# Superdome Management Part 1

Paul Bouchier
Hewlett Packard Co.
3000 Waterview Parkway
Richardson, TX 75080
Ph. 972-497-4349 (W) 972-497-4500 (Fax)
Email: Paul_bouchier@hp.com

March 5, 2001

## Introduction

The Hewlett Packard HP9000 Superdome server can be flexibly configured into partitions, each of which can run a separate instance of the operating system. Partitions provide hardware protection that prevents software or hardware errors in one partition from interfering with another partition. This paper describes the architecture of the server, what partitions are and how they work, and how the Guardian Service Processor interacts with them to enable management of the server and its multiple partitions. This paper provides background information that prepares the reader for the associated paper "Superdome Management Part 2".

The HP9000 Superdome server contains multiple cells, each of which contains processors and memory, and can be connected to I/O. A partition is a grouping of these cells. Partitions are defined by a data structure known as the complex profile. The process by which cells boot and form partitions based on the complex profile contents is described, together with the mechanism that protects partitions from one another.

The Guardian Service Processor plays an important role in managing the server. It provides a command interface that allows the user to connect to partition consoles, display server and partition state, reset partitions, view logs, and many other platform management capabilities. It also manages changes to the complex profile. The user commands that support managing the server from the GSP are described.

## System Architectural Overview

The backbone of the system is a high bandwidth backplane constructed as a crossbar. The crossbar is constructed from an ASIC referred to as "XBC." Each backplane contains ports into which cell modules are inserted. The cells contain 4 processors, memory and support for cable connections to I/O chassis which contain PCI card slots. Pairs of backplanes can be tied together at full bandwidth and with low latency, to build a 16-cell, 64-way system. The backplane supports both current HP PA-RISC cells and the future generation of cells which will support both PA-RISC and Intel/HP IA-64 processors.

At the heart of the processor and memory cell module is an ASIC, referred to as "CC," that integrates both a processor and memory controller to achieve low latency and high bandwidth. The first version of the cell module has four PA-8600 processors and memory capacity of 16 Gbytes. The CC includes a high bandwidth connection to I/O and the backplane crossbar providing low latency between all ports.

Another ASIC, "RIO", is the power behind the I/O PCI chassis.  RIO drives 16 configurable "ropes" that terminate at the "Elroy" ASIC. Elroy is leveraged across HP's entire UNIX product line.  Each Elroy individually controls a single PCI card guaranteeing full bandwidth, fault isolation and serviceability without affecting the operation of other PCI cards.  The ropes can be doubled up to provide twice the bandwidth and lower latency to PCI-4X cards.  The first I/O chassis is configured with eight PCI-2X slots and four PCI-4X slots.

The result is a full featured, modular design ensuring protection of the customer's investment.  The system design consists of four of the world's largest ASICs.


## Supported Systems

HP9000 Superdome servers ship in three varieties; SD32000 – single cabinet 32-way max, SD64000 – dual cabinet 64-way max, and a cost reduced version, SD16000, at a maximum of 16-way. The 32-way and 64-way system cabinets contain eight processor and memory cell modules per cabinet.

 Figure 1 shows a conceptual drawing of the SD32000 backplane.  Two logical XBCs connect to eight cells with two XBC-XBC links in-between.  One XBC-XBC link is built into the backplane and the other is configurable to double the bi-section bandwidth and to allow upgrades to the SD64000 without replacing the backplane.
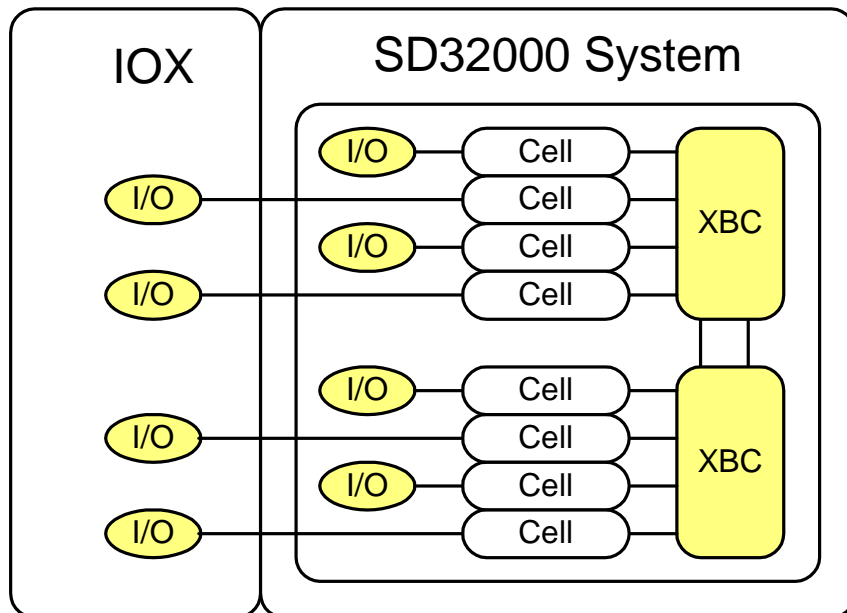


**Figure 1 – SD32000 Single Cabinet 32-way Max**

Figure 2 shows the SD64000 dual-connected cabinet. Cabinet 0 contains the same backplane as the SD32000. Cabinet 1 is a mirror image version of cabinet 0's backplane. Each XBC is directly connected to all other XBCs to achieve lowest latency and reduce bandwidth bottlenecks. Each inter-cabinet XBC-XBC link is implemented as short, high-speed flexible cables. All XBC-XBC links are full bandwidth and have identical latency to ensure minimal queuing delays.
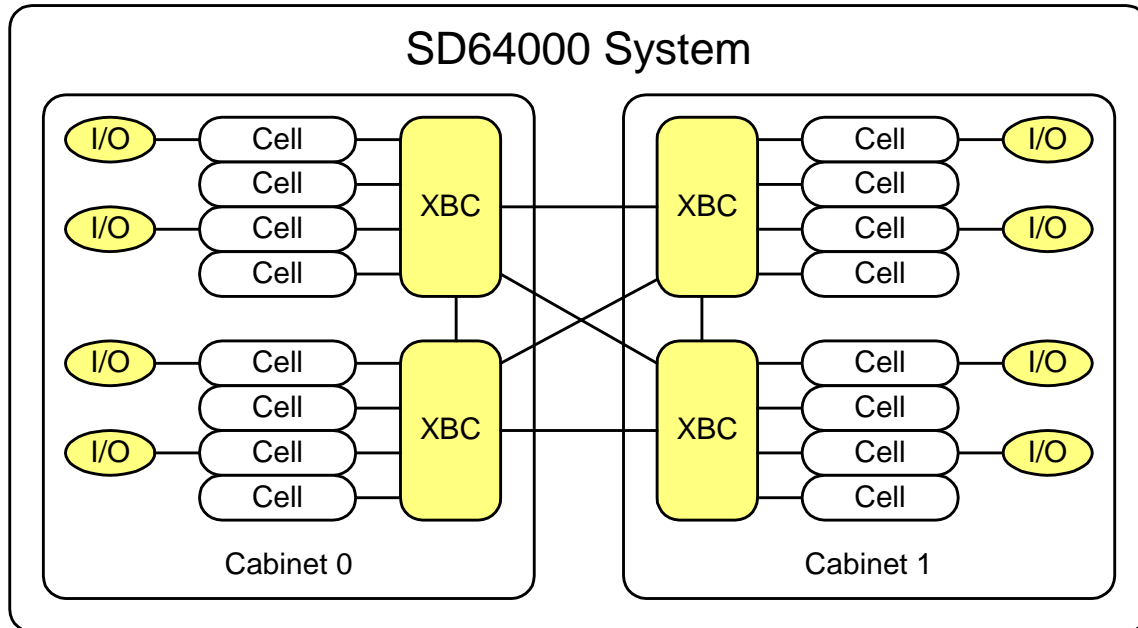


**Figure 2 – SD64000 Dual Cabinet 64-way system**

A maximum of four I/O chassis can be installed in a cabinet. An I/O expansion cabinet can be added to allow connecting the maximum configuration of eight I/O chassis per cabinet.


## The Cell Board

The cell board, see Figure 3, contains CPUs, memory, support hardware including boot ROMs, and the CC and M2 ASICs. The cell board is the most complicated Superdome board. Four processors have point-to-point connections to CC. Each CC has optional I/O and XBC connections. Two replicated memory sub-systems are controlled independently. A data fan-out and directory tag control ASIC, M2, converts the wide DIMM data bus into a narrower high-speed data bus. Extremely low latency is possible since the CC translates requests and sends addresses directly to the DIMMs with data returning through the M2 data buffer

A sub-system called Processor Dependent Hardware, PDH, is included in the cell to guarantee independence and fast boot time. Each PDH contains sufficient resources to boot the cell (ROM and local RAM and miscellaneous registers) as well as an interface to the Guardian Service Processor, or GSP.
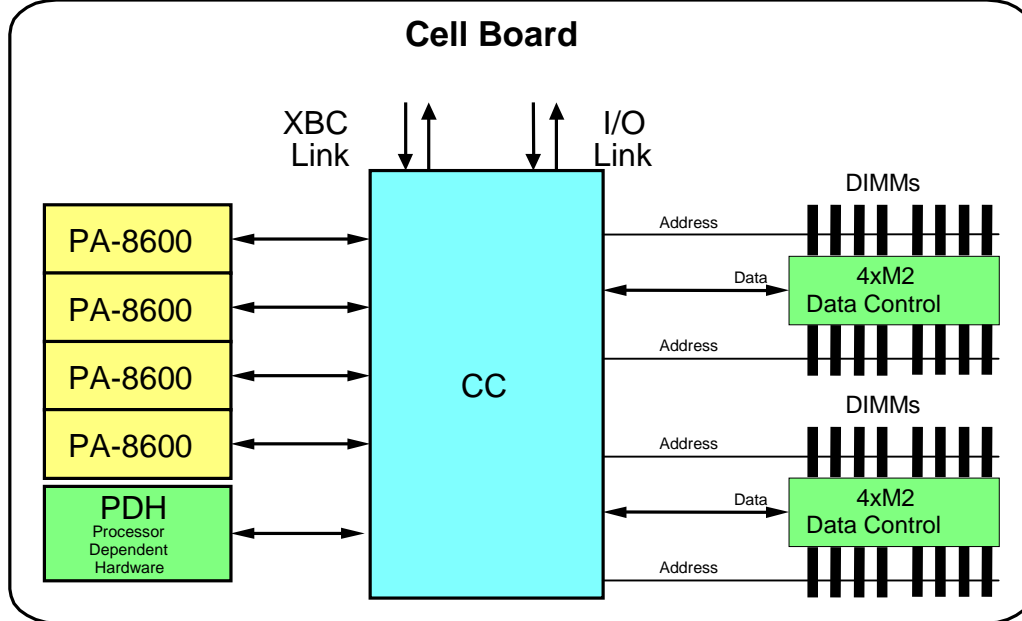
**Cell Board**

To support the exploding database markets, the memory capacity has been increased to 32 DIMMs per cell board. Each DIMM carries 36 SDRAMs with a capacity of 512MB utilizing 128Mb technology. The capacities add up to a total of 16 Gbytes memory on a single cell and 256 Gbytes memory in a 64-way dual-cabinet system. Future SDRAM upgrades up to 1 Gbit are possible.

**Figure 3 - Cell Board**

## CC

The CC is the world's largest ASIC. It provides both processor and memory control as well as connectivity to I/O and the XBC Crossbar. The CC provides the CPUs and I/O with a large, flat, low latency memory system. Memory interleaving across cells provides scaling and performance characteristics common to SMP systems.

HP has patents pending on a number of error strategy and containment features. CC implements a hardware firewall to protect its partition from other partitions. Memory is fully covered by a SEC/DED code that also corrects a single SDRAM "chip-kill" failure. CC implements full SEC/DED ECC on data busses, parity for addresses, and timers to watch over system components. When failures are detected, error containment is guaranteed and data corruption is prevented.

For a full block diagram of the CC, see Figure 4. The core of the ASIC is a crossbar interface called the Central Data Path, (CDP). Connecting to the CDP are four Processor Interfaces, (PI), the two Memory Units, (MU), Processor Dependent Hardware, I/O controller, and XBC crossbar interface.
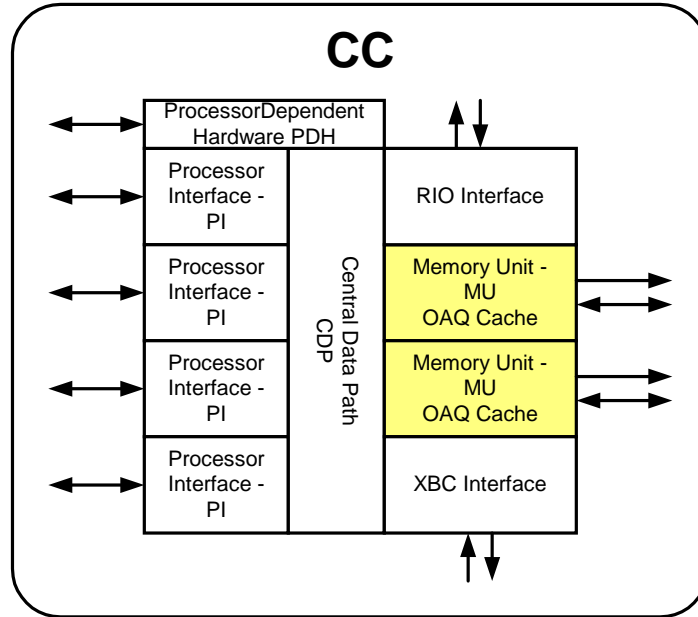
**Figure 4 – CC CPU and Memory Controller ASIC**

**Processor Interface**

The PA-8600 utilizes HP's Runway bus running in Turbo mode.  Turbo mode transfers data on both edges of the Runway clock to lower latency and double bandwidth. Each CC processor interface has it's own routing logic called the "cell map" (patent pending) to quickly decide a packet's final destination.

**Memory Unit**

The two replicated Memory Units provide the master control over both memory and memory coherency.  A directory-based coherency scheme and an innovative internal data cache are implemented with no SDRAM space penalty.  The cache allows low latency resolution of coherency traffic and shared data, such as barriers, without accessing the DIMMs. Coherency is maintained through the use of directory tags stored in main memory with each cache-line.  A directory provides a distinct advantage over a snoopy bus system by issuing snoops to processors only when absolutely necessary.  The snoopy address bus bottleneck prevalent in, for example, the Sun UE10k is prevented. To maintain a simple programming model, all memory in the system is strongly ordered and implements the MESI coherency protocol, where cache lines can take the state of Modified, Exclusive, Shared, or Idle.  A full SEC/DED error correction code covers not only normal transient memory errors but also the failure of a complete SDRAM.

Each MU is tightly integrated with four bit-sliced M2 data fan-out ASICs.  In addition to moving data to and from the DIMMs,  M2 also provides direct local control over the directory tags. For each cache-line access, the directory tag is updated in the M2 and written back to the DIMMs reserving data bus bandwidth to the CC for the more important task of moving data.

The implemented coherency protocol minimizes the number of messages sent between the processors and the memory unit. The vast majority of data requested is fetched into the CPU's cache and marked as either modified or exclusive. The processor is then free to modify the data without any additional memory traffic. After the processor is done with the data, the cache line is cast-out. When cast-outs of modified data are sent to memory, both the data and the "idle" directory states are written simultaneously. Exclusive cast-outs send a short message to the memory controller indicating that the CPU no longer owns the line which in turn sends a message to M2 to update the tag as idle without pulling the cache-line's data back into CC. This conserves bandwidth along the entire path to memory.

Coherency is under control of the Ordered Access Queue (OAQ) caching both the directory tags and data. A fetch that requires a coherency snoop to a third-party owner is held in the OAQ, transferring back to memory only when the transaction is complete. The fully associative OAQ cache resolves multiple requests for the same cache-line quickly and efficiently without cycling the memory SDRAMs. An example of how this improves performance is that it reduces the time to release a semaphore.


## I/O

The HP9000 Superdome I/O chassis, see Figure 5, consists of the RIO I/O controller ASIC, eight PCI-2x and four PCI-4x slots. A 1.6 Gbyte/sec link spawned from the CC provides access to CPUs and memory. The RIO link is covered by both ECC and CRC correcting all single bit and single wire failures. The backside consists of 16 configurable 266 Mbytes/sec "ropes." The rope terminates at the Elroy ASIC fanning out to a PCI slot. Elroy allows each PCI card to be physically and functionally isolated from other cards. A failed card will not interfere with other cards. In addition, a per-slot power controller allows online addition and replacement of failed cards. Elroy is a low cost PCI controller leveraged across HP's 1, 2, 4, and 8-way servers. Two ropes are ganged together to a single Elroy running PCI-4x slots.
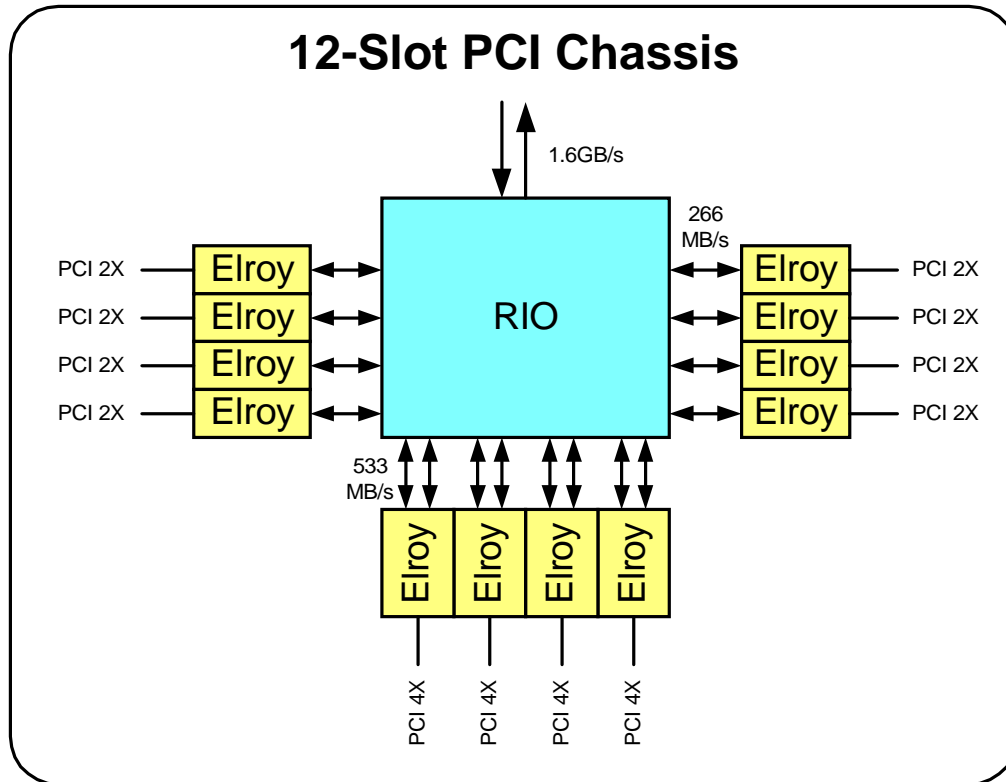
**Figure 5 – 12-Slot PCI Chassis**

## RIO Features

RIO contains two caches, is strongly ordered and participates fully in the cache coherency protocol.  This allows a simple programming model where CPUs can access critical control structures without the special flush and synchronization operations required with non-coherent or buffered I/O.  Each cache has it's own Translation Lookup Buffer, (TLB) for use by 32 bit cards.  Alternatively, a 64-bit PCI card can use Direct Virtual Addressing to issue system physical addresses without requiring the overhead of the TLB.

Error handling features include CRC and ECC on the link providing for single wire recovery with parity protection on all PCI and rope interfaces.  Timers watch over PCI cards and detect downstream system failures.  Configurable response to errors allows smart PCI drivers to recover from failures

RIO is built from lower-level building blocks as is shown in  Figure 6.  At the highest level, RIO is split into a link interface and 2 replicated rope units, driving 8 replicated rope interface blocks.  Each rope unit contains it's own cache and translation buffer (TLB). The cache implements Modified, Exclusive, Idle, and a special read-only Shared cache states.  The special shared state improves performance for optimized drivers that can treat memory as channels.
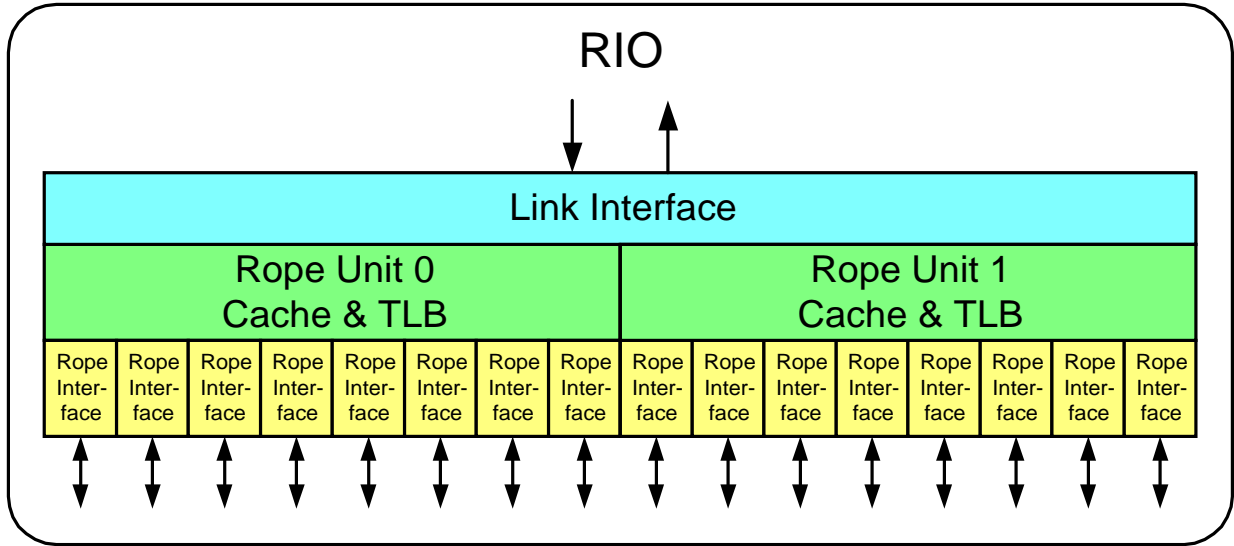
**Figure 6 - RIO I/O Controller**

The eight rope interfaces connect the internal cache to the bi-directional 266 Mbytes/sec ropes. Coupling two ropes together provides a wide 523 Mbytes/sec bus for PCI-4x.

## XBC

The XBC ASIC is an extremely high bandwidth, low latency crossbar. Large queues, that can be bypassed when empty significantly reduce stalls under heavy load conditions.

XBC contains eight identical ports and a register block, see Figure 7, that can talk to either a CC or another XBC. Two physical XBCs are paired together to create one logical XBC. The SD32000 backplane has two-logical, four-physical, XBCs yielding 16 logical ports. Of the 16 ports, eight are used for cell modules, two are permanently connected together, and four are configurable.
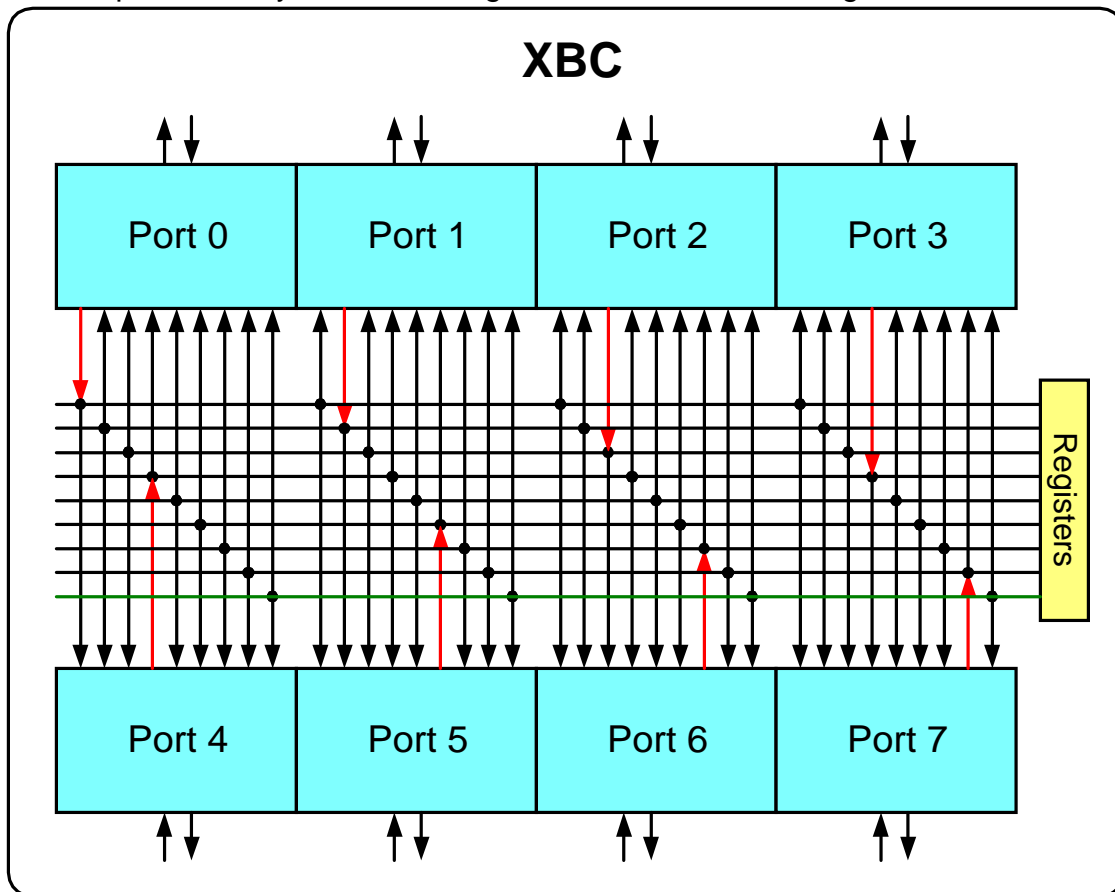


**Figure 7 – XBC Crossbar ASIC**

XBC links are highly reliable with nearly zero bit error rates. To cover connector failures, a two-level SEC/DEC scheme is employed so that any link can withstand a failure of any data wire. An efficient ECC is employed across the entire packet for correction of any errors at the destination. Performance critical routing and packet flow control information is parity duplicated so that each XBC can quickly correct errors on the fly without impacting latency. Additional link parity is used to exactly pinpoint a link with errors.

With a system that supports multiple instantiations of the operating system, it is imperative that XBC provide isolation between individual ports. Each port has independent queues and data paths. Arbitration is performed in lock-step within a XBC pair, with special error handling to ensure that packets arrive at their destination intact. Critical routing tables are protected by a lock-and-key mechanism for protection from stray writes. Extensive error logging is available to isolate failures, allowing firmware to route around, and avoid using, marginal links.

The fully-interconnected mesh of crossbars can be thought of as queues driving queues. Each sending queue has a dedicated receiving queue, reducing delays at arbitration points. The receiving queues tell their sending queues in advance how much empty buffer space is available in the form of credits. The use of credits guarantees the lowest latency for packets and eliminates queuing delays common in handshaking protocols. At chip crossings, the receiving queue is over-sized so that an uninterrupted stream of packets can be transferred without stalls while waiting for credits. Figure 8 details a single XBC port. The input queue is large to handle the delays associated with communicating off chip. The output queues are much smaller since credits are passed quickly on-chip. The use of individual destination queues for each port also minimize blocking delays by allowing the input port to drive a packet to a different destination even when an output port is waiting to win arbitration.
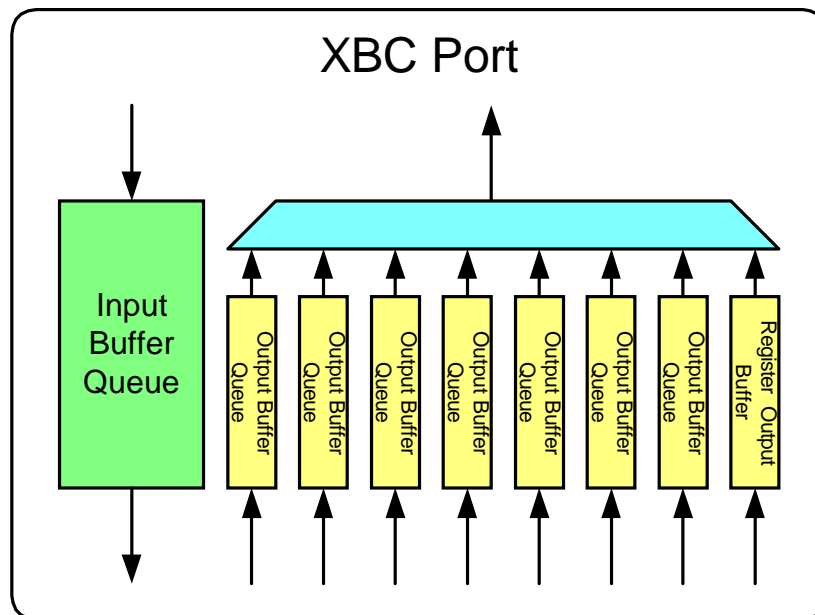


**Figure 8 – XBC Port Block Diagram**

## Performance

The key to HP9000 Superdome's performance is that it is designed as a large memory system with low latency to all memory. Memory is interleaved across cells providing scalability for all applications that work on traditional SMP systems.

Bandwidth is directly proportional to the number of cells in the system, reducing the cost of performance upgrades. To incrementally add bandwidth, simply add another cell to your partition. Overall system bandwidth of the system is matched to processors and memory to keep the latencies low.

Future upgrades will provide even better performance. The PA-8700 upgrade will feature lower latencies, higher frequency, and a larger cache. Memory capacity will increase as new SDRAM technology becomes cost competitive. Upgrades to IA-64 and later PA-RISC processors will have even greater performance with only a cell board upgrade.

## HP9000 Superdome Partitions

The HP9000 Superdome server allows multiple instantiations of the operating system to run simultaneously through a combination of hardware and software techniques. Each running operating system is confined to what is called a partition. The partitions are sometimes called "hard" since CPUs, memory, and I/O are exclusive to the partition and are never shared with other partitions. A partition consists of a collection of cell modules and connected I/O. The CC provides a hardware firewall between partitions. The Guardian Service Processor (GSP) and associated platform manageability infrastructure provide the features necessary for console output from each partition. A data structure called the "complex profile" holds all of the information necessary for software and firmware to build the partitions. Assistance from the GSP subsystem is crucial to broadcasting changes to all of the partitions.

Partitions are physical entities defined as an integral number of cell boards (a cell consists of 4 CPUs , 2 to 16GBs of Memory with 128 MBit DRAMs and a link to 12 PCI I/O Slots (optional)). Partitions hosting a specific application will provide dedicated resources, i.e. hardware isolation from a cell perspective and complete software isolation. A partition's I/O consists of the I/O chassis attached to cells in the partition.

A partition must contain at least one I/O chassis with a core I/O card along with sufficient I/O resources to boot the partition. An instance of an operating system uses the cells and I/O chassis assigned to the partition. Adding/removing cell boards to existing partitions does not impact any other Partition.

Partition configuration changes are made through the Partition Manager GUI or command line I/F.

Hardware firewalls are built into the system ASICs to insulate partitions from one another. Memory and I/O are private to the partition, not shared. The CC ASIC screens out external traffic. The XBC ASIC isolates link failures. A lock and key protects critical XBC ASIC data structures

Plan partition configuration wisely for performance and availability. Proper system configuration eliminates sharing hardware.

Dynamic addition of new partitions is supported. Changes to partitions require a reboot of only the affected partition(s). Follow-on releases of HP-UX will allow adding and deleting cells dynamically, i.e. without reboot.

Figure 9 shows two possible partitioning arrangements of an HP9000
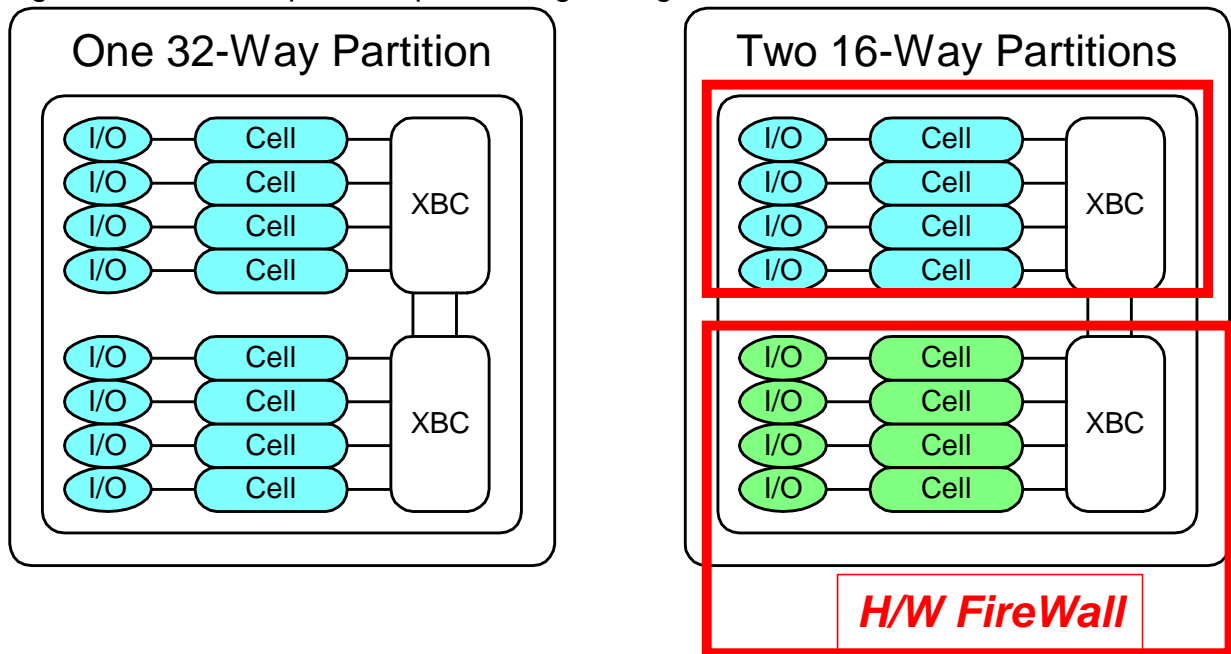


Superdome system. Many other arrangements are possible.

**Figure 9 – Partitioning arrangements**

The partition cell set register in each CC ASIC defines which cells are in the same partition as that cell. CC rejects transactions that originate from cells other than those in its partition.

# Partition Boot Process

HP9000 Superdome partitions go through the following process, from power on to booting HP-UX:

1. Power-On or Reboot. Power is turned on for components of the partition (or for the entire complex). Or the partition is reset or rebooted.

2. System firmware startup. Each cell runs its own copy of the system firmware (PDC).

3. Power-On Self Test (POST). Cells and other system components go through self-test (POST) activities. Early in POST, system firmware sets a flag to indicate to the GSP that shared memory is good, and the GSP can check the complex profile for consistency across cells. The GSP and PDHC push out the current or new copy of the complex profile to the cells if necessary. The check of the complex profile by the GSP happens concurrently with system firmware POST.

4. Boot-Is-Blocked (BIB). The BIB flag is a bit controlled by the GSP, and is used to prevent a partition from booting. There are a number of situations when the platform must delay or prevent a partition from booting:
   1. Partition boot must be delayed until the complex profile has been verified by the GSP.
   2. The BIB flag is used to prevent cells that have booted up as far as boot-is-blocked from proceeding to join into a partition, if the user had commanded a reset-for-reconfig.
   3. A cell that has been deconfigured by removing it from a partition, or by clearing its "use-on-next-boot" flag waits in the boot-is-blocked state.

Boot-is-blocked is an important state – it marks the dividing line between when the complex profile (which defines which cells are in which partitions) may be invalid or can be changed, and when it is known to be valid. System firmware will not use the complex profile until the BIB flag is cleared, but as soon as it is cleared, system firmware reads the complex profile to determine which other cells to rendezvous with in order to form a partition. The boot-is-blocked state marks a transition from cells being independent entities to cells attempting to rendezvous and act as a coordinated partition.

The list of cells that are waiting in the boot-is-blocked state can be examined from the GSP by viewing the Virtual Front Panel (VFP) for a partition. The actual state of the BIB flag on each cell is displayed by the PS command (cell status display). The VFP and PS command are described later in this paper.

Each cell whose use-on-next-boot value is "n" (do not use) remains at boot-is-blocked when the BIB flag is cleared, and does not proceed to rendezvous with the other cells in the partition. Likewise, when a partition is reset to a ready-for-reconfig state, the BIB flag remains set until the user issues the BO command from the GSP, or until a new complex profile is pushed out. This causes all cells in the partition remain at boot-is-blocked.

All cells at BIB are inactive cells that can be powered on or off or removed from the partition without affecting the rest of the partition.

5. Rendezvous. Cells proceed to rendezvous with the other cells in the partition once the GSP clears the BIB flag; they form the partition and proceed to boot to the BCH interface or boot the HP-UX operating system.

The partition rendezvous algorithm allows ten minutes for cells to complete self-test and rendezvous. Any cells that do not rendezvous within 10 minutes are declared missing, and excluded from the partition.

Partition rendezvous completes when either of the following occurs:
1. All cells with "use-on-next-boot" set complete self-test and rendezvous, OR
2. Ten minutes expires since rendezvous began (at that point, any cells that have not participated in rendezvous are not made available to the partition and thus are inactive cells).

Part of rendezvous involves selection of a "core cell" from among those cells that participate in the rendezvous. The core cell has the active core I/O board (CIO) attached to it, which is used as the console. The core cell runs all the remaining code up until HP-UX goes to multiprocessor mode.

6. Boot Console Handler (BCH). The BCH interface provides the main method for interacting with a partition during its boot process. BCH runs on top of SYSTEM FIRMWARE, and it provides menus for getting partition status details, for configuring various partition boot settings, and for issuing HP-UX boot and reboot commands.

When a partition's boot paths and boot actions are set accordingly, a partition can skip past the BCH interface and automatically boot HP-UX.

7. Initial System Loader (ISL) and Secondary System Loader (hpux). In most situations you need not use the ISL and hpux interfaces. However, when using the BCH interface's BO (boot) command you can select to stop at the ISL prompt to perform more detailed booting tasks. For example, you can use the ISL interface to boot HP-UX in single-user or LVM-maintenance mode, or to boot an HP-UX kernel other than /stand/vmunix.

8. HP-UX Operating System. The HP-UX operating system boots on a partition once ISL and the Secondary System Loader (hpux) specify which kernel is to be booted. Normally the /stand/vmunix kernel is selected.

While all of a partition's cells go through their individual self tests, only the core cell is active until HP-UX is booted.

**Boot-Related High Availability Features**

The design of the HP9000 Superdome includes a number of boot-related features that increase the availability of the system by reducing single-points-of-failure:

- Each cell board contains all the resources the cell needs to boot up to the point where it can rendezvous with the partition. This includes boot ROM, RAM, non-volatile storage, etc. The system-level resources necessary to boot have been minimized, and most of those are redundant (cooling, power).
- Cell-local storage of the complex profile (which defines the characteristics of each partition) eliminates dependence on a central resource for boot information. The complex profile data is replicated on each cell in non-volatile memory.
- Cell-local timeout on GSP validation of the complex profile. If the GSP does not contact a cell to validate its complex profile and release the BIB flag within a timeout window, a cell-local controller autonomously releases the BIB flag. This timeout allows partitions to boot even if the GSP is faulty, or if the communication channels between a cell and the GSP have failed.
- A partition can boot with a partial complement of cells. If a cell or cells fail to boot, and do not try to rendezvous, the rest of the cells in the partition will timeout at rendezvous and proceed to form a partition. They can then boot. (Proper configuration of I/O resources is needed to ensure that the loss of any one cell will not render a partition unbootable or unusable.)
- Partitions can be configured with redundant Core I/O boards. Every partition needs a root cell, and a root cell needs an attached Core I/O board. Configuring redundant Core I/O boards ensures that if the intended root cell cannot boot, there is an alternate cell that can become the root cell.

## Complex Profile

The complex profile is a data structure that defines the configuration that the server should boot into. Among other parameters, it defines:

- Which cells are assigned to which partitions, and which cells are unassigned (and therefore which are available to be assigned to any partition).

- The system serial number and other data from which the software ID (SWID, used for software licensing) is derived.
- The boot path for each partition.
- The "use-on-next-boot" flag for each cell

As described previously, the complex profile is replicated on each cell, and consistency between copies is ensured by the GSP.

The GSP creates the very first complex profile (before there are any partitions defined), and thereafter partitions can modify it. System firmware modifies the complex profile when the BCH user changes the boot path. Partition Manager modifies the complex profile under user control.

Neither hardware nor the OS running in a partition rely on the complex profile to keep running. The system design only requires that the complex profile be correct at boot time. Unintended changes in the complex profile cannot disrupt a partition that has booted and is running an OS.

## Console Device

The optimal configuration of console device(s) can depend on a number of factors, including the customer's data center layout, console security needs, customer engineer access needs, and how much an operator must interact with server or peripheral hardware and a partition (e.g. changing disks, tapes). This section gives some guidelines. However the configuration that makes best sense should be designed as part of site preparation, after consulting with the customer's system administration staff and the field engineering staff. It should not be assumed that yesteryear's solutions for single image systems that lacked LAN console capability are appropriate for a multipartition HP9000 Superdome system.

Customer data centers exhibit a wide range of configurations in terms of the preferred physical location of the console device. (The term "console device" refers to the physical monitor/keyboard/mouse system(s) that administrators and customer engineers use to access and control the server.) The HP9000 Superdome server enables many different configurations by its flexible configuration of access to the GSP, and by its support for multiple geographically distributed console devices.

Three common data center styles are:
- The secure site where both the system and its console are physically secured in a small area.
- The "glass room" configuration where all the systems' consoles are clustered in a location physically near the machine room.

- The geographically dispersed site, where operators administer systems from consoles in remote offices.

The three styles are really "dominant usage modes". Combinations of all three are possible. The dominant usage mode can drive different solutions to the console access requirement.

The considerations listed below apply to the design of provision of console access to the server. These must be considered during site preparation.

- The HP9000 Superdome server can be operated from a VT100 or an hpterm compatible terminal emulator. However some programs (including some used by field engineers) have a friendlier user interface when operated from an hpterm.
- LAN console device users connect to the GSP (and thence to the console) using terminal emulators which establish telnet connections to the GSP. The console device(s) can be anywhere on the network. The GSP has 2 network ports, and telnet sessions can connected to either one.
- The system design permits one-to-one, one-to-many, or many-to-one relationships between the number of console devices and the number of HP9000 Superdome servers.
- If an HPUX workstation is used as a console device, an hpterm window running telnet is the recommended way for the CE to connect to the GSP. If a PC is used as a console device, Reflection1 configured for hpterm emulation and telnet connection is the recommended way for the CE to connect to the GSP. Most programs used by system administrators work with a broad range of terminal emulations, and are much less sensitive to terminal type.
- The system administrator is likely to want to run X-windows applications or a browser using the same client that they access the GSP and partition consoles with. (The partition configuration tools, parmgr and sam have graphical interfaces.) The system administrator's console device(s) should have X-window or browser capability, and should be able to access the system LAN of one or more partitions as well as the GSP.
- The GSP currently supports a maximum of 16 telnet-connected users at any one time.
- In order to facilitate rapid repair, it is normally desirable, and sometimes essential for rapid time-to-repair to provide the service engineer with console and GSP access that is physically close to the server, so that he can get immediate access to the results of his actions. Access is also required to the System Management Station (SMS). This may require a console device close to the server. On the other hand, it may be satisfactory to ask the field engineer to carry in a laptop, or to walk to the operations center to access the SMS. Yet another alternative is to designate a system that is there anyway (such as the HAO workstation) for customer engineer use.
- Telnet data is sent between the client console device and the GSP "in the clear", i.e. unencrypted. This may be a concern for some customers, and may dictate special LAN configurations.

- The HP Central Web Console product, J6352AA, has been integrated with HP9000 Superdome and offers strong encryption of console data between the server it is running on and a user's web browser. The diagrams that follow show where to locate the Central Web Console product for best advantage.
- It is advantageous to connect a modem to the SMS, because the HP response center can use it to dial into the customer site and reach as many servers as are accessible from that SMS, and could also run firmware update and scan diagnostics. A modem is provided with certain levels of support contract.
- If a customer has purchased the support level that provides the HAO system, it provides the remote access capability and a modem is unnecessary.

## Sample configuration drawings

The following sections describe some sample network configurations of the HP9000 Superdome server, SMS, console, and HAO. Two styles of configuration are described, others may be appropriate to address particular situations.
1. The local subnet configuration. All network traffic involving console, GSP, scan, firmware update, and system management is carried on a physically isolated subnet. The customer may or may not provide access from other subnets.
2. The general network configuration. Console, GSP, and system management network traffic is carried on the customer's network (which may in fact be an isolated console net with many machines on it). Scan and firmware update data is carried on a physically isolated subnet because of its tendency for adverse interactions with general purpose traffic on heavily loaded networks.

The configuration drawings below require one, or in some cases, two hubs plus cabling to implement. Some customers may want to use existing network infrastructure for these functions. In other cases, customers may want a turnkey solution. This should be understood during site planning, and appropriate equipment should be ordered.

### Console and SMS Local Subnet Configuration
This configuration clusters the GSP and the HP9000 Superdome support station (and HAO if provided) on a local subnet. It is appropriate for:
- Small secure configurations with no outside network connections
- Sites where the customer desires the extra security of requiring users to log in through the test station to the private network.
- Sites where the customer wants to encrypt console data as it travels across the customer site LAN. Encryption between the SMS and a user terminal on the site LAN can be handled in two ways:
  1. The Central Web Console software can be run on the SMS.
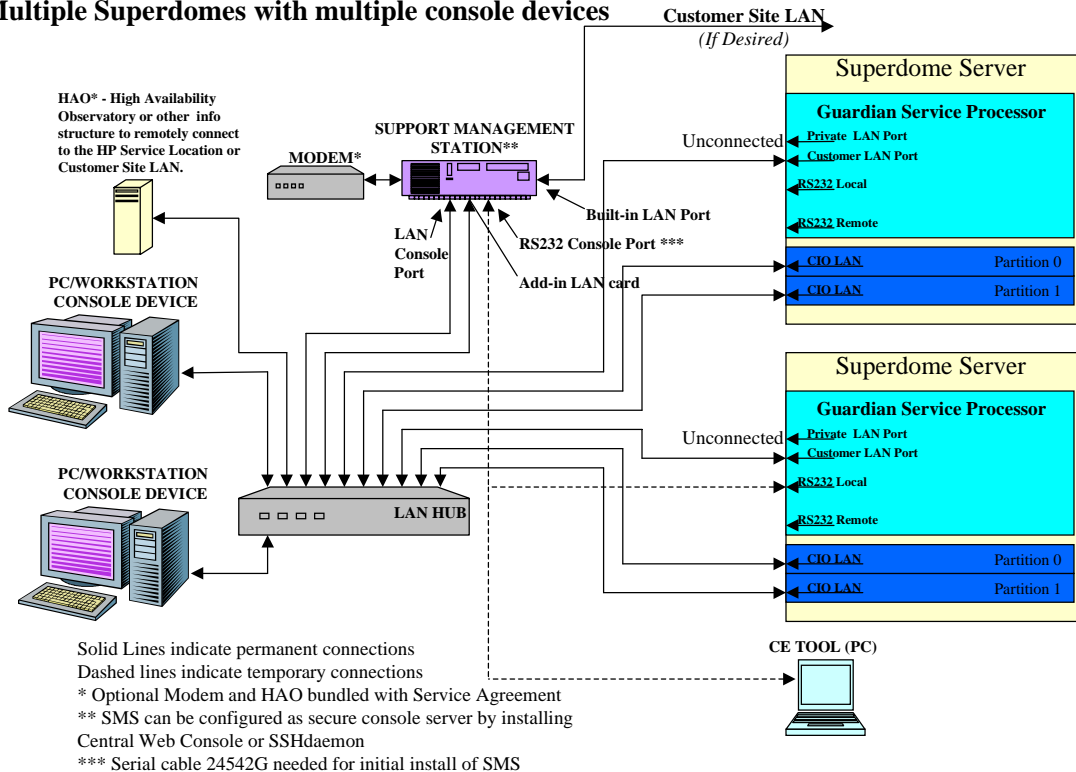  2. An SSH daemon can be run on the SMS.

The Local Subnet Configuration has the following characteristics:

- Console, GSP, scan, firmware update, and system management traffic is confined to a (potentially) physically isolated subnet.

A multi-server configuration is shown below. Two console devices are shown as an example of redundancy, but zero or one could be provided, depending on the design of the console network.

## Console and SMS Local Subnet Configuration
**Multiple Superdomes with multiple console devices**



Solid Lines indicate permanent connections
Dashed lines indicate temporary connections
* Optional Modem and HAO bundled with Service Agreement
** SMS can be configured as secure console server by installing
Central Web Console or SSHdaemon
*** Serial cable 24542G needed for initial install of SMS

Variations on the theme:
- Use the HAO for console access by the customer engineer. Eliminate PC/Workstation console device.
- Redundant or non-redundant console devices.
- Set up SMS to route onto private subnet. This allows graphics from the system (e.g. parmgr X-windows) to be displayed both on the console device and on remote devices.
- Eliminate console access PC/Workstation. Appropriate if the CE can bring his/her laptop on site to use as a console when working close to the server, and minimal interaction with the server is required.
- Install SSH (Secure Shell) on the SMS, thus providing strong encryption for console traffic outside the local subnet.

**General Network Configuration**

This configuration clusters the GSP and the HP9000 Superdome support stations on a general-use customer subnet that may be an administration network. The SMS uses the HP private network for scan data and firmware update only. It is appropriate for:
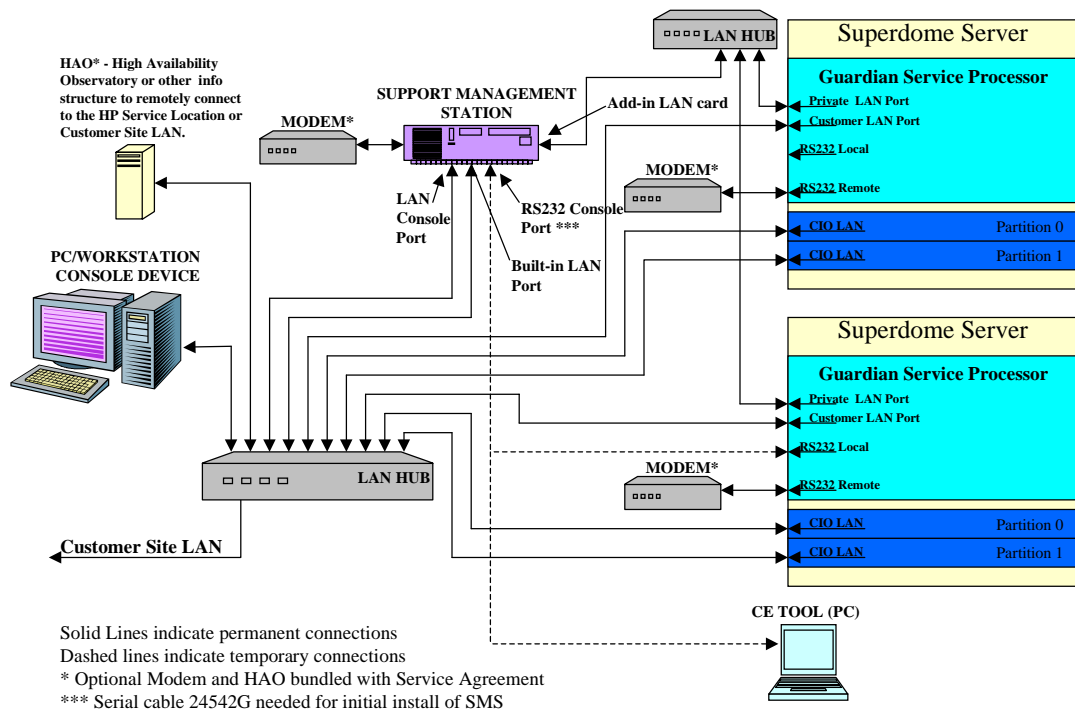
- Sites that have geographically dispersed console devices
- Sites that provide console security by putting all consoles on a dedicated and isolated network.
- Sites with a control center near the machine room (place the console access device in the control center).

It has the following characteristics:

- Console and system management data is available from any console device connected to the customer's network.
- Configurations of more than one HP9000 Superdome server connected to an SMS require an additional hub to connect the private network together. If there is only one server, a crossover ethernet cable can be used.

## Console and SMS General Network Configuration
**Multiple Superdomes with one console device**



HAO* - High Availability Observatory or other info structure to remotely connect to the HP Service Location or Customer Site LAN.

Solid Lines indicate permanent connections
Dashed lines indicate temporary connections
* Optional Modem and HAO bundled with Service Agreement
*** Serial cable 24542G needed for initial install of SMS

Variations on the theme:

- Use the HAO for console access by the customer engineer. Eliminate PC/Workstation console device.
- Redundant or non-redundant console devices.
- Eliminate console access PC/Workstation. Appropriate if the CE can bring his/her laptop on site to use as a console when working close to the server, and minimal interaction with the server is required.

## Console Equipment Ordering Checklist

After planning the console network configuration, the proper hardware and software components must be ordered. The checklist below lists the items discussed in this paper that might be required.

___ Brio PC console device (0 or more)
___ Reflection1 software
___ ReflectionX software
___ B2000 HPUX console device (0 or more)
___ A500 Support Management Station (at least one per site)
___ 16-Port 10/100Base-T Hub or router (0 or more)
___ Ethernet cables (lengths?)
___ Crossover ethernet cable (0 or 1)
___ Central Web Console software


## Utilities Subsystem

The utilities subsystem is the part of the platform that monitors and allows the user to control the platform, but does not run on the system processors. It is centered around the Guardian Service Processor (GSP), and provides:

- Logging of diagnostics emitted by SPU components
- Access to partition consoles via RS232 or telnet
- Interface to system firmware through shared memory and registers in cell PDH
- Support for JTAG scan diagnostics
- Monitoring of power, cooling, and environmental conditions
- Control over power and reset to SPU components
- Initial partition configuration
- Virtual Front Panel
- Firmware update

The components of the Utilities Subsystem are:

- GSP – Guardian Service Processor. The GSP provides core capabilities that are always available, even when no OS is booted.
- CLU – Cabinet Level Utilities, backplane and I/O card cage monitoring and system scan support.
- PM - Module which monitors and controls 48 VDC, monitors cabinet environment (ambient temperature), and controls power to cells and I/O bays.
- PDHC – PDH Controller. The module which provides communication between system firmware running on cells and the GSP.
- CIO – Core I/O card. Provides the console interface for each partition, and a management LAN port from each partition.
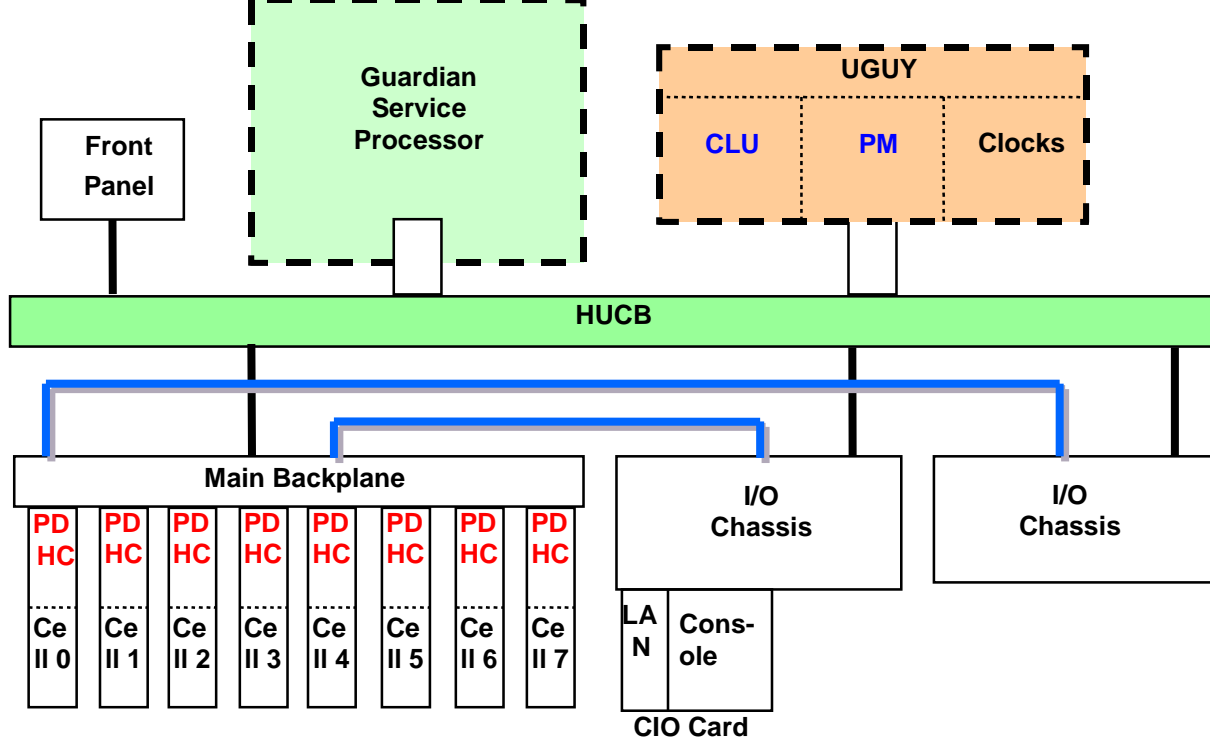
The block diagram below shows the arrangement of the utilities subsystem components.

**Figure 10 – Utilities Subsystem**

A serial bus known as the GSP Bus traverses the HUCB board and cables and connects the GSP to the CLU, PM, each PDHC, and each CIO card in the system. The GSP collects management data from each module, and commands each module to act on its behalf. For example, it commands the PDHC to update the complex profile, or to reset a cell, and it passes console data between users and the active CIO card in each partition.

## Guardian Service Processor

The Guardian Service Processor (GSP) delivers a wide range of platform managability features that enable highly flexible and capable control and monitoring of the HP9000 Superdome server hardware. The GSP is also intimately aware of the configured partitions, and offers many facilities to ease access to, and management of partitions. It delivers value primarily to the system administrator and the customer engineer by giving them a powerful tool to operate the server and its partitions.

The GSP supports multiple partitions, all operating independently. It also supports multiple simultaneous logins (same or different users), each potentially doing different tasks.  Therefore, a system administrator is able to open several independent console connections through the GSP to several partitions while a different administrator is able to simultaneously monitor the run state of partitions using the virtual front panel.

Some of the partition-related capabilities of the GSP are:
- Multiple connections to partitions' consoles allow monitoring and control of firmware and the OS.
- Multiple virtual front-panels, each displaying boot or run status of the cells in selected partitions.
- Reset or TOC a selected partition.
- Display system or partition events live as they are emitted by partitions (for example, while booting), optionally filtered by partition or cell. Display a log of events.
- Automatic reset of a partition if it should hang (configurable).
- Create an initial partition configuration.

Some of the server-related functional capabilities of the GSP are:
- Support the System Management Station as it runs scan diagnostics on the server, or as it updates server firmware.
- Provide the user with control of power to the cabinet, and to cells and I/O chassis.
- Monitor and display the power, temperature, and configuration of the server components.

The GSP works with other system components as part of the high availability design. The GSP is not a single point of failure - it can be reset or replaced if it should fail, without impacting the ability of partitions to reboot or run.

Users interact with the server's consoles and hardware through the GSP. The GSP provides several features that enable remote access:
- Local console physical connection (RS232) - allows connection to one partition at a time or to the GSP.
- Remote console physical connection (RS232 with modem) - allows dialup connection to one partition at a time or to the GSP. If a modem is not connected to the remote RS232 port, it can be used as a second local serial port.
- HP private LAN port – allows multiple telnet sessions to the GSP from the Support Management Station.
- Customer LAN port – allows multiple telnet sessions to the GSP from the customer network (if connected).

Either LAN port can be used for telnet connections, scan diagnostic data from the SMS, and firmware update data from the SMS. However, their behavior varies based on the routing table in the GSP as described below.
- Inbound packets are accepted from either LAN port with no distinction.
- Outbound packets destined for the subnet connected to the private subnet (as determined by comparison of the IP addresses masked with the private port subnet mask) are sent to the private LAN port.
- All other packets are sent out the customer LAN port.

- If the customer and private LAN ports are on the same subnet, all packets are sent out the customer LAN port.

## GSP Login Accounts and Access Levels

To use the HP9000 Superdome service and console interfaces you must have account access to log in to the server's Guardian Service Processor (GSP).

Each HP9000 Superdome complex has its own set of GSP accounts, which are defined for the complex and may differ from accounts on other Superdome servers.

The three GSP access levels are:
1. Administrator. Provides access to all GSP commands, including those that configure the GSP itself and its user accounts, and access to all partition consoles and Virtual Front Panels.
2. Operator. Provides access to the subset of GSP commands that do not reconfigure the GSP, and access to all partition consoles and Virtual Front Panels.
3. Single Partition User. Provides access to the subset of GSP commands that can affect a single partition, and access to a single partition's console and a single partition's Virtual Front Panel. Each single partition user account is tied to a specific partition.

Each GSP login account has a specific login name and password. Also, each GSP account permits multiple concurrent login sessions.

## GSP Main Menu

After logging in, the user is presented with the GSP main menu. Figure 11  below shows the main menu.
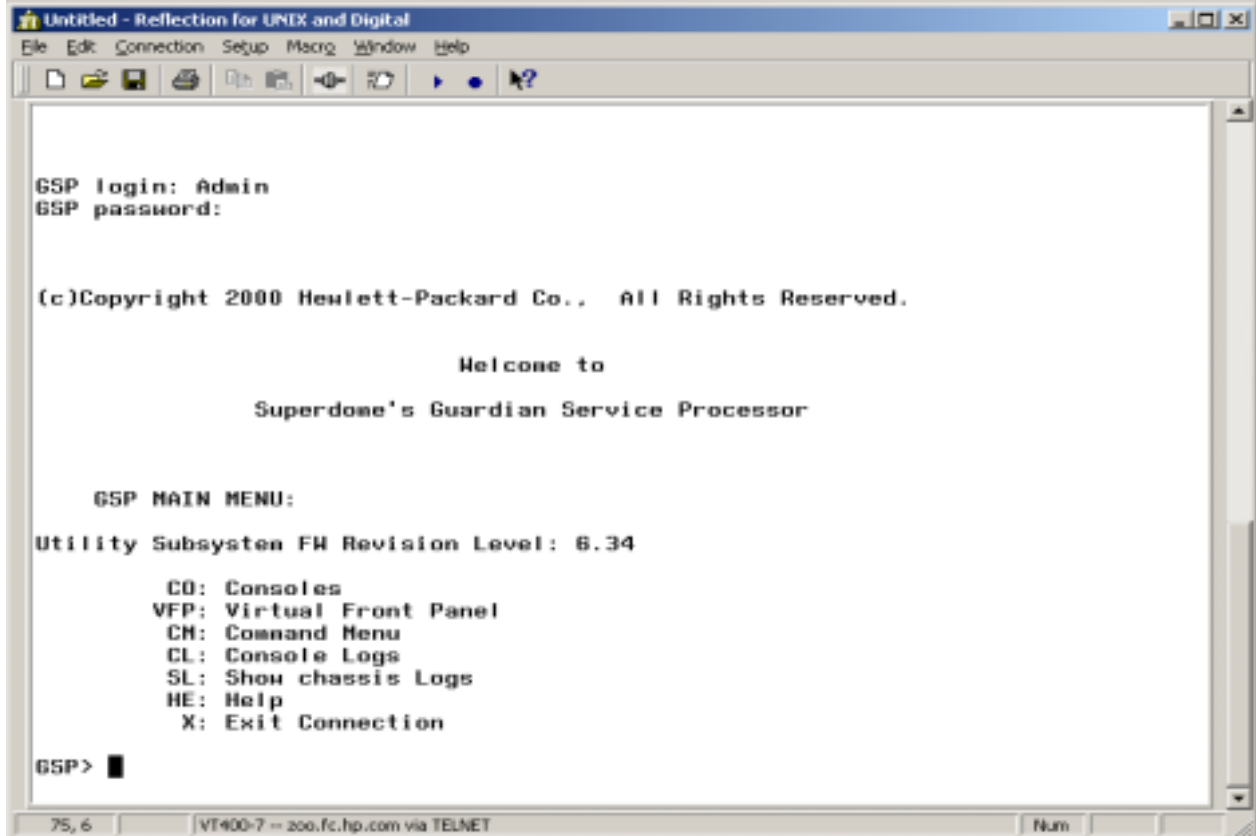
**Figure 11 – GSP main menu**

Note that if no partitions are defined, the CO entry will not appear.

The main menu shows the revision level of the GSP and utilities subsystem firmware. If an unrecognized combination of firmware revisions is present on the boards in the system, the main menu will show that there is an "Unsupported Utility Subsystem FW Revision Combination".

The main menu allows the user to enter the various major subsystems of the GSP. The following sub-menus are available from the GSP Main Menu:

- Console Menu--Provides access to consoles for the HP9000 Superdome server's partitions.
- Virtual Front Panel Menu--Provides a Virtual Front Panel for various partitions or for the entire system.
- Command Menu--Includes service, status, and system access commands.
- Console Log Viewer Menu--Allows access to the HP9000 Superdome server's console logs.
- Chassis Log Viewer Menu--Allows access to the HP9000 Superdome server's chassis code logs.
- Online Help – Provides help on a variety of topics.

These GSP menus provide a central point for managing an HP9000 Superdome complex outside of HP-UX. While the GSP menus provide many tools and details not available elsewhere, additional features are available from HP-UX running in a Superdome partition.

Figure 12 below shows the submenus below the main menu:



**CO — Console Menu**
```
Partitions available:

    #    Name
   ---   ----
   0)   Partition 0
   1)   Partition One
   Q)   Quit

   Please select partition number:
```

**VFP — Virtual Front Panel**
```
Partition VFP's available:
    #    Name
   ---   ----
   0)   Partition 0
   1)   Partition One
   S)   System (all chassis codes)
   Q)   Quit

GSP:VFP>
```

**CM — Command Menu**
```
Enter HE to get a list of available
commands

GSP:CM>
```

**CL — Console Log Menu**
```
Partition Console Logs available:
    #    Name
   ---   ----
   0)   Partition 0
   1)   Partition One
   Q)   Quit

GSP:VW>
```

**SL — Chassis Log Menu**
```
Chassis Logs available:

   (A)ctivity Log
   (E)rror Log
   (L)ive Chassis Logs
   (Q)uit

GSP:VW>
```

**GSP Main Menu**
```
GSP MAIN MENU:

        CO: Consoles
       VFP: Virtual Front Panel
        CM: Command Menu
        CL: Console Logs
        SL: Show chassis Logs
        HE: Help
         X: Exit Connection

GSP>
```
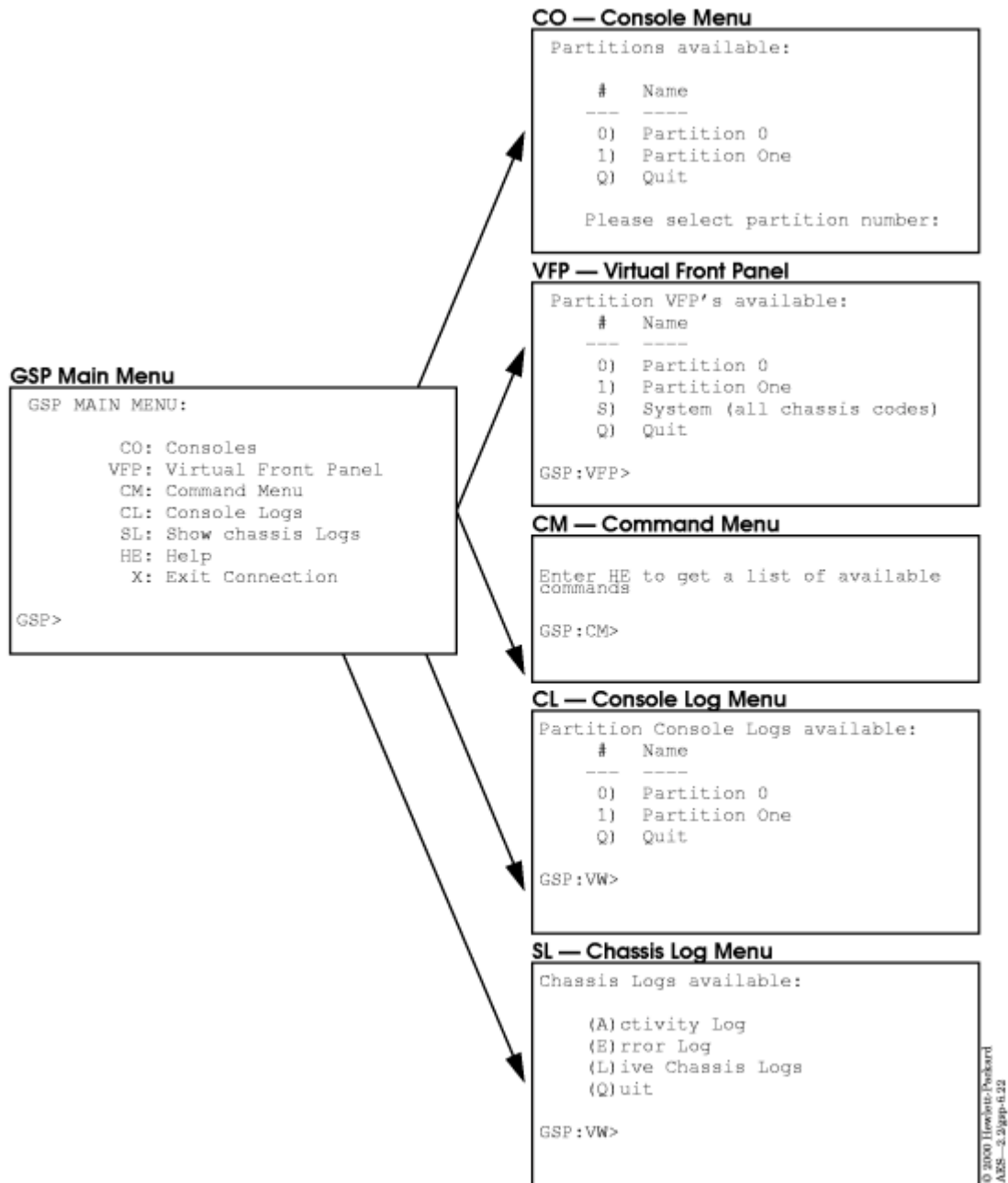
**Figure 12 – GSP submenus below main menu**

Each major area is described below.
**Partition Consoles**

Each partition on an HP9000 Superdome server has its own console. Enter CO at the GSP main menu to access the partition consoles. Each partition's console provides access to:

- Boot Console Handler (BCH) for the partition. The BCH menu is available if the partition has booted but has not yet loaded       or booted the HP-UX operating system.
- HP-UX Console for the partition. This partition console provides console login access to HP-UX and serves as /dev/console for the partition.

To exit the console, type ^b (Control-b).

Multiple users can connect to the same partition's console. Each will see all output, and thus will see what the other users type. Only one of the connected users has the ability to type at any given time. Any connected user can gain control of the console (obtain the ability to type) by entering ^Ecf (Control-E c f). It is important that users not leave themselves logged into the console when they leave a console connection (using ^b) and return to the GSP main menu. Console users are not automatically logged out when they leave a console connection and return to the main menu.

### System Log File Viewers

 Two types of logs are maintained for each HP9000 Superdome complex: (1) partition console logs and (2) complex chassis code logs. Access to these logs is available through the following GSP menus:

- The Console Logs Menu (CL). Enter CL from the GSP main menu to access the console logs menu. To exit the console log, type ^b (Control-b). Each partition has its own console log, which has a history of the partition console's output, including boot output, BCH activity, and any HP-UX console login activity.

    The console log provides a limited history; it is a circular log file that overwrites the oldest information with the most recent.

- The Show Chassis Logs Menu (SL). Three types of chassis code logs are maintained for each HP9000 Superdome complex: (1) activity logs, (2) error logs, (3) live chassis code logs. The activity log and error log provide a snapshot of past chassis codes.

    The live chassis code log provides a real-time view of chassis codes and offers the option of filtering the real-time chassis output to show only a specific cell, or specific partition.

    All logs (activity, error, and live) can be displayed in different formats, including: text format, keywords with hex, and raw hex format.
    At the log menu, type V to change the display format, and type ^b (Control-b) to exit the log viewer.

### The GSP Command Menu

GSP commands are provided for system service, status, access configuration, and manufacturing purposes. To exit the GSP command menu, enter MA to return to the GSP main menu.

### Virtual Front Panels

Virtual front panels provide a real-time display of each partition's boot status and activity. The are functionally equivalent to the LCD display found on many single partition servers. To exit a Virtual Front Panel, type ^b (Control-b).

### Help

The online help subsystem provides help on a variety of topics., including GSP and HP9000 Superdome concepts, and GSP commands. Online help has a copy of the release notes for the GSP firmware.

## Navigating through GSP Menus

This section describes how to navigate through HP9000 Superdome GSP menus and sub-menus.

As Figure 13 below shows, there are various commands and options for returning to the GSP Main Menu and escaping or ending a GSP login session.

The following list has additional tips for navigating through GSP menus and using GSP features:

- Control-b. Exit current console, console log, chassis log, or Virtual Front Panel. When accessing a partition's console, any log files, or any Virtual Front Panel (VFP), you can exit and return to the GSP Main Menu by typing ^b (Control-b).
- Q (or lower-case q). Exit or cancel current menu prompt. Enter Q (or lower-case q) as response to any menu prompt to exit the prompt and return to the previous sub-menu. You can do this throughout the GSP sub-menus, including the console menus, various command menu prompts, and the log and VFP menus.
- From the GSP command menu prompt (GSP:CM>) you must enter MA (not Q) to return to the GSP main menu. However, you can enter Q or q to cancel any command.
- Control-]. Escape the GSP connection and return to the telnet prompt. At any time during your telnet connection to a Superdome server's GSP,

you can type the ^] (Control-right bracket) escape sequence. This key sequence escapes back to the telnet prompt. When at the telnet> prompt you can use the following commands, among others: ? (print telnet command help information), close (close the current connection), and quit (exit telnet). To return to the GSP connection, type enter (or return) one or more times.
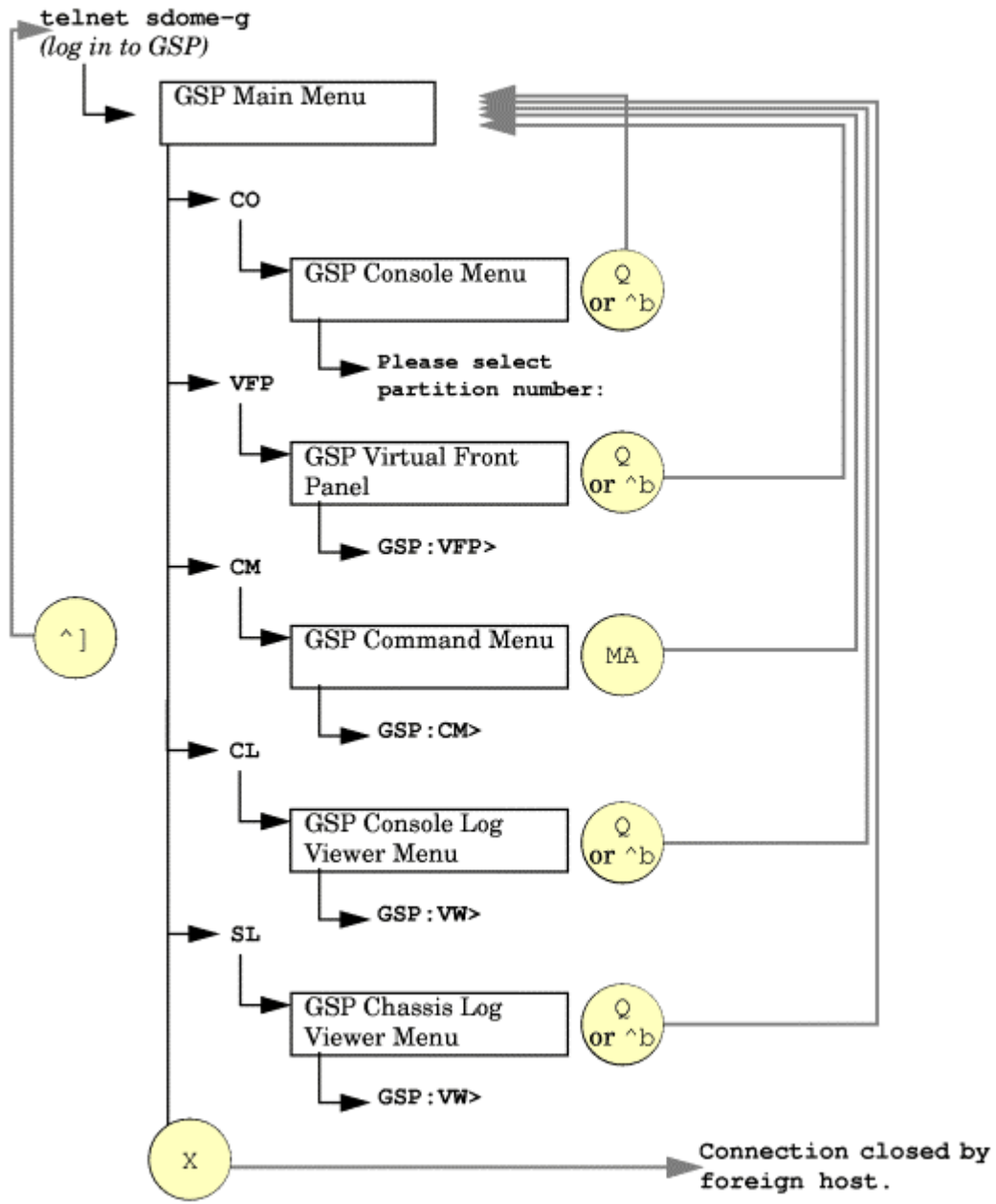


**Figure 13 – Exiting various states**

## Virtual Front Panel

The Virtual Front Panel (VFP) provides ways to monitor the operation of a particular partition or the entire system (all partitions). It is similar to the LCD display found on many single partition servers. It provides the following information:

- Indication that the cells and/or partitions are making forward progress through the boot sequence.
- Indication if cells or partitions get stuck (hang while booting), or restart the boot.
- Indication of approximately how far through the boot each cell and partition has gone.
- Indication of the major state of each cell or partition (e.g. powered off, booting, at BCH, running HP-UX).
- Indication of whether an error was detected during boot.

As partitions boot they emit chassis logs. In addition to saving the logs for later analysis, the GSP analyzes the logs and tracks the major state of each partition. When a user enters the VFP, they are presented with a real-time display (updated every second), showing the analysis of the latest logs received from each cell or the partition (partition view) or from all partitions (system view).

Figure 14 below shows a screenshot of a partition booting.

```
E indicates error since last boot
     Partition 0  state              Activity
     -------------------              --------
     Cell(s) Booting:      710 Logs

  #  Cell state                      Activity
  -  ----------                      --------
  0  Early CPU selftest              Cell firmware test          232  Logs
  1  Early CPU selftest              Processor test              230  Logs
  2  Memory discovery                Physical memory test        242  Logs




















GSP:VFP (^B to Quit) >
```

**Figure 14 – VFP during boot**

Several pieces of information can be gathered from this display.

- The partition has not rendezvoused yet. The cells are still booting independently. This is indicated by the partition state "Cell(s) booting", and the cell state not indicating that cells have joined a partition.
- The major state of cells 0, 1 and 2 is shown in the "Cell State" column, and their major activity is shown in the "Activity" column.

- The number of logs received from each cell since the start of boot is shown on the right of the cell state display. If this number stops incrementing before the cells rendezvous into a partition, it indicates that a cell has hung during boot. If it stays blank, the cell may be unable to start booting (possible hardware failure). If the number goes back to zero, it indicates that the cell reset itself, and is rebooting itself.
- No errors have been detected in the boot so far. If an error is detected, the display shows 'E' next to the cell number. An error is defined as a chassis log with an alert level of 2 or greater.

Figure 15 below shows the VFP display after the cells have rendezvoused into a partition.



**Figure 15 – VFP display after rendezvous**

Several pieces of information can be obtained from this display:
- All 3 cells have rendezvoused and joined the partition.
- The partition is at the BCH prompt, and is therefore waiting for user input.
- The number of logs emitted since the cells rendezvoused into a partition is shown on the right of the partition state. This value will increment while booting proceeds until the boot has completed.
- The partition has not detected an error since it formed (which occurred when the cells rendezvoused). An error would be indicated by an E on the left of the partition state.

Figure 16 below shows a partition in an important state – Boot Is Blocked (aka BIB). In this state all of the cells in a partition have completed POST, the GSP can be used to boot the partition, and the cells at BIB can have their partition assignment changed.

```
X telnet udome                                                    _ □ X
E indicates error since last boot
     Partition 0  state              Activity
     ------------------              --------
     Cell(s) Booting:    1878 Logs

  #  Cell state                      Activity
  -  ----------                      --------
  0  Boot Is Blocked (BIB)           Cell firmware              706  Logs
  1  Boot Is Blocked (BIB)           Cell firmware              583  Logs
  3  Boot Is Blocked (BIB)           Cell firmware              583  Logs




GSP:VFP (^B to Quit) > █
```
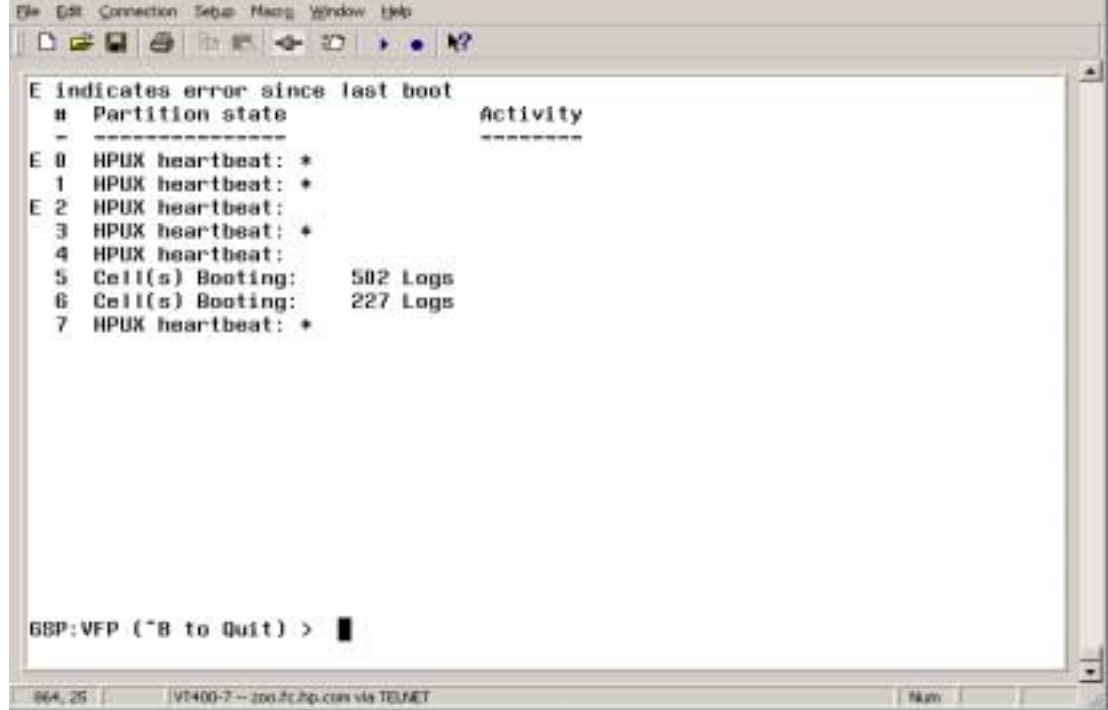
**Figure 16 – VFP showing partition at Boot-is-blocked**

The VFP also has an option to show the status of all partitions. Figure 17 shows 6 partitions running HP-UX and 2 in the process of booting. Several pieces of information are available from this display.
- Partitions 0 and 2 have experienced errors while booting
- Partitions 0 – 4 and 7 are running HP-UX. The asterisk flashes each time an HP-UX heartbeat chassis code is received. This allows the user to keep a system VFP up and see at a glance whether all cells are still running.
- Cells 5 and 6 are still booting and have not rendezvoused yet. When they rendezvous, the state changes from "Cells Booting" to the partition state.
- As in other examples, the log count shows forward progress.

**Figure 17 –
System VFP
showing
booted Oss**

```
 Ble  Edt  Connection  Setup  Macro  Window  Help

  D  ⊟ ⊞  ⊜  ⊞ ⊞  ◆ ⊅  ▸ ●  ▶?

   E indicates error since last boot
     #  Partition state                 Activity
     -  ----------------                --------
   E 0  HPUX heartbeat: *
     1  HPUX heartbeat: *
   E 2  HPUX heartbeat:
     3  HPUX heartbeat: *
     4  HPUX heartbeat:
     5  Cell(s) Booting:      502 Logs
     6  Cell(s) Booting:      227 Logs
     7  HPUX heartbeat: *




   GSP:VFP (^B to Quit) >  ▮

  864, 25        VT400-7 -- zoo.fc.hp.com via TELNET              Num
```

## GSP Command nd Menu

The GSP command menu provides the mechanism for managing the server and its partitions when no OS is booted. It is constructed in a question-and-answer style, and is similar to the service processor UI of previous HP servers (e.g. T-class SP, N-class GSP). It allows the user to:

- Control aspects of the server like reset or Transfer-Of-Control (TOC), system power
- Configure the GSP (users, IP address and other communication parameters)
- Display power & environmental status
- Display system configuration

The GSP command menu is a shared environment. All users in the command menu share the same environment, and each can see what is typed by other users.

Once a user in the command menu types a command name (e.g. RS), they own the interface until either the command completes or they get timed out and the command is aborted. Only the user who owns the interface can type into the command.

There are over 40 commands. The most important ones are listed below:

- RS    Reset partition
- TC    TOC partition
- RR    Reset partition for reconfiguration
- BO    Boot partition
- CC    Create genesis complex profile
- CP    Complex Profile Report
- LS    Display LAN status
- LC    Set LAN configuration
- ND    Enable/disable network diagnostics

- PE    Power enable cabinet
- PS    Display power and configuration status

Each of these commands is described below.

### RS – Reset Partition
Under normal operation, shut down the OS before issuing this command.

Execution of this command irrecoverably halts all system processing and I/O activity and restarts the partition.  The effect of this command is very similar to cycling the system power - the OS is not notified, no dump is taken on the way down, etc. The partition is allowed to reboot.

The user is asked which partition they want to reset.

### TC – TOC partition
Under normal operation, shut down the OS before issuing this command.

This command causes the selected partition to be reset through the INIT (or TOC) signal.

Execution of this command irrecoverably halts all system processing and I/O activity and restarts the partition.  It is different from the RS command in that the processors are signaled to dump state on the way down.

The user is asked which partition they want to TOC.

### RR – Reset partition for reconfiguration
This command resets the specified partition but does not automatically boot it. The Boot-Is-Blocked bit is not cleared until the user runs the BO command, or Parmgr downloads a new complex profile.

The user is asked which partition they want to reset.

### BO – Boot partition
The BO command boots the specified partition. It makes sure that all the cells assigned to the target partition have valid complex profiles, and then releases the Boot-Is-Blocked (BIB) flag on each cell.  This action frees system firmware to boot.

### CC – Create Complex Profile
This command distributes a selected complex profile to all cells in the complex. The user can select either the last complex profile successfully downloaded to the GSP, or a genesis complex profile for distribution.  The genesis complex profile is used when there is no valid complex profile.

The genesis complex profile allows the user to configure partition 0 as a single (user specified) cell. The selected cell must have a Core I/O board connected to it.

In order for the CC command to complete, system firmware boot must have progressed far enough for it to power on the CIO card. The easiest way to ensure this is to use the RE command to reset the selected cell, then wait until its CIO shows up in the PS display, then run the CC command. All remaining cells are left unassigned to any partition.

The CC command makes no assumption about the state of the complex before beginning the configuration. The user must ensure that no partitions are running. The ID command may be required to fully generate a valid complex profile. The CC command does not boot any partitions. Use the BO command to boot the generated partition(s).

The ID command displays information from the complex profile such as the system serial number, model number, as well as other information that uniquely identifies the system. The HP service engineer can change this information if necessary.

## CP – Complex Profile report
This command displays the current assignment of cells to partitions.

## LS – LAN Status
The LS command displays all parameters in the current configuration of the GSP LAN ports.  The LAN parameters are not modified by the execution of this command.

## LC – LAN Configuration
The LC command modifies the LAN configuration for both LAN ports. The user can set the IP address, network name, subnet mask, and gateway. The user must have administrator privileges to use this command.

## ND – Network diagnostics enable/disable
The ND command enables or disables network diagnostics. Disabling network diagnostics is recommended for production machines, because it prevents unauthorized access to the machine by scan diagnostics, the firmware update utility, and the OS debugger. The user must have administrator privileges to use this command.

## PE – Power Enable
For proper system shutdown, shut down the OS before issuing this command.

This command allows the user to switch the power ON or OFF to a cabinet, a cell, or an I/O chassis.

The PS command is the primary source of power and hardware configuration and status information. When invoked, it displays all of the slave microprocessors that it has detected in the system:

- PDH Controller - one on each cell board
- Power Monitor (PM) and Cabinet Level Utilities (CLU), both on the UGUY board. One of each per cabinet.
- Core I/O (CIO)

Pressing the enter key redisplays the detected modules. This can be useful if you're waiting for CIO to come online during SYSTEM FIRMWARE boot, prior to running the CC command or firmware update.

After displaying the detected slave microprocessors, the PS command allows the user to display more detail of the power and configuration status of cabinets, cells, the Core I/O, and the GSP.


## Complete GSP command list

The following list shows all the commands available from the GSP, and their function.

These commands make connections or change mode.

```
CL       : Console Logs.
CM       : Command Menu.
CO       : Server COnsoles.
HE       : Online help, or display the list of available
commands.
MA       : MAin menu.
PD       : Modify default Partition for this login session.
SL       : Show chassis Logs.
TE       : Broadcast a message to all users of the GSP
Command Handler.
WHO      : Display a list of GSP connected users.
VFP      : Virtual Front Panel.
X        : EXit from GSP connection.
```

These commands display status and configuration of the server.

```
CP       : Display partition cell assignments.
DF       : Display FRU information of an Entity.
PS       : Display detailed power and hardware configuration
status.
```

These commands control or configure the server.

```
BO        : Boot a partition.
CC        : Initiate a Complex Configuration.
ID        : Change certain stable complex configuration
profile fields.
PE        : Power entities on or off.
RE        : Reset entity.
RL        : Rekey Complex Profile Lock.
RR        : Reset a partition for reconfiguration.
RS        : Reset a partition.
TC        : Send a TOC signal to a partition.
VM        : Margin the voltage in a cabinet.
```

These commands control or configure the GSP.

```
AR        : Configure the Automatic System Restart
DATE      : Set the time and date
DC        : Reset parameters to default configuration
IT        : Modify command interface inactivity timeout
ND        : Enable/Disable Network Diagnostics (scan, gdb,
firmware update)
SO        : Configure Security Options and Access Control
XD        : GSP Diagnostics and Reset
```

These commands manage GSP ports.

```
CA        : Configure Asynchronous and Modem parameters
DI        : Disconnect Remote or LAN Console
DL        : Disable Lan console access
EL        : Enable Lan console access
ER        : Configure remote/modem port access options
LC        : Configure GSP LAN connections
LS        : Display GSP LAN status
MR        : Modem Reset
MS        : Display the status of the Modem
```

## OS-Present Platform Management

This section provides a brief overview of HP9000 Superdome platform
management capabilities that run under an O/S either on a partition or on a
management station.

### Partition Manager

Partition manager and the command line interface run under an instance of HP-
UX running in a partition. They are used to reconfigure Superdome partitions.
They are described in the paper "Superdome Management Part 2".

### Event Monitoring Service (EMS)

Chassis logs are data records that describe an event or a server condition that
has been discovered by system software or firmware. As previously described,
they can be viewed from the GSP SL command.

Different parts of an HP9000 Superdome server emit chassis logs as they operate, and especially when they discover errors. Some examples of chassis logs are given below, sorted by the entity that creates the log entry.

- Power Monitor logs
    - 48V Bulk Power Suppy failed
    - Cabinet fan failed
- GSP logs
    - Complex profile distributed
    - New device found on GSP bus
- System firmware logs from a partition
    - XBC initialization error
    - PCI slot initialized
    - Cells have rendezvoused into a partition
- HP-UX logs from a partition
    - Boot complete
    - OS Panic

Chassis logs are routed through the system to the GSP, which stores them in its logs. The GSP also turns around all the chassis logs from each source and reflects them back to each partition. A daemon on each running partition gathers the reflected chassis logs, stores them in a log file on the partition, and passes them to the Event Monitoring Service.

The Event Monitoring Service (EMS) is a component of the online diagnostic system, and analyzes each reflected log and decides whether to issue an alert based on the meaning of the log. It also uses MIBs, device APIs, logs, and other interfaces to detect system errors. It issues notifications to registered clients using a variety of protocols, including OpC messages to Openview, SNMP traps, email, etc. Email can cause a paging service to send a message to the operator.

### Openview

The HP9000 Superdome server provides a DMI interface. Openview can access the DMI database.

### HAO and HPEN

The High Availability Observatory (HAO) program situates a workstation at the customer site which connects back to the  HP mission critical support center. It runs a software bundle that monitors the status of HP9000 servers on the customer site. One of the HAO software components is the HP Event Notifier (HPEN). HPEN periodically logs into the GSP and collects errors that the GSP may have logged. HPEN also leverages the event monitoring service (EMS) to obtain notification of hardware and system failures as they occur.

Any events or chassis logs which are transferred to the HP mission critical support center are reviewed by a 24x7 monitor team, which creates workflow cases as required. A workflow case may result in an HP CE being dispatched to the customer site, possibly before the customer even knows they have a problem.

## References

1. Superdome Management Part 2
2. Managing Superdome Complexes: A Guide for HP-UX System Administrators