An Oracle Technical White Paper
April 2012

Sun Blade 6000 I/O and Management Architecture

ORACLE

Introduction

Blade systems play a key role in allowing organizations to derive maximum benefit from their infrastructure, even as their needs change. Many of today's blade systems, however, have limitations that restrict customer choice and flexibility in I/O, networking, or management. Sun Blade 6000 modular systems are designed to work with open and multivendor industry standards without dictating components, topologies, or management scenarios. This enables organizations to acquire and configure systems that closely match their application environments while avoiding vendor lock-in.

Flexible and Expandable Sun Blade 6000 Modular Systems

Sun Blade 6000 modular systems support flexible and expandable configurations, built from a range of standard hot-plug, hot-swap modules, including:

- SPARC T4-1B and T3-1B server modules, Sun Blade X6270 M2, X6270 M3, and X6275 M2 server modules, or Sun Blade Storage Module M2 in any combination up to 10 modules per chassis
- Blade-dedicated PCIe ExpressModules (EM), which support industry-standard PCI Express interfaces and provide each blade with its own unique I/O configuration (just like rack servers)
- Network Express Modules (NEMs), providing access and an aggregated "chassis-wide" common I/O interface to all of the server modules in the Sun Blade 6000 chassis
- Sun Blade Storage Module M2 that expands high-performance, chassis-integrated storage capacity beyond individual server module capacities
- Integral chassis monitoring module (CMM) for transparent management access to server and storage modules
- Hot-swap (N+N) power supply modules and hot-swap redundant (N+1) fan modules

With common system components and a wide choice of compute and I/O modules, organizations can scale capacity of both processing power and I/O throughput with fine or coarse granularity. Sun Blade 6000 modular systems also give organizations a choice of operating systems and virtualization environments. They are optimized to runOracle Solaris, Oracle Linux, and Oracle VM, and are also certified to run Red Hat Enterprise Linux, SUSE Linux Enterprise Server, Windows Server and VMware.

Focus of This Paper

This paper provides an in-depth review of the architecture of the I/O and management components of Oracle's Sun Blade 6000 modular systems. It covers PCI Express ExpressModules (EMs), Network Express Modules (NEMs), Sun Blade Storage Module M2, and the chassis monitoring module (CMM).

This paper is part of a series of white papers about the architecture of Sun Blade 6000 modular systems. The other two white papers in the series include:

- *Oracle's Sun Blade 6000 Modular Systems* — An overview of the Sun Blade 6000 modular systems architecture.
- *Sun Blade 6000 Server Module Architecture* — An in-depth review of the architecture of the SPARC T4-1B, SPARC T3-1B, Sun Blade X6270 M2, X6270 M3, and X6275 M2 server modules.

Complete Separation between CPU, I/O, and Storage Modules

The Sun Blade 6000 modular system design avoids compromises because it provides complete separation between server modules and I/O modules. The following types of I/O modules are supported:

- Up to two industry-standard PCIe EMs can be dedicated to each individual server module.
- Up to two NEMs provide bulk or consolidated "chassis-wide" common I/O for all of the server modules installed in the system.
- Sun Blade Storage Module M2 provides SAS-2 zone-able storage that can be flexibly allocated amongst the server modules in a chassis when connected via appropriate NEMs to expand storage capacity beyond the individual server module capacities.

Through this flexible approach, each server module can be configured with different I/O options depending on the applications hosted. I/O modules are hot-plug and hot-swap capable, and organizations can choose from Sun-branded or third party adapters for networking, storage, clustering, and other I/O requirements.

Leading I/O Throughput

Sun Blade 6000 server modules provide extensive I/O capabilities and a wealth of I/O options, allowing modular servers to be used for applications that require significant I/O throughput:

- Up to 302 Gb/sec of I/O throughput is provided per server module (when using the latest blade server modules). I/O is delivered through 32 lanes of PCIe 2.0 I/O, as well as multiple Gigabit Ethernet (1GbE) and SAS-2 links. Each server module delivers its I/O to the passive midplane and the I/O devices connected to it in the rear of the Sun Blade 6000 chassis.
- 2.5-inch SAS-2 or SATA2 disk drives and SSDs (solid state drives) are supported on some server modules, while others provide Flash Module or USB support.
- Two hot-plug PCIe (EM slots are dedicated to each server module, 20 per Sun Blade 6000 chassis for granular (per blade) I/O configuration support.
- NEMs provide bulk, common I/O across multiple server modules and aggregate I/O functions. Up to two NEMs are supported per Sun Blade 6000 chassis. Each NEM slot features a PCIe x8 or high

bandwidth XAUI connection, a Gigabit Ethernet connection, and two SAS-2 link connections to each server module.

Table 1 shows an overview of the I/O modules and total bandwidth that can be configured on the available Sun Blade 6000 server modules.

TABLE 1. MIDPLANE THROUGHPUT FOR SUN BLADE 6000 SERVER MODULES

LINKS	PCI EXPRESS LINKS TO EMS	PCI EXPRESS LINKS TO NEMS	GIGABIT ETHERNET LINKS	OTHER LINKS	TOTAL BANDWIDTH (Gb/SEC)
SPARC T4-1B Server Module	2 x 8 PCIe 2.0 links @ 64 Gb/sec	2 NEM and 2 XAUI, or 4 NEM links 4 x PCIe 2.0 links @ 64 Gb/sec	2 @ 1 Gb/sec	4 SAS-2 links @ 6 Gb/sec	302 Gb/sec
SPARC T3-1B Server Module	2 x 8 PCIe 2.0 links @ 64 Gb/sec	2 NEM and 2 XAUI, or 4 NEM links 4 x PCIe 2.0 links @ 64 Gb/sec	2 @ 1 Gb/sec	4 SAS-2 links @ 6 Gb/sec	302 Gb/sec
Sun Blade X6270 M2 Server Module	2 x8 PCIe 2.0 links, @ 64 Gb/sec	2 x8 PCIe 2.0 links, @ 64 Gb/sec	2 @ 1 Gb/sec	4 SAS-2 links, @ 6 Gb/sec	282 Gb/sec
Sun Blade X6270 M3 Server Module	2 x8 PCIe 2.0 links, @ 64 Gb/sec	2 x8 PCIe 2.0 links, @ 64 Gb/sec	2 @ 1 Gb/sec	4 SAS-2 links, @ 6 Gb/sec	282 Gb/sec
Sun Blade X6275 M2 GbE Server Module	2 x8 PCIe 2.0 links, @ 64 Gb/sec	None	2 @ 1 Gb/sec	None	65 Gb/sec (per node) 130 Gb/sec (per server module)
Sun Blade X6275 M2 10GbE Server Module	2 x8 PCIe 2.0 links, @ 64 Gb/sec	None	2 @ 10 Gb/sec	None	74 Gb/sec (per node) 148 Gb/sec (per server module)

PCIe ExpressModules

Industry-standard I/O, long a staple of rack-mounted and vertically scalable servers, has been elusive in legacy blade platforms. Unfortunately, the lack of industry-standard I/O has meant that customers often paid more for fewer (proprietary) options, and were ultimately limited by a single vendor's innovation. Unlike legacy blade platforms, Sun Blade 6000 modular systems utilize PCIe EMs, which are a PCI SIG industry standard form factor. This approach allows a wealth of expansion module options from multiple expansion module vendors, and avoids a single-vendor lock on innovation.

The passive midplane implements connectivity between the EMs and the server modules, and physically assigns pairs of EMs to individual server modules. As shown in Figure 1, EMs 0 and 1 (from right to left) are connected to server module 0, EMs 2 and 3 are connected to server module 1, EMs 4 and 5 are connected to server module 3, and so on. Each EM is supplied with an x8 PCIe 2.0 link back to its associated server module, providing up to 64 Gb/sec of I/O throughput. EMs are hot-plug capable according to the standard defined by the PCI SIG, and fully customer replaceable without opening either the chassis or removing the server module.

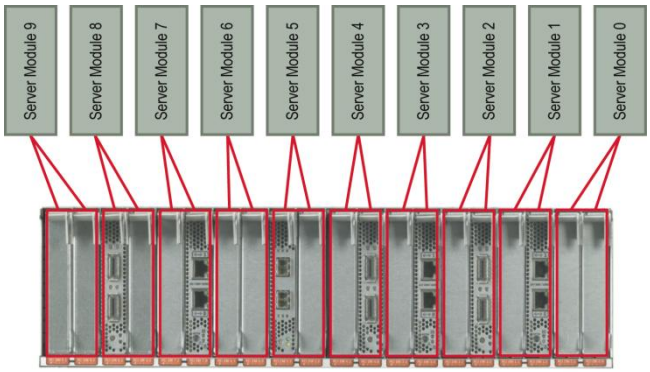

Figure 1. A pair of 8-lane (x8) PCI Express slots allow up to two PCI Express ExpressModules per server module in the Sun Blade 6000 chassis.

With the industry-standard PCIe ExpressModule form factor (Figure 2), EMs are available for multiple types of connectivity, including:

- Gigabit Ethernet
- Combination dual-port GbE, dual-port Fibre Channel
- 10 Gigabit Ethernet (10GbE)
- Fibre Channel Over Ethernet (FCoE)
- Fibre Channel
- InfiniBand
- SAS

For a list of the latest available PCIe ExpressModules, please see: *http://wikis.sun.com/display/SystemsComm/Sun+Blade+Systems+Products*.

Figure 2. Several PCI Express ExpressModules are available for the Sun Blade 6000 chassis.

Network Express Modules (NEMs)

Providing a wide array of I/O access to individual blade servers has always been challenging. Many legacy blade platforms were restrictive in their available options, and many of the options dictated topology and management choices. As a result, data centers often found legacy blade server platforms difficult to integrate into their existing networks, or were resistant to admitting new switch hardware into their chosen network fabrics.

Sun Blade 6000 modular systems address this problem through a specific NEM form factor that provides configurable network I/O for all of the server modules in the system. Connecting to all of the installed server modules through the passive midplane, NEMs represent a space-efficient mechanism for deploying high-density configurable I/O, while providing bulk and consolidated I/O options for the entire chassis.

A selection of NEMs are available for configuration with the Sun Blade 6000 modular systems, providing pass-through access to the Gigabit Ethernet interfaces located on the server modules as well as other I/O protocol capabilities.

Fabric Expansion Modules (FEMs)

A variety of FEMs are offered to provide access to interfaces on Sun Blade server modules. In some cases these FEMs simply convey access to interfaces that reside either on the processors or the server modules themselves. FEMs can also provide access to additional functionality such as alternative I/O via 10GbE networking complementing NEM functionality and connectivity. Through the flexibility of FEMs, different network and I/O fabric interfaces can be exposed to the passive midplane, and on to available Network Express Modules.

As an example of FEM connectivity, the Sun Blade PCIe Pass-Through FEM inserts directly into the FEM socket on the SPARC T4-1B, SPARC T3-1B or Sun Blade X6270 M3 server module (Figure 3). The FEM connects the two x8 PCIe links from the on-board PCIe switches to the passive midplane. Available Network Express Modules can then expose these ports as PCIe interfaces.

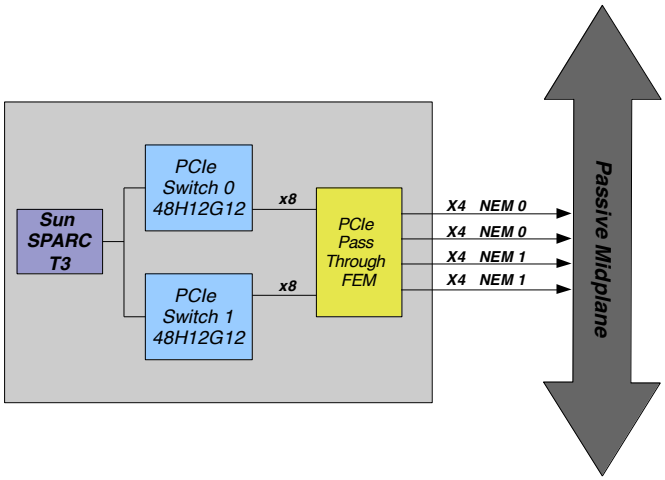

Figure 3. The Sun Blade PCIe Pass-Through Fabric Expansion Module connects 2 x8 PCIe links to NEMs via the passive midplane.

Network Express Modules available as of this writing for the Sun Blade 6000 chassis are described in the sections that follow, along with FEMs required for specific server modules. For the latest Network Express Module information, please refer to *Oracle's Blade Systems Wiki*.

Sun Blade 6000 Chassis NEMs

A variety of NEMs are provided for the Sun Blade 6000 chassis, offering pass-through access to 1GbE and 10GbE interfaces on the various server modules as well as access to on-board SAS links for storage expansion. As of this writing, available Sun Blade 6000 NEMs include:

- Sun Blade 6000 GbE 10-port pass thru NEM
- Sun Blade 6000 Virtualized Multi-Fabric 10GbE M2 NEM
- Sun Blade 6000 Ethernet Switched NEM 24p 10GbE
- Sun Blade 6000 Virtualized 40 GbE NEM
- To access advanced networking functionality, some NEMS require that Fabric Expansion Modules (FEMs) be installed in the server modules. Part numbers for FEMs for connecting the various server modules to the Sun Blade 6000 NEMs are listed in Table 2.

To access the SAS-2 storage fabric zoning capabilities RAID Expansion Modules (REMs) may also be required as shown in Table 3. For the latest NEM and FEM information, please refer to *Oracle's Blade Systems Wiki*.

TABLE 2. FABRIC EXPANSION MODULE REQUIREMENTS

SERVER MODULE	SUN BLADE 6000 ETHERNET SWITCHED NEM 24P 10GbE	SUN BLADE 6000 VIRTUALIZED MULTI-FABRIC 10GbE M2 NEM	SUN BLADE 6000 VIRTUALIZED 40 GbE NEM
SPARC T4-1B	X4871A-Z	7100633	7100633
SPARC T3-1B	X4871A-Z, X5735A	X4263A-N	7100633
Sun Blade X6270 M2 and X6270 M3 Server Modules	X4871A-Z	X4263A-N	7100633

Sun Blade 6000 chassis RAID Expansion module (REM) SAS connectivity using NEMs is shown in Table 3.

TABLE 3. RAID EXPANSION MODULE REQUIREMENTS

SERVER MODULE	SUN BLADE 6000 ETHERNET SWITCHED NEM 24P 10GbE	SUN BLADE 6000 VIRTUALIZED MULTI-FABRIC 10GbE M2 NEM	SUN BLADE 6000 VIRTUALIZED 40 GbE NEM
SPARC T4-1B Server Module	SGX-SAS6-REM-Z	SGX-SAS6-REM-Z	SGX-SAS6-REM-Z
SPARC T3-1B Server Module	SGX-SAS6-REM-Z	SGX-SAS6-REM-Z	SGX-SAS6-REM-Z
Sun Blade X6270 M2 and X6270 M3 Server Modules	SGX-SAS6-REM-Z	SGX-SAS6-R-REM-Z SGX-SAS6-REM-Z	SGX-SAS6-REM-Z

Sun Blade 6000 GbE 10-port pass thru NEM

For applications that don't require 10GbE networking or extended blade storage functionality, the Sun Blade 6000 GbE 10–port Pass Thru NEM offers an effective solution. A single Sun Blade 6000 GbE 10-port Pass Thru NEM (Figure 4) provides ten 1GbE pass-through ports that provide access to one of two 1GbE interfaces on all server modules. In order to support both GbE links that are integrated in the respective server modules, two of the 10-port Pass Thru NEMs need to be installed.

Figure 4. The Sun Blade 6000 GbE 10-port Pass Thru NEM provides ten Gigabit Ethernet pass-through ports.

Sun Blade 6000 Virtualized 40 GbE NEM

The Sun Blade 6000 Virtualized 40 GbE NEM delivers virtualized 40 GbE or 10GbE access with near zero management. Designed to seamlessly integrate into the Sun Blade 6000 chassis, Oracle's Sun Blade 6000 Virtualized 40 GbE Network Express Module is industry's first fully-virtualized 40GbE network aggregation for all ten blade server modules in the Sun Blade 6000 chassis. It is an ideal network interface for connecting multiple virtualized blades that require scalable I/O throughput for workloads such as Web servers, application servers and database servers.

The Sun Blade 6000 Virtualized 40 GbE NEM is a multifunction PCIe 2.0 connectivity module that combines virtualized 40GbE or 10GbE network connectivity, support for SAS-2 storage connectivity to storage modules in a single chassis, and 1GbE pass-through. This NEM is a unique platform for simplifying datacenter networks without adding the extra cost of switches to manage. It utilizes an Oracle-designed 40GbE NIC ASIC that virtualizes the 40GbE or 2x10GbE network connectivity across ten server modules in the Sun Blade 6000 chassis, thus simplifying the networks by reducing cables by 10:1—and without adding the extra switching layer. Figure 5 illustrates the hardware architecture of the NEM.

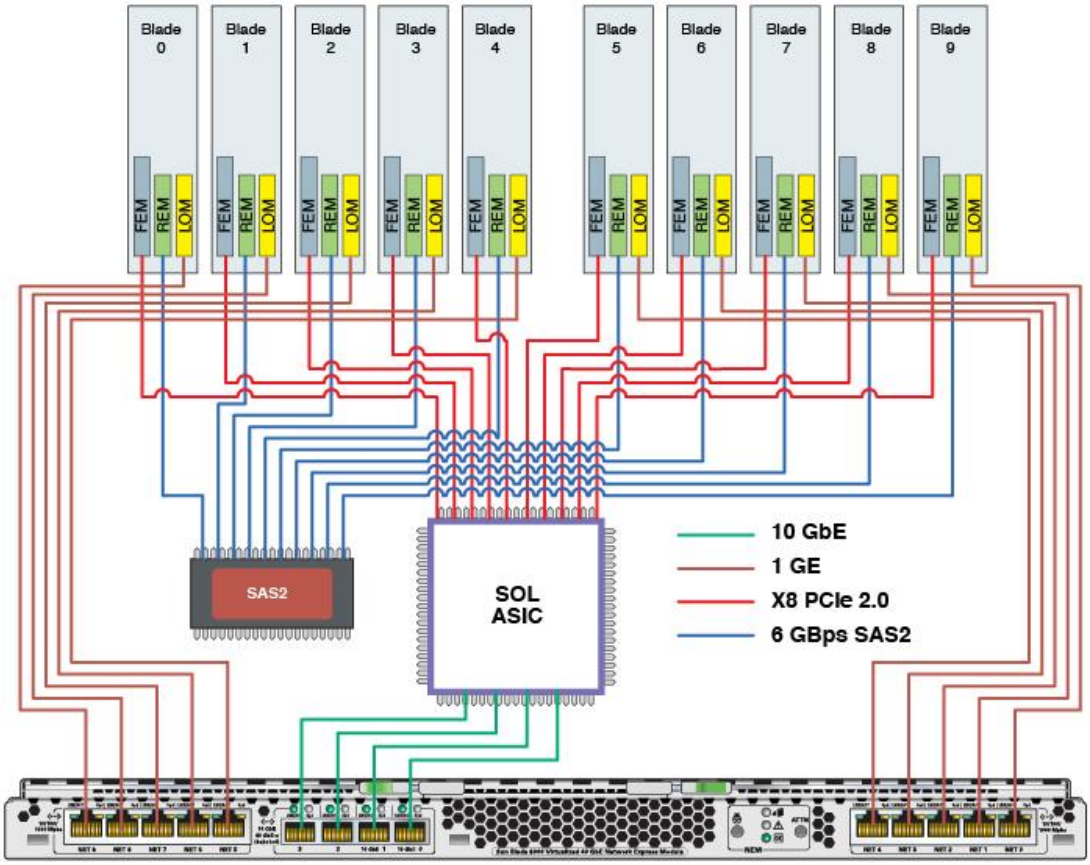

Figure 5. The Sun Blade 6000 Virtualized 40 GbE NEM provides aggregation and optional redundancy at the chassis level.

The Sun Blade 6000 Virtualized 40 GbE NEM offers configurable options, making it ideally suited for different types of workloads ranging from applications just needing high intra-blade communications to those needing higher virtualized bandwidth. Examples of some of these workloads include those that move smaller amounts of data and require high-speed network for lower latency such as Web servers and application servers, as well as workloads that move a lot of data intermittently such as enterprise applications and databases. By default, the blades share the available bandwidth in equal amount and are in privacy mode so that unicast traffic from one blade is not seen by other blades. More deterministic bandwidth can be configured for any or all blades in increments of 1% of the total bandwidth. Ingress filtering and VLANs supported by the NEM provide security options for both the unicast and broadcast traffic from the blades.

The Sun Blade 6000 Virtualized 40 GbE NEM supports connection to external devices either through small form-factor pluggable (SFP)+ ports, or 10/100/1000 twisted-pair Ethernet (TPE) ports. Internal Sun Blade Storage Module M2 components connect through the internal SAS-2 expanders that connect the server modules in the Sun Blade 6000 chassis with the storage modules in the same chassis. Up to 2 SFP+ ports can be used as 2x10GbE. A 40GbE link is provided when 4 SFP+ ports are used (with QSFP to 4 SFP+ splitter cables).

The Sun Blade 6000 Virtualized 40 GbE NEM form factor provides a method of deploying bulk remote I/O that allows tool-less installation/removal and packs more performance and functionality in a smaller space, while delivering higher network throughput. It makes efficient use of data center real estate by reducing the number of cables.

This NEM is easy to install and manage. The flexible architecture of the Sun Blade 6000 modular system is based entirely on the hot-pluggable components for I/O, processing, system management, and chassis infrastructure. All critical components, including the NEM modules, are hot-swappable and redundant, providing enterprise-class reliability, availability and serviceability (RAS) features. The Sun Blade 6000 Virtualized 40 GbE NEM helps to boost data center efficiency and uptime and lower total cost of ownership (TCO).

Pass-thru PCIe 2.0 Fabric Expansion Module (Marketing Part # 7100633) is required on the server blade for its connectivity to the SFP+ ports on the NEM. Raid Expansion Module (Marketing Part # SGX-SAS6-REM-Z or SGX-SAS6-R-REM-Z) is required on the blade server for connectivity to the Sun Blade Storage Module M2 through the NEM.

Sun Blade 6000 Virtualized Multi-Fabric 10GbE M2 NEM

Many organizations are grappling with massive levels of cabling complexity, and are looking for ways to consolidate their networking infrastructure along with their computing infrastructure. The Sun Blade 6000 Virtualized Multi-Fabric 10GbE NEM addresses this need by providing cable aggregation at the chassis level. The NEM combines a next generation SAS-2 expander and passive Gigabit Ethernet interfaces to each server module with dual ASICs for virtualized I/O. In addition to the SAS-2 fabric connectivity and pass-through GbE interfaces, two 10GbE SFP+ optical interfaces are provided on the back panel as shown in Figure 6.

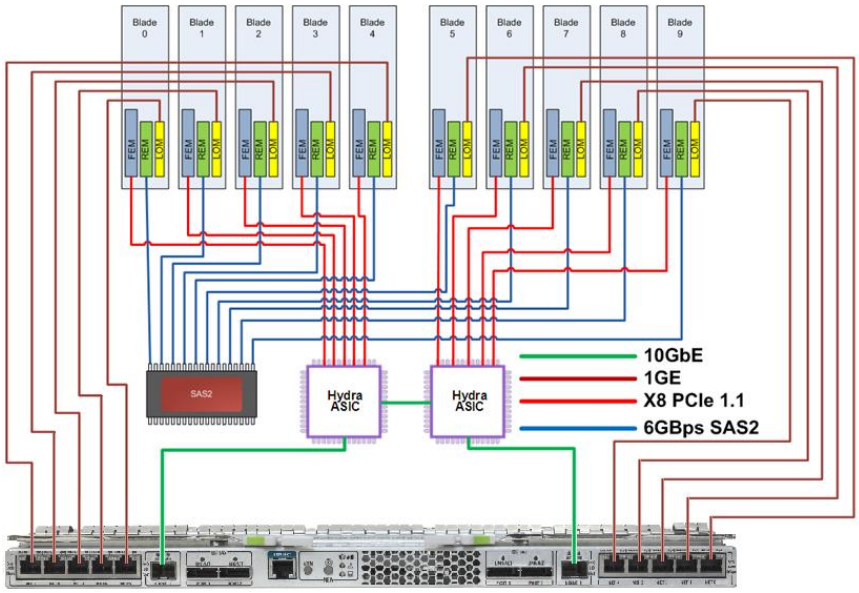

Figure 6. The Sun Blade 6000 Virtualized Multi-Fabric 10GbE M2 NEM provides aggregation and optional redundancy at the chassis level.

The NEM features 10 virtual NICs that connect to each server module via a PCIe connection, giving each server module the impression that it is connected to its own 10 Gb link. Each Server module is also provided with dual SAS-2 disk ports from an integrated SAS-2 fabric expander to allow very flexible SAS storage zoning within the Sun Blade 6000 chassis. The NEM's external SAS-2 ports are not supported at this time.

The virtual NICs are connected via a virtual Ethernet segment that is internal to the NEM. The interface appears as a 10GbE interface to the server module, and installed server modules share the available bandwidth, up to 10 Gb/sec for each external interface. Server modules on the same Ethernet segment share a single physical 10GbE uplink port, and can communicate directly via an internal Loopback path.

The Sun Blade 6000 Virtualized Multi-Fabric 10GbE M2 NEM provides near-zero administration for configuration and maintenance. MAC addresses remain with the server module slot, so that individual server modules can be swapped out without affecting addressing or configuration. In addition, the NEM is configurable in one of three modes, via the Intra ASIC Link (IAL):

- If a single 10 Gb SPF+ cable is inserted, the NEM will be in "Connectivity Mode", where all server modules in a Sun Blade 6000 chassis can share the same Ethernet segment across the IAL. This configuration produces more contention for the 10GbE uplink bandwidth, but can be configured for link-level redundancy across the dual interfaces (via Oracle Integrated Lights Out Manager).
- If two 10 Gb SPF+ cables are inserted, the NEM will be automatically configured into "Bandwidth Mode", where the two ASICs operate as separate Ethernet segments, each with interfaces to five server modules exported through the two external 10 Gb connectors. In this configuration, redundancy can be provided by installing two Sun Blade 6000 Virtualized Multi-Fabric 10GbE NEMs.
- Finally for private communications between blades, only within the chassis, a "Private Mode" is provided, enabling the Intra-ASIC Link to allow all blades to communicate on a single Ethernet segment. Private Mode disables both SFP+ ports to ensure that communications are not passed outside of the Sun Blade 6000 chassis and no external communications can affect in-chassis communications. This is an ideal scenario for private, secure inter-blade communication configurations.

Sun Blade 6000 Ethernet Switched NEM 24p 10GbE

Beyond providing aggregation of 10 Gigabit Ethernet, the Sun Blade 6000 Ethernet Switched NEM 24p 10GbE offers an active non-blocking, low-latency 10GbE switch compatible with the Sun Blade 6000 chassis (shown in Figure 7). Based on a 24-port 10GbE switch chip, the NEM is a full featured Layer 2 / Layer 3 edge switch, allowing dense, non-blocking 10GbE fabrics to be constructed in combination with suitable rack and enterprise-level switches. Providing a full switch within a standard single-height NEM form factor allows considerable consolidation and savings. A block-level diagram of the NEM aligned with a photo of the back panel is provided in Figure 7.

The switch is designed to support:

- Sun Blade 6000 server modules via 10GbE connections that attach through the Sun Dual 10GbE PCIe 2.0 Fabric Expansion Module
- Server module connections to storage modules via two SAS-2 links to each server module slot in the chassis through the passive midplane. The NEM's external SAS-2 ports are not supported at this time.

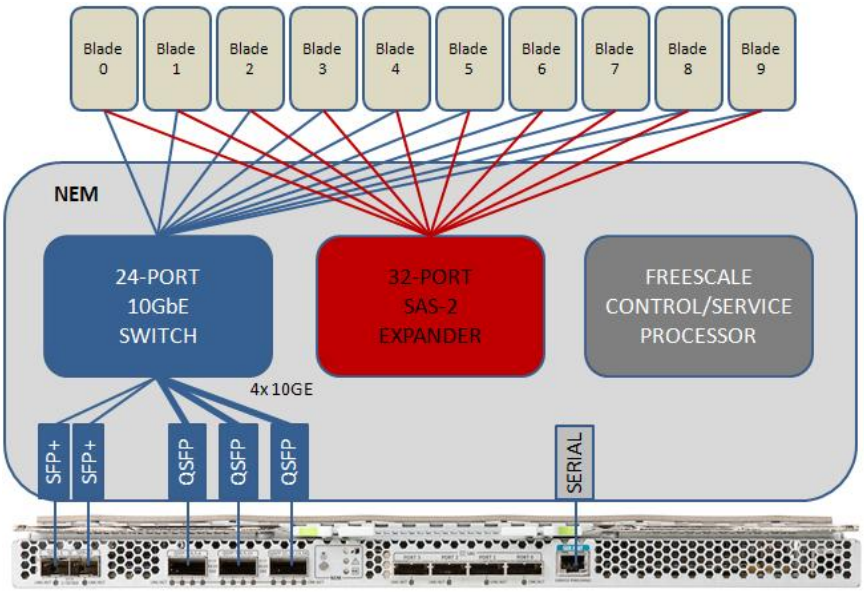

Figure 7. The Sun Blade 6000 Ethernet Switched NEM 24p 10GbE provides a 24-port 10GbE switch for non-blocking 10 Gb access to each server module.

Each Sun Blade 6000 Ethernet Switched NEM 24p 10GbE delivers a 10GbE connection to each server module installed in the chassis. Redundant 10GbE connections to each server module can be provided by installing two NEMS into the Sun Blade 6000 chassis.

In order to offer non-blocking throughput, and scale-out connectivity to other chassis or to existing datacenter networks, each NEM provides a total of 14 external 10GbE connections through the back panel:

- Two SFP+ connectors
- Three 4x 10GbE QSFP (quad SFP) connectors

In addition to switch consolidation offered by the NEM, the three 4x 10GbE QSFP connectors can be split for connectivity to 4 SFP+ or used as an aggregated 40GbE interconnect, providing considerable cable consolidation when connecting to a compatible rack or enterprise switch. In addition, the two SFP+ ports can be configured to operate at either 1GbE or 10GbE for cable aggregation and legacy interconnect. An integrated 36-port SAS-2 expander provides two SAS-2 connections to each slot in the chassis, allowing server modules to connect to Sun Blade Storage Module M2 components for in-chassis storage expansion. The external SAS-2 ports are not supported as of this writing.

Consistent with Oracle's modular design principals, the NEM is easy to manage, and offers standard interfaces and network protocols, including:

- Unified chassis management
- A Web browser interface and a standard command-line interface (Oracle ILOM shell)
- Multiple user privileges
- Single sign on
- Oracle ILOM support via the chassis monitoring module
- Environmental monitoring
- An industry standard compatible L2 / L3 network stack
- CLI and command set.

Several server modules utilize the Sun Dual 10GbE PCIe 2.0 FEM to connect to the NEM (see Table 2). The FEM provide dual 10GbE interfaces to the server module, and is pictured in Figure 8.

Figure 8. The Sun Dual 10GbE PCIe 2.0 FEM provides two 10GbE interfaces.

Extending 10GbE server networks beyond a single shelf

The Sun Blade 6000 Ethernet Switched NEM 24p 10GE provides 10GbE connectivity within a single chassis of server modules but can be further linked to other Sun Blade 6000 chassis to significantly increase the size of the server module deployment. Multiple chassis of Sun Blade 6000 Ethernet Switched NEM 24p 10GbE equipped server modules can be interconnected using only a few QSFP-to-QSFP cables to provide high bandwidth low latency, high node count configurations of servers. To extend even further, to very large node counts, many 10GbE equipped chassis of server modules can be connected using the Sun Network 10GbE Switch 72p top-of-rack (ToR) 1RU switch. The Sun 10GbE ToR switch is equipped with 16 (4x10GbE) QSFP ports and 8 (1/10GbE) SFP+ switch ports providing wire rate, low latency, cut through switching on all 72 ports and delivers full, non-blocking bandwidth to each blade or rack server. The Sun ToR has enough density to build a sizable fabric that operates at full bandwidth without the need for spanning tree protocol. The Sun ToR supply of 8 SFP+ uplinks can cost-effectively join 64 servers to a tier 2 switch, obviating the need for complex tier 1 inter-connectivity. Some simple, example network configurations are illustrated in Figure 9.

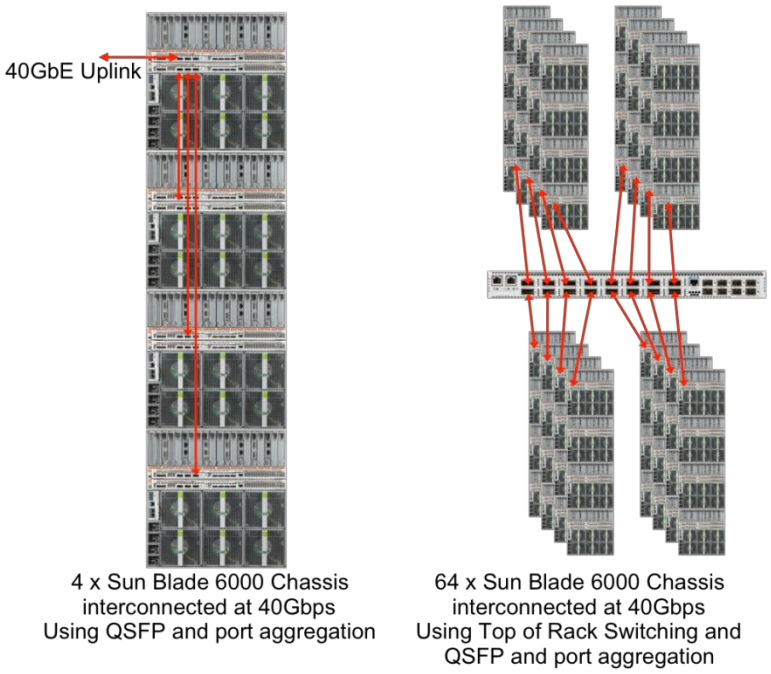

Figure 9. A rack full of Sun Blade 6000 chassis using the 24p NEM switch interconnected using a single QSFP uplink to one NEM, and utilizing a 72p Switch to connect 16 racks of servers from 24p NEMs to a single 72p Switch. Further bandwidth improvement could be achieved through port aggregation configurations.

In-Chassis Storage and Storage Expansion

A number of storage expansion options are available for server modules. Individual server modules have their own internal expansion options, including hard disk drives (HDDs) and flash memory devices, including solid-state drives (SSDs), Sun Flash Modules (FMODs), and USB. The Sun Blade Storage Module M2 allows storage to be expanded beyond the individual server module. Supported storage options for the various server modules are shown in Table 4. Note that an appropriate REM and NEM combination is required in some circumstances.

TABLE 4. SERVER MODULE STORAGE OPTIONS				
SERVER MODULE ARCHITECTURE	**SERVER MODULE**	**HDDS/SSDS**	**SUN FLASH MODULE**	**SUN BLADE STORAGE MODULE M2**
SPARC T4	SPARC T4-1B server module	2	---	Yes (Full zoning capability)
SPARC T3	SPARC T3-1B server module	4	—	Yes (Full zoning capability)
Intel Xeon 5600 Series	Sun Blade X6270 M2 server module	4	—	Yes (Full zoning capability)
Intel Xeon 5600 Series	Sun Blade X6275 M2 server module	0	Yes (1 per node)	—
Intel Xeon Processor E5-2600 Product Family	Sun Blade X6270 M3 server module	4	—	Yes (Full zoning capability)

Server Module Hard Disk Drives

A choice of (optional) hot-swappable 2.5-inch SAS or SATA HDDs are supported with all server modules except for the diskless Sun Blade X6275 server modules which utilize Flash Modules for local storage requirements.

- Serial Attached SCSI (SAS) drives provide high performance and high density. Drives are 10,000 rpm and available in various capacities. These drives provide enterprise-class reliability with up to 1.6 million hours mean time between failures (MTBF).
- Serial ATA (SATA) drives are 7200 rpm and available in various capacities.

Innovative Enterprise Flash Technology

Modern servers are driving throughput levels that can rapidly outpace the capabilities of traditional HDD storage solutions. While many servers can achieve processing capabilities in excess of one million I/O operations per second (IOPS), today's fastest HDDs are only capable of about 300 to 400 IOPS. To match throughput more closely to server performance and to address the challenging demands of data-intensive applications, many datacenters implement large pools of high-speed disk drives, often utilizing only the fastest outer portions of their storage media, wasting capacity. In some cases, a large buffer of expensive DRAM is also deployed so that the application's working set can be stored in memory to reduce latency. Flash technology provides a more economical alternative that can dramatically enhance application I/O performance while also operating with significantly better energy efficiency than conventional HDDs. Recent advances in the flash technology have made enterprise

levels of reliability, life expectancy and manageability available from flash based storage devices, in the form of both Flash Modules and SSDs. These now become an effective and reliable solution for enterprise storage. Flash technology contains no moving parts, avoiding the seek times and rotational latencies inherent with traditional hard disk drive technology and removing the single largest bottleneck to disk drive performance. Oracle Flash technology can offer instant availability of increased performance, power and cost savings, and reliability benefits over rotational media, including:

- 97% lower cost per I/O operation per second (IOPS)
- 99% lower Watts/IOPS
- 66% lower Watts/GB
- High reliability of up to 7x24x3 years with a 100% write duty cycle[1]

Solid-state Drives (SSDs)

Because SSDs offer low latency and are significantly less expensive than an equivalent amount of DRAM storage, they balance cost and performance in a manner that can provide significant value for I/O intensive workloads. SSDs are offered in conventional disk drive form factor (Figure 10).

Figure 10. Solid sate drives (left) provide enterprise flash technology in a standard disk drive form factor. FlashFire technology (right) provides optimized performance in a considerably smaller form-factor.

Sun Flash Technology

Engineered by Oracle, the Sun Flash Module (FMOD) is a storage innovation that combines NAND flash components and a flash memory controller (FMC) to provide an industry-standard SATA device in a compact, highly-efficient form factor. The Sun Blade X6275 M2 server modules can support installation of up to two 24 GB Sun Flash Modules. In Sun Blade X6275 M2 server modules one Sun Flash Module is available to each compute node on the server module. These flash modules are ideal for hosting a boot environment for the compute node, and can also serve to deliver a high-capacity

[1] Statistics for the Sun Flash Module

and high-performance storage option for I/O sensitive applications. Similar to SSDs, Sun Flash Modules are viewed by the operating system and BIOS as standard SATA drives, removing the burden of OS specific drivers and special configurations.

Sun Blade Storage Module M2

The Sun Blade Storage Module M2 (Figure 11) is a high-density storage device for the Sun Blade 6000 chassis, allowing flexible storage expansion beyond the capabilities of individual server modules. Like the server modules, Sun Blade Storage Module M2 devices do not contain fans or power supplies, deriving power, cooling, and management from the chassis. Each Sun Blade Storage Module M2 provides storage capacity of up to eight SAS-2 hard drives. Up to 9 Sun Blade Storage Module M2 components can be installed in a single Sun Blade 6000 chassis.

Each disk slot in the storage module is Dual Ported allowing redundant paths to SAS-2 storage devices that support this feature. All communications with the server modules are via Serial Attached SCSI (SAS-2) links routed through the chassis midplane and supported NEMs. The Sun Blade Storage Module M2 can act as a just-a-bunch-of-disks (JBODs) device attached to one or more server modules. The configuration and mapping of the available storage devices to the server modules is carried out via the chassis management module (CMM) Oracle ILOM interface, which provides an easy-to-use and visual Web-based GUI, or alternatively via a CMM CLI interface.

Figure 11. The Sun Blade Storage Module M2 provides up to eight SAS-2 drives which may be shared with other server modules within the Sun Blade 6000 chassis.

The disk module offers the ability to flexibly provide storage capacity to a server module within the chassis, with up to eight SAS-2 hard drives — providing up to 4.8 TB of disk storage capacity (using 600 GB drives). The bandwidth between the server module and the disk module is well balanced, as

each server and disk module has up to four SAS-2 links via the midplane and NEMs to the storage module. The total bandwidth available through four SAS-2 links is 24 Gb/sec.

As shown in Figure 12, each pair of SAS-2 links connects a NEM slot to one of two 36-port SAS-2 Storage expanders in the Sun Blade Storage Module M2. All disk drives are connected to both storage expanders providing dual paths. Two NEMs must be installed in the chassis in order for all four links to be available to both the server module and the disk module and to take advantage of dual paths. If a single NEM is installed, only two of the SAS links from each server module will be able to communicate and to access the storage module.

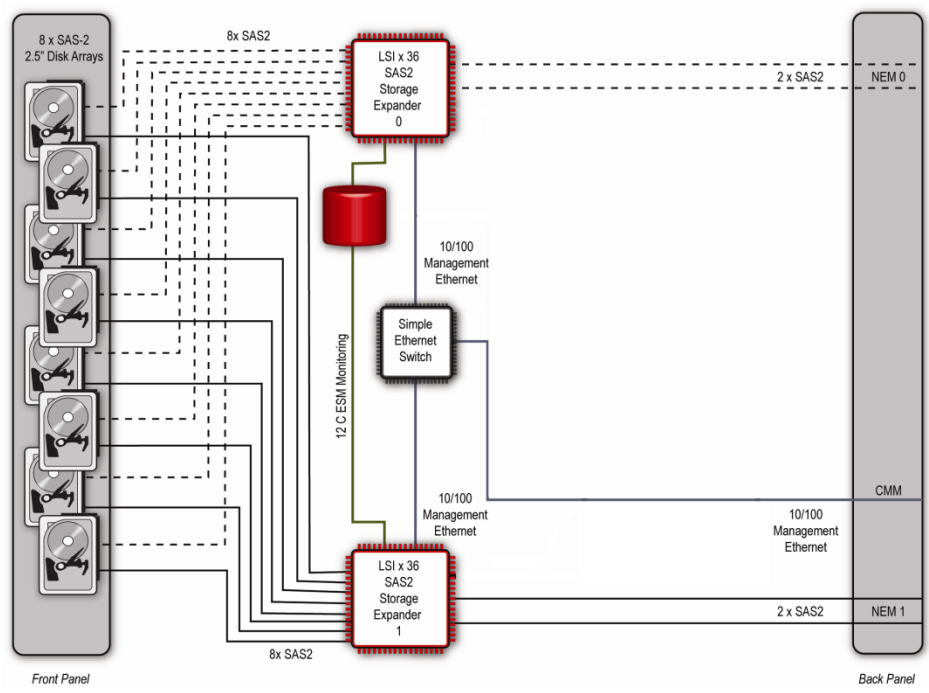

Figure 12. Each Sun Blade Storage Module M2 includes two 36-port SAS-2 storage expanders and each storage expander connects to all of the storage devices as well as to one of the NEM slots, providing dual paths to each storage device if both NEMs are populated.

In order for the server module to communicate to the disk module, it must be equipped with an appropriate RAID expansion module (REM) and the chassis must have at least one SAS-2 storage expander-equipped NEM installed. If the server module supports on-board disk drives, then its REM will provide communication to the on-board hard disk drives in addition to the devices zoned for access from the storage module. This feature makes management of the additional hard disk drives extremely simple, as only one storage controller has to be managed to use and configure all the hard drives visible to the server module, including those in RAID volumes.

A single Sun Blade Storage Module M2 can provide flexible storage to an entire Sun Blade 6000 chassis of server modules via the easy-to-use Web GUI and CLI provided for zoning the storage devices. Options are provided to quickly set up simple configurations such as spreading the available storage devices across the chassis on a per disk, per storage blade, per adjacent disk, or per adjacent storage module basis. These options set up simple zoning that can either be used as provided or further

modified by choosing the "full resource control" setup option to map individual devices to individual blade modules. Some examples of the Oracle ILOM provided Web GUI for zone configuration are shown in Figure 13.

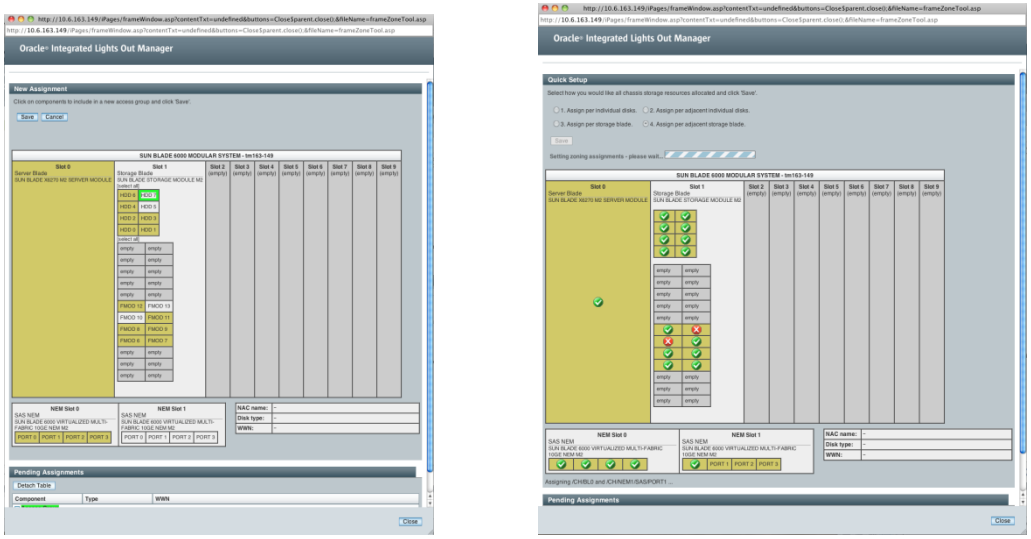

Figure 13. Server modules can be zoned easily through the CMM Zoning GUI.

Full SAS-2 connectivity and zoning capability between server modules and the Sun Blade Storage Module M2 is provided by the following Sun Blade 6000 NEMs:

- Sun Blade 6000 Virtualized Multi-Fabric 10GbE M2 NEM
- Sun Blade 6000 Ethernet Switched NEM 24p 10GbE

Sun Storage 6Gb SAS RAID Expansion Module (REM) HBA

The Sun Storage 6Gb SAS RAID REM is a mezzanine card that can be fitted on the rear of the motherboard for mirroring, or striping of the disk storage devices as well as offering access for simple JBOD arrays. The REM currently provides four SAS-2 (6 Gb/sec) or SATA (3 Gb/sec) links to the four internal disk drives as well as exporting four SAS links to each NEM via the midplane for off-blade storage expansion. Integrated striping can be used in conjunction with integrated mirroring or integrated mirroring enhanced. The RAID configuration is self-sufficient, and does not require intervention from the host CPU or OS. All standard SPARC T4-1B and SPARC T3-1B server modules support one SAS REM and the card is also supported on Sun Blade X6270 M2/M3 server modules.

Transparent and Open Chassis and System Management

Management in legacy blade platforms has typically either been lacking, or administrators have been forced into adopting unique blade or platform-specific management infrastructure. To address this issue, the Sun Blade 6000 modular systems provide a wide range of flexible management options.

Chassis Monitoring Module (CMM)

The chassis monitoring module (CMM) is the primary point of management for all shared chassis components and functions, providing a set of management interfaces. Each server module contains its own service processor, giving it similar remote management capabilities to other Sun servers. Through their respective Oracle ILOM service processors, individual server modules provide IPMI, HTTPs, CLI (SSH), SNMP, and file transfer interfaces that are directly accessible from the Ethernet management port on the CMM. Each server module is assigned an IP address (either manually, or via DHCP) that is used for the management network.

CMM Network Functionality

A single CMM module is provided with each Sun Blade 6000 modular system, and is configured with an individual IP address assigned either statically or dynamically via DHCP. The CMM provides complete monitoring and management functionality for the chassis (or shelf) while providing access to server module management functions. In addition, the CMM supports HTTP and CLI "pass-thru" interfaces that provide transparent access to each server module. The CMM also provides access to each server module via a single serial port through which any of the various LOM interfaces can be configured. The CMM's management functions include:

- Implementation of an IPMI satellite controller, making the chassis environmental sensors visible to the server module's BMC functions
- Direct environmental and inventory management via CLI and IPMI interfaces
- CMM, Oracle ILOM, and NEM firmware management
- Pass-through management of blades using IPMI, SNMP, and HTTP links along with command line interface (CLI) SSH contexts

The management network internal to the CMM joins the local management processor on each server module to the external management network through the passive midplane.

CMM Architecture

A portion of the CMM functions as an unmanaged switch dedicated exclusively to remote management network traffic, letting administrators access the remote management functions of the server modules. The switch in the CMM provides a single network interface to each of the server modules and to each of the NEMs, as well as to the service processor located on the CMM itself. Figure 14 provides an illustration and a block-level diagram of the Sun Blade 6000 CMM.

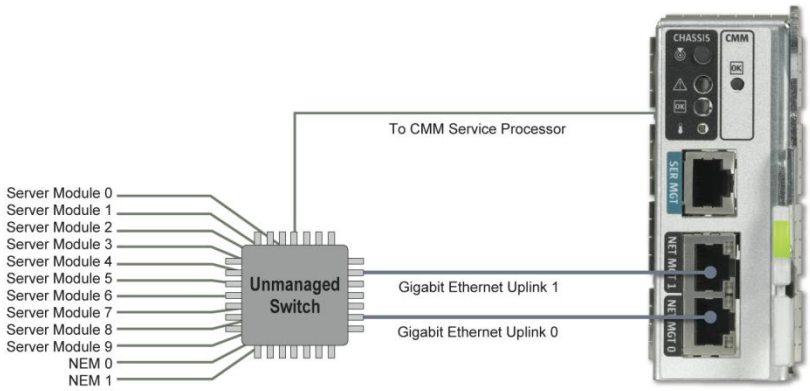

Figure 14. The CMM provides a management network that connects to each server module, the two NEMS, and the CMM itself.

The CMM's functionality provides various management functions, including power control of the chassis as well as hot-plug operations of infrastructure components such as power supply modules, fan modules, server modules, and NEMs. The CMM acts as a conduit to server module LOM configuration, allowing settings such as network addresses and administrative users to be configured or viewed. In a chassis equipped with one or more Sun Blade Storage Module M2 components and an appropriate NEM with SAS-2 fabric expansion capabilities, the CMM is also the controller for the SAS-2 storage zoning configuration. This is done through the WEB GUI provided by the CMM as well as by a CLI command set.

Oracle Enterprise Manager Ops Center

Beyond local and remote management capabilities, data center infrastructure needs to be agile and flexible, allowing not only fast deployment, but also streamlined redeployment of resources as required. Oracle Enterprise Manager Ops Center technology provides an IT infrastructure management platform for integrating and automating management of thousands of heterogeneous systems. To improve lifecycle and change management, Oracle Enterprise Manager Ops Center supports the management of applications and the physical and virtual servers on which they run, including the Sun Blade 6000 modular systems.

Further information about Oracle Enterprise Manager Ops Center is provided in the Sun Blade 6000 modular systems architecture overview white paper and is available on the Web at *http://www.oracle.com/us/products/enterprise-manager/opscenter/*.

Conclusion

Oracle's innovative technology and open-systems approach make modular systems attractive across a broad set of applications and activities — from consolidating infrastructure through virtualization to deploying dynamic enterprise applications or cloud computing. Oracle's Sun Blade 6000 modular system provides the promised advantages of modular architecture while retaining essential flexibility for how technology is deployed and managed.

Oracle's standard- and open-systems-based approach yields choice and avoids compromise — providing a platform that benefits from widespread industry innovation. Industry-standard I/O and multiple storage options provide flexibility and leading throughput for individual server modules. Transparent networking and management means that the Sun Blade 6000 modular systems fit easily into an existing network and management infrastructure.

For More Information

For more information on Sun Blade 6000 modular systems, please visit *http://www.oracle.com/goto/blades*, which also contains links to additional Sun Blade 6000 modular systems white papers. The Web sites listed below also provide more specific references:

- Sun Blade systems OTN page: *http://www.oracle.com/technetwork/server-storage/sun-blade/documentation/index.html*
- Blades power calculator*: http://www.oracle.com/sun-power-calculators/calc/6000chassis*
- Oracle Enterprise Manager Ops Center: *http://www.oracle.com/us/products/enterprise-manager/opscenter/*
- Oracle Solaris operating system*: http://www.oracle.com/technetwork/server-storage/solaris/overview/*

Oracle is committed to developing practices and products that help protect the environment

Sun Blade 6000 I/O and Management Architecture
April 2012 Version 1.4

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

Worldwide Inquiries:
Phone: +1.650.506.7000
Fax: +1.650.506.7200
oracle.com

Hardware and Software, Engineered to Work Together