software, which provides unique and powerful software technology, including its Exadata I/O Resource Manager and Exadata Storage Indexes features, as well as these Oracle Exadata features: Exadata Smart Scan, Exadata Smart Flash Cache, Exadata Smart Flash Logging, and Exadata Hybrid Columnar Compression. The hardware components of each Oracle Exadata Storage Server (also referred to as an Oracle Exadata cell) are carefully chosen to match the needs of high-performance database processing. Exadata Storage Server Software is optimized to take the best possible advantage of the hardware components and Oracle Database. Each Oracle Exadata cell delivers outstanding I/O performance and bandwidth to the database. When used in Oracle SuperCluster, Oracle Exadata Storage Server can be accessed only by Oracle Database 12*c* and Oracle Database 11*g* Release 2 running in the database domain. They cannot be used for any other purpose beyond Oracle Database 12*c* and Oracle Database 11*g* Release 2.

Figure 5. Oracle Exadata Extreme Flash and High Capacity Storage Server

Oracle Exadata Storage Server delivers exceptional database performance by way of the following critical technologies:

- **Exadata Smart Scan**—processes queries at the storage layer, returning only relevant rows and columns to the database compute nodes. As a result, much less data travels over fast 40 Gb InfiniBand interconnects—dramatically improving the performance and concurrency of simple and complex queries.
- **Exadata Smart Flash Cache**—addresses the disk random I/O bottleneck problem by transparently caching "hot" frequently accessed data to fast solid-state storage. It provides up to a 30x improvement in response time for reads compared to regular disk and up to 20x more write performance—a hundred-fold improvement in IOPS for reads compared to regular disk. And, it is a less expensive higher-capacity alternative to memory.
- **Exadata Smart Flash Logging**—takes advantage of the flash memory in Oracle Exadata storage to speed up log writes.
- **Exadata Hybrid Columnar Compression**—can reduce the size of data warehousing tables by an average of 10x and of archive tables by 50x. This offers significant savings on disk space for primary, standby, and backup databases, and dramatically improves the performance of data warehousing queries.

There are two Oracle Exadata Storage Server configuration options. The first is a high-capacity option with both flash storage and 12 high-capacity SAS disks. The second is an extreme flash option using eight NVMe PCI flash drives, providing maximum performance for the most demanding workloads.

While Oracle Exadata Storage Server can be accessed only by a production Oracle Database 12*c* or Oracle Database 11*g* Release 2 database running in a database domain, an Oracle Database 12*c* or Oracle Database 11*g*

Release 2 database can be run in an application domain on Oracle Solaris 11 for testing and development purposes. The Oracle ZFS Storage ZS3-ES appliance then can store the database. Exadata Hybrid Columnar Compression software technology also can be used on the Oracle ZFS Storage ZS3-ES appliance to test Exadata Hybrid Columnar Compression functionality before moving the database to the database domain and using Oracle Exadata Storage Server.

Oracle ZFS Storage Appliance

For shared file storage, Oracle SuperCluster includes an appliance from the Oracle ZFS Storage Appliance family, specifically the Oracle ZFS Storage ZS3-ES appliance, which features a common, easy-to-use management interface and the industry's most comprehensive analytics environment. To deliver high performance using cost-effective components, the Oracle ZFS Storage Appliance file system, Oracle Solaris ZFS, seamlessly optimizes access to the different types of media in the Hybrid Storage Pool technology. Oracle Solaris ZFS is designed to automatically recognize different I/O patterns and place data in the best storage media for optimal performance.

Oracle ZFS Storage Appliance in Oracle SuperCluster features a two-node cluster configuration that enables high performance and high availability (HA) to maximize business productivity. This is the configuration used in Oracle SuperCluster.

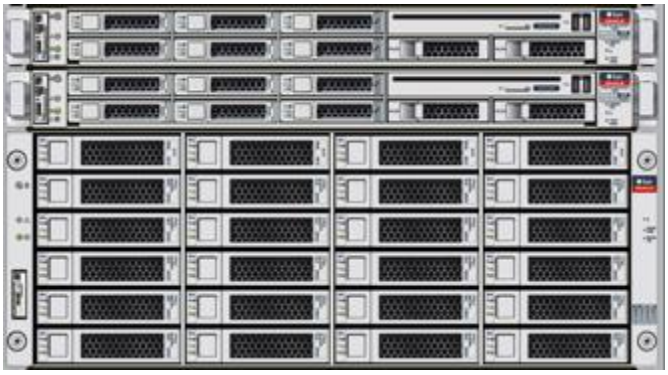

Figure 6. Oracle ZFS Storage ZS3-ES appliance

Figure 7. Real-time analytics help administrators increase storage optimization and reduce down time.

For Oracle SuperCluster M7, Oracle ZFS Storage ZS3-ES appliance is installed in the main Oracle SuperCluster M7 rack. For Oracle SuperCluster M6-32, Oracle ZFS Storage Appliance is installed in the Oracle SuperCluster M6-32 storage rack. Oracle ZFS Storage ZS3-ES appliance can be made accessible from the database domain and used for Oracle Recovery Manager (Oracle RMAN) backups or flat file staging. Oracle RMAN is a feature of Oracle Database.

Networking—Oracle's Sun Datacenter InfiniBand Switch 36

InfiniBand technology has emerged as an attractive fabric for building large supercomputing grids and clusters. As an open standard, InfiniBand presents a compelling choice compared to proprietary interconnect technologies that depend on the success and innovation of a single vendor.

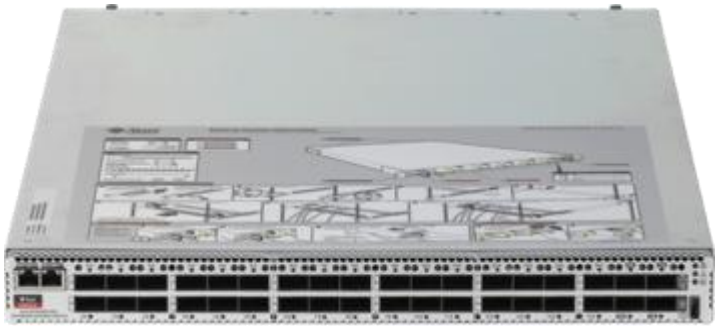

Figure 8. Sun Datacenter InfiniBand Switch 36

Oracle SuperCluster M7 contains either two or three Sun Datacenter InfiniBand Switch 36 switches. Oracle SuperCluster M6-32 contains three Sun Datacenter InfiniBand Switch 36 switches. The first two switches are always leaf switches and the third is used as a spine switch to connect multiple SuperCluster and other engineered system racks together. The two leaf switches are connected to each other to provide redundancy in case one of the two leaf switches fail. In addition, each SPARC compute node, Oracle Exadata Storage Server, and each controller on the Oracle ZFS Storage ZS3-ES appliance has connections to both leaf switches to provide redundancy in the InfiniBand connections in case one of the two leaf switches fail.

Oracle Solaris

Oracle Solaris provides key functionality for virtualization, optimal use, high availability, unparalleled security, and extreme performance for both vertically and horizontally scaled environments. Oracle SuperCluster supports both Oracle Solaris 10 and Oracle Solaris 11 on Oracle SuperCluster M6-32. On the Oracle SuperCluster M7, Oracle Solaris 11 is required for all domains while Oracle Solaris 10 can be used in a zone. Database domains exclusively use Oracle Solaris 11 on both SuperCluster platforms.

Oracle Solaris includes the following features that make it the right operating system for Oracle SuperCluster:

- **Advanced reliability**—Uptime is enhanced through comprehensive testing across an integrated solution stack and features such as Predictive Self Healing for hardware and software faults, data integrity with Oracle Solaris ZFS, and live observability with DTrace Analytics.
- **Superior performance**—Oracle Solaris is optimized for throughput and scalability for the latest SPARC processor technologies and has achieved record-setting benchmarks for TPC-H, TPC-C, Oracle's PeopleSoft applications, Oracle Business Intelligence Suite, Enterprise Edition, and many others.
- **Built-in virtualization**—Oracle Solaris Zones and Oracle VM Server for SPARC (previously known Oracle's Sun Logical Domains) along with other OS and network virtualization capabilities, enable efficient consolidation for flexibility and performance without significant overhead.
- **Pervasive security infrastructure**—Oracle Solaris provides the compartmentalization and control needed for multitenancy environments and enables governments and financial institutions to meet their strict requirements.
- **Committed support**—Oracle offers sustaining support for Oracle Solaris releases for as long as customers operate their systems, making it possible to keep software infrastructures in place for as long as it makes business sense.

Fault Management and Predictive Self-Healing

Oracle Solaris provides an architecture for building and deploying systems and services capable of fault management and predictive self-healing. The Predictive Self Healing feature in Oracle Solaris automatically diagnoses, isolates, and recovers from many hardware and application faults. As a result, business-critical applications and essential system services can continue uninterrupted in the event of software failures, major hardware component failures, and even software misconfiguration problems.

- **Fault Manager**—The Fault Manager feature in Oracle Solaris collects data relating to hardware and software errors. This facility automatically and silently detects and diagnoses the underlying problem, with an extensible set of agents that automatically responds by taking the faulty component offline.
- **Service Management Facility**—The Service Management Facility feature in Oracle Solaris creates a standardized control mechanism for application services by turning them into first-class objects that administrators can observe and manage in a uniform way. These services then can be automatically restarted if an administrator accidentally terminates them, if they are aborted as the result of a software programming error, or if they are interrupted by an underlying hardware problem.

Predictive Self Healing and Fault Manager can offline processor threads or cores in faults, retire suspect pages of memory, log errors, or faults from I/O operations, or deal with any other issue detected by the system.

Virtualization on Oracle SuperCluster

Oracle SuperCluster leveraged several different virtualization technologies to deliver a robust, customizable environment to meet a wide variety of application and database deployment requirements. These technologies can be layered, or used together, to achieve various levels of security, isolation, performance, and availability when consolidating both applications and databases onto a single Oracle SuperCluster. Oracle's virtualization technology is provided as a part of the system, for no additional cost.

Oracle Solaris Zones

Oracle Solaris provides a unique partitioning technology, a feature called Oracle Solaris Zones, which enables creation of zones. A zone is a virtualized operating system environment created within a single instance of Oracle Solaris. Oracle Solaris Zones can be used to isolate applications and processes from the rest of the system. This isolation helps enhance security and reliability since processes in one zone are prevented from interfering with processes running in another zone.

Oracle Solaris Zones also can be used to run older versions of the Oracle Solaris operating system, databases, and applications on the latest-generation SPARC servers. Hardware migrations and software updates can be deferred until times when both are more convenient. Legacy applications can benefit from faster hardware by migrating to new systems while running on an older OS revision. Resource management tools provided with Oracle Solaris enable administrators to allocate resources such as CPUs to specific applications or zones. CPUs in a multiprocessor system (or threads in a multicore processor) can be logically partitioned into processor sets and bound to a resource pool, which in turn can be assigned to a zone. Resource pools provide the capability to separate workloads so that consumption of CPU resources does not overlap. They also provide a persistent configuration mechanism for processor sets and for scheduling class assignment. In addition, the dynamic features of resource pools enable administrators to adjust system resources in response to changing workload demands.

With Oracle SuperCluster, Oracle Solaris Zones can be deployed within a physical domain or Oracle VM Server for SPARC. Within the Oracle SuperCluster architecture, there are application zones and database zones. Each type of zone follows specific best practices for deployment on Oracle SuperCluster.

- **Database zone**—an Oracle Solaris zone dedicated to running Oracle Database 12*c* and Oracle Database 11*g* Release 2, using Oracle Exadata Storage Server systems for database storage.
- **Application zone**—a domain dedicated to running applications on either Oracle Solaris 11 or Oracle Solaris 10. With the Oracle SuperCluster M6-32, application zones running Oracle Solaris 10 also support the use of Oracle Solaris legacy zones for applications that require either an Oracle Solaris 8 or Oracle Solaris 9 environment. Oracle SuperCluster M7 does not support legacy zones for Oracle Solaris 8 or Oracle Solaris 9, and it will only support legacy zones for Oracle Solaris 10.

Oracle VM Server for SPARC

Oracle SuperCluster systems also support Oracle VM Server for SPARC (also known as logical domains) virtualization technology. Oracle VM Server for SPARC provides full virtual machines that run an independent operating system instance and contain virtualized CPU, memory, networking interfaces, storage, console, and

cryptographic devices. Within the Oracle VM Server for SPARC architecture, operating systems such as Oracle Solaris 11 are written to the hypervisor, which provides a stable, idealized, and virtualizable representation of the underlying server hardware to the operating system in each domain. Each domain is completely isolated.

With the release of Oracle SuperCluster Software 2.0, support for Oracle VM Server for SPARC is expanded to include the I/O Domain Creation Tool, a new management tool that is designed to greatly simplify deployment and lifecycle management of Oracle VM for SPARC virtualization technology on Oracle SuperCluster. This tool provides a new browser user interface (BUI) to create and manage both application and database virtual machines (LDoms). In addition, this tool supports Oracle VM for SPARC Templates. With this combination of technologies on Oracle SuperCluster, IT administrators to capture fully configured and approved O/S and application settings, patch levels, security configurations, network configurations, and other tasks into a single template that can be deployed by users through the I/O Domain Creation Tool. As a result, virtual machines can be deployed and operational in minutes with fully preconfigured and fully isolated Oracle Solaris OS and customer applications and achieve complete compliance with IT business requirements for application deployments.

- **Database domain**—a domain dedicated to running Oracle Database 12*c* and Oracle Database 11*g* Release 2, using Oracle Exadata Storage Server for database storage. This domain must run Oracle Solaris 11.
- **Application domain**—a domain dedicated to running applications on either Oracle Solaris 11 or Oracle Solaris 10 with Oracle SuperCluster M6-32. Application domains running Oracle Solaris 10 also support the use of Oracle Solaris legacy zones for applications that require either an Oracle Solaris 8 or Oracle Solaris 9 environment. For Oracle SuperCluster M7, all application domains must run Oracle Solaris 11. Application domains running Oracle Solaris 11 also support the use of Oracle Solaris 10 Zones. Exalogic Elastic Cloud Software runs only in an application domain on Oracle Solaris 11.

Physical Domains (PDoms)

Both Oracle SuperCluster M7 and Oracle SuperCluster M6-32 support a third type of virtualization, called Physical Domains (PDoms). They are also known as compute nodes. While Oracle Solaris Zones provides process isolation and Oracle VM for SPARC (that is, logical domains) provides OS isolation, PDoms provide hardware isolation. SPARC M7 can provide up to two PDoms per server chassis. And the SPARC M6-32 server can provide either two or four PDoms per server rack. These PDoms are essentially two or four individual servers sharing redundant power supplies, redundant fans, and redundant service processors. Each Physical Domain can be either a database domain or an application domain and can be virtualized using Oracle Solaris Zones and Oracle VM for SPARC to increase overall resource utilization for each PDom. For the purpose of simplifying terminology, and because a PDom is an independent server, any future references to PDoms will be as compute nodes.

Secure Multitenancy on Oracle SuperCluster

The Oracle SuperCluster platform is designed to support a diverse set of customer workloads and use cases. One of the most challenging deployment scenarios supported is multitenant operation, whereby multiple independent entities are permitted to operate and manage their own services on a single, physical Oracle SuperCluster platform. To accomplish this formidable task, a collection of security controls must be correctly implemented to enforce key security requirements such as secure isolation, strong access control, comprehensive data protection, and trustworthy auditing. For this engineered system, Oracle experts provide a recommended set of technical security controls that, when used together, help to create a secure and reinforcing multitenant architecture built upon the Oracle SuperCluster platform.

Out-of-the-Box Security Controls

Oracle SuperCluster M7 delivers the industry's most advanced security in every layer of the stack: application to chip. An engineered system, Oracle SuperCluster M7 is a complete stack of applications, database, server, operating system, networking, storage, and system management that is integrated, preconfigured, pretuned, and pretested by Oracle experts. In each layer of the solution stack, there are industry-leading security features that can be leveraged to make Oracle SuperCluster M7 the most secure platform available for enterprise workloads. With Oracle SuperCluster M7, Oracle experts have preconfigured many of these security settings to make Oracle SuperCluster M7 more secure at time of installation, reducing the workload for system administrators and security officers to implement security measures across the many technologies that need to be configured to ensure secure operations. Out-of-the-box security for Oracle SuperCluster M7 enables organizations to realize the value of pre-engineered, pretested, and preverified systems, enabling organizations to have peace of mind when it comes to security.

Oracle Enterprise Manager Ops Center

Oracle Enterprise Manager Ops Center delivers a converged hardware management solution for Oracle SuperCluster that integrates management across the infrastructure stack. With advanced virtualization management and reporting capabilities, application-to-disk management, intelligent configuration management, and more, Oracle Enterprise Manager Ops Center helps IT managers reduce complexity and streamline and simplify infrastructure management. The inclusion of Oracle Enterprise Manager Ops Center with every Oracle SuperCluster enables data center administrators to monitor and manage the storage, network, servers, Oracle Solaris, and virtualized environments from a single interface. This improves operational efficiency and lowers operational costs.

Figure 9. Management of the Oracle technology stack

Using Oracle Enterprise Manager Ops Center 12*c*, Oracle SuperCluster is managed as a specific engineered solution, not just as a rack of servers and storage.

Oracle Enterprise Manager Ops Center is the most comprehensive management solution for the Oracle SuperCluster hardware infrastructure. Offering a single console to manage multiple server architectures and myriad

operating systems, Oracle Enterprise Manager Ops Center can manage the components in Oracle SuperCluster using asset discovery, provisioning of firmware and operating systems, automated patch management, patch and configuration management, virtualization management, and comprehensive compliance reporting.

Oracle Enterprise Manager Ops Center automates workflow and enforces compliance via policy-based management—all through a single intuitive interface. With Oracle Enterprise Manager Ops Center, IT staff can implement and enforce data center standardization and best practices, regulatory compliance, and security policies while efficiently deploying infrastructure to meet business requirements. Figure 10 shows the intuitive browser-based user interface for Oracle Enterprise Manager Ops Center.

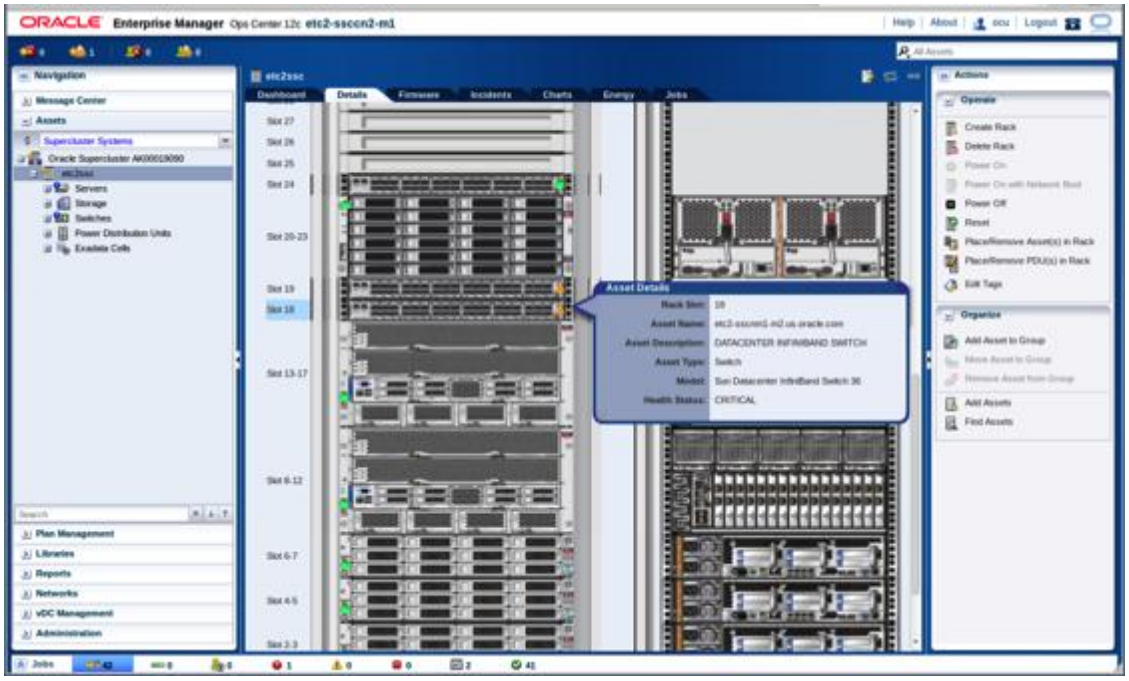

Figure 10. The Oracle Enterprise Manager Ops Center interface

There are other tools that can be used to manage Oracle SuperCluster. Some of these tools are native to the hardware (for example, Oracle ILOM and Oracle ZFS Storage Appliance), while others are part of the management software (for example, Oracle Enterprise Manager Ops Center and Oracle Solaris Cluster Manager). Some of the tools require installation and setup after Oracle SuperCluster is installed.

Oracle ILOM exchanges data with Oracle Enterprise Manager Ops Center for highly functional and easy-to-use management. If customers wish, they can access Oracle ILOM directly.

The other tools for managing Oracle SuperCluster include the following:

- Oracle ILOM (that is, the service processor) on SPARC M7 and SPARC M6-32 servers:
  - Two interfaces are provided: a command-line interface (CLI)—over serial or Ethernet—and a browser user interface (BUI) for using a browser for data center management.
  - Once the server is set up, there should be little need to connect directly to Oracle ILOM, except to connect to the Oracle Solaris domain console.
  - SNMP—The MIB should be given to the existing data center management tool in the event that Oracle Enterprise Manager Ops Center is not used to monitor Oracle SuperCluster.

- Oracle ILOM on the Oracle ZFS Storage Appliance controller:
  - Two interfaces are provided: a command-line interface (CLI)—over serial or Ethernet—and a browser user interface (BUI) for using a browser for data center management.
  - There should be very little reason to connect to Oracle ILOM for platform management. Almost all management should be done with a browser connecting to the Oracle ZFS Storage Appliance software.
- InfiniBand switches and Ethernet switch—After Oracle Enterprise Manager Ops Center is properly set up during post-installation, it can be used to monitor the InfiniBand switches and the Cisco switch using SNMP. If Oracle Enterprise Manager Ops Center is not used at the customer site, then the MIBs for the switches should be added to the data center management tool currently being used.
- Oracle Solaris domains and applications—It is highly recommended that Oracle Enterprise Manager Ops Center be used to manage all Oracle Solaris instances. Oracle Solaris does provide a built-in SNMP MIB for basic monitoring of events.

The major element that needs to be installed and set up separately from the installation of Oracle SuperCluster is the specific application software that is unique to a given deployment.

Oracle Solaris Cluster

To limit outages due to single points of failure, mission-critical services need to be run in clustered physical servers, or Oracle Solaris Zones, that efficiently and smoothly take over the services from failing nodes or zones with minimal interruption to the user experience. While Oracle SuperCluster is designed with full redundancy at the hardware level, Oracle Solaris Cluster provides the best HA solution for Oracle's SPARC servers running Oracle Solaris and applications. Tightly coupled with Oracle Solaris, Oracle Solaris Cluster detects failures without delay (zero-second delay), provides much faster failure notification, application failover, and reconfiguration time.

Significantly reducing recovery time for services results in much faster resumption of IT services. Oracle Solaris Cluster speeds this process on Oracle SuperCluster by doing the following:

- Integrates tightly with the Predictive Self Healing framework and supports applications controlled by the Service Management Facility in Oracle Solaris Zones and logical domains
- Makes extensive use of Oracle's storage management and volume management capabilities
- Supports Oracle Solaris ZFS as a failover file system and as a boot file system, allowing the use of ZFS storage as the single file system type used
- Leverages ZFS features such as pooled storage, built-in redundancy, and data integrity
- Uses I/O multipathing (MPxIO) in Oracle Solaris to represent and manage devices that are accessible through multiple I/O controller interfaces within a single instance of Oracle Solaris
- Supports network IP multipathing to enhance resiliency and throughput in a clustered environment
- Integrates with Oracle Enterprise Manager Ops Center
- Offers secure administrative capabilities through role-based access control (RBAC) capabilities in Oracle Solaris to enhance security

Virtualization with Oracle Solaris Cluster

Oracle Solaris Cluster works seamlessly with Oracle's virtualization technologies to consolidate multiple applications within the same cluster of physical servers, optimizing resource use, ensuring availability of mission-critical services, and improving data integrity. A cluster can be a mix of whole SPARC T5-8 or SPARC M6-32 compute nodes, PDoms, logical domain guests, or Oracle Solaris Zones.

Multiple options are available when using Oracle Solaris Zones. The “failover” approach treats zones as “black boxes,” which can be easily restarted or moved among cluster nodes. This solution supports Oracle Solaris 8 Zones, Oracle Solaris 9 Zones, and Oracle Solaris 10 Zones.

Oracle's Exalogic Elastic Cloud Software (Optional)

Exalogic Elastic Cloud Software on Oracle SuperCluster utilizes the SPARC M7 and SPARC M6-32 servers, flash, InfiniBand I/O fabric, and storage technology with the Oracle Solaris 11 operating system to provide a tested, tuned, optimized, and factory-assembled platform for Oracle Fusion Middleware and Oracle’s business applications portfolio. Exalogic Elastic Cloud Software is the unique set of software components and tools that are optimized for Oracle Fusion Middleware and business applications. As stated earlier, Exalogic Elastic Cloud Software can be run only in an Oracle Solaris 11–based application domain.

Because Oracle currently offers Oracle’s Exalogic Elastic Cloud X4-2 hardware, it is important to note the differences when deploying Exalogic Elastic Cloud Software on Oracle SuperCluster. There are really only a few key points of difference:

- Oracle SuperCluster utilizes SPARC server nodes and supports applications on the Oracle Solaris 8, 9, 10, and 11 releases. Exalogic Elastic Cloud 43-2 utilizes x86 server nodes and supports Oracle Linux (UEK 5.5 64-bit only) and Oracle Solaris 11 Express (or Oracle Solaris 10 applications in Oracle Solaris 10 Zones).
- Oracle SuperCluster incorporates instances of Oracle Exadata Storage Server and is a recommended platform for Oracle RAC, supporting data warehouse and OLTP use cases. Exalogic Elastic Cloud X4-2 provides no optimization for Oracle Database and is designed to be deployed in conjunction with Oracle Exadata.

It is important to note the functional differences between the Exalogic Elastic Cloud X4-2 solution and running the Exalogic Elastic Cloud Software on Oracle SuperCluster. Here are the features of Exalogic Elastic Cloud Software that are immediately available on Oracle SuperCluster:

- Oracle WebLogic Server Java Database Connectivity (JDBC) and data source optimizations
- Oracle WebLogic Server cluster state replication optimizations
- Oracle WebLogic Server socket direct protocol (on the Oracle SuperCluster InfiniBand stack only)
- Coherence messagebus API support (on the Oracle SuperCluster InfiniBand stack only)
- Oracle Traffic Director

The ability to combine the application tier with the database on a single, integrated engineered solution demonstrates one of the greatest values of Oracle SuperCluster. By adding Oracle Enterprise Manager and Oracle Enterprise Manager Ops Center to manage the entire software and hardware stack, Oracle SuperCluster M7 and Oracle SuperCluster M6-32 provide a complete solution.

In many customer environments, it is highly desirable that all data services be configured in a high-availability configuration. The Oracle Solaris Cluster software addresses this requirement for a variety of customer deployment scenarios. Having deployed a cluster, the next step is for platform management and monitoring. Oracle Enterprise Management Ops Center is the recommended package to manage the servers, network, and storage. This can be used alongside Oracle Enterprise Manager, which manages the Oracle application software offerings.

Oracle SuperCluster M7 Configurations

Oracle SuperCluster is a complete solution that has been designed, tested, and integrated together to deliver industry-leading performance, availability, and security. The basic building block of an Oracle SuperCluster M7 solution is the SPARC M7 server chassis, which contains up to two configurable compute nodes. Compute nodes are also referred to as PDoms. The basic configuration starts with an Oracle SuperCluster rack that can contain one or two SPARC M7 server chassis as well as Oracle Exadata Storage Server, an appliance from the Oracle ZFS Storage Appliance family, InfiniBand switches, and an Ethernet magagement switch. Both Oracle Exadata storage cell types, the high capacity and extreme flash, may be configured within a single Oracle SuperCluster rack. Additional Oracle Exadata Storage Servers can be configured as part of the storage expansion rack. Multiple Oracle SuperCluster racks may be connected together via the InfiniBand interconnect.

Figure 11 shows the components of Oracle SuperCluster M7. The numbered items in the figure are identified in the table that immediately follows the figure.

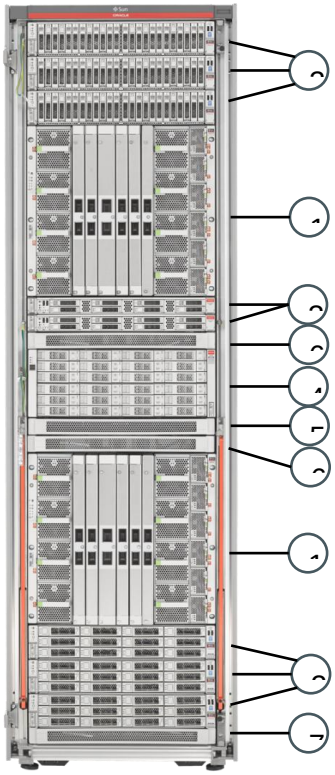

Figure 11. Full Oracle SuperCluster M7 rack

TABLE 1. MAPPING OF ORACLE SUPERCLUSTER M7 COMPONENTS

1 SPARC M7 chassis (2)	5 Ethernet management switch
2 Oracle Exadata Storage Servers (6)	6 Sun Datacenter InfiniBand Switch 36 leaf switches (2)
3 Oracle ZFS Storage ZS3-ES appliance: controllers (2)	7 Sun Datacenter InfiniBand Switch 36 spine switch
4 Oracle ZFS Storage ZS3-ES appliance: disk shelf	

TABLE 2. COMPONENT CONFIGURATIONS

Component	Minimum Configuration	Maximum Configuration
SPARC M7 Compute Chassis	» 1 x SPARC M7 compute node (only when configured with two SPARC M7 compute chassis per rack) » Redundant Oracle ILOM service processors » 6 x 3,000 watt AC power supplies (N+N) » 8 x redundant hot-swap fan modules	» 2 x SPARC M7 compute Nodes » Redundant Oracle ILOM service processors » 6 x 3,000 watt AC power supplies (N+N) » 8 x redundant hot-swap fan modules
SPARC M7 Compute Node	» 1 x 32-core SPARC M7 processor (4.1GHz) » 512 GB of DDR4 memory » 1 x 10 GbE SFP+ PCIe 2.0 PCIe low-profile card with pluggable transceivers (two ports) » 1 x Quad Data Rate (QDR) InfiniBand PCIe low-profile host channel adapter (two ports) » 1 x GbE adapter (for management network)	» 4 x 32-core SPARC M7 processor (4.1 GHz) » 2 TB of DDR4 memory » 4 x 10 GbE SFP+ PCIe 2.0 PCIe low-profile card with pluggable transceivers (2 port) » 4 x Quad Data Rate (QDR) InfiniBand PCIe low-profile host channel adapter (two ports) » 1 x GbE adapter (for management network) » (Optional) 3 x Fibre Channel adapters
Oracle ZFS Storage ZS3-ES Appliance	Two controllers, each containing the following: » 2 x Intel® Xeon® processor E5-2658 (8-core, 2.1 GHz CPU) » 256 GB of memory (16 x 16 GB DDR3) » 1 x two-port InfiniBand HCA » 2 x 900 GB SATA disks » 2 x 1.6 TB read-optimized SSDs » 1 x SAS HBA » 2 x dual-port 10 GbE cards » 1 x cluster card One disk shelf containing the following: » 20 x 8 TB SAS disks at 7,200 rpm » 4 x 200 GB write-optimized SSDs	Same as Minimum Configuration
High-Capacity Oracle Exadata Storage Cell	» 2 x ten-core Intel® Xeon® processor E5-2630 v4 2.2 GHz » 128 GB of memory (8x16 GB 2,133 MHz DIMM) » Disk controller HBA with 1GB cache » 12 x 8 TB 7,200 RPM SAS HDD and 4 x 3.2 TB NVMe PCIe flash drives » Dual port InfiniBand QDR (40 Gb/sec) HCA (PCIe 3.0)	Same as Minimum Configuration
Extreme Flash Oracle Exadata Storage Cell	» 2 x ten-core Intel® Xeon® processor E5-2630 v4 2.2 GHz » 128 GB of memory (8 x 16 GB 2,133 MHz DIMM) » 8 x 3.2 TB NVMe PCIe flash drives » Dual port InfiniBand QDR (40 Gb/sec) HCA (PCIe 3.0)	Same as Minimum Configuration
InfiniBand Switch	» 36 port QDR (40 Gb/sec) InfiniBand switch	Same as Minimum Configuration
Ethernet Management Switch	» Ethernet switch provides 48 Ethernet ports, and each port has a wire speed of 10/100/1000 BASE-T	Same as Minimum Configuration

TABLE 3. RACK CONFIGURATIONS

Rack Configurations	Minimum Configuration	Maximum Configuration
Small Rack Configuration	» 1 x SPARC M7 compute chassis » 2 x SPARC M7 compute node (minimum one processor per node) » 1 x Oracle ZFS Storage ZS3-ES appliance » 3 x Oracle Exadata storage cell (either high capacity or extreme flash) » 2 x InfiniBand switches » 1 x Ethernet management switch » 2 x redundant power distributions units (PDUs)	**Compute-Intensive Rack:** » 2 x SPARC M7 compute chassis » 4 x SPARC M7 compute node » 6 x Oracle Exadata storage cell (high capacity and/or extreme flash) **Storage-Intensive Rack:** » 1 x SPARC M7 compute chassis » 2 x SPARC M7 compute node » 11 x Oracle Exadata storage cell (high capacity and/or extreme flash) **Common in Both Storage or Compute Rack:** » 1 x Oracle ZFS Storage ZS3-ES appliance » 3 x InfiniBand switches » 1 x Ethernet management switch » 2 x redundant power distributions units (PDUs)
Elastic Rack Configuration	» 1 x SPARC M7 compute chassis » 2 x SPARC M7 compute node (minimum two processor per node) » 1 x Oracle ZFS Storage ZS3-ES appliance » 3 x Oracle Exadata storage cell (either high capacity or extreme flash) » 3 x InfiniBand switches » 1 x Ethernet management switch » 2 x redundant power distributions units (PDUs)	**Compute-Intensive Rack:** » 2 x SPARC M7 compute chassis » 4 x SPARC M7 compute node » 6 x Oracle Exadata storage cell (high capacity and/or extreme flash) **Storage-Intensive Rack:** » 1 x SPARC M7 compute chassis » 2 x SPARC M7 compute node » 11 x Oracle Exadata storage cell (high capacity and/or extreme flash) **Common in Both Storage or Compute Rack:** » 1 x Oracle ZFS Storage ZS3-ES appliance » 3 x InfiniBand switches » 1 x Ethernet management switch » 2 x redundant power distributions units (PDUs)
Exadata Storage Cell Expansion Rack	» 4 x Oracle Exadata storage cell (high capacity and/or extreme flash) » 2 x redundant power distributions units (PDUs)	» 19 x Oracle Exadata storage cell (high capacity and/or extreme flash) » 2 x redundant power distributions units (PDUs)

Oracle SuperCluster M6-32 Configurations

Oracle SuperCluster M6-32 provides the same complete solution as Oracle SuperCluster M7.

The basic building block of an Oracle SuperCluster M6-32 solution is the SPARC M6-32 server rack, which contains two configurable compute nodes. Compute nodes are also referred to as PDoms. The basic configuration starts with an Oracle SuperCluster server rack that contains the SPARC M6-32 server chassis. A second rack contains the other required components: the Oracle Exadata Storage Server, an appliance from the Oracle ZFS Storage Appliance family, InfiniBand switches, and an Ethernet magagement switch. Both Oracle Exadata storage cell types, the high capacity and extreme flash, may be configured within a single Oracle SuperCluster storage rack. Additional

Oracle Exadata Storage Server can be configured as part of the storage expansion rack. Multiple Oracle SuperCluster racks may be connected together via the InfiniBand interconnect.

Figures 12 through 14 show the components of Oracle SuperCluster M6-32. The numbered items in the figures are identified in the tables that immediately follow the figures.

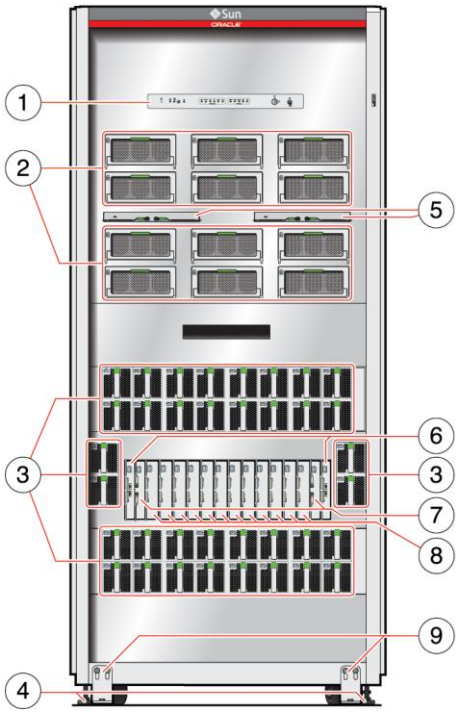

Figure 12. Full Oracle SuperCluster M6-32 rack: front view

TABLE 4. MAPPING OF ORACLE SUPERCLUSTER M6-32 COMPONENTS

1 Operator panel	6 Service processors
2 Fans	7 Clock boards
3 Power supplies	8 Scalability switch boards
4 Leveling feet	9 Mounting brackets
5 Power supply distribution boards	

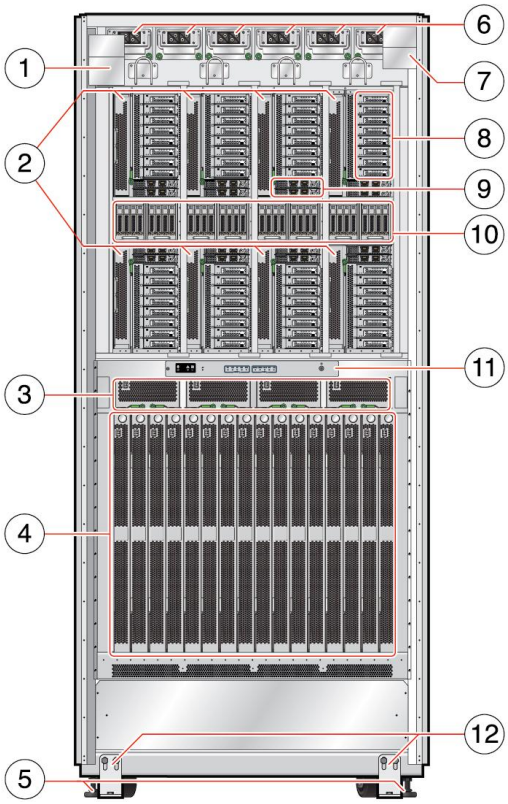

Figure 13. Full Oracle SuperCluster M6-32 rack: rear view

TABLE 5. MAPPING OF ORACLE SUPERCLUSTER M6-32 COMPONENTS

1 Left cable management bracket	7 Right cable management bracket
2 I/O units (IOUs)	8 PCIe hot-pluggable carrier for low-profile PCIe cards
3 Service processor proxies	9 Express Module SAS (EMS) for disks
4 CPU memory units (CMUs)	10 Hard disk drives
5 Leveling feet	11 Second operator panel
6 AC input filters (3+3)	12 Mounting brackets

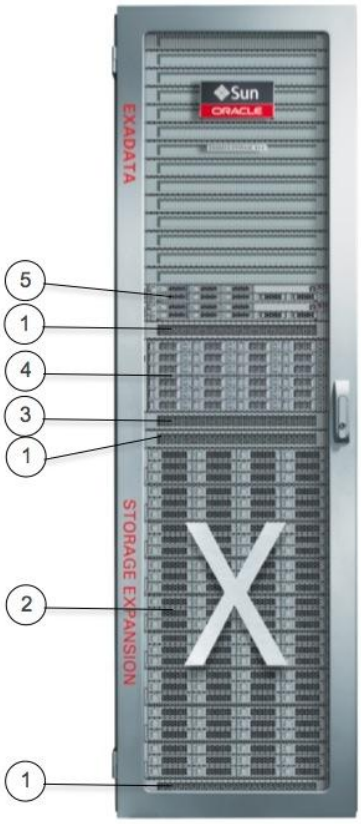

Figure 14. Oracle SuperCluster M6-32 storage rack

TABLE 6. MAPPING OF ORACLE SUPERCLUSTER M6-32 COMPONENTS

1	InfiniBand switch
2	Oracle Exadata Storage Server instances
3	GbE management switch
4	Oracle ZFS Storage ZS3-ES appliance: disk shelf

TABLE 7. COMPONENT CONFIGURATIONS

Component	Configuration
SPARC M6-32 Server	» *Minimum of eight SPARC M6 processors, each with 12 cores (eight threads per core), operating at 3.6 GHz with 48 MB of L3 cache. Maximum of 32 SPARC M6 processors.* » *512 GB (16 GB DIMMs) or 1 TB (32 GB DIMMs) of memory per SPARC M6 processor* » *8 x 1.2 TB10,000 RPM internal SAS disk drives per IOU* » *Four or eight base I/O SAS controller cards (with dual 10 GbE ports) per IOU* » *GbE management card per IOU* » *Redundant Oracle ILOM service processors* » *Redundant clock boards* » *Redundant scalability switches*
Oracle ZFS Storage ZS3-ES Appliance	Two controllers, each containing the following: » *2 x Intel® Xeon® processor E5-2658 (8-core, 2.1 GHz CPU)* » *16 x 16 GB DDR3 1,600 MHz DIMMS (256 GB total)* » *1 x two-port InfiniBand HCA* » *2 x 900 GB SATA disks* » *2 x 1.6 TB read-optimized SSDs* » *1 x SAS HBA* » *2 x dual-port 10 GbE cards* » *1 x cluster card* One disk shelf containing the following: » *20 x 8 TB SAS disks at 7,200 rpm* » *4 x 200 GB write-optimized SSDs*
Nine Instances of Oracle Exadata Storage Server	High capacity option: » 2 x ten-core Intel® Xeon® processor E5-2630 v4 2.2 GHz » 128 GB of memory (8x16 GB 2,133 MHz DIMM) » Disk controller HBA with 1GB cache » 12 x 8 TB 7,200 RPM SAS HDD and 4 x 3.2 TB NVMe PCIe flash drives » Dual port InfiniBand QDR (40 Gb/sec) HCA (PCIe 3.0) Extreme flash option: » 2 x ten-core Intel® Xeon® processor E5-2630 v4 2.2 GHz » 128 GB of memory (8x16 GB 2,133 MHz DIMM) » Disk controller HBA with 1GB cache » 12 x 8 TB 7,200 RPM SAS HDD and 4 x 3.2 TB NVMe PCIe flash drives » Dual port InfiniBand QDR (40 Gb/sec) HCA (PCIe 3.0) »
Three Instances of Sun Datacenter InfiniBand Switch 36	Each 1U switch provides 36 QSFP InfiniBand ports.
Ethernet Management Switch	This 1U switch (currently a Cisco 4948 switch) provides 48 Ethernet ports, and each port has a wire speed of 10/100/1000 BASE-T.

Configuration Notes

For both Oracle SuperCluster M7 and Oracle SuperCluster M6-32, the card slots are assigned to minimize slot, bus, and controller bottlenecks. The only optional component for the SPARC M7 servers are the Fibre Channel (FC) cards—a maximum of three FC HBAs per SPARC M7 compute node. For SPARC M6-32 servers, Fibre Channel (FC) cards can be placed anywhere there is an open PCIe slot in a domain.

The SPARC server configurations are designed to address a wide range of database and application services. For both Oracle SuperCluster M7 and Oracle SuperCluster M6-32, configurations always require a minimum of two

compute nodes to ensure application and database high availability can be achieved thorugh either Oracle Solaris Cluster or Oracle RAC. Base configurations support either two or four compute nodes.

Multiple database and application domains of each type (application and database) are supported in a single compute node. Oracle Solaris Zones is supported in the database domain. This allows multiple Oracle RAC and/or single-instance databases to be supported per database domain.

Multiple database and application domains of each type (application and database) are supported in a single compute node. Oracle Solaris Zones is supported in the database domain. This allows multiple Oracle RAC and/or single-instance databases to be supported per database domain.

For customers already using an Oracle Exadata and are placing Oracle SuperCluster in the same data center as Oracle Exadata Database Machine, the Oracle Exadata Storage Server instances in Oracle SuperCluster can be managed from the existing Oracle Exadata Database Machine. Oracle Exadata Database Machine connects to Oracle Exadata Storage Server instances over a shared InfiniBand fabric.

InfiniBand Connections

There are specific cabling guidelines for Oracle SuperCluster. Accuracy is ensured since all the components are assembled and cabled up at the factory before shipping the system to its final destination. All InfiniBand connectivity is internal to Oracle SuperCluster. If up to 17 Oracle SuperCluster racks or Exadata Storage Expansion Racks are connected, then the InfiniBand fabric is extended from Oracle SuperCluster to those other racks.

10 GbE Connections

Just as with the InfiniBand cables, there are specific 10 GbE cabling guidelines for Oracle SuperCluster. Cabling is between specific ports on the switches and the servers. 10 GbE connectivity is the only supported interface through which clients can access Oracle SuperCluster.

Oracle Database 12*c* and Oracle Database 11*g* Release 2

The database domain on Oracle SuperCluster provides an optimal solution for all database workloads, ranging from scan-intensive data warehouse applications to highly concurrent online transaction processing (OLTP) applications.

With its combination of smart Exadata Storage Server Software from Oracle, complete and intelligent Oracle Database software, and the latest industry-standard hardware components, the database domain on the Oracle SuperCluster system is designed to deliver extreme performance in a highly available, highly secure environment. Oracle provides unique clustering and workload management capabilities so that the database domain is well suited for consolidating multiple databases into a single grid. Delivered as a complete preoptimized and preconfigured package of software, servers, and storage, the database domain can be implemented quickly and can support large-scale business applications without time-consuming configuration and tuning.

The database domain on the Oracle SuperCluster system does not include any Oracle software licenses. Appropriate licensing of the following software is required when the Oracle SuperCluster system is used as a database server:

» Oracle Database

» Exadata Storage Server Software for Oracle Exadata Storage Server only

In addition, Oracle recommends that the following software be licensed:

» Oracle Real Application Clusters (Oracle RAC)
» Oracle Partitioning

Conclusion

Oracle SuperCluster M7 and Oracle SuperCluster M6-32 are designed to help IT organizations consolidate multiple workloads in an environment that has been optimized for performance and availability. By incorporating Oracle SuperCluster into their IT infrastructures, IT organizations can expect to:

» Meet SLAs consistently due to the performance and availability characteristics of Oracle SuperCluster.
» Reduce IT management costs by managing a consolidated infrastructure that comes with an integrated management environment.
» Deploy applications and databases rapidly without spending days or weeks optimizing and testing the hardware and software infrastructure.

Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

ORACLE

CONNECT WITH US

blogs.oracle.com/oracle
facebook.com/oracle
twitter.com/oracle
oracle.com

Integrated Cloud Applications & Platform Services

A Technical Overview of Oracle SuperCluster
April 2016
Author: Gary Combs