research

HP **rp7400** **hp servers**

System Architecture and Design Guide

This white paper describes the design goals and system architecture of the hp rp7400 server and provides recommendations for implementing business-critical solutions.

September 2001

Table of Contents

Introduction

Overview of the rp7400 Server

The next-generation, midrange rp7400 continues a long line of binary compatible servers from Hewlett-Packard. The rp7400 provides high-end performance at midrange prices, thus expanding HP's unrivaled midrange performance, price/performance, performance density, support, and system reliability.

The rp7400 is a symmetrical multiprocessing (SMP) server supporting up to eight high-performance, 64-bit PA-8700 or PA-8600 processors. The rp7400 uses state-of-the-art system architecture, providing the benefits of multiple fast, independent buses. The design provides leadership in multiprocessing, memory, and I/O performance, along with superior scalability and reliability. And with online replacement and addition of PCI I/O cards, disks, N+1 power supplies and cooling, this next-generation server provides a quantum leap in availability and serviceability.

Figure 1. Internal View of Major rp7400 Features
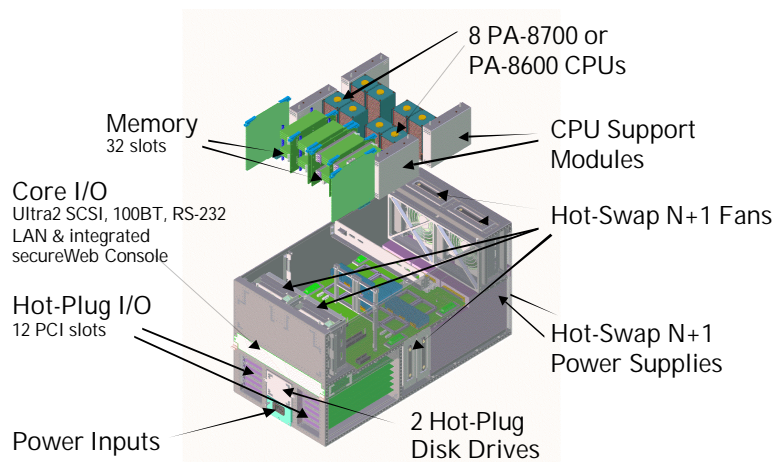
rp7400 Features At-a-Glance

- 1 to 8 64 bit CPUs, PA-8700(650MHz & 750MHz) or PA-8600 (550MHz)
- Up to 32GB of memory
- 12 hot-plug (66MHz × 64-bit) PCI I/O slots
- Independent channels for I/O slots
- N+1 power and cooling
- 2 hot-plug disk drives
- 4.3GB/s system bus bandwidth
- 6.4GB/s I/O bus bandwidth
- 8.5GB/s memory bus bandwidth
- 32-way memory interleaving
- Supported by the 64-bit HP-UX 11 & 11i operating system
- High-density 10 EIA unit, 19" rackmount package

The HP Enterprise Server Product Line

The rp7400 is a successful member of the HP Enterprise Server product line from HP and addresses the major computing challenges customers face today in online transaction processing (such as ERP, CRM, and SCM), electronic commerce and Internet/intranet serving, technical applications, business intelligence, and data warehousing.

Supporting the Enterprise Server product line are the hp server A-, L-, rp7400, and SuperDome platforms, shown in Figure 2. Through the use of common HP architecture components such as HP's PA-RISC processor, these platforms provide complete binary compatibility and application portability. The systems fit neatly into existing environments and provide compute and file services for mission-critical OLTP, e-commerce, technical, data warehousing, and server consolidation applications.

HP Enterprise Server systems feature:

- **Leading uni-processor performance**—PA-RISC processors always among the fastest in the industry.

- **Unmatched scalability**—providing consistent price/performance over the entire range and protecting customer investments in both hardware and software.

- **Investment protection through complete binary compatibility**—up and down the product line, permitting users to select the appropriate platform at the appropriate price-point without concern for application availability. This is because all systems, regardless of the performance levels, present the same application programming environment, which greatly increases the number of compatible "off-the-shelf" third-party applications and reduces porting and development costs.

Figure 2. The Industry's Most Scalable Range of UNIX servers
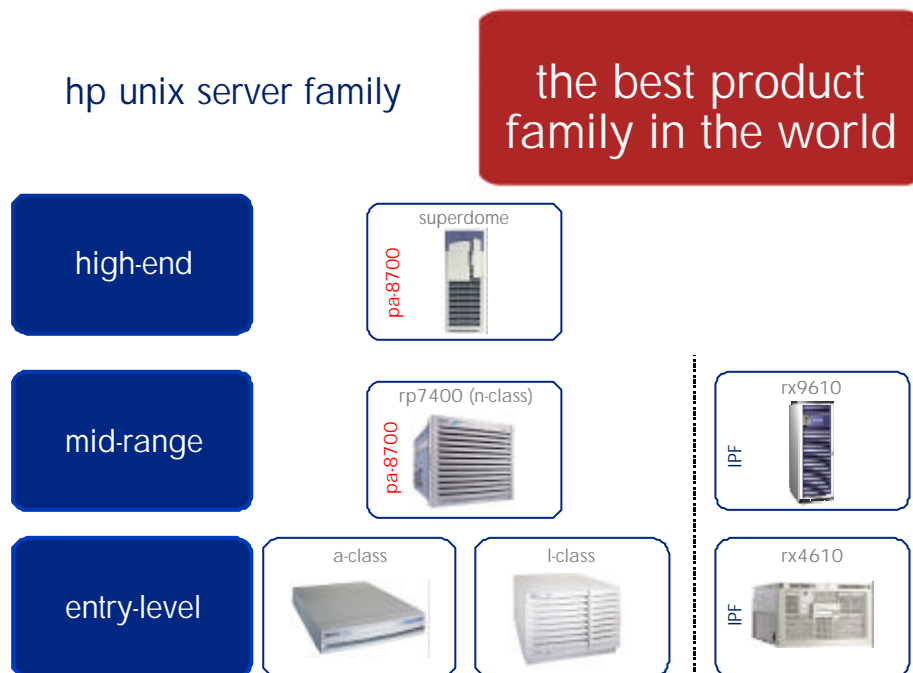
Performance

The design philosophy for the rp7400 was simple: lead in all areas of performance, scalability, manageability, and reliability.

The architecture is a uniquely balanced design, providing huge amounts of processor, memory, and I/O bus bandwidths. Unlike many systems, these huge bandwidths are coupled with very low latencies between CPUs, memory, and I/O. This "no-compromise" balance of bandwidth and low latencies provides leadership performance in any workload, as proven by the open industry benchmarks shown in Table 1 below. With this ability to adapt and still deliver high performance, the rp7400 provides customers with the confidence that their server can easily be re-deployed.

The balanced performance is provided by:

- Up to eight high-speed PA-8600 or PA-8700 processors
- 4.3GB/s system bus bandwidth, shared across two system buses
- Up to 6.4GB/s aggregate I/O bandwidth, shared across 24 266MB/s I/O channels
- Up to 8.5GB/s memory bus bandwidth, shared across up to four memory buses

Table 1. rp7400 Benchmark Results

hp rp7400 Enterprise Server Performance								
	550MHz			650MHz, Estimated			750MHz, Estimated		
	1-way	4-way	8-way	1-way	4-way	8-way	1-way	4-way	8-way
SPECint_2000	379			465			551		
SPECfp_2000	369			446			522		
SPECint_rate2000		17.0	32.7			40		24.6	46.7
SPECfp_rate2000		14.4	23.0			28		20.5	32.1
OLTP[1]	1.0	1.81	2.92	1.15	2.08	3.29	1.25	2.27	3.58

[1] HP relative performance across multiple OLTP workloads as compared to 550Mhz 1-way

Design Philosophy—Balanced High Performance

From the outset, the design philosophy for the rp7400 was to create a no-compromise, easy-to-understand, high-performance server with lasting value. Therefore, unlike other servers, the rp7400 does not force customers to make tradeoffs between CPU and memory capacity or CPU and I/O capacity, as illustrated in Figure 3. Here are some of the performance features that have resulted from this philosophy:

Simultaneous support of maximum CPU, memory, and I/O capacities—no system slot trade-offs

- High-performance, 64-bit PA-8600 (550MHz) and PA-8700 (650 & 750MHz)
- Sync-DRAM memory
- 64-bit × 66MHz PCI
- Huge CPU, memory, and I/O bus bandwidths to reduce system performance bottlenecks
- Multiple independent buses to eliminate bandwidth contention
- Robust and scalable HP-UX 11.0 & 11i operating system

Figure 3. HP's No-Compromise Configuration

Maximum Configurations

8 CPUs
32GB of memory
12 PCI cards
all at the same time

12 PCI Slots

32GB of Memory

No
Tradeoffs

8 CPUs
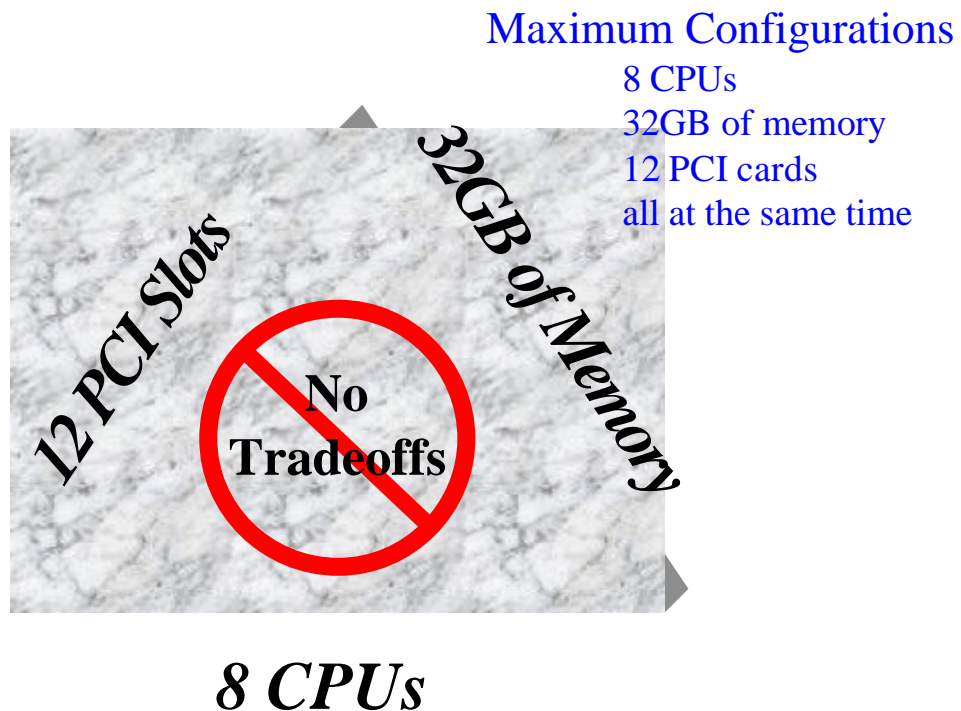
Binary Compatibility

The rp7400 server supports the 64-bit HP-UX 11i & 11.0 operating systems. With HP-UX 11i & 11.0, HP maintains its long-standing tradition of providing the industry's best record of investment protection. HP-UX provides forward binary compatibility, in which a fully bound application developed on an earlier version of HP-UX is ensured to run smoothly on HP-UX 11. Thus, current 32-bit and 64-bit applications can run without requiring recompilation.

Architecture

 Figure 4 shows the relationship of the rp7400's main blocks and the buses that connect them.

- **Memory controller.** The memory controller is central to the architecture and is supported by up to four memory carriers and from 1 to 16 pairs of memory DIMMs. With 1GB and 2GB DIMMs, main memory ranges from 1 GB to 32GB. The memory controller consists of three VLSI chips and handles all transactions between main memory and system buses.

- **System buses.** Each system bus connects the memory controller to two runway bus converters and one I/O controller. Each bus converter is a single VLSI chip supporting two PA-8600 or PA-8700 CPUs via their own dedicated Runway bus. In rp7400s configured with PA-8600 CPUs (550MHz) & PA-8700 CPUs (650 & 750 MHz), the system bus runs at 133MHz

- **I/O controllers.** Each I/O controller consists of a master I/O controller and six to eight slave I/O controllers. Each PCI Turbo and Twin Turbo slot uses a single slave I/O controller while the core I/O utilizes two slave I/O controllers. The core I/O is a multifunctional card with the commonly needed I/O functions including Ultra2 SCSI, 100Base-T LAN, RS-232, integrated secure web console and console connections.

Figure 4. rp7400 Architecture
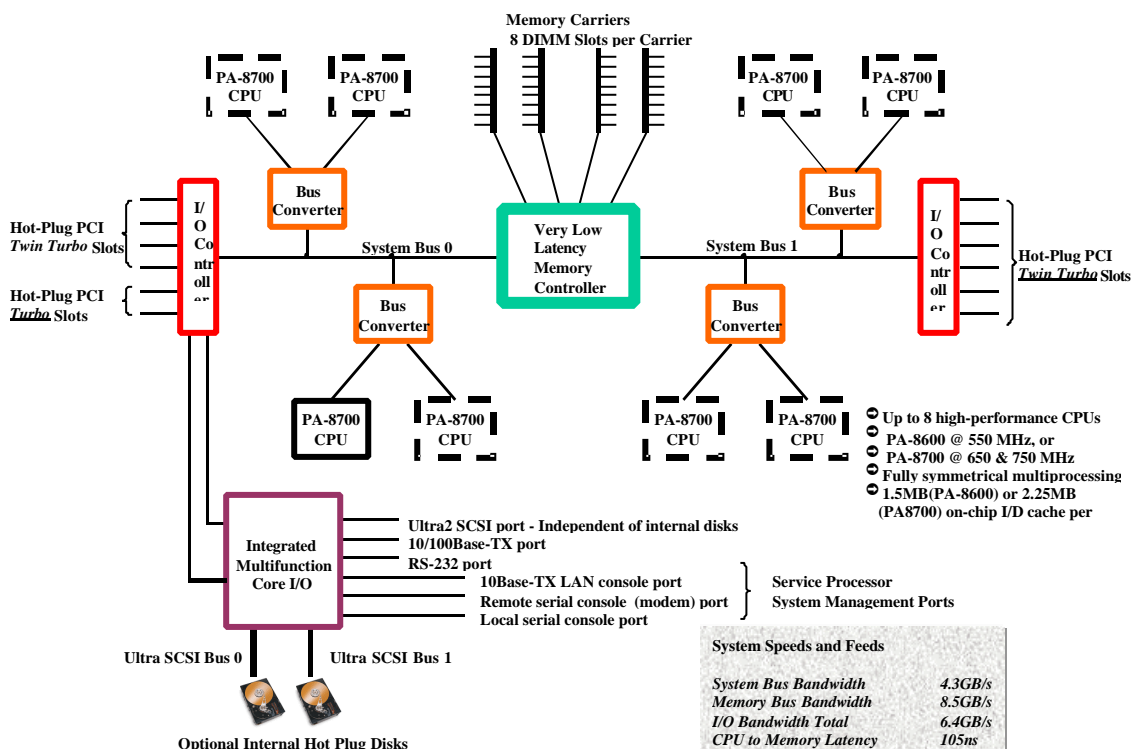
Twin system Bus

Central to the rp7400 system are twin system buses. All processor, memory, and I/O traffic transfers over one of these two buses. Both the number of buses and the choice of bus are significant for the performance and lasting value of the rp7400 system.

Advantages of Twin Buses

An eight-processor system could be built around a single bus, but only by sacrificing performance. For example:

- **Competition for resources.** Having eight processors share a single bus has several negative impacts on system performance. On a single-bus system, processors compete for access to the single critical resource, the bus. Contention over this bus ends up adding directly to the memory latency. Further, the limited bandwidth of the bus must be divided among the total number of processors on the bus.

- **Bus frequency.** An equally important problem is bus frequency. Physical length and the number of loads on the bus determine the maximum bus frequency. Many processors on a single bus means the bus will be long and have many loads. This translates into a bus that runs slowly. This vicious cycle of adding processors to a single-bus system can quickly limit system performance. Each extra processor causes bus frequency to drop and further divides the limited bandwidth. A single-bus, eight-processor system could be designed for a single processor. However, upgrades to next-generation processors would not reach full potential with performance left unrealized due to limitations in the system.

However, multiple buses can also cause problems. Having too many buses increases cost and quickly makes the system overly complex. Although each bus frequency may be high, overall memory latency is dominated by the response time of the worst-case bus/processor. For all of these reasons, the rp7400 was designed around two independent system buses. This provides the best performance now with sufficient headroom in the system to fully utilize the next generations of processors. Simply put, two buses are necessary for an eight-processor system for world-class performance now and in the future.

Advantages of the system Bus

The system bus was chosen as the central system bus for several reasons. The bus is state-of-the-art both in features and performance. Many of its features were designed to provide a world-class solution to the dominant challenges of maintaining reliability and data integrity, performance, and lasting value of the system.

- **Reliability.** Reliability is ensured through the use of error checking and correcting (ECC), which ensures data integrity, not simple parity. This allows continual operation through transient single-bit bus errors or even a failed bit on the bus.

- **Performance.** A limited number of electrical loads and short physical distance allow very high bus frequency for outstanding performance. Combined with a wide data transfer, the bandwidth available is sufficient not only for the current generation of processors, but also for several more to come. Independent address and data buses help provide very low memory latencies. Read data can be returned to a processor while cache coherent snoop queries continue with uncontested access to the address bus.

  In terms of raw performance, compare the last generation of HP servers with the new rp7400 server. The HP K-Class server was designed around the Runway bus with a total raw bandwidth of approximately 960MB/s. Each bus in the rp7400 has a raw bandwidth of approximately 2.1GB/s, combining to a total of 4.3GB/s.

Very Low Latency Memory Access

The rp7400 Main Memory Subsystem consists of the main memory controller, one to four memory carriers, and two to thirty-two DIMMs loaded in pairs. The memory subsystem is implemented with highly parallel address and data paths. These paths are driven and controlled by HP-designed memory controller VLSI chips, consisting of one address controller and two data controllers.

Within the main memory subsystem, the address controller drives the address paths. To increase the parallelism, the address lines are buffered three times: once on the system board to drive each memory carrier, once on the memory carrier to drive banks of DIMMs, and again on each DIMM before driving the memory components.

The time taken for data to transfer between CPU and memory after a read or write request is critical to overall system performance. Long CPU-to-memory latencies can cause the CPU to stall and suffer huge performance penalties. The rp7400 memory subsystem is designed with an extremely low latency of 105ns/135ns. This is less than half that of the K-Class architecture, which is still a very competitive design.

Each of the four memory carriers has a separate 64-bit-wide data path. The two memory data controllers both provide two of these data paths and perform data multiplexing/de-multiplexing. Each memory carrier handles the data path with two VLSI chips, which buffers the data for the memory DIMMs. DIMMs use either 512 MB or 1024 MB SDRAM (Synchronous Dynamic Random Access Memory) memory technology.

Minimum and Maximum Memory Configurations

The rp7400's minimum memory is 1024MB, configured as a pair of 512MB DIMMs on a single memory carrier. In this configuration, the theoretical maximum bandwidth is 1.92GB per second with an average sustainable bandwidth of 1.09GB per second. To achieve the maximum memory bandwidth, install four memory carriers and distribute DIMMs evenly across each carrier.

Today, the rp7400's maximum memory consists of thirty-two 1GB DIMMs across four memory carriers. This represents an impressive 32GB. In this configuration, the theoretical maximum bandwidth is 8.512GB per second and the average sustainable bandwidth is 4.8GB per second. This represents 50 percent more bandwidth than the system buses; hence, the rp7400 has extra memory bandwidth to support CPU upgrades and provide lating value to investments in memory DIMMs.

Speeds and Feeds

Table 2 shows the theoretical maximum bandwidth for various system buses. This is defined as the bus width multiplied by the frequency and number of buses.

Table 2. Maximum Bandwidth for RP7400 System Buses

PA-8700/8600 Systems	# of Buses (or Controllers)	Maximum Bus Bandwidth	Aggregate Bus Bandwidth
Twin Turbo PCI slots	10	532MB/s	5.32GB/s
Turbo PCI slots	2	266MB/s	532MB/s
I/O subsystem	2 (controllers)	3.19GB/s	6.38GB/s
System buses	2	2.13GB/s	4.26GB/s
Memory subsystem	4	2.13GB/s	8.51GB/s
CPU buses	8	2.13GB/s	17.0GB/s

Although theoretical maximum bandwidth is an interesting calculation for comparing systems, sustainable bandwidth is more representative of a system's capability. Sustainable bandwidth will vary depending on workload and control algorithms. Sustainable bandwidths for various system buses are given in Table 3.

Table 3. Sustainable Bandwidth for rp7400 System Buses

	Sustainable Aggregate Bus Bandwidth in PA-8700 and PA-8600 Systems
I/O subsystem—average	2.2GB/s
Inbound	1.6GB/s
Outbound	2.8GB/s
System buses	3.4GB/s
Memory subsystem	4.8GB/s
CPU buses	13.6GB/s

All buses and controllers have design margin, which HP may choose to exploit for faster product upgrades.

I/O Subsystem Design

Figure 5 shows the design of the rp7400's I/O subsystem. There is one master I/O controller chip on each of the two system buses. Each master I/O controller has 12-byte-wide buses to the slave I/O controllers. Each Turbo slave I/O controller uses one bus, while Twin Turbo slave I/O controllers use two buses. Therefore, each PCI slot has its own slave I/O controller and its own interface through a master I/O controller to one of the system buses.

Each PCI slot has hot-plug capability. With the assistance of the system's software (enabled with next release of HP-UX 11) and firmware, the slave controller and associated hardware can allow any PCI card to be safely replaced or added while the operating system is running.

All rp7400's PCI slots are keyed to accept 5-volt and Universal PCI cards; therefore, each slot is compatible with all available PCI cards. The rp7400 uses HP-developed Adaptive Signaling technology to allow each slot to operate at the full speed of the PCI card. The speed capability of PCI cards is sensed and the signaling speed set to the fastest possible for each card.

Figure 5. rp7400 I/O Subsystem
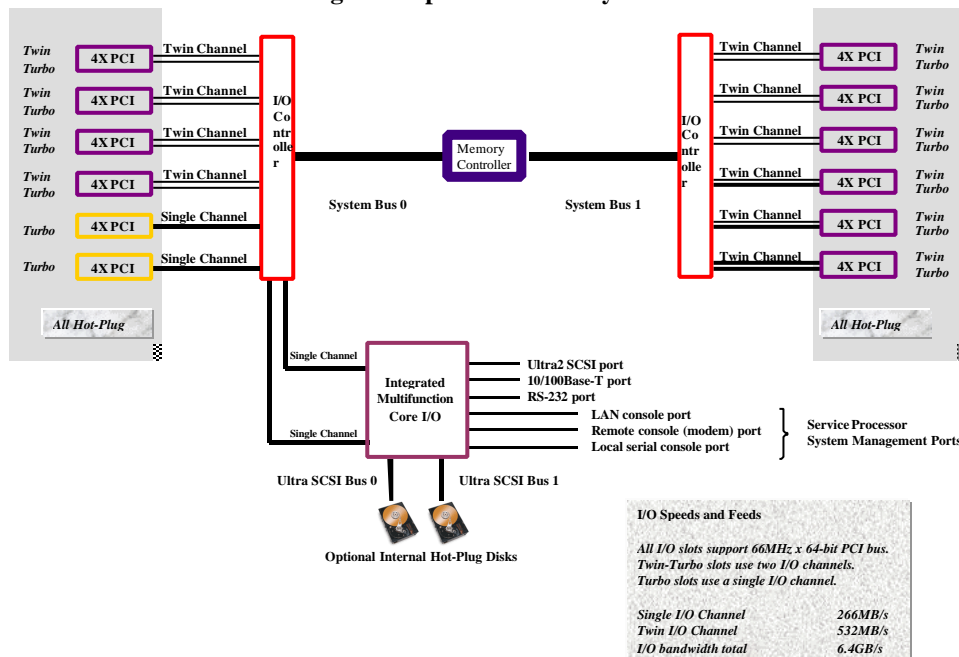
Independent PCI I/O Buses

PCI is the optimized, industry-standard I/O bus. As shown in Table 4, the rp7400 system has twelve PCI slots. Two of the slots are designated as Turbo slots and the other ten are designated as Twin Turbo. The core I/O uses the equivalent of two additional Turbo PCI slots.

Table 4. rp7400 PCI I/O

PA-8700 and PA-8600 Systems	# of Slots	Hot-Plug	Bandwidth per Slot	Bus Width	Signaling Speed	Slot Keying	Adaptive Signaling
TwinTurbo	10	Yes	532MB/s	64 bits	66MHz & 33MHz	5 volts	Yes
Turbo	2	Yes	266MB/s	64 bits	66MHz & 33MHz	5 volts	Yes

The rp7400 is the first application of HP's high-performance I/O, with one I/O card per PCI bus. This is a key differentiator for the rp7400. Exceptional single-card I/O performance together with outstanding aggregate system I/O performance with multiple I/O cards is a highlight for the rp7400 compared to systems from all other vendors.

In fact, the rp7400 is a substantial improvement over current midrange platforms. For example, the K-Class offers two HP-HSC+ buses and five HP-HSC+ card slots with an aggregate peak I/O bandwidth of approximately 608MB per second. The rp7400's I/O subsystem implementation supports 6.4GB/s of peak bandwidth in 12 PCI slots, plus the core I/O. The rp7400 is designed with a modular split I/O controller architecture, enabling easy enhancements to higher-performance I/O buses when they become available. The physical system layout allows simple in-field upgrades of the I/O backplanes without replacing expensive system boards.

The high I/O bandwidth available in each PCI slot permits the possibility of external I/O expander bays based on PCI bridges. Though not available now, the rp7400 is easily capable of expansion like this in the future.

Scalability

rp7400 is designed without tradeoffs in CPU, memory, or I/O expandability to offer the best scalability in the market.

- **CPU upgrades.** From an entry-level configuration of one CPU scaling in single-CPU increments up to eight PA-8700 (650MHz & 750MHz) or PA-8600 (550MHz) 64-bit processors, the rp7400 offers great flexibility to cover a wide range of performance points.

- **Memory upgrades.** The rp7400 memory subsystem is also designed for scalability. With 16 increments of 1GB or 2GB board pairs and with no dependencies between CPU, memory, and I/O loading, the rp7400 is designed for great price/performance from minimum to maximum configurations. With the 2GB memory board pair, the rp7400 supports a maximum of 32GB of memory. That's enough memory for even the most demanding applications running on an eight-processor system.

The HP-UX operating system ensures that the rp7400 system hardware scalability translates directly into useable system performance scalability, proven by leadership performance from the 64-processor SuperDome server. HP-UX is the state-of-the-art UNIX operating system, preserving your investment by ensuring that you get the most performance out of each added processor.

High Availability

System availability is the percentage of time the system is up during a year. The formula reflects unplanned hardware and software downtime, but not external causes such as an operator's choice to take the system down. Downtime is the average number of downtime events times the average time to bring the system back online. The formula for determining availability is:

$$\text{Percent Availability} = 100 * \{8766 \text{ hours} - \text{annual downtime in hours}\}/8766$$

Typical systems classified as highly available achieve about 99.8% total system availability (18 hours of downtime per year) with 99.95% hardware availability (4 hours of hardware downtime per year).

The rp7400 has raised the competitive bar for 8-way midrange UNIX systems by introducing advanced high availability features. These features, which improve the availability level of the total system, are introduced in this chapter.

Redundant, Hot-Swap Power Supplies

HP power supplies have a long history of excellent reliability, and the redundant power supply option increases HP's commitment to even higher reliability and availability.

The rp7400 uses three 1220-watt power supplies in a hot-swappable N+1 redundant configuration. The rp7400 requires two power supplies to support its maximum configuration of CPUs, memory, I/O, and disks. The third power supply adds redundancy so that any one power supply can fail and the rp7400 will continue to operate. Chassis codes indicating the failure will be sent to the console. In the unlikely event that two supplies fail, the remaining supply will try to sustain the rp7400's operation. In this situation, the rp7400's configuration, workload, environmental conditions, and AC input will greatly influence the uptime.

Exchanging a power supply in a running system involves removing the rp7400 front cover and EMI filter. The failed power supply is easily identified and removed. The power supply is exchanged with a good one and this procedure is reversed. The system will log another chassis code to indicate that redundancy is re-enabled. It is that simple.

With this redundant, hot-swap power supply feature, there are several new and unique advantages to using the rp7400 server. For example, the rp7400 can be purchased with only two power supplies during application development, and then the third supply can be added later to ensure that production and mission-critical applications have power redundancy. This approach defers the cost of the redundant power supply until it is needed.

This is another advantage for those customers with rigorous preventive maintenance programs. While the rp7400 continues to operate, the power supplies can be removed one at a time so that built-up dust can be vacuumed using proper ESD (electrostatic discharge) procedures.

Redundant Power Input Protection

Figure 6 contains a diagram of the rp7400 power subsystem. This section explains how customers can utilize these capabilities to achieve different levels of power input protection.

Figure 6. rp7400 Power Subsystem
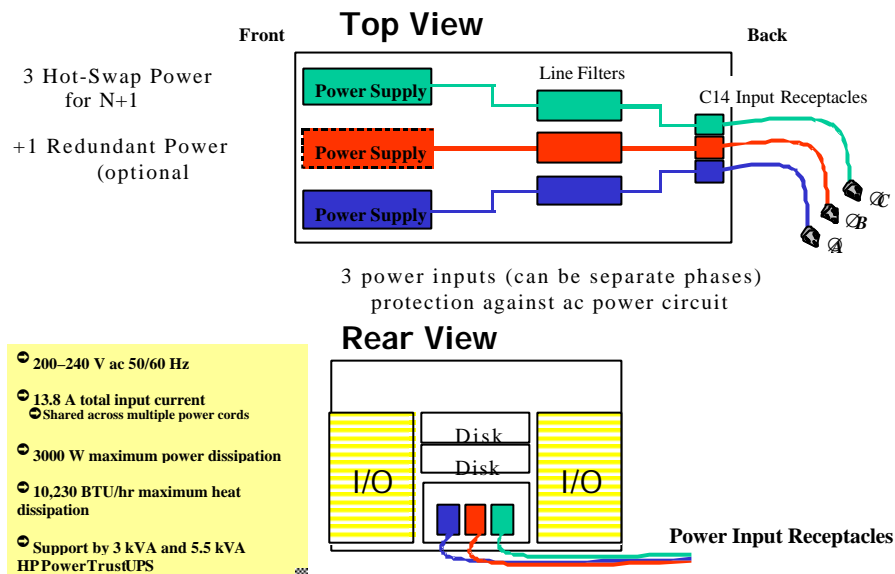
The rp7400 has three ac input line cords to greatly reduce single points of failure. Each line cord supplies power to one of the three internal power supplies. The system is designed to operate on nominal 200, 208, 220, 230, or 240 V ac rms and 50Hz or 60Hz power without line select switches. Each power supply can draw up to 6.9 A.

Because the rp7400 will continue to operate with two of the three supplies functioning, many possibilities exist for the customer to configure the ac input, depending on the level of protection desired.

- If the site has very stable ac power, all three line cords could be plugged into the same power grid. For additional protection, a single uninterruptible power supply (UPS) could be utilized to supply power to all three cords should primary ac power fail.

- The next higher level of protection is to have three-branch ac circuits, one for each ac input. This reduces the dependencies on single-point breaker failures and common wiring. Additional protection to this configuration would utilize three smaller UPSs.

- The highest level of protection is to have three electrical utilities each supply a branch circuit. This approach is expensive but does greatly reduce single points of failures. Large sites with many rp7400 systems may find this configuration cost-effective. For ultimate protection of large sites, install a large UPS on each branch circuit.

For a total power protection solution, include external disks, modems, LAN infrastructure, and other related equipment.

The redundant power input gives the rp7400 additional flexibility in the customer's environment. While maintaining, optimizing and expanding computer room environments, there is often a need to move and untangle cabling or plug the line cords into different ac outlets. With the redundant third power supply and line cord, the rp7400 allows one line cord at a time to be unplugged and rerouted. By relocating the third line cord, a UPS and its branch ac circuit can be taken off online for preventative maintenance or reconfiguration. Small changes like these to a computer room setup can often defer more intense shutdowns, which impact the end-use applications.

Redundant, Hot-Swap Cooling

The rp7400 server utilizes eight large fans to provide cooling that is more than adequate for the rp7400 and all foreseeable product upgrades. In fact, any one fan can fail and the server will continue to operate with margin. This is made possible via a unique fan-speed control, fan monitoring, and highly reliable heat-sink technology, which is the result of years of experience and research at HP with cost-effective air-cooled servers.

The rp7400 monitors ambient temperature and the power consumed within the box to determine the desired fan speed. The actual speed is determined by sensing the tachometer outputs from each fan. Digital Phase Locked Loop (DPLL) circuitry is used to adjust the speed of each fan individually to the desired common speed.

This smart algorithm reduces unnecessary fan noise, power consumption and wear, while producing a very clear indication of a working cooling subsystem. In the unlikely event of a fan failure, that fan will drop out of lock with the DPLL. The rp7400 signals a fan failure via chassis codes to the console and will light an LED on the failed fan assembly. Since all eight fans are the same, HP field service personnel or the system administrator needs just one common fan to repair any fan failure, while the rp7400 continues to operate with no downtime.

This is another advantage for those customers with rigorous preventive maintenance programs. While the rp7400 continues to operate, the fans can be removed one at a time so that built-up dust can be vacuumed using proper ESD (electrostatic discharge) procedures.

Main Memory—Advanced ECC and Parity

Data stored in the rp7400's main memory is protected by error checking and correcting (ECC) and address/control parity. The rp7400 ECC design will tolerate an arbitrary failure of a single SDRAM on each DIMM pair. This functionality, combined with memory scrubbing and memory page deallocation, will allow the rp7400 to tolerate typical hard single-bit SDRAM failures without requiring DIMM replacement. In fact, the design is robust enough to tolerate a single SDRAM failing on each DIMM pair.

The data controllers generate ECC bits and store these ECC bits with the data in the DIMMs. The 512MB and 1GB DIMMs use x4 SDRAMs to store each bit of a word, including its ECC bits, in a different SDRAM within the DIMM pair. When reading the data back, the data controllers are able to detect and correct single-bit data errors. Double-bit errors are also detected but are not correctable. Double-bit data errors are highly unlikely because the data and ECC bits are stored one bit per SDRAM and multiple SDRAMs would have to be involved in the error. Hence, a single SDRAM could fail within each DIMM pair and the system would still function.

The system also detects address and control parity errors to prevent data corruption from reading or writing to the wrong location in main memory. The address controller and each address buffer generate address and control parity. Each address buffer detects address and control parity problems and reports it back to the address controller. There are three levels of address buffers as the address lines fan out. These address buffers are located on the system board, on each memory carrier on each DIMM.

Hot-Plug Disk Drives

The rp7400 has two embedded SCSI disks accessible from the back of the server. These disks can be removed and inserted while the rp7400 continues to operate. This operation is called "hot-plug," and it is different from "hot-swap."

During both hot-plug and hot-swap operations, the power remains on and the system continues to function. However, hot-swap means that the assembly can be removed, added, or replaced without informing the system. Hot-plug requires the assembly to be de-configured before removal and re-configured before the system can utilize the newly inserted assembly. Because disks have unique information stored on them, hot-plug methods are used. Fans and power supplies in the rp7400 are hot-swap assemblies.

The rp7400 contains circuitry to properly control the disk's power and reset during the hot-plug operation. Either SAM (System Admin Manager) or online diagnostic software (MESA) can be utilized to effectively de-configure and re-configure the disk.

This is another advantage for those customers with rigorous security programs, which require disk mechanisms to be completely removed and isolated in a disaster- or theft-safe environment.

Hot-Plug PCI I/O Slots

The rp7400's ability to hot-plug PCI cards offers excellent flexibility for adding, reconfiguring, and maintaining I/O functions while the rp7400 continues operations. No reboot is required. This section explains how to take advantage of this capability.

In order to optimize performance and packaging, the rp7400's PCI expansion I/O was split into two separate bays. Each of these bays supports up to 6 PCI cards. Access to the I/O slots is achieved by sliding the rack-mounted rp7400 backwards approximately 15 inches. Special features on the rp7400, along with custom rack rails, allow the unit to move safely and smoothly during online service with all cables still attached. Once the system is slid into service position, the left and/or right I/O bay covers can be removed to gain side access to the PCI cards.

The PCI boards are spaced on a .9" pitch to allow for special hot-plug features and increased PCI reliability. Extra airflow holes between bulkheads have more than doubled PCI airflow. Between the PCI slots, PCB separators prevent electrical shorting and exposure to hazardous energy during hot-plug installation and removal. Locking features are designed into the main chassis to eliminate the need for individual PC board bulkhead screws, thus eliminating a potential electrical safety hazard.

The hot-plugging of I/O cards has both hardware and software elements. The hardware requirements are met by the electronics on the I/O backplanes and by mechanical design in the I/O cardcage. Bus idling, slot-to-slot electromechanical isolation, per-slot power and reset control, and visual indicators are all components of the total hot-plug hardware solution. With associated software, any single I/O card in any PCI I/O slot can be removed, replaced, or added without power cycling, rebooting the system, or impacting the operation of other I/O transactions.

Dynamic Processor Deallocation and Resilience

Every multi-CPU rp7400 with properly loaded HP-UX 11.0 or 11i have the capability for Dynamic Processor Deallocation and Resilience. Incorporated into HP's version 11.0 & 11i of HP-UX is the capability to take a processor out of service while the system is running without interruption to applications. This technology is referred to as Dynamic Processor Deallocation. Once a processor is deallocated, the HP-UX operating system will migrate all application processes that are currently scheduled on that processor to other active processors. Note that if the processor has been assigned to handle interrupts for any I/O drivers, it will continue to do so while it is deallocated.

The HP PA-RISC processors have the ability to detect and correct single-bit cache errors. The embedded EMS (event monitoring service) monitors the rate of correctable errors in each PA-RISC processor's on-board cache. These errors are manifested as Low Priority Machine Checks (LPMCs). While occasional correctable errors are to be expected in the on-board cache, too many of these errors in a short period of time indicate an increased likelihood that a non-correctable cache error could occur. The EMS LPMC monitor will continuously monitor the rate at which LPMCs are occurring and dynamically deallocate a processor, using the Dynamic Processor Deallocation facility. This technology is referred to as Dynamic Processor Resilience.

Quality by HP Design

HP servers have a long history of excellent reliability. Each new product and upgrade is designed, verified, and manufactured by continuously improved processes. The rp7400 continues this tradition by increasing HP's commitment to even higher quality levels.

The rp7400 is expected to achieve greater than 99.8% single system availability (less than 18 hours of downtime per year) with greater than 99.95% hardware availability (less than 4 hours of hardware downtime per year). Downtime is the average number of downtime events times the average time to bring the system back online. This high level of availability requires very high quality levels in order to reduce the number of downtime events and make it possible to recover quickly from these events. Quick recovery times are dependent on preventing data corruption faults from propagating and requiring recovery from the last backup.

The features designed into the rp7400 to prevent downtime events include:
- Redundant, hot-swap power supplies
- Redundant power input protection
- Redundant, hot-swap cooling fans
- Main memory error correction of single-bit errors
- Hot-plug internal disk drives
- Hot-plug I/O PCI slots
- CPU cache error correction of single-bit errors

External uninterruptible power supply support includes:
- Online firmware updates
- HP MC/ServiceGuard support

Data corruption prevention features in the rp7400 are:
- Detection of address parity errors in main memory
- Parity and/or error correction on all system buses
- Memory scrubbing and page de-allocation routines
- Internal power and temperature checking

The following features reduce the rp7400's recovery time:
- Minimized boot times
- Remote console support
- Automatic response center notification
- Automatic boot through alternate console path
- Fault isolation to the replaceable subassembly with offline diagnostics
- Error logging with access via offline error log analysis tools
- Automatic de-allocation of processors and memory that fail self-test

CPUs

The rp7400 uses PA-8700 and PA-8600 processors with three available performance points based on the speed of the processor. The first speed option uses a PA-8600 running at a core frequency of 550MHz. In the second option, a PA-8700 processor core runs at 650MHz, for an increase in raw compute performance of approximately 12 percent. The highest-speed version runs at 750MHz, for another increase in raw compute performance of approximately 23 percent. In-field upgrades are easily performed by replacing one processor type with another.

PA-RISC—PA-8600 and PA-8700

The PA-8600 and PA-8700, illustrated in Figure 8, are the newest members of the PA-RISC family of processors. The design is based on the PA-8500 processor but is implemented in a
0.18-micron (PA-8700) process. The new process allows the large first-level caches to be moved on-chip so that the frequency can be boosted without the need to add tightly coupled second-level caches. Improvements have also been made in the branch prediction hardware to allow a single branch prediction structure to seamlessly take advantage of static and dynamic branch prediction techniques. These improvements and others allow the PA-8600 and PA-8700 to deliver industry-leading application performance.

Figure 8. Die Photo of PA-8600/PA-8700 CPU
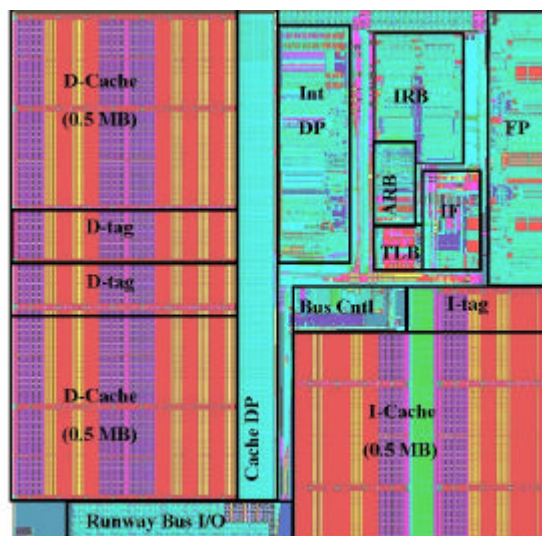
PA-8700 and PA-8600 Design

The design goals of the PA-8700 are to:

- Lead the industry in application performance.
- Provide full binary compatibility with all existing PA-RISC binaries.
- Extend the robust PA-RISC design to even higher reliability levels.
- Deliver maximum performance on binaries tuned for the PA-8600.

The PA-8700 design builds upon the foundation established in the PA-8600. With its large complement of functional units, dual-ported data cache, and sophisticated instruction reordering hardware, the 64-bit PA-8600 effectively exploited instruction-level parallelism to set new standards of performance. The PA-8600

design improved on this by increasing the sizes of the caches, Branch History Table (BHT), and Translation Lookaside Buffer (TLB), as well as increasing operating frequency.

To deliver higher levels of performance, the PA-8700/PA-8600 design emphasizes increased frequency over improvements to Cycles Per Instruction (CPI), making the most of the infrastructure established in the PA-8600. Higher frequency is achieved by implementing the PA-8700 in a 0.18-micron process, which allows the inclusion of enough SRAM to implement effective first-level on-chip caches. Although priority has been given to increasing the frequency of the PA-8700 and PA-8600, CPI improvements have been incorporated. Most notable is an enhancement to the branch prediction hardware which allows static and dynamic branch prediction techniques to be combined seamlessly.

Primary Caches: the Larger the Better

Some benchmarks run well using small, one-cycle, on-chip caches backed by a larger level two (L2) cache. Many on-chip caches are too small to hold the data necessary to complete complicated tasks; more demanding applications frequently suffer from this arrangement. Commercial applications such as transaction processing and Web serving require a large instruction cache with high bandwidth. Implementing large, two-cycle primary caches in the PA-8x00 family has effectively improved performance across a wide range of workloads by significantly reducing first-level cache misses.

In instances of high cache-miss rates, L2 cache latency and bandwidth are a significant impediment to performance. Level-two caches cycle slower than level-one (L1) caches and have lower data bandwidth. The latency incurred by accessing the slower L2 cache is compounded by the need to retrieve a full cache line to fill the L1 cache rather than just the current data. Large L1 caches avoid the penalties of a multilevel cache structure by providing a high-bandwidth instruction and data paths directly to a large store of data. Members of the PA-8x00 family have always used large high performance caches. The advantages of large caches are obvious. Small on-chip level one (L1) caches backed by larger and much slower off-chip level two (L2) cache may perform adequately for small benchmarks. But when faced with more demanding applications such as data warehousing, on-line transaction processing, or technical computing, such caches usually result in disappointing performance. The PA-8000 and PA-8200 both utilized off-chip caches because the processes they were implemented in did not support the level of integration necessary for adequate on-chip caches. Starting with the PA-8500, integration levels reached the point where large on-chip L1 caches could be implemented. The PA-8500 set an industry first by including 1.5 Mbytes of on-chip cache. The move to on-chip cache provided several benefits. The expensive and power-hungry external RAM chips used by the PA-8000 and PA-8200 (at least 12 parts) were eliminated, slashing and dramatically reducing power. The move to on-chip cache also reduced the PA-8500's I/O count from almost 1100 signals to 550 signals. This improved system reliability while making it possible to run at much higher frequency than would be possible with off-chip caches. The PA-8000 and PA-8200 utilized direct-mapped caches—meaning that for any given address, there is only one location where the data can possibly reside. In contrast, an n-way set-associative cache has *n* possible locations where an address' data can reside. An n-way set-associative cache performs better than a direct-mapped cache of the same size. The larger the value of *n* in an n-way set-associative cache, the better the performance. Unfortunately, an n-way set-associative cache also requires *n* times as many signals to the cache compared with a direct-mapped cache. As a result, implementing an off-chip set-associative cache would have been cost prohibitive for the PA-8000/PA-8200 due to the number of I/Os required. By moving the cache on chip, a set-associative cache becomes practical, as the large number of signals to the cache are now wires on a chip instead of I/Os on a package. The PA-8500 as well as the PA-8600 and PA-8700 use 4-way set-associative caches for both the instruction and data caches. The PA-8700 increases the amount of cache by 50% over the PA-8600 to provide a total of 2.25 Mbytes of on-chip L1 cache. As a result, cache hit rates for even demanding applications are significantly improved on the PA-8700 leading to even higher levels of performance. The data cache is a four-way set-associative pipelined cache and is constructed of two independent banks. One bank holds data for even double-word addresses and one bank holds data for odd double-word addresses. Each 0.75MB data cache bank is implemented as four arrays which independently provide a double word of data plus error correction bits. Each array provides one way of associativity. Data is organized within the arrays so a full cache line can be addressed at a time, or four ways of associativity can be addressed together. The cache line tags are held in a smaller and separately addressable RAM array. In this way the data and tag can either be accessed together for a data read, or independently for a data store at the same time another store is accessing its cache line status. The

PA-8700 instruction cache is a 0.75MB four-way set-associative pipelined cache that can provide 4 instructions per cycle to the instruction fetch unit. The instruction cache is implemented as four arrays which each provides 128 bits of instruction plus pre-decode bits and parity bits. Each array provides one way of associativity.

Figure 9. PA-8700/PA-8600 Primary Caches
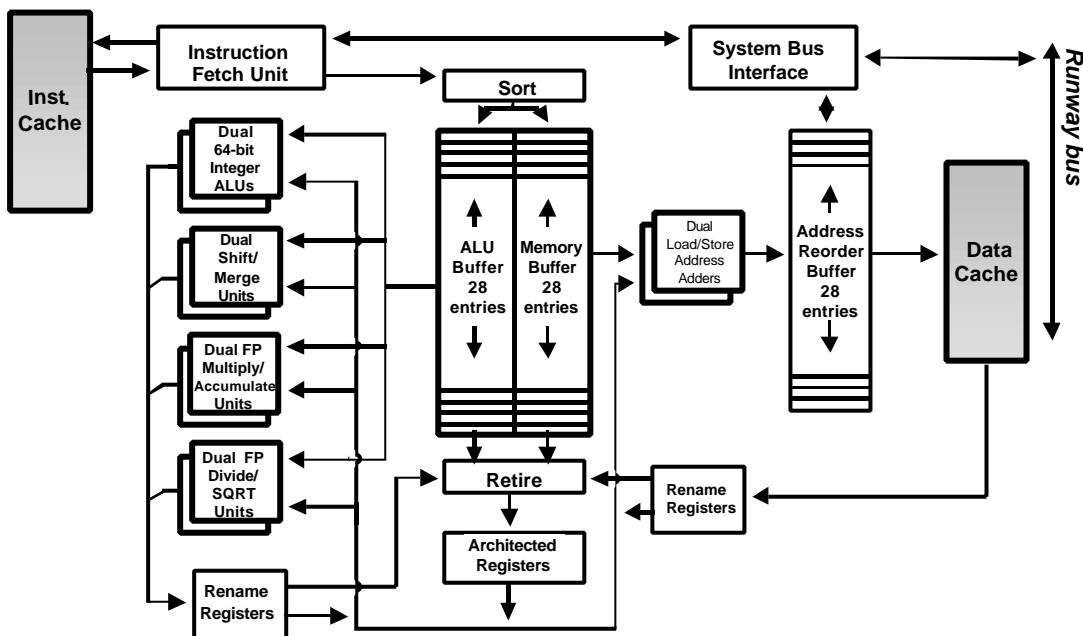
The PA-8700 is the next processor in the PA-8x00 line and leverages the industry award winning PA-8500 design (Microprocessor Report, 1998). Notable features provided by the PA-8700 include:

• 50% increase in the size of the data cache to 1.5 Mbytes

• 50% increase in the size of the instruction cache to .75 Mbytes

• Data pre-fetching capability

• Quasi LRU replacement policy for the data cache—in addition to the instruction cache quasi LRU carried over from the PA-8600

• Support for 44-bit physical addresses—enough to address 16 Terabytes.

With its introduction in systems in 2001, the PA-8700 will once again continue the industry-leading performance in Hewlett-Packard's enterprise servers and technical computing products. The advanced micro-architecture of the PA-8700 aggressively executes as many instructions as possible each cycle to maximize performance. The processor exploits techniques such as out of order execution, speculative execution, and non-blocking caches. The Instruction Fetch Unit can fetch 4 instructions each cycle from the instruction cache. From there, the instructions are forwarded to the Sort unit which places instructions into either the 28 entry ALU Reorder Buffer or the 28 entry Memory Reorder Buffer depending on the

instruction type. The Reorder Buffers constantly scan their contents looking for instructions that are ready to execute. The Reorder Buffers track all dependencies between instructions and allows data flow execution across all instructions in the buffers. Instructions are no longer constrained to execute in program order. Instead, as soon as an instruction is identified as ready to be executed, and an appropriate functional unit is available, the instruction is dispatched to the functional unit to begin execution. Each cycle, up to four instructions can be dispatched. The PA-8700 includes 10 functional units to maximize the number of instructions that can be executed each cycle. These functional units consist of 4 integer units (2 Arithmetic/Logic Units and 2 Shift Merge Units), 4 floating point units (2 Multiply and Accumulate Units and 2 Divide/Square Root Units), and 2 Load/Store Units (one for even double-word addresses and one for odd double-word addresses). After instructions have completed execution in the various functional units, they return to the Retire Unit to update the architected state of the processor and thus complete execution.

Data Integrity - Error Detection and Error Correction Capabilities

The PA-RISC family has always incorporated error checking and error correction features in their caches to protect customer data. For obvious reasons, data integrity is absolutely critical in creating highly available systems, and Hewlett-Packard's commitment begins with the processor. All data stored in PA-8700 caches are protected from single-bit errors. For the instruction cache, simple single-bit parity checking is sufficient, because its contents are always clean (data in cache that hasn't been modified is referred to as 'clean'; data which has been modified by a store instruction is referred to as 'dirty'). Any time an instruction access encounters an error, the access is treated as a cache miss. Cache lines with corrupted data are invalidated, and the correct data is re-accessed from memory.

It takes more effort to protect the data cache, because error correction is required if a dirty line in cache becomes corrupted. The PA-8700 provides six extra bits per word to enable single-bit error correction to protect the cache data. The correction, however, is not directly in the cache access path, where it would lengthen the critical cache access latency. Errors are instead detected by parallel error checking logic. If an error is detected, the corrupted data is forced out of the cache. The data is corrected in the copy out path if the line is dirty. If the line is clean, it is invalidated. The access is then re-executed, causing the line to be brought back into cache with the corrected data. The data cache's tag contains the physical address of the line and its status. This information is needed to know whether a line needs to be copied back to memory and where in memory it belongs. Therefore, the data cache's tag RAMs are also correctable. The PA-8700 maintains two copies of each cache line's tag to allow two accesses to be serviced simultaneously. Each is parity-protected. If an error is detected in either tag, a copy-out of the line is started. During the copy-out, the tag array that does not have the parity error provides the physical address for memory and the correct status. If the status indicates the line was dirty, the line is copied back to memory. If the line was clean, the cache line is invalidated. Once the cache has been scrubbed in this fashion, the access is executed again to bring the error-free line back into cache.

Static and Dynamic Branch Prediction - Faster Program Execution

Today's high-frequency, superscalar processors pay a high cost every time they incorrectly predict program execution paths. This penalty is high due to increased pipeline depth. But just as importantly, mispredicted branches in program execution prevent the processor from effectively exploiting instruction-level parallelism across branches. There are many approaches to the problem of reducing branch penalties. These can be categorized into static and dynamic methods. Each has its strengths and weaknesses.

• **Static methods**—Static methods rely on the compiler to optimize branch performance. Because the optimizations are performed when the program is compiled, the optimizations do not change as the program runs. These optimizations may be based on execution profile information gathered during a training run (Profile Based Optimization or PBO) or they may be based on the structure of the source code by using heuristics. These optimizations include eliminating certain branches altogether using predication, and encoding hints in branch instructions of the likely branch direction. For branches, which are either almost always taken or not taken, compiler hints are very accurate and produce excellent results. The principal

advantages of static compiler branch optimizations is that they do not consume chip area and do not have the capacity limits that dynamic branch prediction encounters.

• **Dynamic methods**—In contrast, dynamic branch prediction predicts the future behavior of a branch based on the past behavior of that branch and/or other branches during program execution. Dynamic branch prediction performs better under many circumstances, such as predicting branches based on runtime options and when branches change over time. The main weakness of dynamic branch prediction, is a limited resource for collecting branch statistics. Because both static and dynamic branch prediction methods have their advantages under different circumstances, the design of the PA-8x00 accommodates both techniques. At compile time, the application developer can mark a binary to be run in one mode or the other depending on which performs better for that program. The PA-8700's Branch History Table (BHT) is an array of two-bit saturating counters. The counters record whether the branch went in the direction indicated by the compiler's static hint. If the static hint was wrong, the counter is incremented; if correct, the counter is decremented. Each time a branch is fetched, the BHT is consulted and if the counter is zero or one, the static hint encoded in the instruction is followed. If the counter is two or three, the hardware predicts that the branch will go in the opposite direction. These enhancements enable the PA-8700 to combine the advantages of static and dynamic branch prediction methods and are backward compatible with all previous members of the PA-8x00 family. The PA-8700 has a 56-entry instruction reorder buffer. Instructions are fetched 4 at a time and each instruction is placed in either the ALU buffer or the memory buffer depending on the type of instruction it is. The instruction reorder buffer constantly scans the instructions it contains and as it finds instructions that can execute, they are dispatched to one of the 8 execution units or one of the two load/store pipes. Instructions are no longer constrained to execute in program order. The PA-8700 is also capable of speculative execution. When branch instructions are encountered, the PA-8700 uses sophisticated branch prediction algorithms involving its branch hardware table to guess whether or not the branch will be taken. The machine then speculatively executes instructions from the predicted branch. If the branch prediction was correct the speculatively executed instructions are allowed to retire normally. If the branch was incorrectly predicted all speculatively executed instructions are purged and the processor begins fetching from the correct location.

Profile-Based Optimization – Increase Application Run-time Performance

The radical design of the PA-8000 opened up new opportunities for compiler optimizations. Profile-based optimization uses trial runs of the code with representative data sets to gather statistics about how the code normally executes. This information can be used to optimize the code to improve performance. By leveraging the PA-8000 core micro-architecture, the PA-8700 protects investment in compilers and application tuning. This enables the compiler to perform optimizations not possible without knowledge of the dynamic behavior of the program. PBO also enables the compiler to more accurately encode static branch prediction hints. The investments made by software developers in PBO will pay off more and more as compilers improve their ability to exploit the increasing potential of future processors.

44-Bit Addressing – Connect to Larger Memory

Previous members of the PA-8x00 family have implemented 40-bits of physical address—enough to address 1 Terabyte of physical memory. While no system that Hewlett-Packard currently ships can hold this much memory, system needs continue to increase, driven by larger data-sets. In the coming years, some customers may have a requirement for more than 1 Terabyte of physical memory. The PA-8700 implements 44-bit physical addresses—enough to address 16 Terabytes of physical memory. The PA-8700 can also operate in 40-bit mode for capability with existing systems.

Data Cache Pre-fetch Capabilities – Data Ready Before It's Needed

Pre-fetching data before it is needed improves the cache hit rate (the probability that the data you want is in the cache when you need it) and thus boosts application performance. The PA-8700 supports pre-fetch capability for the data cache. When a data cache miss occurs, the line that misses is fetched and placed into the data cache. If data pre-fetch is enabled, an adjacent line is also fetched if it is not already in cache. This

pre-fetching of a line likely to be used in the near future, allows the line to be brought into the cache pro-actively rather than waiting for a costly cache miss to occur to retrieve the line.

Quasi LRU Cache Replacement Algorithm for Data and Instruction Cache

Complex processors like the PA-8700 need to maximize the probability that the required data is present when an access to the data or instruction cache is made. If the data is not present, then the processor must fetch the needed data from main memory, wasting precious microseconds. The PA-8x00 family incorporates a quasi LRU (Least Recently Used) algorithm to improve the odds that the most frequently used data will remain in the cache. Whenever new data is moved into a four-way set associative cache, there are four possible locations where the new data can be placed. The replacement algorithm used by the cache determines which of these four locations will have its existing data replaced by the new data. Ideally, the cache would pick the oldest data least likely to be used in the future and replace it with the new data. The PA-8500 uses a 'round-robin' replacement algorithm to determine which data to replace. This replacement selection occurs without any consideration of whether or not a line has been recently used and is thus likely to be used again. An LRU algorithm takes data usage history into account. It identifies which of the four locations in cache was least-recently-used and selects that location for replacement with the new data. In most cases, the data least-recently- used is least likely to be used again in the future and is therefore the best candidate to be replaced with new data being brought into the cache. Unfortunately, a true LRU policy is expensive to implement. Instead, the PA-8700 provides a quasi Least Recently Used (quasi LRU) algorithm. Performance simulations show that the PA-8700 quasi LRU algorithm performs almost as well as a true LRU algorithm while requiring only a fraction of the hardware resources to implement. The use of a quasi LRU replacement algorithm boosts cache hit rates by making better choices about what cache data to replace, and thus allows applications to perform better. The PA-8600 implemented a quasi LRU replacement algorithm for the instruction cache that provides similar advantages for instruction cache performance. The instruction quasi LRU replacement algorithm is also implemented on the PA-8700.

Lock-step Functionality – Hardware-implemented Fault Tolerance

The PA-8700 supports a lock-step mode. In lock-step mode, two or more processors can operate in parallel. On a state-by- state basis, it is possible to compare the results produced by each processor, pin-to-pin. Any discrepancy in results indicates that an error has occurred. By operating two parallel chips in lock-step mode, it is possible to detect a failure by either of the processors. Three parallel chips operating in lock-step mode make it possible to correct a failure by any one of the three processors or detect an error by any two of the processors. This provides the capability to build fault tolerant systems. By paralleling even more chips it is possible to provide even greater levels of fault tolerance. If a discrepancy occurs, it is up to the system to determine what action to take to respond to the fault.

Manageability and Support

The rp7400 has many features to minimize the effort required to manage one system or an entire computer room. rp7400 simplifies system management in several aspects: event notification, automatic error handling, power monitoring, and the user interface to system management.

Event Notification

For an operator that is physically present, there are two ways to receive event notification. The simplest and easiest are the status LEDs on the front of the system. These five LEDs allow rapid verification of system status via a quick glance at the system. The five status LEDs each have a specific meaning:

- Power: power is present and on, and power supplies are OK
- Remote: remote console is enabled
- Run: system is up and running
- Attention: occurrence of a non-catastrophic event; e.g., failure of an N+1 component
- Fault: occurrence of a catastrophic system event

In addition to the five specific meanings of the LEDs, related system status is encoded based on whether the LED is simply on (*solid*) or is blinking (*flashing*); for example, unexpected reboot system recovered, operating system not running, or operator intervention required.

Event Monitoring Service (EMS)

HP EMS is a system monitoring application designed to facilitate real-time monitoring and error detection for HP products in the enterprise environment. This framework provides centralized management of hardware devices such as the rp7400 server and system resources, and it provides immediate notification of hardware failures and system status. HP EMS can receive data on unusual activity, add information on the problem's source, and provide recommendations on problem resolution.

HP EMS consists of a set of system and network monitors within a monitoring environment. This monitoring framework has an easy-to-use interface and provides a mechanism for monitoring resources, registering monitoring requests, and sending notification when resources reach user-defined critical values.

The monitors can also poll hardware, disks, clusters, network interfaces, and system resources and send information to the framework. An event can be simply defined as something you want to know about—for example a disk failure or file space dropping below a predefined level.

The primary HP EMS benefits include:

- Enabling efficient and effective system monitoring within a single, comprehensive framework
- Delivering the ability to tailor the monitoring system to fit specific needs
- Providing a wide variety of notification methods through multiple protocols (SNMP traps, TCP, UDP, OPC messaging)
- Providing immediate alerts if a component fails, enabling proactive replacement
- Interacting with MC/ServiceGuard and MC/LockManager to provide a complete high availability solution

Service Processor

The rp7400 employs a dedicated processor to aid system management and diagnosis. The Service Processor can diagnose a system failure even in the unlikely event that the system is unable to execute code. The Service Processor allows system power to be remotely turned on or off and has battery backup allowing diagnosis even of power failures. The Service Processor interfaces to key system components via an I^{2}C (Inter-Integrated Circuit) bus to continually monitor the status of system fans, temperature, and power supplies; and it signals the operator for any significant system events.

Major features and benefits of the Service Processor include:

- System console redirection
- Console mirroring
- System configuration for automatic restart
- Viewing history log of system events
- Viewing history log of console activity
- Setting inactivity timeout thresholds
- Remote system control
- Power control—remote power on and off
- Viewing system status logs
- Configuring virtual front-panel display
- Event notification to system console, e-mail, pager, and/or HP Response Centers
- Alert-level notification to pager, even if N4000 is off-line (modem required)
- Ability to connect to another Service Processor
- Auto system restart
- Virtual front-panel display

System Power Monitor

Closely integrated to the Service Processor is the rp7400's system power monitor. The system power monitor controls and monitors system power and cooling. Aspects controlled/monitored by the system power monitor are:

- Power supply status and temperature
- System supply voltages—including remote system power on and off
- Total system power consumption
- Individual Processor Support Module status
- External ambient air temperature
- Individual fan speed and status

Various temperatures are monitored to control the system fans, provide thermal warnings, and prevent permanent damage from overheating by graceful shutdown in over-temp situations. (Note that the system fans are run only as fast as necessary to keep the system cool. The fans are kept in sync with each other, turning at exactly the same rate. This intelligent fan control allows the rp7400 to generate as little noise as possible while maintaining an optimum operating environment to maximize reliability.)

The power monitor senses the presence of power supplies and power consumption of system components to determine if the system is in an N or N+1 power configuration; it can determine:

- Number of bulk power supplies
- Number of CPUs
- Amount of memory present
- Number and power consumption of each installed PCI I/O card

System configuration and health is tracked by the system power monitor and passed via a dedicated I^2C bus to the Service Processor. This information can be processed as follows: simply displayed on the system console, logged to an event file, or used to trigger an alert based on a specific threshold (system temperature, fan status, or power supply status, for example).
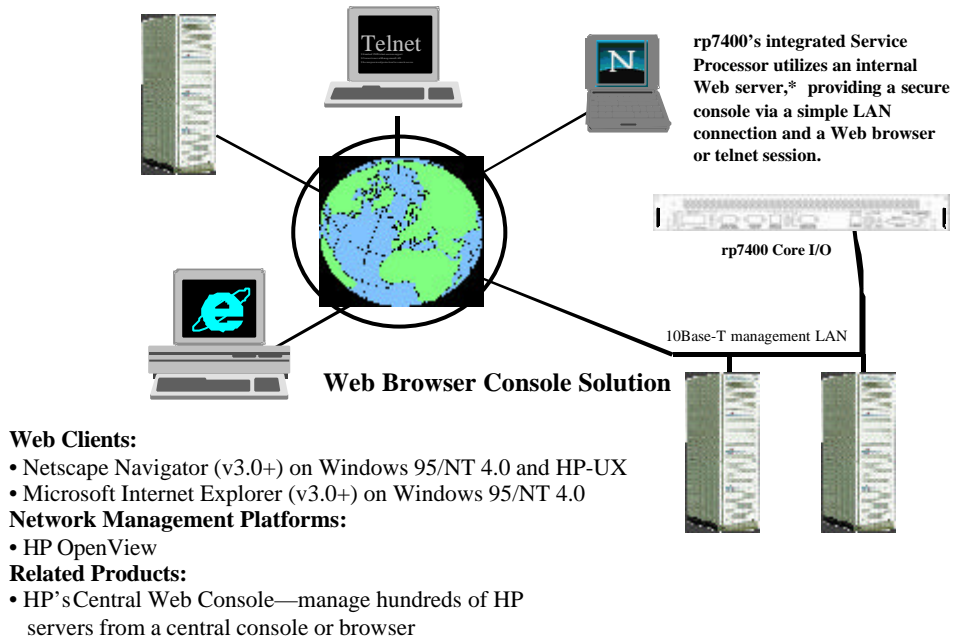
Built-in Web Console

The rp7400 includes HP's SecureWeb Console to allow management of many systems from a single Internet browser. As illustrated in Figure 10, the HP SecureWeb Console allows any Netscape or Internet Explorer browser to be used as a system console, giving total system access to authorized system administrators anywhere, just as if they where at an ASCII console. A high level of password protection is used to control access to the Web console.

Major features of HP SecureWeb Console include:

- System management over the Internet or intranet
- Mirrored access—up to four operators can simultaneously share the same screen and keyboard
- Security—built-in password encryption, data scrambling, and Java™ download protection
- Universal browser-based support for Netscape v.3.0+ and Microsoft Internet Explorer v.3.0+ Web browsers
- Easy updates of Web console software over the network
- Easy installation—just connect the rp7400 console port to a LAN; no client software to install
- Support for HTTP, FTP, TFTP, and other key Internet standards

Figure 10. HP's Web Browser Console Solution

Web Console Support

rp7400's integrated Service Processor utilizes an internal Web server,* providing a secure console via a simple LAN connection and a Web browser or telnet session.

rp7400 Core I/O

10Base-T management LAN

Web Browser Console Solution

Web Clients:
• Netscape Navigator (v3.0+) on Windows 95/NT 4.0 and HP-UX
• Microsoft Internet Explorer (v3.0+) on Windows 95/NT 4.0
Network Management Platforms:
• HP OpenView
Related Products:
• HP's Central Web Console—manage hundreds of HP
   servers from a central console or browser

LAN Console

The rp7400 also provides a secure web console interface using industry-standard telnet connections. Like the Web console, the LAN console can be used remotely for managing many systems from a single control center. The telnet interface allows scripts to be used to vastly simplify multiple system management. Password protection provides a high level of security to control access to the LAN console, ensuring that only authorized personnel perform system management.

ASCII Consoles

As many customers may not initially be comfortable with Web or LAN console concepts, the rp7400 provides an RS-232 port to use for ASCII terminal console connections. Any VT100-capable terminal or emulator can be used as a local system console.

Figure 11. ASCII Console Options

Local Console Using ASCII Terminal   **Remote Console Via External Modem**
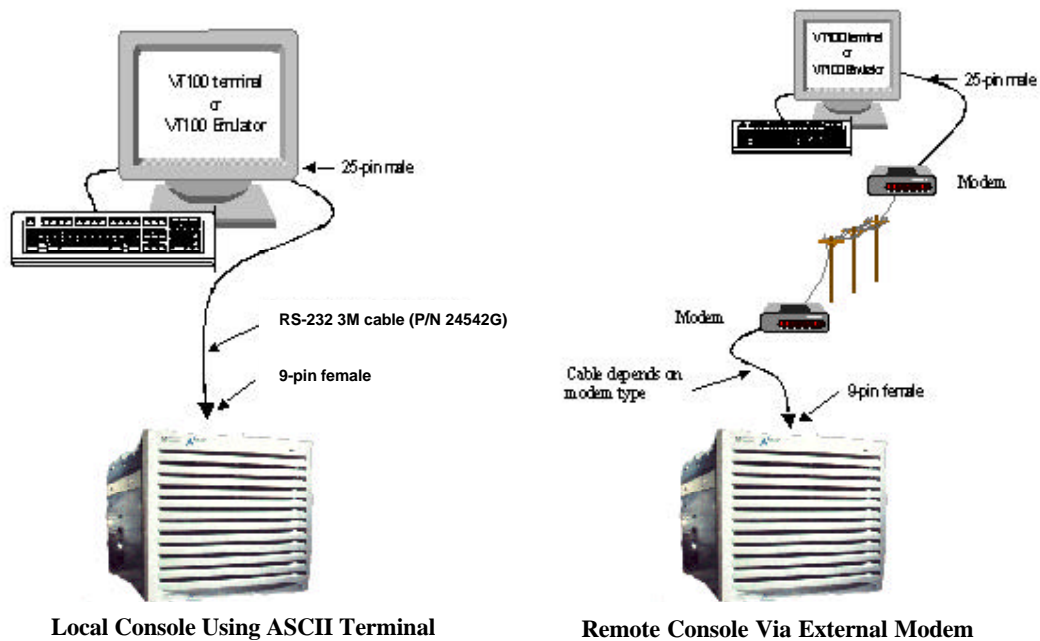
Remote Access

As with previous HP server systems, the rp7400 provides an RS-232 interface for a remote console that is useful for obtaining help from HP service experts. Customers need only add a modem to allow remote access via the phone; security is ensured by having to explicitly enable remote console access, which is protected with a password, and via dial-back phone verification.

Self-Diagnosis

Many features have been designed into the rp7400 system to maximize system uptime. There are several aspects to maximizing uptime: eliminating common single points of failure, allowing the system to continue running after some errors, and allowing quick identification and servicing of hardware faults if they do occur.

Besides traditional diagnostic software, the rp7400 also continuously monitors system health with the Sentinel monitor. Knowing a failure has occurred that reduces N+1 protection is important. It is important to minimize the risk of downtime by replacing a failed component soon as possible to get back to the safety of an N+1 configuration. To enable this, the rp7400 provides several methods of event notification.

Like its predecessor, the K-Class server, the rp7400 has extensive firmware-based self-tests. These diagnostics are evoked on power-up or reset. The self-tests check for correct system operation prior to booting the operating system. The firmware diagnostics first check the processors, then processor caches and memory, and finally I/O devices. Testing complexity increases as more of the system is proven good and more pieces of the system can be relied upon to increase test coverage on the remaining parts. Self-test failures are reported to the system console and the support processor, along with failure specifics and corrective action.

With the addition of an independent support processor, the rp7400 can provide diagnostic information even in the unlikely event of a complete system failure that prevents the processors from correct execution. This feature allows quick fault isolation and helps ensure that field engineers can bring the system back online on the first visit.

Online and Offline Diagnostics

The rp7400 also offers traditional online and offline diagnostics to validate system health and provide extensive system fault coverage.

With online diagnostics, the system is tested while the operating system and applications continue to run. This allows basic testing of system components that are not currently being used, or where testing does not prevent continued use of the operating system and applications.

Offline diagnostics provide increased coverage of system components for improved fault isolation and intensive system testing before returning to production.

Self-Healing

The rp7400 was designed to minimize single points of failure and has many features that allow it to continue operating after experiencing errors. Internal system buses (the system bus, for example) provide error checking and correcting (ECC) that allows the system to continue operation after a transient single-bit error or a hard failure of a single data bit.

Similarly, memory employs an ECC scheme that allows a transient or hard failure on any single bit per DIMM pair. Error correction in memory does not end there. The same ECC scheme in memory also allows an entire DRAM to fail on each DIMM pair without interrupting correct operation. With a total DIMM count of 32, the rp7400 could experience 16 complete DRAM failures without experiencing any downtime.

For a server the size of the rp7400, this self-healing is very important due to the potentially large memory configurations. With the high number of DRAMs possible in a large memory configuration, the possibility of a single DRAM failing is too significant not to protect against. To keep memory errors from causing downtime, the rp7400 employs memory scrubbing, which periodically scans all the memory for transient single-bit errors and deallocates them before a double-bit error occurs.

Memory page de-allocation will attempt to allow the system to continue running even after an uncorrectable memory error has occurred. Page deallocation works in two different ways:

- At system power on (or reboot), memory is tested, and any memory pages with errors are marked so that they are not used.
- Once the operating system is running, an uncorrectable memory error in user space causes only the affected jobs to be terminated. The bad page is marked as unusable and the system continues running. Uncorrectable failures in the operating system memory will cause a system reboot, removing the bad page in the reboot process.

Processor failures are very similar to memory page de-allocation. A processor failure that occurs while the operating system is running will cause a crash and reboot. On reboot, the processor that fails self-test will be deallocated, and the system will return to operation without the failed processor until it is replaced.

rp7400 Industrial Design and Packaging

The rp7400 has been designed to be an attractive addition to any data center. The industrial design is coordinated with other HP servers and peripherals for a consistent appearance. However, the rp7400 industrial design provides much more than just good looks.

High Density

The rp7400 system is designed to provide unprecedented performance density that easily adapts to different data center environments. The rp7400 is primarily designed to mount in a 19-inch HP rack. At 10U high, four rp7400 systems can be racked into one 2-meter rack. With eight of the highest-performance processors available in the industry per system and four systems per rack, that's an impressive 32 processors in a single rack. With the high cost of computer room floor space, rp7400's small footprint dramatically lowers total cost of ownership.

The rp7400 industrial design and packaging was designed to allow easy service access to the system. This allows all parts of the rp7400 system to be completely serviced without having to remove it from the rack. Only front and rear access is required to service the system. This allows side-by-side racks of systems to be completely supported without having to lose floor space for side access to the system.

The rp7400 provides rack slides to enable system service while the system is operating. Types of service that can be done while the rp7400 is operational include the removal and addition of PCI cards, cooling fans, power supplies, and internal disks.

Designed for Growth

Upgrades can be performed to the rp7400 while it is still in the rack. Although these upgrades require shutting down power to the rp7400, they do not require removing the rp7400 from the rack, which speeds up the process to quickly return the upgraded system to operation.

Designed for Strength

The rp7400 was not only designed for robust data center environments, but it also meets the rigorous standards required for telecom central office (CO) compliance. NEBS (Bellcore's Network Equipment Building System) certification requires that the system must pass several aggressive environmental and safety tests that ensure its safe and reliable operation within the telecom central office. The standard rp7400 meets these requirements without modification, providing an extremely robust platform without the high price normally associated with telecom equipment.

For More Information

HP product information and technical documentation is available online at www.eproducts.hp.com In addition, configuration tools and pricing information allow registered users to place orders online. For registration, please contact your Hewlett-Packard sales representative.

Contact any of our worldwide sales offices or HP Channel Partners (in the U.S., call 1-800-637-7740) or visit our HP Enterprise Servers Web site at www.hp.com/go/rp7400.

For more information, contact any of our worldwide sales offices or HP Channel Partners.

For the location of the nearest sales office call:

United States of America: +1 800 637 7740

Canada:
Hewlett-Packard Ltd.
5150 Spectrum Way
Mississauga, Ontario L4W 5G1
+1 905 206 4725

Japan:
Hewlett-Packard Japan, Ltd.
Japan Country H.Q.
3-29-21, Takaido-Higashi, Suginami-ku,
Tokyo, 160-8585 Japan
+81 3 3331 6111

Latin America:
Hewlett-Packard
Latin American Region Headquarters
Waterford Building, 9th Floor
5200 Blue Lagoon Drive
Miami, Florida 33126 USA
+1 305 267 4220
Refer to country phone numbers

Australia/New Zealand:
Hewlett-Packard Australia Ltd.
31-41 Joseph Street
Blackburn, Victoria 3130
Australia (A.C.N. 004 394 763)
+61 3 9272 2895

Asia Pacific:
Hewlett-Packard Asia Pacific Ltd.
17-21/F, Shell Tower
Times Square
1 Matheson Street
Causeway Bay
Hong Kong
+8522 599 7777

Europe/Africa/Middle East:
Hewlett-Packard S.A.
150, Route du Nant-d'Avril
CH-1217 Meyrin 2
Geneva, Switzerland
+41 22 780 81 11
European Multicountry: +41 22 780 81 11
Middle East and Africa: +41 22 780 71 11
European Headquarters: +41 22 780 81 81
Refer to country phone numbers

For direct country contact call:

Argentina: +541 787 7145

Austria: +43 1 25 000 0

Belgium and Luxembourg: +32 2 778 31 11

Brazil: +5511 7296 8000

Chile: +562 203 3233

Colombia: +571 629 5030

Denmark: +45 45 99 10 00

East Central Europe, CIS, and Yugoslavia:
+43 1 25 000 0

Finland: +358 9 887 21

France: +33 1 69 82 60 60

Germany: +49 7031 140

Greece: +30 1 689 644

Hungary: +36 1 252 7300

Iceland: High Performance Systems hf.
+354 1 67 10 00

Ireland: +353 1 615 8200

Israel: Computation and Measurement Systems (CMS) Ltd.
+972 3 5380 333

Italy: +39 2 92122770

Mexico: +525 326 4600

Netherlands: +31 20 547 6911

Norway: +47 22 7356 00

Poland: +48 22 608 77 00

Portugal: +351 1301 7343

Russia and the CIS, excl. Ukraine: +7 095 923 5001

Slovenia: +38 61 55 84 72

Spain: +34 1 631 1600

Sweden: +46 8 444 2000

Switzerland: +411 735 7111

South Africa: Hewlett-Packard South Africa (Pty) Ltd.
+27 11 806 1000

Turkey: +90 212 224 5925

United Kingdom: +44 1344 369231

Venezuela: +582 239 4133