

# Meet the HP 9000 Superdome

A white paper from Hewlett-Packard Company



Executive summary.....	3
The HP enterprise server product line .....	3
The need for “big iron” .....	4
Introducing the HP 9000 Superdome enterprise servers.....	5
Top-of-the-line performance and capacity .....	5
High-availability features for greater uptime .....	6
Flexibility with partitions.....	6
Investment protection with support for PA-RISC or Intel® Itanium® processors.....	6
Consistent manageability .....	6
HP-UX 11i, the industry-leading 64-bit operating environment.....	6
HP 9000 Superdome hardware overview .....	7
HP 9000 Superdome hardware.....	9
The HP 9000 Superdome cabinet .....	9
HP 9000 Superdome system architecture .....	10
Modular architecture .....	11
The cell board .....	12
Processors .....	14
Cell controller ASIC .....	14
Memory controller ASIC.....	15
The crossbar backplane.....	15
Crossbar ASIC .....	15
Memory and backplane latencies.....	16
The I/O subsystem .....	17
The HP 9000 Superdome: performance and scalability.....	18
System bandwidth.....	18
Scalability.....	18
The HP 9000 Superdome: high availability.....	19
System reliability.....	19
Supportability .....	19
Event Monitoring Service.....	20
Support Management Station (SMS) .....	20
Instant Support Enterprise Edition (ISEE).....	20

Repairability .....	21
N+1 CPUs and cell boards with Instant Capacity .....	21
Hot-swap N+1 fans, power supplies, and backplane DC/DC converters .....	21
Dual power sources .....	21
Online addition and replacement (OLAR) of cards and chassis .....	22
Online addition and replacement of cell boards .....	22
Online addition of hard partitions (nPars) .....	22
HP Serviceguard HA Cluster-in-a-Box .....	22
Uptime Institute certification .....	22
Support services .....	22
The HP 9000 Superdome: investment protection .....	23
Binary compatibility across PA-RISC .....	23
Binary compatibility for the Intel Itanium Processor Family .....	24
Upgrading an existing HP 9000 Superdome system .....	24
The HP 9000 Superdome: server virtualization .....	24
Benefits of HP 9000 hard partitions (nPars) .....	26
Increased system utilization .....	26
Increased flexibility via support of multiple environments .....	27
Increased uptime .....	27
Increased scalability .....	27
The HP 9000 Superdome: ease of management .....	27
System management features .....	27
The HP 9000 Superdome Support Management Station .....	29
SMS installation .....	29
SMS functional capabilities .....	29
The HP 9000 Superdome: HP-UX 11i v1 .....	30
Key features of HP-UX .....	30
HP-UX 11i operating environments .....	30
HP-UX 11i entry-level operating environment .....	31
HP-UX 11i Enterprise operating environment .....	31
HP-UX 11i Mission Critical operating environment .....	31
The rest of the HP 9000 Superdome story .....	32
The Adaptive Enterprise and the HP 9000 Superdome .....	32
Mass storage and the HP 9000 Superdome .....	32
Conclusion .....	32
For more information .....	33

## Executive summary

The HP 9000 Superdome, the proven high-end server in the HP Server product family, is now faster and more scalable. With the release of the PA-8800 and subsequent PA-8900 processors, the HP 9000 Superdome and HP-UX 11i offer twice the performance density of the previous HP 9000 Superdome generations. A great system just got better!

The HP 9000 Superdome delivers outstanding performance, superior scalability, and best-in-class solution deployment—a level of functionality and value well beyond anything previously offered.

The HP 9000 Superdome is designed to deliver the performance, availability, capacity, security, and manageability needed for a variety of applications, from mission-critical to compute-intensive. This white paper will focus on the architecture, system software, and management tools of the new HP 9000 Superdome.

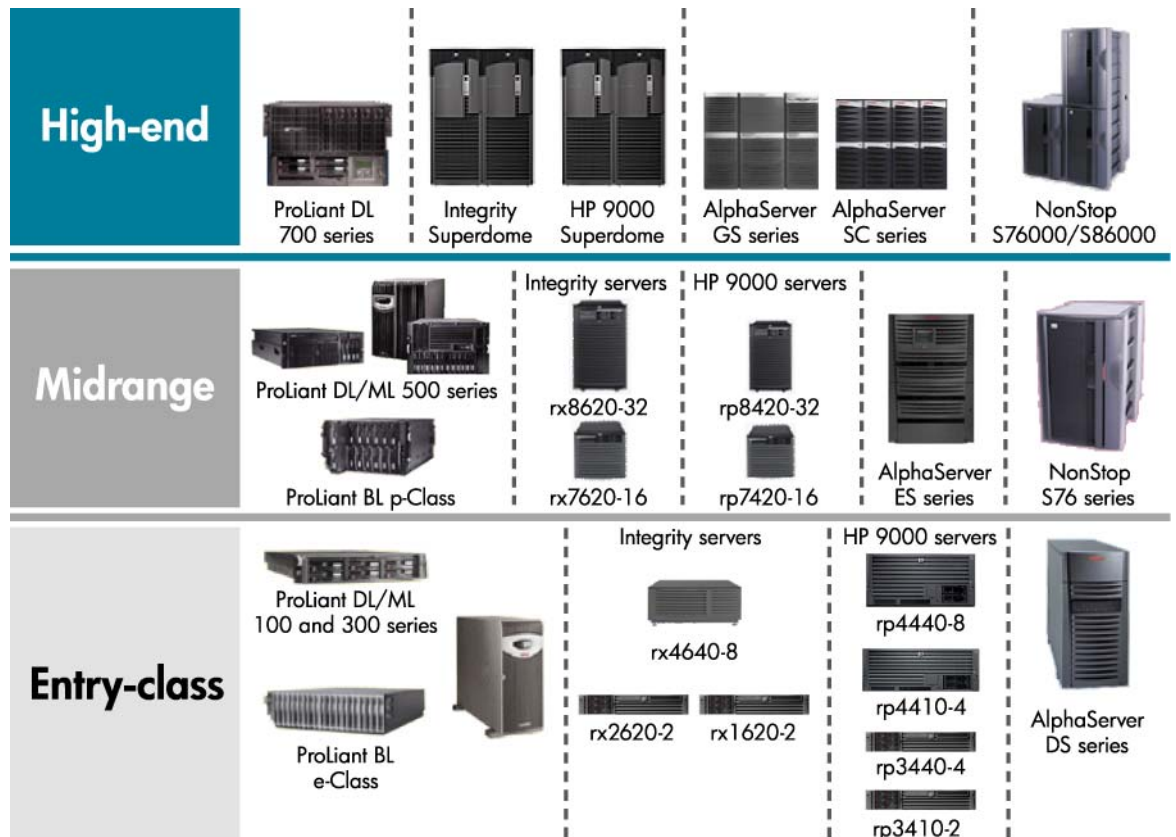
The new HP 9000 Superdome is based on the HP Super-Scalable Processor Chipset sx1000, which supports the PA-8900 and PA-8800 processors, the 1.5 GHz Intel® Itanium® 2 processor and the HP mx2 Dual-Processor Module. The HP sx1000 Chipset provides higher CPU and memory bandwidth plus greater memory capacity. Fast, low-level error correction further enhances overall system availability. With the PA-8900 or PA-8800 processors and the HP sx1000 Chipset for the HP 9000 Superdome, the server can now scale to 128 processors, delivering the raw power you need to perform any compute-intensive task. The PA-8900 and PA-8800 processors can accommodate two CPU cores on a single chip and in the same socket, boosting speed, density, and performance. As a result, the high-end scalability of the HP 9000 Superdome provides customers with more than twice the compute power in the same box. There is now more power in each partition, enabling a much higher level of server consolidation—resulting in both cost savings and a lower total cost of ownership. In addition, the power and scalability of the PA-8900 or PA-8800-based HP 9000 Superdome provide an exceptional ability to handle planned and unexpected growth in high-end compute loads. For existing HP 9000 Superdome customers, new levels of performance can be achieved through in-box upgrades to the new PA-8900 processor, bringing significant IT capability improvements at an incremental cost. An existing HP 9000 Superdome can be upgraded to a PA-8900-powered HP 9000 Superdome by simply removing the PA-8600, PA-8700, PA-8700+, or PA-8800 cell boards, transferring the memory to the new PA-8900-based cell boards, plugging the new PA-8900 cell boards back into the cabinet, updating the software, and then resuming operation.

## The HP enterprise server product line

First, let's look at the overall HP Server product lineup. These servers address the major computing challenges that customers face today. Running the powerful HP-UX operating environment (HP's version of UNIX®), HP 9000 enterprise servers are especially suited to the demands of mission-critical applications, the Internet, online transaction processing (OLTP), system consolidation, large decision support systems, and compute-intensive technical applications.

As shown in Figure 1, the HP Server line is comprehensive, with products ranging from entry-class to high-end. And now, HP continues its strong enterprise market presence by offering upgrades to the HP Integrity Superdome and the PA-8900-based HP 9000 Superdome, providing even more computing power, support for multiple operating systems, and other new capabilities. The PA-8900-based HP 9000 Superdome is positioned as the HP 9000 family's most powerful high-end platform.

Figure 1: The HP Server family



The HP Server product line represents the most scalable range of computing available today. These systems feature the following:

- **Application compatibility and portability** across the product line, permitting users to select the best platform at the appropriate price point without worrying about whether applications will run on it
- **Exceptional performance**, providing reduced time-to-solution through parallelism and increased throughput through multiprocessing in all environments
- **Unmatched scalability**, providing the best price/performance over the entire range of products and protecting customers' investments in hardware and software
- **A consistent programming model**, presenting the same application programming environment regardless of the systems' performance levels; this greatly increases the number of available "off-the-shelf" third-party applications and reduces porting and development costs
- **Leading-edge utility pricing models** (pay per use, capacity on demand, and pay per forecast), allowing customers who run HP-UX to better align their IT costs with revenues as their technology usage fluctuates over time

## The need for "big iron"

Some applications simply need the maximum amount of performance, scalability, and availability. For instance, in the data center, where transaction processing is handled and where the backbone of mission-critical applications runs, you'll want the best servers available to keep those applications running and meeting quality-of-service goals on a 24 x 7 basis. Workloads such as online transaction

processing (OLTP) with database engines, data mining, high-performance technical computing, and customer relationship management (CRM) are best handled by “big iron”—top-of-the-line computers that also combine security, scalability, flexibility, capacity, and manageability. That’s where the HP 9000 Superdome fits in.

## Introducing the HP 9000 Superdome enterprise servers

HP 9000 Superdome is a proven high-end UNIX server that provides customers with superiority in all of the following areas:

- Performance
- Scalability
- Availability
- Flexibility (partitioning options)
- Utility technology and pricing
- Connectivity
- Manageability
- Investment protection

As a high-end platform, HP 9000 Superdome is uniquely able to address customer requirements, even as those requirements change over time. This capability is based on service-level agreement (SLA) features, hard and soft partitions, support of up to 128 processors, and multiple operating environments (OEs). What’s more, when the time is right for them, HP 9000 Superdome customers can upgrade to HP Integrity Superdome with Intel Itanium 2 processors by simply swapping out the cell boards, thus retaining their investment in the cabinet, I/O chassis, backplane, and memory DIMMs.

### Top-of-the-line performance and capacity

The best-choice HP 9000 Superdome greatly extends the capabilities of the HP Server line at the high end. The HP 9000 Superdome offers the following performance-enhancing features:

- High-performance RISC processor (PA-8900)
  - Dual cores to support two CPUs on a single chip, boosting speed, density, and performance
  - State-of-the-art 130nm IC process
  - Separate L1 cache for each CPU
  - Large (64 MB), unified L2 cache with sophisticated cache controller
  - 1.1 GHz core frequency
  - Large on-chip cache (6 MB)
  - Upgradeable to Intel Itanium 2 processors
- Increased memory subsystem performance and capacity
  - Increased peak memory bandwidth to 256 GB/s per 64-way system
  - Doubled memory capacity (1024 GB overall)
  - Higher-density DIMM support (512 MB, 1 GB, and 2 GB DIMMs)
  - Configurations available from 4 to 128 CPUs
- Support of both PCI-X and PCI (legacy HP 9000 Superdome only) I/O
  - 64-bit, 33 MHz (2X), industry-standard PCI
  - 64-bit, 66 MHz (4X), industry-standard PCI
  - 133 MHz, industry-standard PCI-X

- 16 2.0 GB/s I/O channels (32 GB/s per 64-way system)
- 192 PCI slots

## High-availability features for greater uptime

Mission-critical applications are no problem for the HP 9000 servers. This family of servers has the increased system reliability and availability necessary to provide the right level of uptime for the mission-critical environment based on the organization's requirements. It also offers improvements in system quality, resiliency, and fault management, and it provides support for HP Serviceguard.

## Flexibility with partitions

The HP 9000 Superdome supports the HP Partitioning Continuum, providing both hard and soft partitions. In fact, in the future, with a new release of the HP-UX 11i operating system, the HP 9000 Superdome can be configured as one large symmetric multiprocessor (today, the maximum size of a single hard partition is 64 processors). In addition, the HP 9000 Superdome can be configured as several independent hard partitions (also known as nPars). These nPars can provide hardware isolation, full electrical isolation, and complete software isolation. Virtual or "soft" partitions (vPars) can be dynamically resized for the highest degree of flexibility. The combination of hard partition, soft partition, and resource partitioning capabilities enable the HP 9000 Superdome to reach new heights in flexibility, availability, and scalability.

## Investment protection with support for PA-RISC or Intel® Itanium® processors

The HP 9000 Superdome supports PA-RISC processors, with in-the-box upgrades to the HP Integrity Superdome. The HP 9000 Superdome with 750 MHz or 875 MHz PA-8700 or PA-8800 processors can be upgraded in the box to PA-8900 processors. And, when the time is right, the HP 9000 Superdome can easily be upgraded to an HP Integrity Superdome. This makes the HP 9000 Superdome the logical choice—not only for present and future high-end computing needs, but also for investment protection in the years to come.

## Consistent manageability

HP is committed to making the management of the HP 9000 Superdome as simple and cost-effective as possible. HP's proven management features, coupled with tested management software and technologies, are all designed to increase control while reducing administrative overhead. Provided tools, such as HP Systems Insight Manager, Partition Manager, and HP-UX Workload Manager, make the HP 9000 Superdome very easy to manage.

## HP-UX 11i, the industry-leading 64-bit operating environment

HP-UX 11i v1 is specifically focused on the needs of both online and traditional economy concerns. And the HP 9000 Superdome runs this industry-leading operating environment.

HP-UX 11i is a 64-bit operating environment that enjoys from a wide variety of independent software vendors. In fact, there are more than 16,000 supported applications to choose from, including native 64-bit versions of all major databases and leading enterprise resource planning (ERP) applications. The widely supported 64-bit HP-UX operating system is compatible with the full HP Server line.

The HP 9000 Superdome running the PA-RISC chip family supports HP-UX 11i v1. The HP Integrity Superdome with Intel Itanium 2 processors supports HP-UX 11i v2. Plans are in place to integrate support for both PA-RISC and Intel Itanium 2 processors in a future version of HP-UX.

## HP 9000 Superdome hardware overview

The HP 9000 Superdome is available in several different configurations—with capabilities from 4-way multiprocessing all the way to 128-way. Depending on the model, the HP 9000 Superdome comes in one or two cabinets. The following table is a summary of the HP 9000 Superdome hardware.

	HP 9000 Superdome 16 sockets (32 processors)	HP 9000 Superdome 32 sockets (64 processors)	HP 9000 Superdome 64 sockets (128 processors)	HP 9000 Superdome I/O expansion cabinet (IOX) (see note 1)
<b>Number of PA-8900 processors</b>	4–32	4–64	12–128	
<b>8-CPU cell boards<sup>1</sup></b>	1–4	1–8	3–16	
<b>Memory (with 512 MB, 1 GB, or 2 GB DIMMs)</b>	2 GB–256 GB	2 GB–512 GB	6 GB–1 TB	
<b>Max. memory per partition</b>	256 GB	512 GB	1 TB	
<b>Max. processors per partition</b>	32	64	128	
<b>Virtual partitions</b>	32	64	128	
<b>Physical (hard) partitions</b>	4	8	16	
<b>12-slot I/O card cages</b>	1–4	1–4 (1–8 with IOX)	1–8 (1–16 with IOX)	6
<b>Hot-swap PCI-X I/O slots</b>	48 slots (32 slots @ 533 MB/s, 16 slots @ 1066 MB/s)	48 slots (12–96 with IOX: 64 @ 533 MB/s, 32 @ 1066 MB/s)	96 slots (12–192 with IOX: 128 @ 533 MB/s, 64 @ 1066 MB/s)	72 slots
<b>Hot-swap redundant power supplies (N+1 included)</b>	4	6	12	2
<b>I/O fans</b>	6	6	12	4 per I/O chassis enclosure (max. 12)
<b>Hot-swap redundant blowers or fans (N+1 included)</b>	4	4	8	
<b>Bandwidth:</b>				
Standard PCI-X bus	533 MB/s			
High PCI-X bus	1066 MB/s			
2X PCI I/O bus	266 MB/s			
4X PCI I/O bus	533 MB/s			
Cell controller to I/O subsystem (peak)	2.0 GB/s			
I/O (peak)	8 GB/s			
Crossbar (peak)	16 GB/s	16 GB/s	32 GB/s	
Memory (peak)	16 GB/s	32 GB/s	64 GB/s	
Cell controller to memory subsystem (peak)	16 GB/s	32 GB/s	64 GB/s	

	<b>HP 9000 Superdome 16 sockets (32 processors)</b>	<b>HP 9000 Superdome 32 sockets (64 processors)</b>	<b>HP 9000 Superdome 64 sockets (128 processors)</b>	<b>HP 9000 Superdome I/O expansion cabinet (IOX) (see note 1)</b>
<b>Avg. latency, memory load- to-use:</b>				
8 CPUs	246 ns	246 ns	246 ns	
16 CPUs	330 ns	330 ns	330 ns	
32 CPUs	371 ns	371 ns	371 ns	
64 CPUs	N/A	417 ns	417 ns	
128 CPUs	N/A	N/A	440 ns	
<b>Operating system</b>	HP-UX 11i v2, HP-UX 11i v1			
<b>Number of cabinets</b>	1 (left)	1 (left)	2 (1 left, 1 right)	1 Rack System E expansion
<b>Dimensions:</b>				
Height	77.2 in. (1960 mm)	77.2 in. (1960 mm)	77.2 in. (1960 mm)	77.2 in. (1960 mm) or 63.0 in. (1600 mm)
Width	30 in. (762 mm)	30 in. (762 mm)	60 in. (1524 mm)	24 in. (610 mm)
Depth	48 in. (1220 mm)	48 in. (1220 mm)	48 in. (1220 mm)	48 in. (1220 mm)
<b>Weight</b>	1,102 lb (500 kg)	1,318 lb (598 kg)	2,636 lb (1,196 kg)	

Note 1: For I/O expansion up to 168 slots, one expansion cabinet (holds up to 6 I/O card cages) is required. For I/O expansion up to 192 slots, two expansion cabinets are required. Any remaining space in the I/O expansion cabinets can be used to store peripherals. The I/O expansion cabinet is available in two heights: 77.2 in. (1.96 m) and 63 in. (1.6 m).



## HP 9000 Superdome hardware

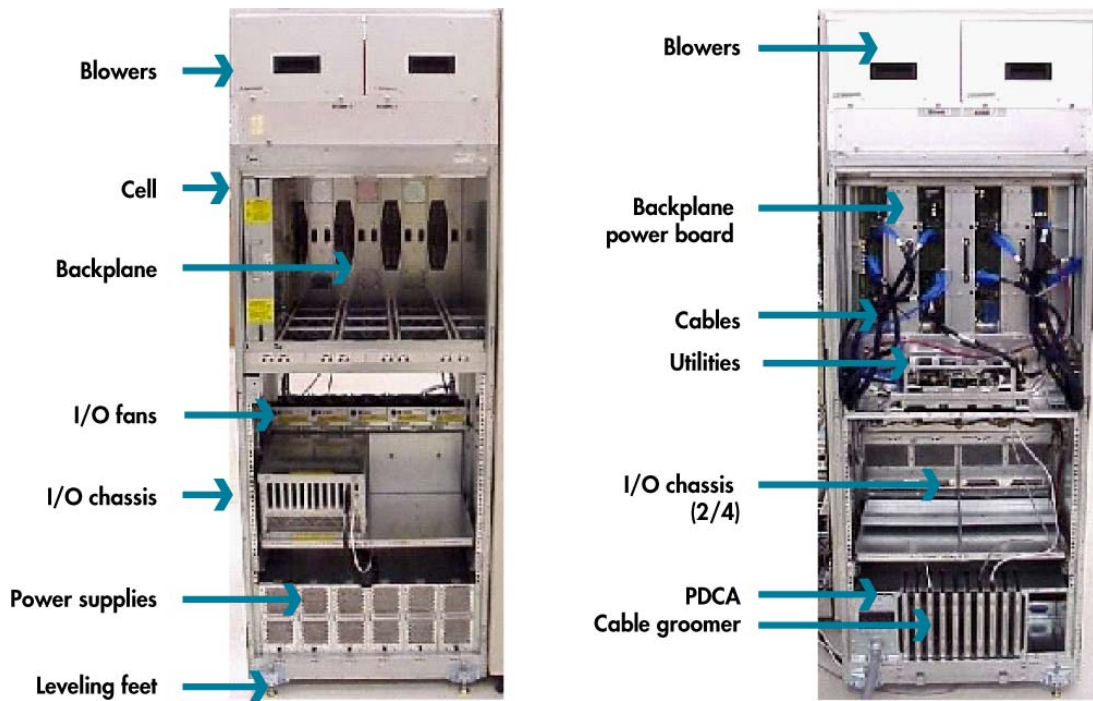
For maximum flexibility in configuring the data center, an HP 9000 Superdome system offers several different cabinet assemblies:

- HP 9000 Superdome 32-way and 64-way systems are single-cabinet configurations with one left cabinet. The cabinet contains all of the processors, memory, and core devices of the system. It also houses up to 48 I/O cards or 4 I/O chassis.
- The HP 9000 Superdome 128-way system is a dual-cabinet configuration with one left cabinet and one right cabinet. Together, the two cabinets contain all of the processors, memory, and core devices of the system. The dual-cabinet configuration houses up to 96 I/O cards or 8 I/O chassis.
- An optional 1.6-meter or 2-meter I/O expansion cabinet may be added if the required number of I/O cards exceeds the number of I/O cards that can be accommodated in a 64-way or a 128-way configuration.

## The HP 9000 Superdome cabinet

Figure 2 shows the physical layout of an HP 9000 Superdome 32-socket cabinet. At the top of the cabinet, hot-swappable main fans are installed after the cabinet arrives at the customer site. Below these fans is a cage for the eight cell boards on which processors and memory DIMMs reside. In a future release of HP-UX 11i, these cell boards will support hot-swap capability, so they can be replaced without bringing down the system.

**Figure 2:** Front and rear views of the HP 9000 Superdome 32-socket cabinet



Directly below the cell boards is the main air intake, and below that are two I/O chassis. Each I/O chassis holds 12 PCI-X I/O cards. Redundant power supplies are at the bottom of the cabinet. The HP 9000 Superdome family does not use electric plugs; instead, 48 Vdc power is hard-wired to each cabinet. There are two redundant power inputs, so the system can be powered up via two different power grids. An opening in the side of the cabinet allows two HP 9000 Superdome 64-way cabinets to be cabled together as an HP 9000 Superdome 128-way.

In order to differentiate between the existing HP 9000 Superdome systems (based on the PA-8600 and/or PA-8700+) and new HP 9000 Superdome systems with dual cores (based on the HP sx1000 Chipset and the PA-8900 or PA-8800 processors), the cabinet skin color will no longer be “quartz,” but rather “graphite.” The I/O expansion cabinet associated with the HP 9000 Superdome will also be graphite in color.

Features of the HP 9000 Superdome cabinets include:

- **Front and rear servicing**—Because all HP 9000 Superdome cabinets can be serviced from the front or rear, they can be arranged in the traditional “row” orientation found in most computer rooms.
- **Sized for easy installation**—Installation is no problem; the width of the cabinet allows it to pass through most doorways without disassembly.
- **Hot-swappable filters**—The intake air to the cell boards is filtered, helping to keep these boards clean and operational longer. The filter can be removed for cleaning or replacement while the system is fully operational.
- **External status display**—Status displays on the outside of each cabinet’s front and rear doors enable maintenance personnel and HP field engineers to determine the basic status of each cabinet without opening any cabinet doors.

## HP 9000 Superdome system architecture

The HP 9000 Superdome is a cache-coherent, Non-Uniform Memory Access (ccNUMA) system. And the HP 9000 Superdome presents a symmetric multiprocessing (SMP) programming model to the operating system by enabling any processor to have access to any byte of memory anywhere in the system. In fact, it is the first mission-critical UNIX system to exploit distributed processors and memory. Here’s why:

- Usable bandwidth scales with system size, thanks to two important design features:
  - **HP 9000 Superdome’s coherency scheme**—Unlike some UNIX systems that rely on snoop-based coherency, which results in bottlenecks (large, flat latencies and bandwidth starvation) in high-end configurations, the HP 9000 Superdome uses a specialized directory coherency scheme that scales to 16 cells with very low coherency bandwidth overhead.
  - **HP 9000 Superdome’s topology**—HP 9000 Superdome implements a point-to-point global packet switch for a communication fabric that is very well balanced across processor, memory, and I/O traffic.
- **Large physical memory with extremely low latency**—The HP 9000 Superdome supports 512 MB, 1 GB, and 2 GB DIMMs, yielding a maximum memory of 1 TB. Even with these high amounts of memory, the latency growth from 4 to 64 processors is only 1.8X (80% growth), which is extremely flat.
- **High I/O bandwidth and connectivity**—HP 9000 Superdome provides a high degree of I/O connectivity while preserving bandwidth. There are 192 I/O cards in the system, each with its own dedicated I/O bus; at the adapter level, this aggregates to 64 GB/s of raw bandwidth for PCI and 128 GB/s for PCI-X. The I/O bandwidth available between the system core (processors and memory) and the I/O controllers is roughly 30 GB/s for PCI and 32 GB/s for PCI-X.
- **Large number of high-performance processors**—The HP 9000 Superdome provides up to 128 PA-8900 processors. The new PA-8900 architecture is a dual-core technology with a clock speed of 1.1 GHz. There is 64 MB of unified L2 cache.

- **True hardware isolation of system resources**—Processors, memory, and I/O resources are truly isolated from each other in order to provide flexibility in system usage. This means:
  - The system has great SMP performance to attack large single workloads.
  - Hardware-enforced isolation of resources, coupled with great single-system high availability and manageability features, provides a strong consolidation platform.

## Modular architecture

The HP 9000 Superdome's architecture is modular, with components that can be used to construct several different server-class computer products based on a common architecture with the PA-8900 processors and the HP sx1000 Chipset. Processor upgrades from earlier-generation HP Superdomes to PA-8900 and PA-8800-based systems and beyond require a new version of the cell board. The same cell board can support either PA-RISC (the new PA-8900 or PA-8800) processors or the Intel Itanium 2 processor (Itanium 2 6M, Itanium 2 9M, or mx2 Dual Processor Module), but not both in the same HP 9000 Superdome. (In the second half of 2005, HP will support the both PA-RISC and Itanium 2 processors in the same Superdome; however, they will have to be in separate hard partitions.) Such modularity helps provide investment protection—you can simply add or change components as system requirements change.

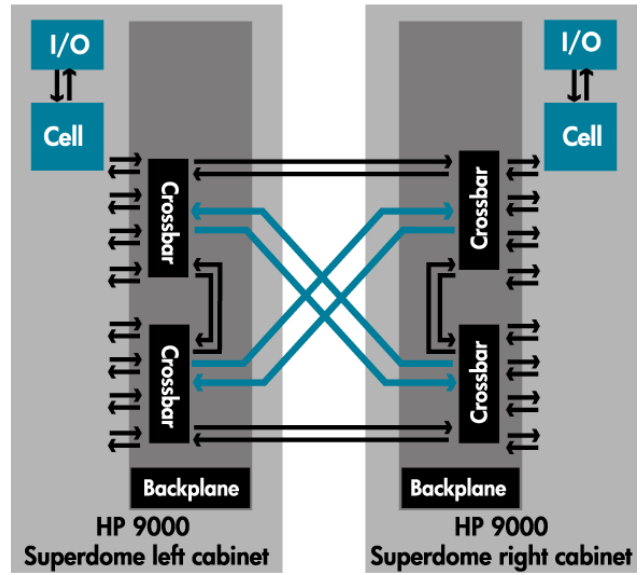
The HP 9000 Superdome platform supports a variety of system configurations ranging from 4 to 128 processors. (There are three models to choose from—the largest is a 12- to 128-processor system with a scalable memory and I/O system.) The system is extremely flexible, and its cell-based hierarchical crossbar architecture can be configured as one large symmetric multiprocessor or as several independent hard partitions (nPars). In an HP 9000 Superdome system, any processor on any nPar can directly address any byte of memory on any nPar through processor-issued load or store instructions.

There are three basic components in the HP Integrity Superdome system architecture: the cell or cell board, the crossbar backplane, and the PCI-X-based I/O subsystem (the PCI-based I/O subsystem is supported when an existing HP 9000 Superdome is upgraded to a PA-8900 or PA-8800-based HP 9000 Superdome or an HP Integrity Superdome). Figure 3 is an illustration of this architecture.

**Figure 3:** HP 9000 Superdome hierarchical crossbar architecture

**The crossbar mesh: interconnect fabric**

- Fully connected crossbar mesh
  - Four crossbars
  - Four cells per crossbar
- Equal bandwidth and latency for all links
  - Lower latency
  - More usable bandwidth
- Point-to-point packet filtering and routing network implementation
  - Allows hardware isolation of all faults
- 16 cells interconnected with 3 latency domains
  - Cell local
  - Crossbar local
  - Remote crossbar



**Industry-leading, flexible elegance in a high-end UNIX server**

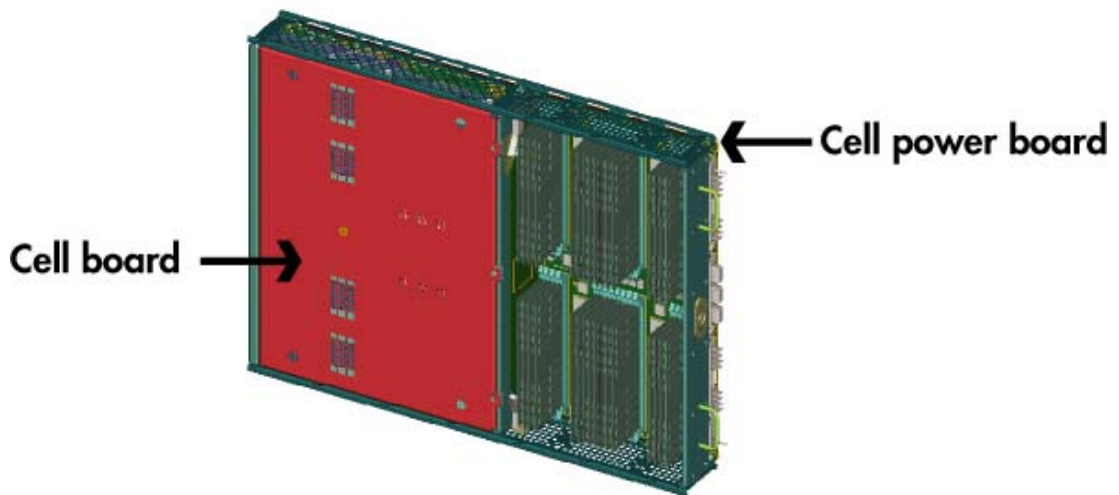
## The cell board

A cell, or cell board, shown in Figure 4, is the basic building block of an HP 9000 Superdome system. Each cabinet can contain up to eight cell boards, which are plugged into the backplane of the cabinet.

In the original HP Superdome release, the cell board frame contained both the cell board and associated cell power board. In the HP 9000 Superdome Servers with the new PA-8900 or PA-8800 Processors, the cell and cell power board are in separate assemblies to improve handling. The same cell power board is used with all varieties of cell boards, including those for Itanium 2–based HP Integrity Superdome Servers.

When upgrading an existing HP 9000 Superdome to a PA-8900–based HP 9000 Superdome or an HP Integrity Superdome, the existing cell boards are removed from the system. The DIMMs from these boards are removed and placed into the new HP 9000 Superdome cell boards, which are then loaded into the system along with the cell power boards. When upgrading an existing HP 9000 Superdome to a PA-8900-based HP 9000 Superdome or an HP Integrity Superdome, the existing cell boards are removed from the system. The DIMMs from these boards are removed and placed into the new HP 9000 Superdome cell boards, which are then loaded into the system along with the cell power boards. Existing HP 9000 Superdome cell boards cannot coexist in the same complex with PA-8900–based or PA-8800–based HP 9000 Superdome cell boards or HP Integrity Superdome cell boards. In addition, in 2005, PA-RISC (PA-8800 and PA-8900) and Intel Itanium 2 9M processors will be able to co-exist in the same Superdome, but they must be in different hard partitions. Any existing cell blockers in the original HP 9000 Superdome system must also be replaced with new cell blockers.

**Figure 4:** Cell board layout of an HP 9000 Superdome with 1.1 GHz PA-8900 processors



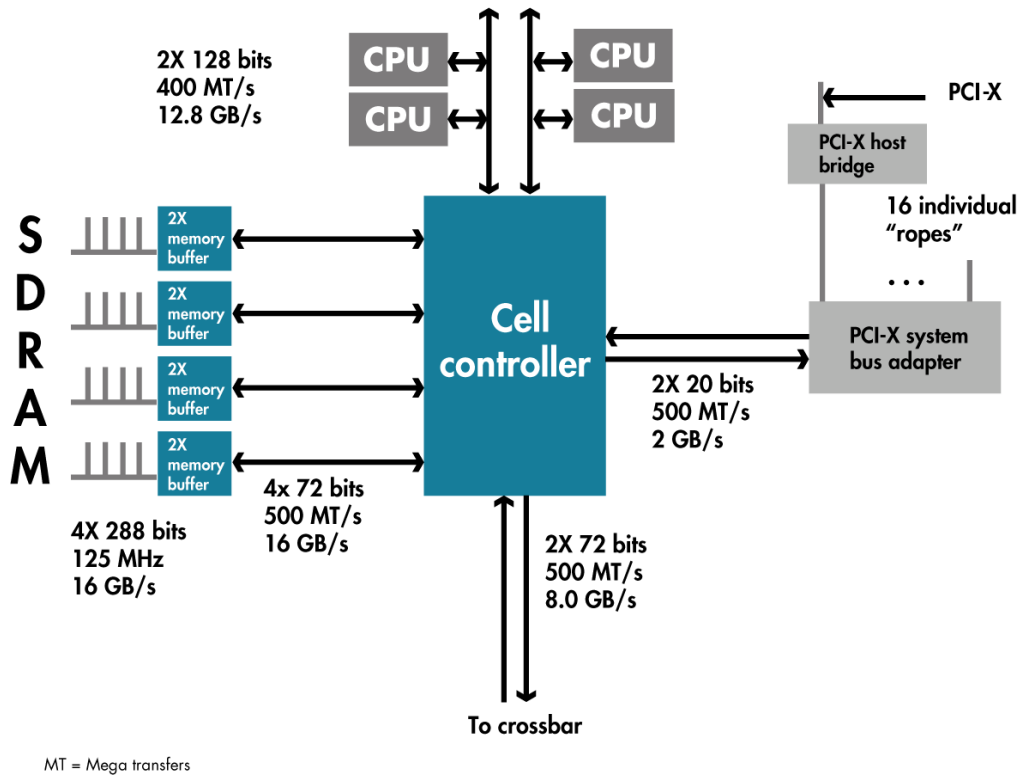
Each cell board is a self-contained unit, with a symmetric multiprocessor (SMP), main memory, and all necessary hardware:

- CPUs (up to 8 processors per board—4 sockets with 2 CPUs per socket)
- Cell controller ASIC (application-specific integrated circuit)
- Main memory DIMMs (up to 32 DIMMs per board in four-DIMM increments, using 512 MB, 1 GB, or 2 GB DIMMs or some combination of the three)
- Voltage regulator modules (VRMs)
- Data buses
- Optional link to 12 PCI-X I/O slots

Figure 5 illustrates the cell board architecture. The cell has a peak memory bandwidth of 16 GB/s. A connection to a 12-slot PCI-X card cage is optional for each cell, and the peak bandwidth of this link is 2 GB/s. Bandwidth to the crossbar is 8 GB/s per cell.

Error checking and correcting (ECC) exists on all fabric paths, on all memory paths, and on CPU cache. Parity protection exists on all CPU and I/O links. Single-wire correction exists on fabric and I/O paths.

**Figure 5:** HP Integrity/HP 9000 Superdome cell board and interconnect architecture



This release of HP 9000 Superdome with PA-8900 processors supports 128-megabit and 256-megabit SDRAM.

### Processors

The HP 9000 Superdome uses PA-8900 processors with a core frequency of 1.1 GHz and 64 MB of L2 cache.

### Cell controller ASIC

Residing on the cell board, the cell controller ASIC (application-specific integrated circuit) is part of the HP 9000 Superdome system chipset, the HP sx1000 Chipset. It coordinates traffic between the major components of a cell board and determines if a request requires communication with another cell or with the I/O subsystem. HP believes that this cell controller ASIC is the largest ASIC in the world—it has approximately 41 million transistors.

The cell controller ASIC has five major interfaces:

- Four memory subsystems
- Two ports to processors (one dedicated port per two processors)
- Crossbar interface, through which all communication to other cells flows
- Processor-dependent hardware (PDH)
- I/O interface, which connects the cell to an I/O subsystem

In addition to the interface logic, the cell controller ASIC maintains cache coherency throughout the system. The cell controller ASIC supports the PA-8900, the PA-8800, and the Intel Itanium 2 processor.

The processor-dependent hardware (PDH) module provides the cell-local resources required to reset a cell and bring it up to a point where it can join other cells and boot the operating system. The module contains the system boot firmware, which is also used at runtime.

### **Memory controller ASIC**

The memory controller ASIC is also part of the HP sx1000 Chipset. Its primary function is to multiplex and de-multiplex data between the cell controller ASIC and the SDRAMs in the memory subsystem. When the cell controller ASIC issues a read transaction on the memory interface command bus, the memory controller ASIC buffers the DRAM read data and returns it as soon as possible. When the cell controller ASIC issues a write transaction, the memory controller ASIC receives the write data from the cell controller ASIC and forwards it to the DRAMs.

Note that only the data portion of the memory subsystem goes through the memory controller ASIC. All address and control signals to the DIMMs are generated by the cell controller ASIC and sent directly to the DIMM via the memory interface address bus.

The memory subsystem is a quad-ported implementation. It supports memory DRAM fault tolerance: a discrete SDRAM chip can fail without compromising data integrity. The memory subsystem provides 16 GB/s of peak bandwidth to the cell controller ASIC and significantly reduces the overhead typically associated with directory coherency. What's more, the memory subsystem enjoys a very low latency for cell-to-local-memory access: as low as 245 ns average idle latency at load-to-use.

## **The crossbar backplane**

Each crossbar backplane contains two sets of two crossbar ASICs that provide a non-blocking connection between eight cells and the other backplane. Each backplane cabinet can support up to eight cells or 64 processors (an HP 9000 Superdome 64-way in a single cabinet). Two backplanes can be linked together with flex cables to produce a cabinet that can support up to 16 cells or 128 processors (HP 9000 Superdome 128-way in dual cabinets).

### **Crossbar ASIC**

The crossbar ASIC is yet another part of the HP sx1000 Chipset. It implements a high-performance 8-port non-blocking crossbar and the 500 MHz crossbar link protocol. All ports are functionally and electrically identical. The HP 9000 Superdome's fabric is a fully connected crossbar mesh, with four pairs of crossbar ASICs and four cells per crossbar ASIC pair. Like the cell controller ASIC, the crossbar ASIC has truly impressive scale. HP believes it, too, to be one of the world's largest ASICs, with 18 million transistors and six layers of metal, using copper technology.

A very important aspect of the crossbar mesh is that all links have the same bandwidth and latency—which is key to increasing overall bandwidth and decreasing overall latency. Cell-to-crossbar and crossbar-to-crossbar communication occur at the same speed—there are no excessive latency penalties for going remote. In addition, the HP 9000 Superdome's memory interleaves across cells first, then across memory banks. This interleaving scheme tends to balance out memory traffic across all the links.

The crossbar mesh implements a global point-to-point packet-filtering network. This mesh features an extremely high level of integrity, with each crossbar port fully independent. The crossbar mesh has dedicated paths for data and control. Each port can be reset, assigned, or reconfigured fully independently of other ports. The HP 9000 Superdome's crossbar mesh is an excellent foundation for resource isolation.

The crossbar ASIC offers several features that contribute to the HP 9000 Superdome's high performance:

- Support for scaling up to a 128-way coherent shared memory system (with PA-8900 and PA-8800 processors)
- 250 MHz operation
- 500 mega transfers/second (MT/s) link speed
- Support for two interleaved channels on link protocol
- Support for double-length data packets for Intel Itanium Processor Family mode
- Performance counters to enable software tuning

Each port on the crossbar ASIC has 8 GB/s of peak bandwidth. These ports provide a high-throughput path to cells and other crossbar ASICs:

- Four ports that connect to the four cells residing on the crossbar ASIC (one port per cell)
- Three ports that connect to the remaining three crossbar ASICs (in an HP 9000 Superdome 128-way)

The total crossbar bandwidth for each the HP 9000 Superdome model is calculated as follows:

$$(\text{Number of cells} \times \text{peak crossbar bandwidth per cell}) \div 2 \text{ ports}$$

Crossbar bandwidth for the different HP 9000 Superdome systems is excellent:

- 16 GB/s for the HP 9000 Superdome 32-way
- 32 GB/s for the HP 9000 Superdome 64-way
- 64 GB/s for the HP 9000 Superdome 128-way

### **Memory and backplane latencies**

The HP 9000 Superdome has been designed to reduce memory and backplane latencies, enhancing performance. There are three types of memory latencies within the HP 9000 Superdome system:

- **Memory latency within the cell** refers to an application running on a hard partition (nPar) that consists of a single cell board.
- **Memory latency between cells on the same crossbar** refers to an nPar consisting of up to four cells that reside on the same crossbar. For example, if there are four cells in the nPar,  $\frac{1}{4}$  of the requests go to the memory of the cell board in which the processor resides, and  $\frac{3}{4}$  of the requests go to the memory of the other three cell boards.
- **Memory latency between cells on different crossbars** refers to an nPar consisting of cell boards that do not all reside on the same crossbar. For example, if there are 16 cells in the nPar,  $\frac{1}{16}$  of the requests are to the memory of the cell board in which the processor resides.  $\frac{3}{16}$  of the requests go to the memory of the other three cell boards on this same crossbar. Finally, the remaining  $\frac{12}{16}$  of the requests transit two crossbars.



The HP 9000 Superdome memory latency depends on the number of CPUs and the location of their corresponding cell board. Assuming that there is equally distributed traffic to all memory controllers and that cell boards are installed to reduce latency, the average idle load-to-use memory latency is shown below.

Number of cell boards	Number of CPUs	Average idle load-to-use memory latency
1	8	246 ns
2	16	330 ns
4	32	371 ns
8	64	417 ns
16	128	440 ns

## The I/O subsystem

Each HP 9000 Superdome cell has an optional link to an I/O chassis: this enhances modularity and means there are no trade-offs for scaling of processors, memory, and I/O. Each cell connects to its remote I/O chassis through an I/O cable link.

The HP 9000 Superdome's I/O subsystem has plenty of capability for today and expansion for tomorrow. Each I/O module consists of 12 PCI-X connections, divided among eight standard PCI-X and four high-bandwidth PCI-X slots, with an I/O controller ASIC and power. Each PCI-X slot has its own PCI-X bus—the standard PCI-X slot has 533 MB/s bandwidth and the high-bandwidth PCI-X slot achieves a bandwidth of 1066 MB/s. The point-to-point connectivity allows the earliest detection, containment, and recovery from errors.

Providing a dedicated bus for each PCI-X card has the following advantages:

- **High availability**—Unlike a shared bus, if the bus or the PCI-X card has problems, only that PCI-X card is impacted.
- **Throughput**—Having a dedicated bus, the PCI-X card can take advantage of the entire bandwidth of the bus.
- **Investment protection**—Since the bus is dedicated to the PCI-X card, future faster PCI-X cards or PCI-X cards with multiple ports can be utilized without any impact to the system.

Any I/O module can support a core I/O card (required for each independent nPar). The HP 9000 Superdome 32-way, 64-way, and 128-way systems can accommodate four, four, and eight I/O modules, respectively, within the system cabinet. This provides a total of 48, 48, and 96 PCI slots, respectively.

An I/O expansion cabinet can also be added. On the HP 9000 Superdome 64-way system, this yields an additional 48 PCI-X slots, giving a maximum connectivity of 96 PCI-X slots for 8 cells. For an HP 9000 Superdome 128-way system, the I/O expansion cabinet provides an additional 96 PCI-X slots, for a maximum connectivity of 192 PCI-X slots for 16 cells.

Note: If an existing HP 9000 Superdome system with an I/O expansion cabinet is being upgraded to an HP sx1000 Chipset-based HP 9000 Superdome with PA-8900 or PA-8800 processors, the side panels must be replaced and upgraded.

A system configured with 16 cells—each with its own I/O module and core I/O card—can support up to 16 independent nPars. Note that cells can be configured without I/O modules attached, but I/O modules cannot be configured in the system unless they are attached to a cell.

The I/O subsystem bandwidth is 2.0 GB/s per cell. This makes the total I/O subsystem bandwidth for HP 9000 Superdome 32-way, 64-way, and 128-way systems 8.0 GB/s, 16.0 GB/s, and 32 GB/s, respectively.

## The HP 9000 Superdome: performance and scalability

The HP 9000 Superdome offers a host of features that are especially important in the OLTP, business intelligence, and technical markets. And bandwidth and scalability are two important areas of HP 9000 Superdome leadership.

### System bandwidth

	HP 9000 Superdome 32-way	HP 9000 Superdome 64-way	HP 9000 Superdome 128-way
<b>Crossbar bandwidth (peak)</b>	16 GB/s	32 GB/s	64 GB/s
<b>Cell controller to I/O subsystem bandwidth (peak)</b>	2.0 GB/s	2.0 GB/s	2.0 GB/s
<b>I/O bandwidth (peak)</b>	8 GB/s	16 GB/s	32 GB/s
<b>Memory bandwidth</b>	64 GB/s	128 GB/s	256 GB/s

### Scalability

	HP 9000 Superdome 32-way	HP 9000 Superdome 64-way	HP 9000 Superdome 128-way
<b>8-CPU cell boards</b>	4	8	16
<b>Maximum number of PA-8900 processors</b>	32	64	128
<b>Maximum memory</b>	256 GB	512 GB	1 TB
<b>12-slot I/O card cages</b>	4	8	16
<b>Hot-swap PCI/X I/O slots</b>	48 slots (32 std. PCI-X slots, 16 high BW PCI-X slots)	96 slots (64 std. PCI-X slots, 32 high BW PCI-X slots)	192 slots (128 std. PCI-X slots, 64 high BW PCI-X slots)
<b>Maximum number of hard partitions (nPars)</b>	4	8	16

Note 1: For I/O expansion up to 168 slots, one expansion cabinet (6 I/O card cages) is required. For I/O expansion up to 192 slots, two expansion cabinets (two I/O card cages) are required. Any remaining space in the I/O expansion cabinets can be used to store peripherals.

For more information about performance and scalability, please refer to the HP 9000 Superdome pages on the HP Web site ([www.hp.com/go/Superdome](http://www.hp.com/go/Superdome)).

# The HP 9000 Superdome: high availability

HP 9000 Superdome offers unprecedented single-system availability in the following areas:

- System reliability
- Supportability
- Repairability
- HP Serviceguard HA Cluster-in-a-Box

## System reliability

The HP 9000 Superdome product family is bolstered by design enhancements and production techniques that substantially enhance system reliability. Among these enhancements are the following:

- Memory DRAM fault tolerance—recovery of a single SDRAM failure
- DIMM address parity protection
- Dynamic memory resilience—page de-allocation of bad memory pages during operation
- Dynamic processor resilience
- CPU cache ECC protection and automatic de-allocation
- CPU bus parity protection
- Full single-bit error detection and correction on crossbar and I/O links
- I/O error recovery and system resilience to I/O card failures
- I/O cards fully isolated from each other
- Prevention of silent corruption of data going to I/O
- Recovery of a single I/O-to-cell controller link failure
- Recovery of a single crossbar wire failure
- Localization of crossbar failures to the partitions using the link
- Automatic de-allocation of bad crossbar link upon boot
- ASIC full burn-in and high-quality production process, resulting in 10X improvement in ASIC failure rates
- Full “test to failure” and accelerated life testing on all critical HP 9000 Superdome assemblies
- Strong emphasis on quality for multiple nPar single points of failure (SPOFs)
- System resilience to Guardian Service Processor (GSP)
- Isolation of nPar failures—failure of one nPar will not bring down another partition
- Protection of nPars against spurious interrupts or memory corruption

## Supportability

Supportability is another key feature of the HP 9000 Superdome product line. Here are just a few of the supportability features of the HP 9000 Superdome:

- Event Monitoring Service (EMS)
- Support Management Station (SMS)
- Instant Support Enterprise Edition (ISEE)

### **Event Monitoring Service**

For the HP 9000 Superdome, HP Event Monitoring Service (EMS) aids in system management by keeping track of the system's vital signs. EMS is installed and activated by default. It lets you monitor virtually all of the hardware on the system, including these components and events:

- Mass storage
- Memory
- Fibre Channel components (multiplexer, switch, card, fabric, etc.)
- I/O cards
- ECC errors on the main system bus
- ECC errors in the CPU cache
- System temperature
- Support processor problems
- Processor-dependent hardware problems
- Processor-dependent hardware low-battery conditions
- Cabinet fans
- Cabinet power supplies
- Chassis code logging failure
- Processor-dependent code (firmware) problems

The system hardware configuration and selected kernel parameters are also monitored.

When a hardware monitor detects a problem, it generates an event that can be reported to you via SNMP to HP Systems Insight Manager or HP OpenView Operations, e-mail, page, syslog, console, or a selected text log file. Each event contains a complete description of the problem, a severity classification (informational, warning, serious, critical), and text that shows the probable cause and the suggested action to take.

When excessive processor cache errors are detected, the processor is automatically removed from use until it is replaced. If the customer has available Instant Capacity processors in the partition, then EMS will automatically activate one of these processors to replace the processor that has been removed from use. The failed processor is also marked so it will be removed from use on the next system reboot.

### **Support Management Station (SMS)**

To decrease the possibility of unscheduled downtime still further, the HP 9000 Superdome Support Management Station (SMS) is a separate server that provides diagnosis and test capabilities for the HP 9000 Superdome system. The SMS includes ASIC-level scan tools that can be used by HP support engineers to diagnose the HP 9000 Superdome systems in your data center. (For more details on the SMS, see "The HP 9000 Superdome Support Management Station" on page 29.)

### **Instant Support Enterprise Edition (ISEE)**

The HP 9000 Superdome systems can be supported by HP Instant Support Enterprise Edition (ISEE), a support solution designed to keep mission-critical environments up and running. The ISEE is a leading-edge remote support solution that combines processes and the expertise of HP's qualified support personnel to provide powerful support capabilities for systems covered by mission-critical HP support contracts. The HP ISEE solution improves the availability of HP 9000 Superdome systems by

- Conducting frequent, automated data collection of the customer's mission-critical environment
- Providing a secure, high-bandwidth link for HP support personnel to remotely access the customer's environment to conduct diagnostic tests and potentially solve reported problems
- Providing reactive and proactive support for HP mission-critical customers

HP's Mission-Critical Support Center (MCSC) plays a key part in ISEE's ability to reduce the number, duration, and impact of outages on covered systems. The MCSC resides inside HP. Its job is to act as the primary support center, collecting and tracking customer information and performing proactive problem analysis. It also provides a control point in the ISEE remote support architecture. To maintain the security of your environment, the MCSC allows only authorized high-availability (HA) support engineers to access the isolated LAN.

ISEE collects configuration data from your mission-critical environment on a weekly basis for both reactive and proactive support purposes. HP qualified support engineers interpret the data in order to expedite the problem resolution process. The HP account support team will also analyze the data as part of providing the following proactive services:

- **HP-UX Patch Analysis**—a detailed review of your systems' current patch level, with recommendations to install, update, or remove patches to align your environment with your pre-selected patch management strategy
- **HP-UX Operating System, Hardware, and Firmware Analysis**—a detailed assessment of your HP-UX environment (can also be applied to HP Serviceguard clusters) that identifies opportunities to increase your environmental availability through planned configuration change management

HP network support customers who are entitled to ISEE Advanced Configuration will also receive the benefits of HP Network Support Tools (NST). Network Support Tools provide HP's network support engineers with powerful diagnostics and mapping tools for network discovery, configuration collection, and troubleshooting to resolve your network problems quickly. NST takes advantage of the ISEE platform, features, and single remote access point. HP performs remote support only with customer authorization.

## Repairability

The HP 9000 Superdome servers are endowed with a substantial number of features that make them easier to repair with little or no downtime. Among these features are:

- N+1 CPUs and cell boards with Instant Capacity
- Hot-swap N+1 fans, power supplies, and backplane DC/DC converters
- Dual power sources
- Online replacement of PCI-X I/O cards
- Online addition/replacement of cell boards (available in a follow-on release to HP-UX 11i)
- Online addition of hard partitions (nPars)

### **N+1 CPUs and cell boards with Instant Capacity**

The HP 9000 Superdome has support for Instant Capacity with HP-UX 11i. This feature lets you bring additional CPUs or cell boards (CPUs plus memory) online without a system reboot. Because no reboot is needed, the system has no loss of availability, even while you're adding capacity.

Another advantage of Instant Capacity is that it allows a system to be set up with N+1 CPUs or cell boards for top single-system availability. If one CPU or one cell board (CPU or memory) fails, another is already running and prepared to take its place automatically without disturbing system operation.

### **Hot-swap N+1 fans, power supplies, and backplane DC/DC converters**

The HP 9000 Superdome servers are equipped with N+1 fans, power supplies, and backplane DC converters for higher availability of these components.

### **Dual power sources**

Dual power sources in the HP 9000 Superdome servers mean the power supply can be protected against becoming a single point of failure.

### **Online addition and replacement (OLAR) of cards and chassis**

OLAR enables the online addition and replacement of I/O cards on HP 9000 Superdome systems. The system hardware uses per-slot power control combined with operating system support for the I/O card OLAR feature to allow the addition of a new card and the replacement of an existing card without affecting other components or requiring a reboot.

### **Online addition and replacement of cell boards**

The HP 9000 Superdome and a future HP-UX release after 11i will support online addition and replacement of cell boards, allowing these critical components to be repaired or maintained without bringing the system down.

### **Online addition of hard partitions (nPars)**

nPars may be added to the HP 9000 Superdome without affecting other running nPars. Such dynamic reconfiguration is another reason why the HP 9000 Superdome provides superior uptime. And the HP 9000 Superdome supports dynamic cell board migration, which will be fully supported with the next release of HP-UX 11i.

## **HP Serviceguard HA Cluster-in-a-Box**

In order to increase the uptime of applications within an HP 9000 Superdome system, nPars can be configured in an HP Serviceguard HA cluster so that the cluster membership is within the system. Upon detection of a failure within an nPar, the HP Serviceguard HA cluster software fails-over the application to another nPar within the HP 9000 Superdome system.

## **Uptime Institute certification**

The interaction of a server with its physical environment is a critical component of maintaining the availability of the environment. As such, the HP 9000 Superdome demonstrates its mission-critical capabilities by being the first server to become fully certified by The Uptime Institute's Fault Tolerant Power Compliance Certification Version 1.2. The certification verifies that all computer and communication hardware maintain full functionality through various power-testing manipulations. The HP 9000 Superdome server qualified as "fault tolerance power compliant" based on requirements collectively developed by the 48 members of the Uptime Institute's Site Uptime Network.

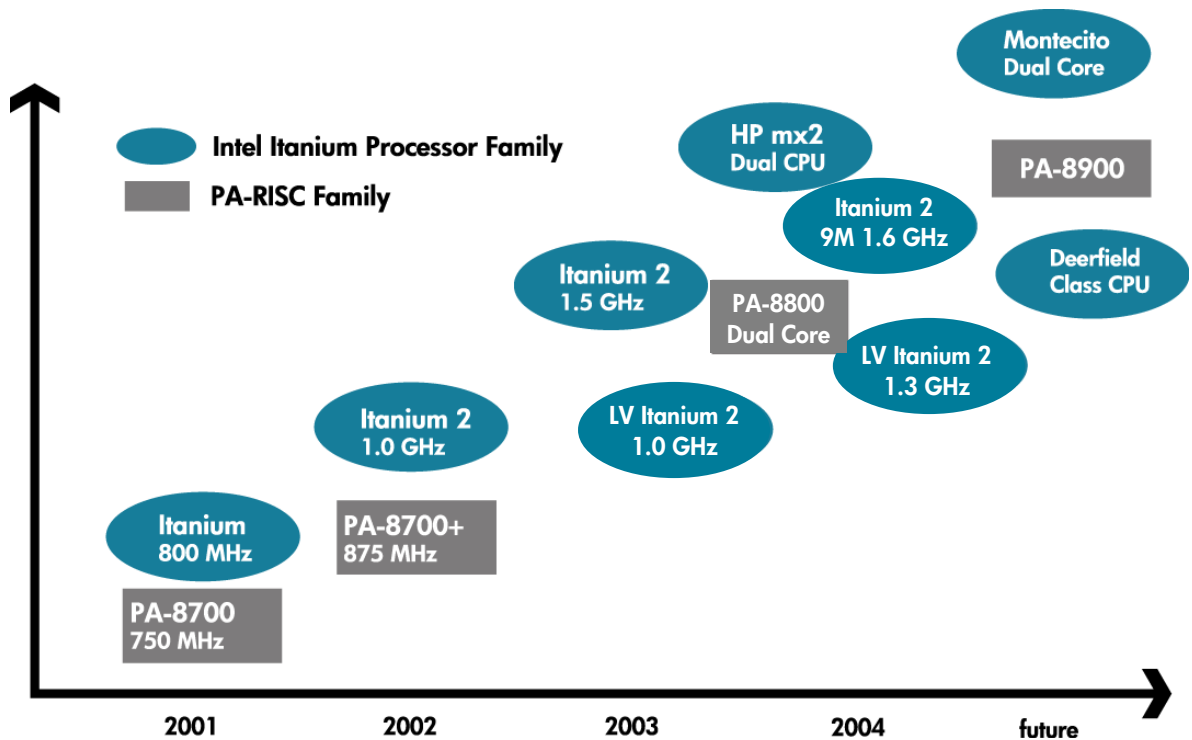
## **Support services**

There is much more to consider in order to provide high availability than just the hardware capabilities of the HP 9000 Superdome. For example, customers should keep in mind not only the overall HA solutions being offered by HP products (delivering both single-system and cluster HA), but also support services (such as Business Continuity Support and Critical Systems Support), IT process support (consulting, disaster-recovery services, etc.), and partnerships with third parties (e.g., Oracle®, Cisco, SAP) for fully integrated high-availability solution offerings. When HP Support is purchased, fault management events can be forwarded to the HP Support Organization. HP can then monitor, filter, and interpret the events and take action on items that need attention.

## The HP 9000 Superdome: investment protection

The HP 9000 Superdome family provides outstanding customer investment protection and lasting value, thanks in part to a system infrastructure designed to accommodate several generations of processor upgrades. For example, existing HP 9000 Superdomes are board-upgradable to the PA-8900 and PA-8800 processors, to 1.5 GHz Intel Itanium 2 processors, and to the HP mx2 Dual-Processor Module. And all major system components other than the cell board and power conversion card remain the same—these processors even use the same memory DIMMs. PA-8900 processors and PA-8800 processors can be combined in the same complex, but not in the same partition.

Figure 6: HP microprocessor roadmap



The processor roadmap shows not only HP's leading line of RISC processors but also the binary-compatible Intel Itanium processors. Notice that HP has continued its investment in PA-RISC enhancements since the introduction of Itanium-based systems, so customers can move to the new Intel Itanium architecture when they're ready, not when they're forced to by a vendor. Superior microprocessors and binary compatibility help make HP 9000 Superdome an ideal high-end server choice for customers today—and the safest investment for the future.

### Binary compatibility across PA-RISC

Customers can continue to rely on binary compatibility across the PA-RISC family, enabling seamless interoperability with legacy applications on HP systems. Binary compatibility protects investments, enabling rapid growth and adoption of new technology infrastructures. For performance improvements, existing applications and operating systems can be used with new or more advanced processor technology as it develops.

## Binary compatibility for the Intel Itanium Processor Family

HP will continue to support binary compatibility through the introduction of Itanium-based systems (HP Integrity servers). As a result of HP's work with Intel on the Explicitly Parallel Instruction Computing (EPIC) architecture—the technology foundation for the Intel Itanium architecture—today's HP-UX, Windows®, and Linux® applications will run unchanged on Itanium-based systems. And, for higher performance, applications can be recompiled without source changes.

## Upgrading an existing HP 9000 Superdome system

Upgrading an existing HP 9000 Superdome system to a PA-8900 or PA-8800-based HP 9000 Superdome is simple:

1. Remove the cell boards.
2. Transfer all memory to the new cell board.
3. Plug the new cell board and the separate power board into the cabinet.
4. Replace the existing cell blockers if the upgraded system requires cell blockers.
5. Replace the side panels on the system cabinet if the upgraded system is going to be used with an I/O expansion cabinet.
6. Change the system cabinet skins and I/O expansion cabinet skins to the graphite color.

Note: The original HP 9000 Superdome cells and the new HP 9000 Superdome cells supporting the PA-8900 or PA-8800 processor cannot coexist in the same complex, even in different partitions.

To stay ahead of tomorrow's performance demands, today's HP 9000 Superdome is already prepared for the next generation of processors. The HP 9000 Superdome delivers investment protection through in-chassis upgrades.

## The HP 9000 Superdome: server virtualization

Server virtualization enables administrators to optimize the usage and simplify the management of single- or multiple-server environments by configuring them as reusable pools of resources. Virtualization means that the physical resources are separated from the logical view of the server infrastructure. HP offers a number of solutions for vertically scaled environments that are available on the HP 9000 Superdome.

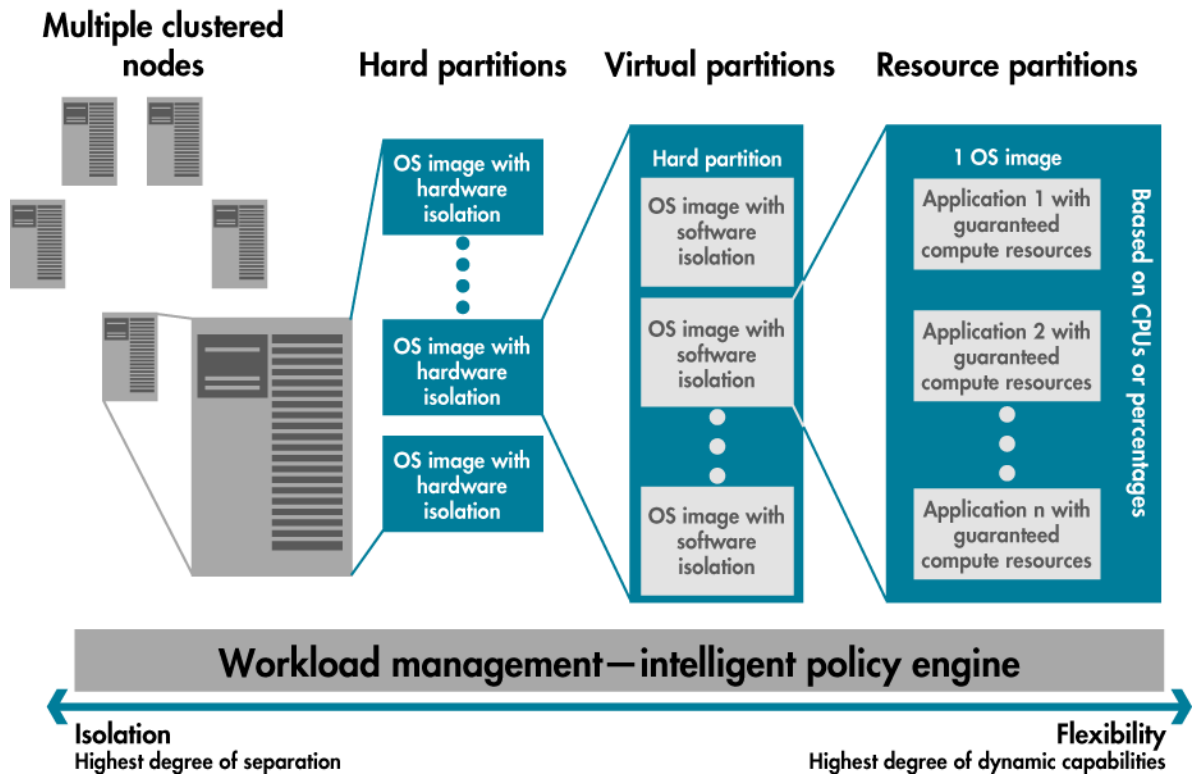
Vertical scaling typically means the use of a few consolidated high-performance servers to run many concurrent complex applications. This environment offers major opportunities for cost reduction and better utilization of existing assets. HP solutions for server partitioning and resource optimization can help administrators drive server utilization from a typical 15–50% to over 90% without service-level degradation. The combination of a virtualized server environment with Utility Pricing solutions such as Instant Capacity and Pay per use allows customers to activate additional capacity only when it is needed and to make payments for server resources based on actual usage.

The HP Virtual Server Environment (VSE) allows enterprises to achieve a greater return on their IT investment by realigning server resource utilization in real time based on business priorities. At the heart of HP VSE is the intelligent policy engine, HP-UX Workload Manager (HP-UX WLM), which orchestrates the virtual server resources. HP-UX WLM enables a virtual HP-UX server that automatically grows and shrinks based on the service-level objectives (SLOs) for each application it hosts. HP-UX WLM performs real-time assessment of the resource usage and then advises and acts in accordance with the defined SLOs and business priorities. To create a VSE, HP-UX WLM is tightly integrated with HP's server virtualization offerings, such as resource management groups, partitioning, clustering, and capacity-on-demand.



The HP Partitioning Continuum provides a range of hard, virtual, and resource partitioning tools that offer resource virtualization at the server or partition level, improve overall system and subsystem utilization, and lower costs in consolidated environments. Partitions are physical or logical mechanisms for isolating operational environments within single or multiple servers. Partitioning provides IT managers with the flexibility to dynamically resize an application’s resource usage while protecting all applications against disruptive events that could cause service interruption or performance degradation.

**Figure 7:** The HP Partitioning Continuum combines high isolation with excellent flexibility



HP offers several different partitioning solutions, each designed to support a different balance of application isolation and resource realignment:

- **Hard partitions (nPars)**—Hard partitions within a node are designed to isolate application environments from single points of failure (SPOFs). This means that applications running within hard partitions are not affected by hardware OR software events occurring in other partitions. On the HP 9000 Superdome, hard partitions are within the node and are called nPartitions or nPars. An HP 9000 Superdome server can support anywhere from one to 16 nPars, each supporting its own operating system, applications, peripherals, and networks. Each nPar can host its own independent version of the HP-UX operating environment.

With nPars, the HP 9000 Superdome can be configured as one large symmetric multiprocessor (currently up to a maximum of 64 processors) or as several independent nPars (a maximum of 64 processors per nPar). Using independent nPars has advantages in several areas: increased system utilization, increased flexibility with support for multiple environments, increased uptime, and increased scalability.

How do nPars work? Within the HP 9000 Superdome, cell boards are grouped into physical structures. An nPar consists of one or more cells that communicate coherently over a high-

bandwidth, low-latency crossbar fabric. Special programmable hardware in the cell boards defines the boundaries of an nPar in such a way that isolation from the actions of other nPars is enforced.

Each nPar runs its own independent operating system, and different nPars can be executing the same or different versions of an OS. In an HP Integrity Superdome, they can even be executing different operating systems altogether (such as HP-UX, Linux, OpenVMS, and Windows).

Each nPar has its own independent set of CPUs, memory, and I/O resources. You can use system management commands to move resources from one nPar to another without having to physically change the hardware. In addition, dynamic addition of new nPars is supported.

- **Virtual partitions (vPars)**—Virtual partitions provide complete software fault isolation within an individual server or a hard partition. This means that any application or OS-related failure impacts only the virtual partition in which it is executing—without having any effect on other vPars running on the same system. In a system with vPars, each OS instance is completely independent of all others. OSs in different vPars can be different versions or have different tuning parameters. Therefore, vPars are useful for testing new OS versions or applications without the need to replicate the deployment environment. In addition, vPars also let you dynamically adjust partition size (without rebooting) by the dynamic addition and deletion of CPUs, effectively moving processors from one vPar to another.

Finally, vPars provide even more portioning granularity. The HP 9000 Superdome supports up to 16 nPars and up to 128 vPars. So a system with 16 nPars can have as many as eight vPars in each nPar—or, to put it another way, one vPar per CPU. HP recommends a maximum of eight vPars per nPar. Each vPar runs its own image of the HP-UX 11i operating system and can full host its own applications.

- **Resource partitions**—Resource partitions address the need for the dynamic allocation, within a single OS instance, of dedicated resources between competing applications in order to avoid resource contention. Resource partitions run within hard partitions or virtual partitions and are controlled by workload management functions within the operating environment. The HP-UX Workload Manager (WLM) and Process Resource Manager (PRM) software dynamically creates resource partitions for applications that need guaranteed amounts of dedicated resources such as CPU cycles, memory, or disk I/O. The HP Process Resource Manager enables system administrators to control the amount of resources that applications, users, or groups may use during peak system load. Applications with specific goals (service levels), such as response time, can use HP's goal-based HP-UX WLM to automatically and dynamically allocate the necessary resources to applications or user groups within hard partitions or virtual partitions. With resource partitions and workload management tools, varying service-level objectives can be met every time!

As part of the Virtual Server Environment, partitioning can be complemented by **HP Utility Pricing solutions** to provide a number of options for implementing the infrastructure hardware and software required for on-the-fly expansions and reductions in computing power. Specific solutions for HP servers with HP-UX are Instant Capacity and Pay per use. Instant Capacity allows customers to activate processors within a partition or server when they are needed. (Processors can also be activated temporarily to meet short-term demand, then deactivated and held again in reserve.) Pay per use solutions from HP are usage-based leasing solutions—customers pay only for the actual resources that are used. In terms of virtualization, HP Utility Pricing solutions match expenditures directly to resources that are actually consumed by a specific IT service and can be activated only when they are really needed.

## Benefits of HP 9000 hard partitions (nPars)

### Increased system utilization

Businesses today require that multiple applications running on the same server deliver full performance without conflicting or interfering with one another. It is critical that the actions of one

application do not impact other applications or result in denial of service to them. Furthermore, a single failure in the operating system or hardware components on the server supporting the set of applications must not result in the loss of multiple applications running on that server.

nPars are the answer. With the HP 9000 Superdome nPars, each nPar is isolated from the others. The dynamically configurable hardware provides an effective firewall between applications so that there is virtually no possibility of interference between nPars. The result is increased system utilization with security—the perfect answer for system consolidation scenarios.

### **Increased flexibility via support of multiple environments**

Business requirements and competitive pressures can dictate the need for a large set of applications, both old and new. Often applications require different versions or revisions of operating systems, different system parameter settings, and even different patch levels. Performance can suffer from the interactions of applications with incompatible behaviors.

HP 9000 Superdome nPars can provide multiple independent environments for these applications. For instance, with nPars a portion of the compute and I/O resources can be configured to run an old copy of the operating system, while the rest of the machine is updated to run the newer software. Or you can use nPars to set up a test environment where new software and operating systems can be installed, leaving production applications to continue operation outside the test nPar without impact. For an even higher degree of flexibility, you can run virtual and resource partitions within an nPar—and change them on the fly—to even further improve your resource utilization.

The HP Integrity Superdome provides the added ability to run different operating systems, such as HP-UX, Linux, OpenVMS, and Windows, on the same HP Superdome platform.

### **Increased uptime**

nPars provide independence of failure or operation. If you have to reboot one nPar or take it down to perform an operating system upgrade, the other nPars in the system are not affected.

Unlike a traditional large symmetric multiprocessor, nPars allow a server to be configured into a cluster of independent systems. Since nPars provide failure independence, high-availability clusters consisting of two or more nPars can be configured within the same server to improve single-system availability. In addition, you can configure HP Serviceguard to fail-over applications from one nPar to another nPar within the same HP Superdome system or to another HP server (“Cluster-in-a-Box”).

### **Increased scalability**

Some applications may not have the ability to scale linearly with large numbers of CPUs. The HP 9000 Superdome’s nPars allow you to tailor the processor, memory, and I/O resources to match application scaling characteristics. What’s more, as new application or operating-system releases provide improved scaling characteristics, nPars allow the resources to be readjusted to boost application performance.

## **The HP 9000 Superdome: ease of management**

In spite of all its power and flexibility, the HP 9000 Superdome system is remarkably easy to manage, making your life easier with a host of management features and technologies. These include both system-level and workload-management tools as well as the HP 9000 Superdome Support Management Station—an external, PC-driven station that may be used to scan the HP 9000 Superdome.

### **System management features**

System-level and workload-management tools for the HP 9000 Superdome include the following:

- **HP Systems Insight Manager** is an easy-to-use multi-system management solution with Web-enabled and command line interfaces. HP Systems Insight Manager delivers multi-system access to all key system administration tools for fault monitoring, configuration, and workload management. HP Systems Insight Manager replaces HP Servicecontrol Manager. It is available for download from the Web now and will be included in the box soon.

HP Systems Insight Manager integrates with many other HP-UX-specific system management tools, including the following:

- **Ignite-UX** addresses the need for HP-UX system administrators to perform fast deployment for one or many servers. It provides the means for creating and reusing standard system configurations, enables replication of systems, permits post-installation customizations, and is capable of both interactive and unattended operating modes.
- **Software Distributor (SD)** is the HP-UX administration tool set used to deliver and maintain HP-UX operating systems and layered software applications. Delivered as part of HP-UX, SD can be of help in managing the HP-UX operating system, patches, and application software.
- **System Administration Manager (SAM)** is used to manage accounts for users and groups, perform auditing and security, and handle disk and file system management and peripheral device management. HP Systems Insight Manager enables these tasks to be distributed to multiple systems and delegated using role-based security.
- **HP-UX Kernel Configuration** is used for self-optimizing kernel changes. The new HP-UX Kernel Configuration tool allows users to tune both dynamic and static kernel parameters quickly and easily from a Web-based GUI to improve system performance. This tool also sets kernel parameter alarms that notify the user when system usage levels exceed thresholds.
- **Partition Manager (parmgr)** provides control and management of the HP 9000 Superdome nPars. It can be launched as a GUI or directly from the command line. It is aware of Instant Capacity and Pay-per-use processors. With it the user can display server status; create, delete, and modify nPars; display a complete hardware inventory; display the status of key server components; check for problems or unusual server conditions; and manage power to cells and I/O chassis.
- **Security Patch Check** determines how current a system's security patches are, recommends patches for continuing security vulnerabilities, and warns administrators about recalled patches still present on the system.
- **Event Monitoring Service (EMS)** keeps the administrator of multiple systems aware of system operation throughout the cluster and notifies the administrator of potential hardware or software problems before they occur. HP Systems Insight Manager can launch the EMS interface and configure EMS monitors for any node or node group that belongs to the cluster, resulting in increased reliability and reduced downtime.
- **Process Resource Manager (PRM)** controls the resources that processes use during peak system load. PRM can manage the allocation of CPU, memory resources, and disk bandwidth. It allows administrators to run multiple mission-critical applications on a single system, improve response time for critical users and applications, allocate resources on shared servers based on departmental budget contributions, provide applications with total resource isolation, and dynamically change configuration at any time—even under load.
- **HP-UX Workload Manager (WLM)**, a key differentiator in the HP-UX family of management tools, provides automatic CPU resource allocation and application performance management based on prioritized service-level objectives (SLOs). In addition, WLM allows administrators to set real memory and disk bandwidth entitlements (guaranteed minimums) to fixed levels in the configuration. The use of workload groups and SLOs improves response time for critical users, allows system consolidation, and helps manage user expectations for performance.
- **HP's Management Processor** enables remote server management over the Web regardless of the system state. In the unlikely event that none of the nPars are booted, the Management Processor can be accessed to power cycle the server, view event logs and status logs, enable console redirection, and more. The Management Processor is embedded into the server and does not take an I/O slot.

And, because secure access to the Management Processor is available through SSL encryption, customers can be confident that its powerful capabilities will be available only to authorized administrators.

In addition to the system and workload management tools listed above, the following HP OpenView products are available to manage the HP 9000 Superdome:

- **HP OpenView Operations Agent** collects and correlates OS and application events.
- **HP OpenView Performance Agent** determines OS and application performance trends.
- **HP OpenView Glanceplus Pak** shows real-time OS and application availability and performance data to diagnose problems.
- **HP OpenView Data Protector (Omniback II)** backs up and recovers data.
- **HP OpenView Network Node Manager (NNM)** management station automatically discovers, draws (maps), and monitors networks and the systems connected to them. The NNM will run on PA-RISC-based HP-UX servers. NNM helps network managers evaluate network performance, pinpoint problem sources, and proactively manage their networks and network availability.

Additional HP OpenView management tools, such as HP OpenView Operations, Service Desk, and Service Reporter, will be able to collect and process information from the agents running under HP-UX on the HP 9000 Superdome.

## The HP 9000 Superdome Support Management Station

The Support Management Station (SMS) is used to run HP 9000 Superdome scan tools. The HP 9000 Superdome uses a PC device as the SMS. Current HP 9000 Superdome users can continue to use the existing SMS device with the new HP 9000 Superdome running PA-8900 processors or with the HP Integrity Superdome, as long as the SMS software is updated. The scan tools are used to enhance diagnosis and testability of the platform throughout the HP 9000 Superdome development and manufacturing, and they are now available to HP field engineering organizations. This translates into better management, as well as faster and easier upgrades and hardware replacement.

### SMS installation

One SMS is installed per customer site (or data center) and connected to each HP 9000 Superdome platform via a private LAN. Ideally, the SMS is installed near the associated HP 9000 Superdome because the HP customer engineer will then be able to run the scan tools and immediately make any necessary hardware repairs. The physical connection from the platform is a private Ethernet connection.

The SMS can support two LAN interfaces:

- The dedicated connection to the systems to be supported
- The connection to interface with the customer's general LAN

These two LAN connections allow SMS operations to be performed remotely. More than one SMS can exist on a private LAN, but only one SMS should be actively using the LAN.

### SMS functional capabilities

Here's what the SMS can do:

- Allow remote access via customer LAN (no modem access)
- Be disconnected from the HP 9000 Superdome platform(s) without disrupting their operation
- Be connected to a new HP 9000 Superdome platform and be recognized by scan software
- Support multiple, heterogeneous HP 9000 Superdome platforms with scan software capability
- Scan one HP 9000 Superdome platform while other HP 9000 Superdome platforms are connected, without disrupting the operational platforms

- Run the scan software tools
- Run up to four scan processes concurrently, each on a different HP 9000 Superdome platform
- Support utility firmware updates

## The HP 9000 Superdome: HP-UX 11i v1

HP-UX 11 is HP's complete 64-bit UNIX operating environment that delivers significant scalability and performance for demanding applications. And the latest version of HP-UX is HP-UX 11i. When teamed with HP's leading server systems, HP-UX 11i provides the power of supercomputing at a fraction of the cost.

HP-UX 11i is ranked the #1 UNIX in the world according to D.H. Brown's 2002 UNIX Function Review. HP-UX 11i garnered the top spot in all five of the report's categories against competitive offerings. HP-UX 11i led the D.H. Brown scorecard in reliability, availability, and serviceability; Internet and Web application services; directory and security services; systems management; and scalability.

### Key features of HP-UX

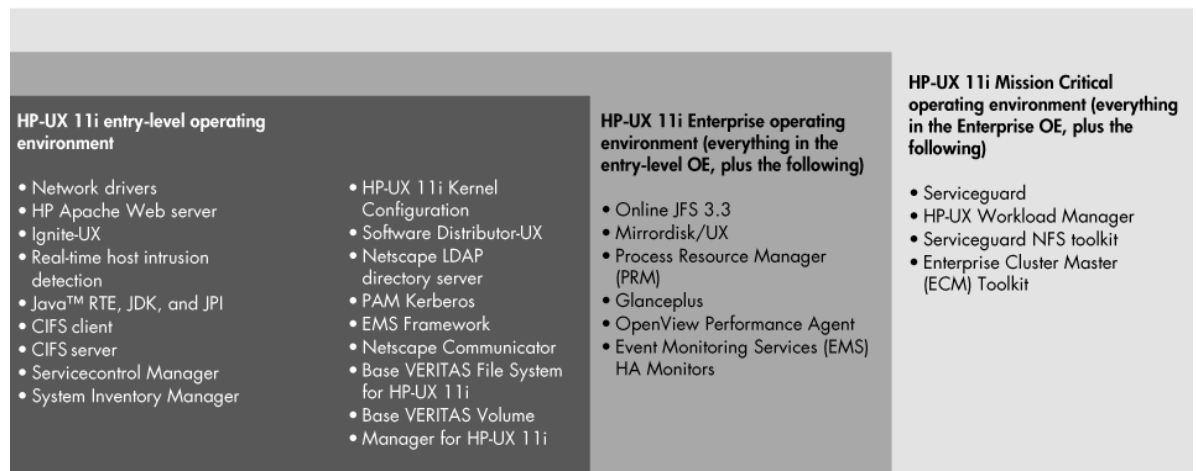
- Excellent compatibility with Windows and Linux
- Comprehensive Internet-critical functionality
- Ultimate performance and scalability, including 64-way SMP support for HP 9000 Superdome
- Industry-leading investment protection through binary compatibility
- HP 9000 Superdome management features, including partition management
- Support for hard and virtual partitions
- Excellent high-availability features, such as online addition and replacement of I/O and networking cards, dynamic patching, dynamically tunable kernel parameters, and up to 70 percent faster rebooting
- Forward compatibility of middleware and application software from HP-UX 11.0

HP-UX 11i delivers an end-to-end, scalable, manageable, and secure infrastructure for developing, deploying, and brokering mission-critical applications. It has features targeted at Internet and mission-critical applications, and it contains functionality for managing HP 9000 Superdome servers.

### HP-UX 11i operating environments

With HP-UX 11i, ordering and configuration have been vastly simplified. HP-UX 11i is available in three operating environments. Each consists of basic HP-UX plus a collection of layered software products addressing specific needs for availability, manageability, and security. These environments are simple to install and maintain, and they are a great value to purchase.

**Figure 8:** HP-UX 11i operating environments



### **HP-UX 11i entry-level operating environment**

The HP-UX entry-level operating environment offered with every server addresses key needs from a Web and small application perspective. And new HP Superdome-specific functionality, such as Partition Manager, is tightly integrated with the base HP-UX. The core functionality of the new virtual partition capability is also included.

### **HP-UX 11i Enterprise operating environment**

The HP-UX 11i Enterprise operating environment is targeted at larger application servers that are typically served by midrange systems such as the HP 9000 rp7420-16 server. Customers in this environment usually need to make the most of system resources and do proactive capacity planning with higher uptime, so the Enterprise operating environment includes tools such as Process Resource Manager, Glanceplus Pak, and Mirrordisk/UX.

### **HP-UX 11i Mission Critical operating environment**

The HP-UX 11i Mission Critical operating environment is highly recommended for most HP Superdome systems because it specifically addresses the needs of large application and database server environments. This operating environment includes host intrusion detection to protect the system from outside attackers, MC/Serviceguard for the highest degree of availability, and HP-UX Workload Manager for automatic resource allocation based on set service levels. Goal-based resource allocation is a must in today's mission-critical data center, and HP is the *only* UNIX vendor today offering such sophisticated resource allocation capability.

## The rest of the HP 9000 Superdome story

This white paper has barely touched the surface of HP 9000 Superdome's capabilities. Here are some other areas of interest.

### The Adaptive Enterprise and the HP 9000 Superdome

Because no HP 9000 Superdome stands in isolation, you can expect even more from HP. The HP 9000 Superdome can be part of an Adaptive Enterprise—one in which IT and business are synchronized to capitalize on change—meeting the needs of the total IT environment throughout the life of the solution.

The HP 9000 Superdome is delivered with a foundation configuration of services that ensure customers get the right configuration for their solution the first time—and in record time. The services begin with up-front assessments of business/IT alignment, IT technical skills, and environment and mission-critical readiness. Customers get a detailed architecture design based on their solution; factory pre-integration for software, peripherals, and middleware; and an HP solution manager assigned to ensure their satisfaction. They'll also be able to add mission-critical support services to meet the demanding stability and availability requirements so characteristic of the mission-critical data center.

### HP StorageWorks solutions: putting information to work

HP StorageWorks helps businesses put information to work. Our industry-leading portfolio of network storage solutions makes it easier for you to keep the right information at your fingertips, quickly adapt to change and exploit new opportunities, and increase information availability while radically cutting costs. For more information on StorageWorks, please contact your HP sales representative or visit [www.hp.com/storage](http://www.hp.com/storage)

## Conclusion

In this era of 24 x 7 applications and mission-critical operations, businesses need a bulletproof infrastructure. That infrastructure is powered by HP's industry-leading, high-performing enterprise servers and the robust HP-UX 11i operating environment.

The HP 9000 Superdome with PA-8900 processors represents another giant step forward for the HP Server line and a continued commitment to the PA-RISC product line.

- **Availability**—With industry-leading multi-system availability, single-system reliability, and superior hard partition resilience, the HP 9000 Superdome is designed to meet the needs of large enterprises, service providers, and established online economy companies.
- **Capacity**—The HP 9000 Superdome provides a computing utility infrastructure unmatched by any competitive offering. The HP 9000 Superdome's superior system and processor architecture provides scalability that outstrips the competition.
- **Connectivity**—The HP 9000 Superdome offers massive amounts of I/O connectivity and has high-speed connectivity, too.
- **Manageability**—HP's suite of management products is ready to make HP 9000 Superdome management easier. And with its dynamic capabilities as part of the HP Partitioning Continuum, the HP 9000 Superdome provides the best combination of system flexibility and convenient management.



## For more information

Looking for more information about HP 9000 Superdome?

For more information on the HP 9000 Superdome line of servers, go to:

[www.hp.com/products1/servers/scalableservers/superdome/index.html](http://www.hp.com/products1/servers/scalableservers/superdome/index.html)

© Copyright 2003, 2004, 2005 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel, Itanium, and Intel Itanium Processor Family are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Java is a U.S. trademark of Sun Microsystems, Inc. Linux is a registered trademark of Linus Torvalds. Windows is a U.S. registered trademark of Microsoft Corporation. Oracle is a registered U.S. trademark of Oracle Corporation, Redwood City, California. UNIX is a registered trademark of The Open Group.

5982-4000EN Rev. 2, 04/26/2005

