HP nPartition Servers

Architecture and Operations

Paul Bouchier

Hewlett-Packard Corp.

July 21, 2001

Reference Web sites

The Motherlode:

- http://docs.hp.com/hpux/pdf/B2355-90752.pdf

Platform-specific documentation

- **http://docs.hp.com/hpux/onlinedocs/hw/superdome/**

- **http://docs.hp.com/hpux/onlinedocs/hw/rp8400/**

- **http://docs.hp.com/hpux/onlinedocs/hw/rp7410/**

Agenda

¤ Superdome Hardware and Architecture

¤ rp8400/rp7410 Hardware and Architecture

¤ Partitions

¤ Manageability Architecture

¤ SP Management Capabilities

Superdome Cabinet Packaging

Blowers

Cell

I/O Fans (N+)

I/O chassis

BPS
(N+)

Blowers
(N+)

Utilities

I/O Bays

Cables

Feet
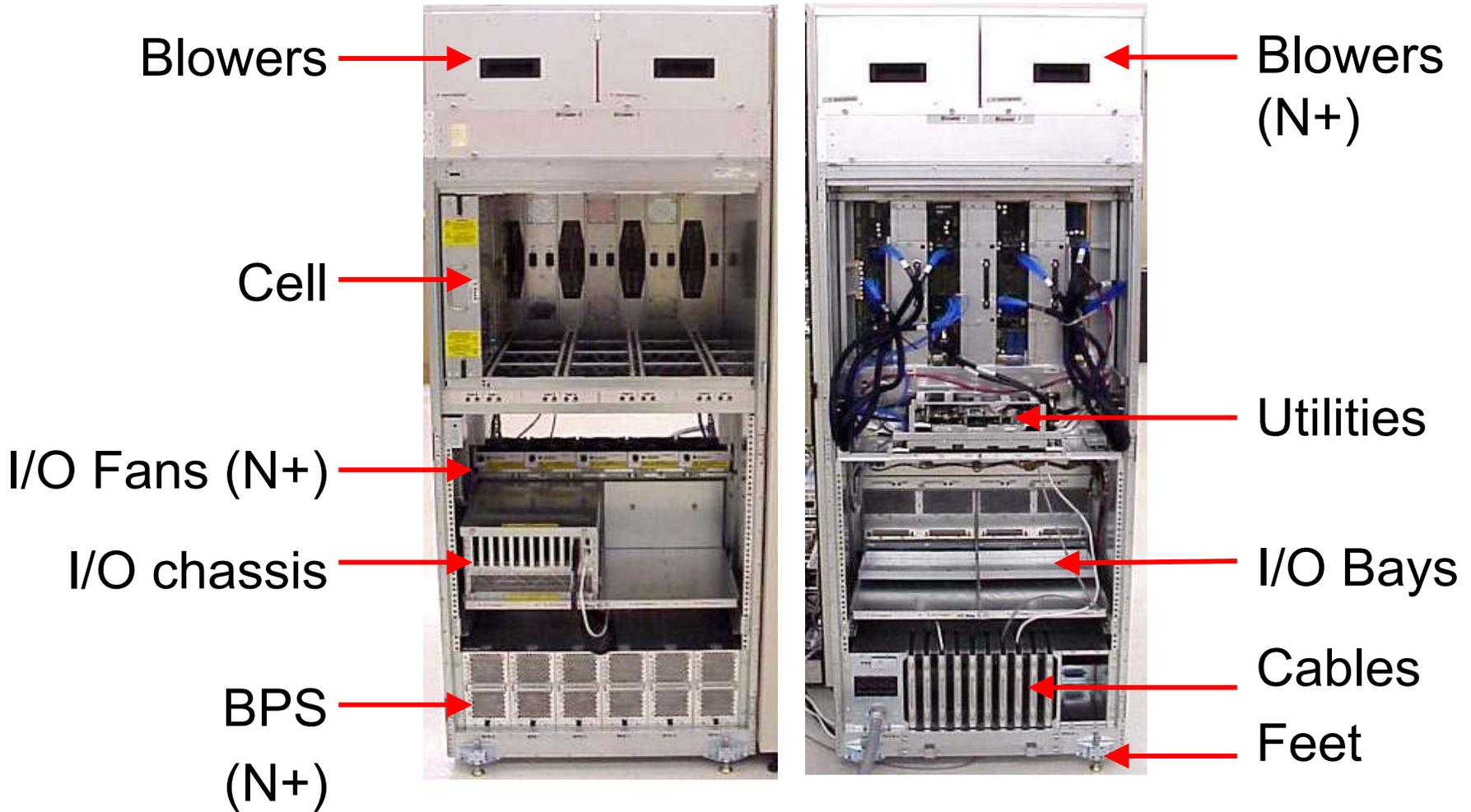
SuperDome Cell Board
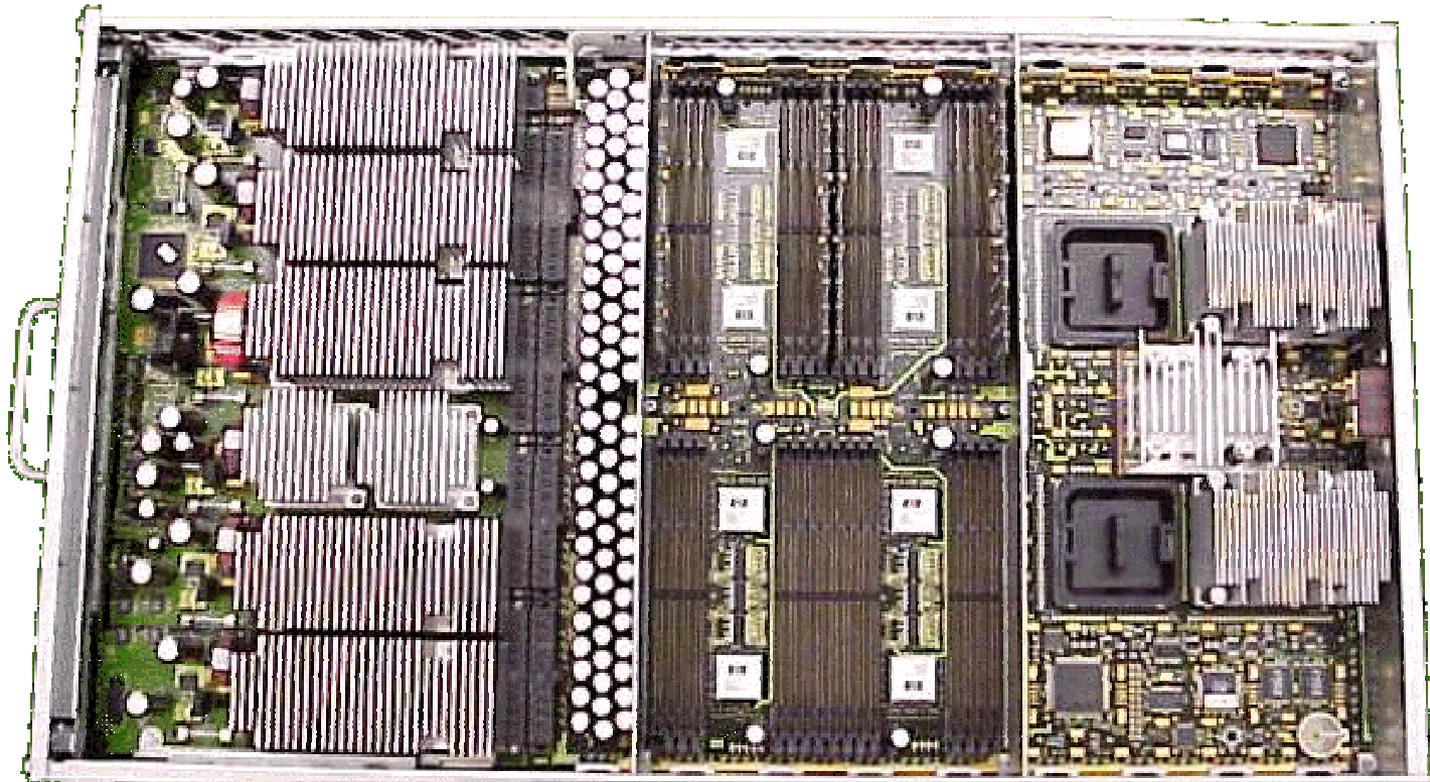
Superdome I/O Chassis
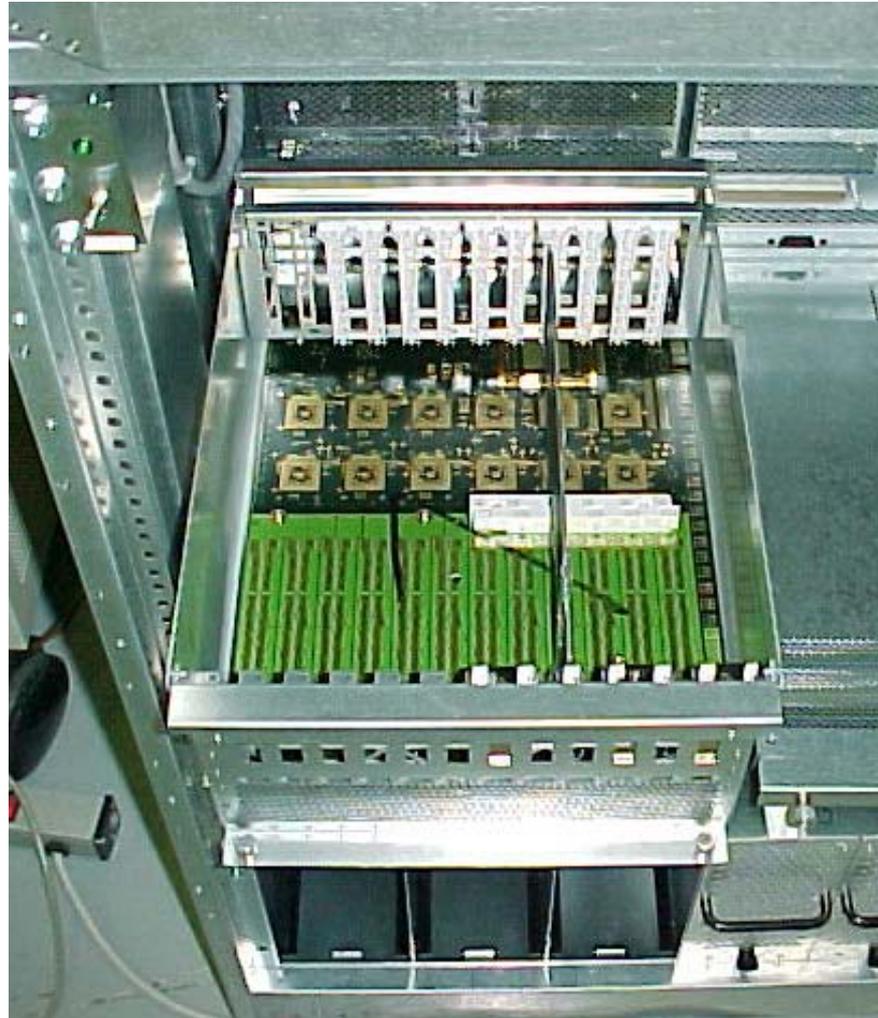
Superdome Service Processor
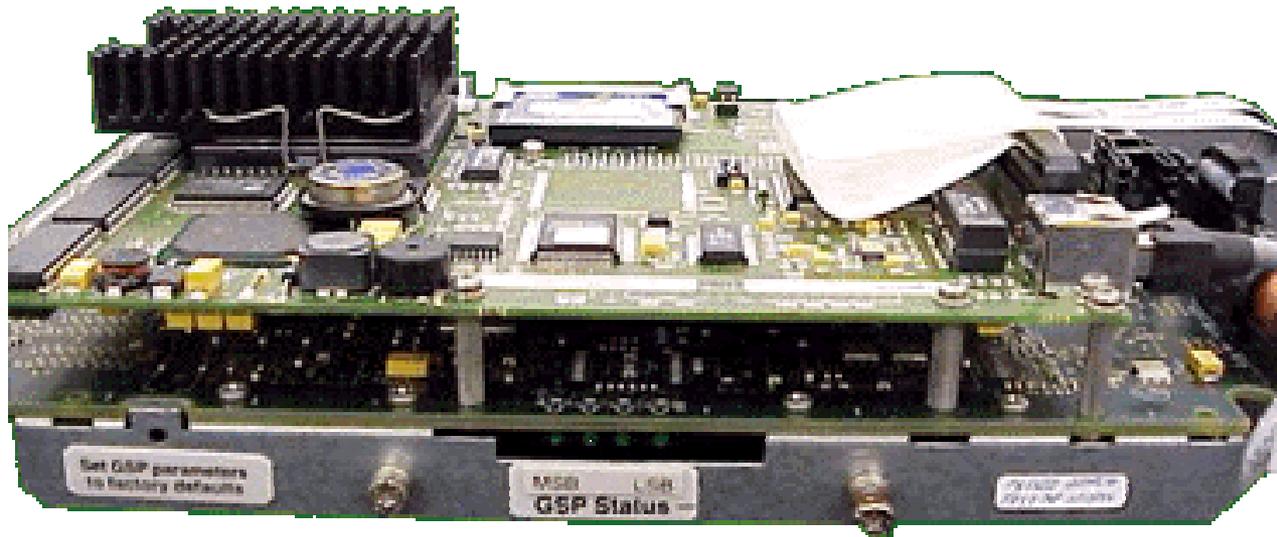
Superdome Cells

Cell Board

MEMORY

Processor Dependent Hardware

PA-8600
PA-8600
PA-8600
PA-8600

Cell Controller

I/O Chassis (optional)

Crossbar
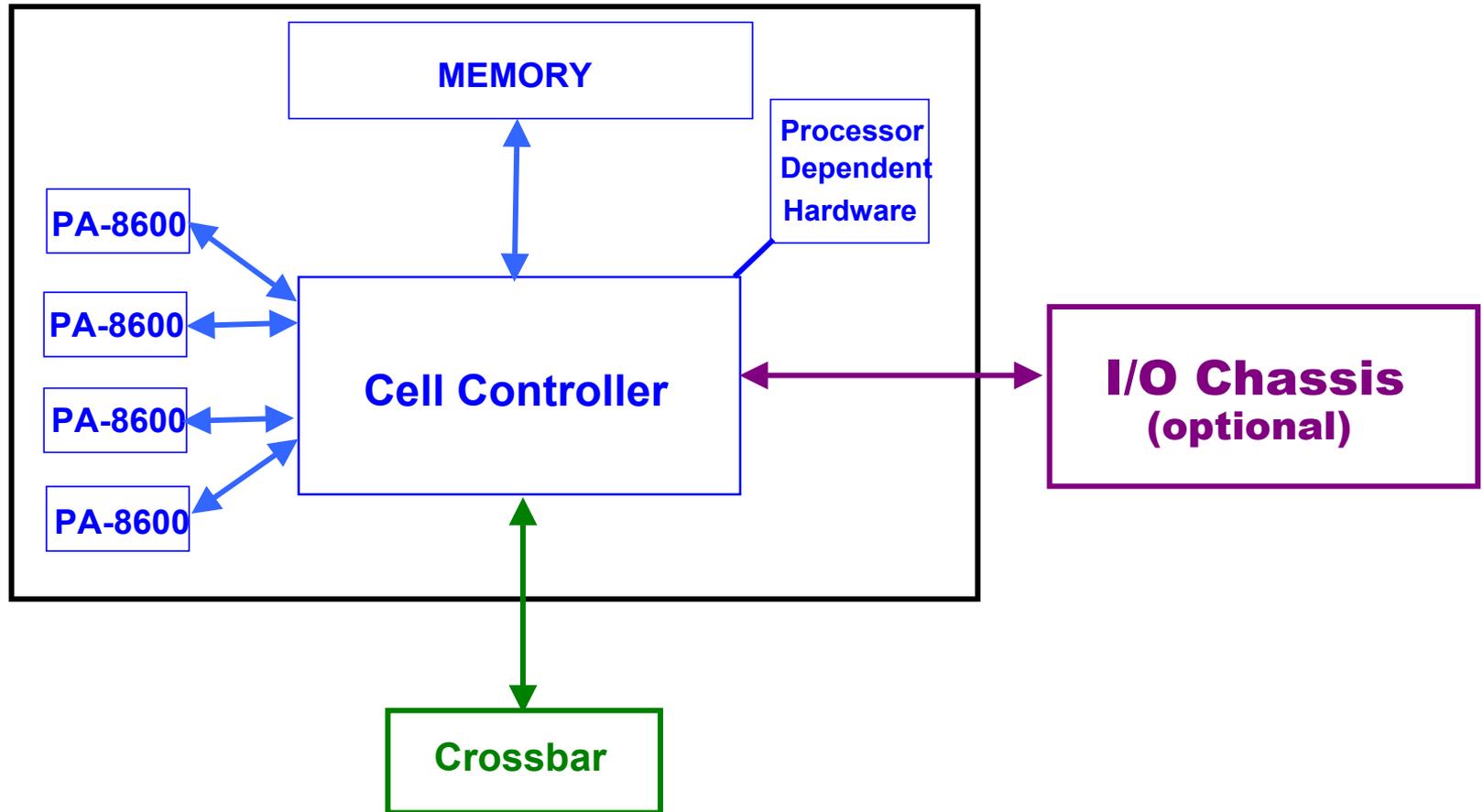
Superdome I/O Chassis

Chassis contains:

- System Bus Adapter (SBA)

- One Local Bus Adapter

(LBA) per slot

- slots 0-3 & 8-11 are PCI-2x

- slots 4-7 are PCI-4x

Local Bus Adapters

Cell Controller

System Bus Adapter

I/O Bus   **I/O Bus**

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
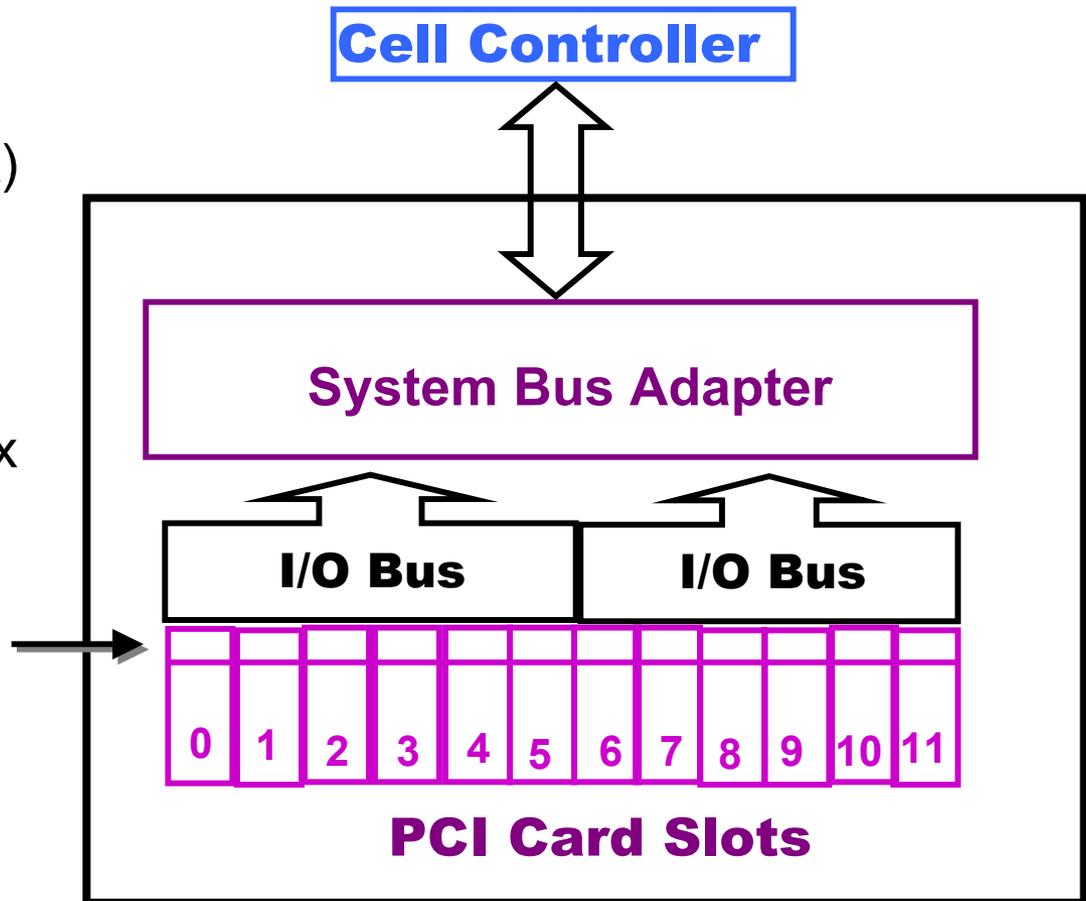
PCI Card Slots

Superdome System Architecture
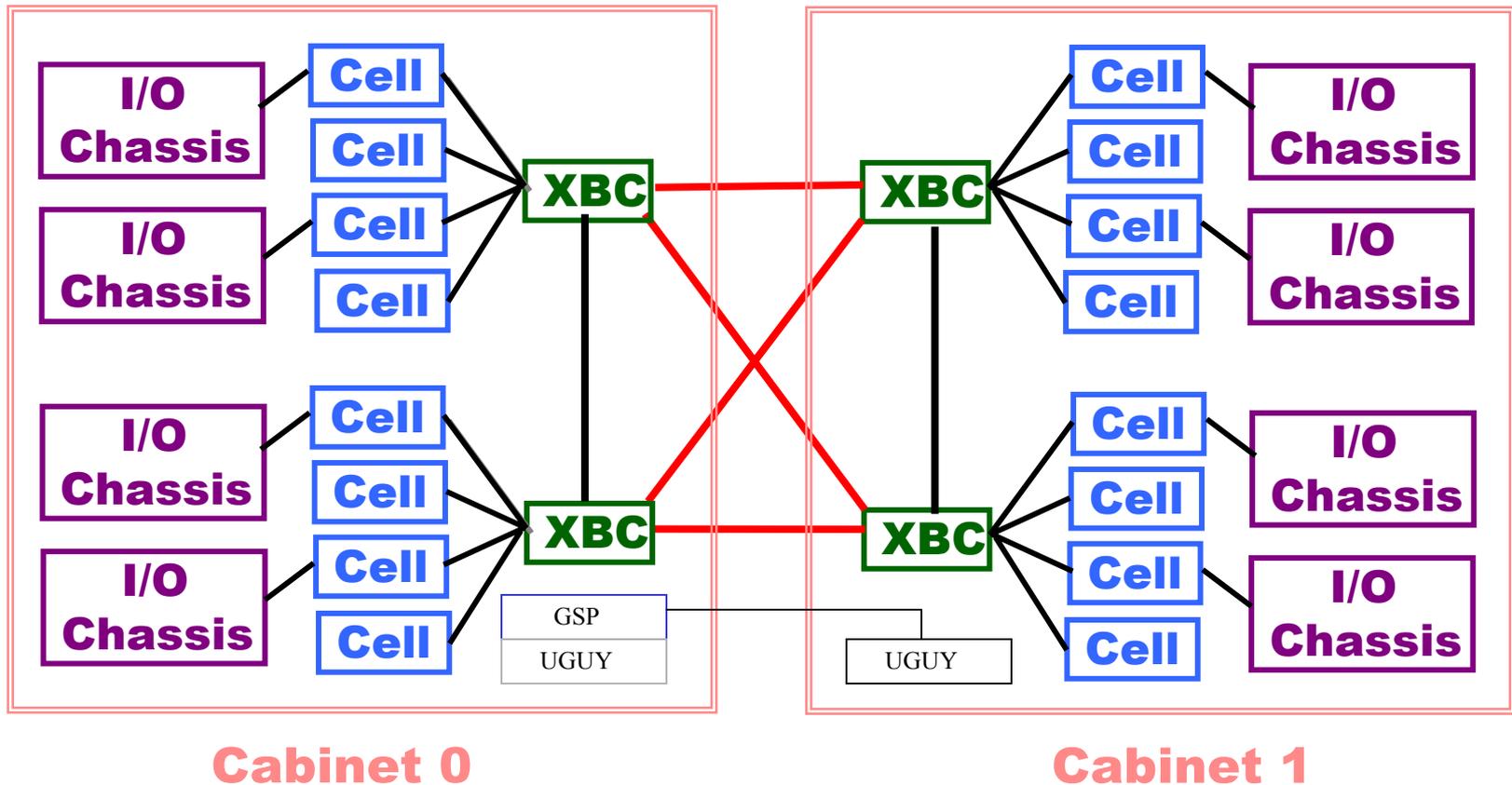
SD64000

rp8400 cabinet

Front

Rear
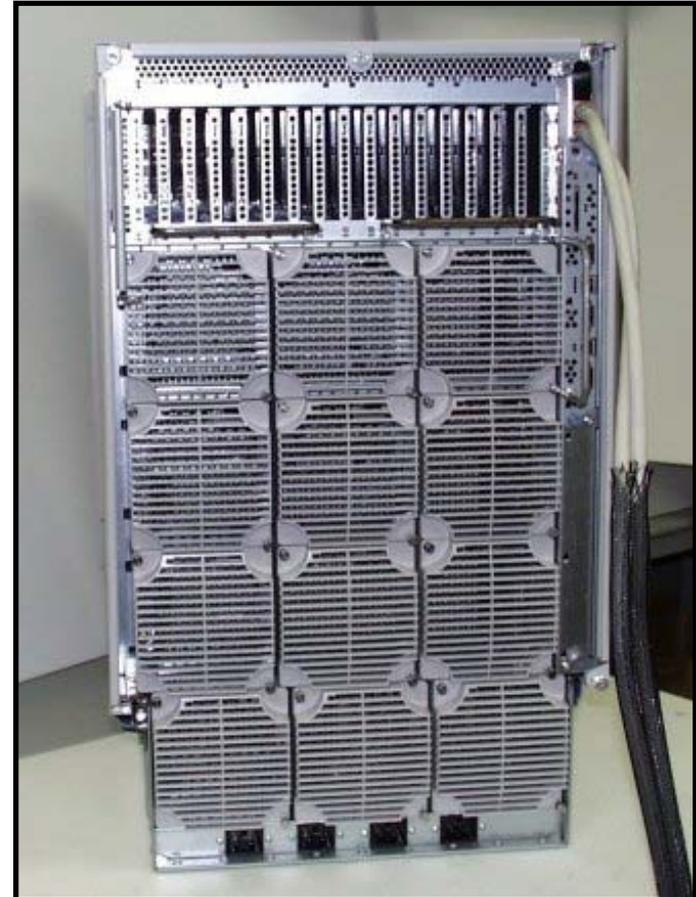
rp8400 Cell Boards

Cell 0

Cell 1

Cell 2

Cell 3

RP8400 Cell Board
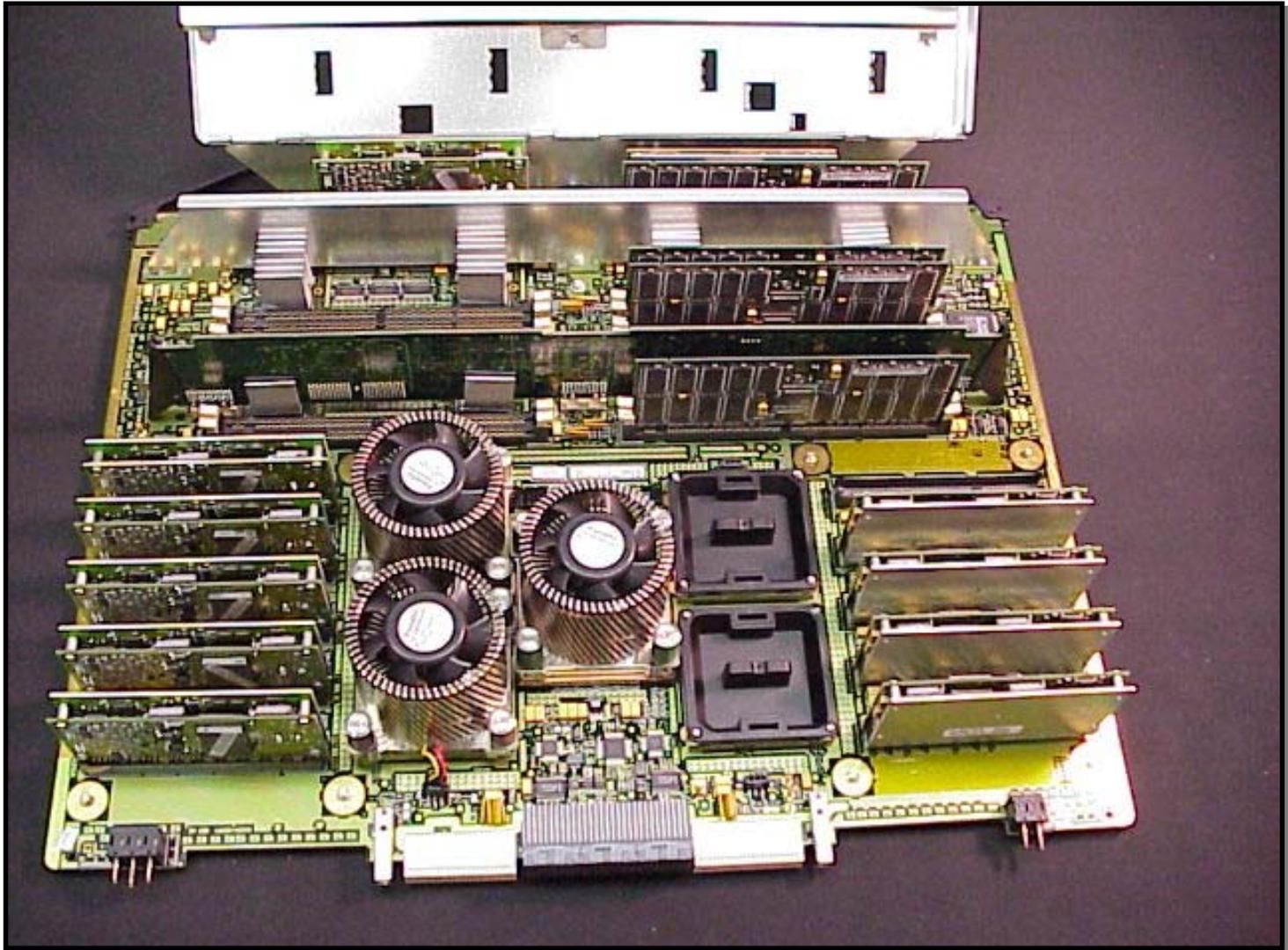
PCI Card Cage
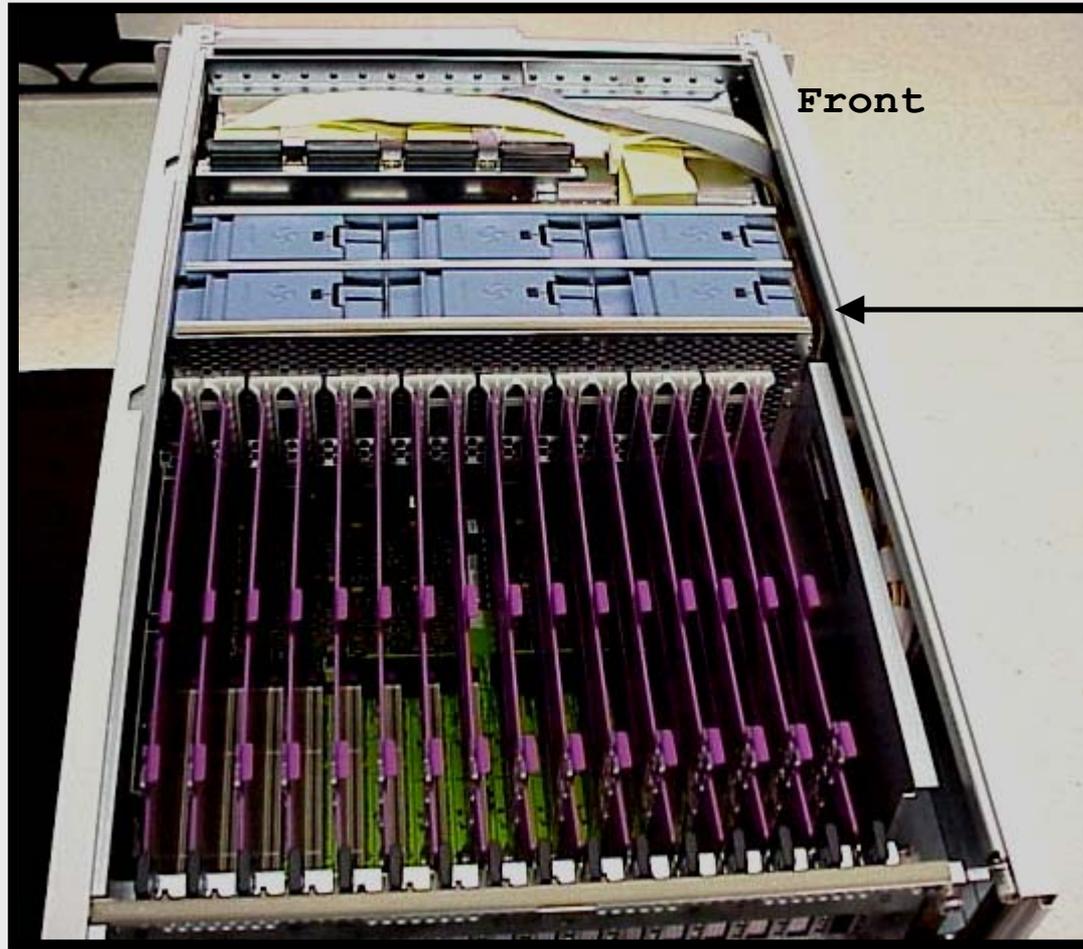
Front

I/O Fans

Slots 1 - 8   Slots 1 - 8

Cell 0   Cell 1

Multi-function Core I/O Board
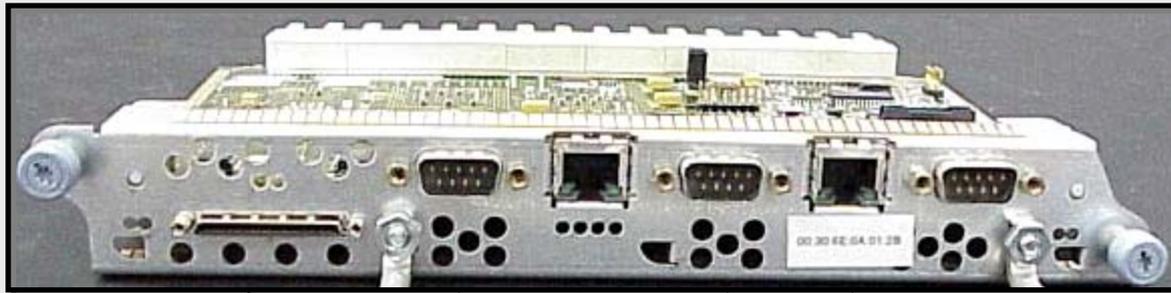
SCSI   UPS   Remote   Local
                          Console   Console

        System LAN   MP LAN

System Block Diagram
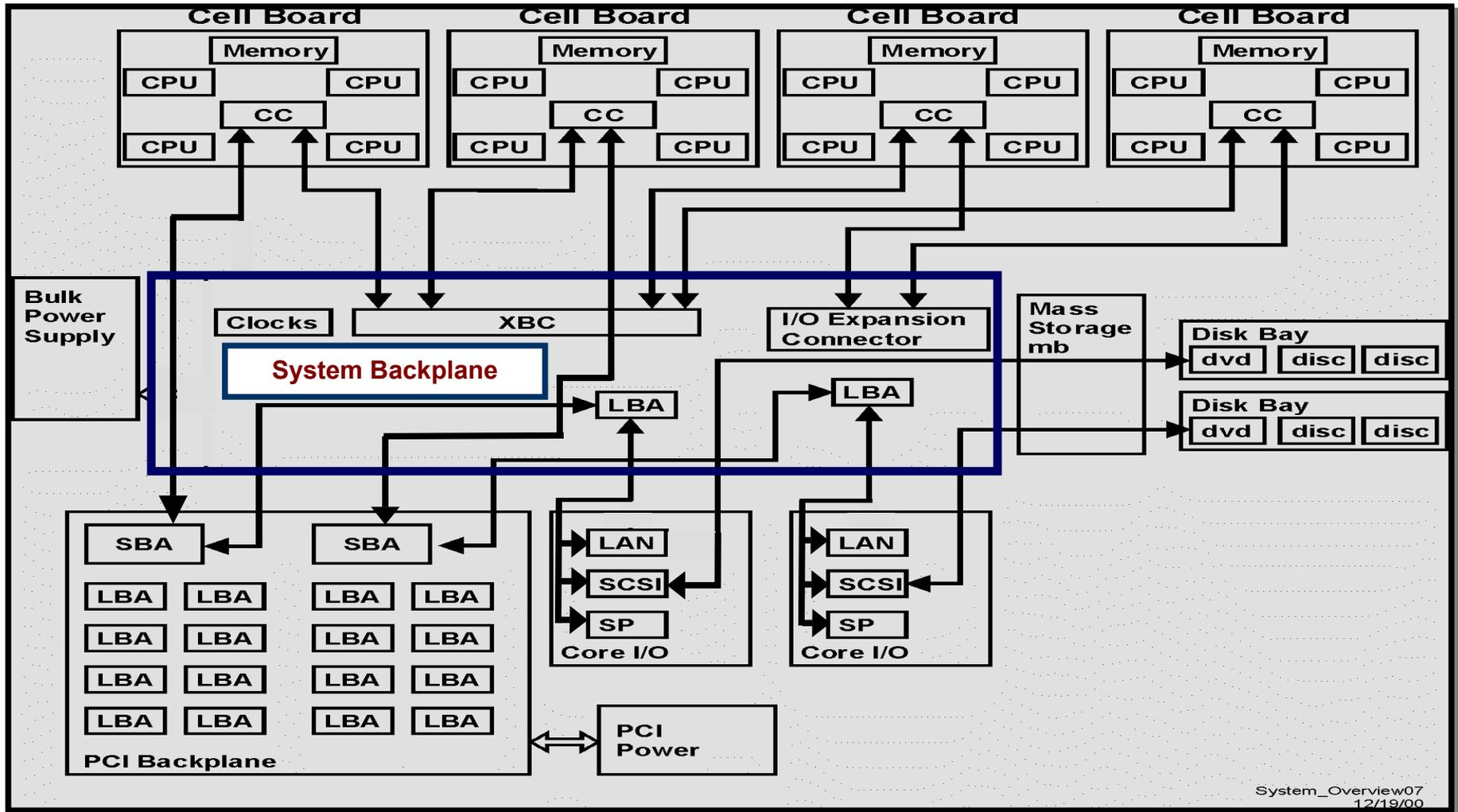
I/O Subsystem Overview

OV84-011
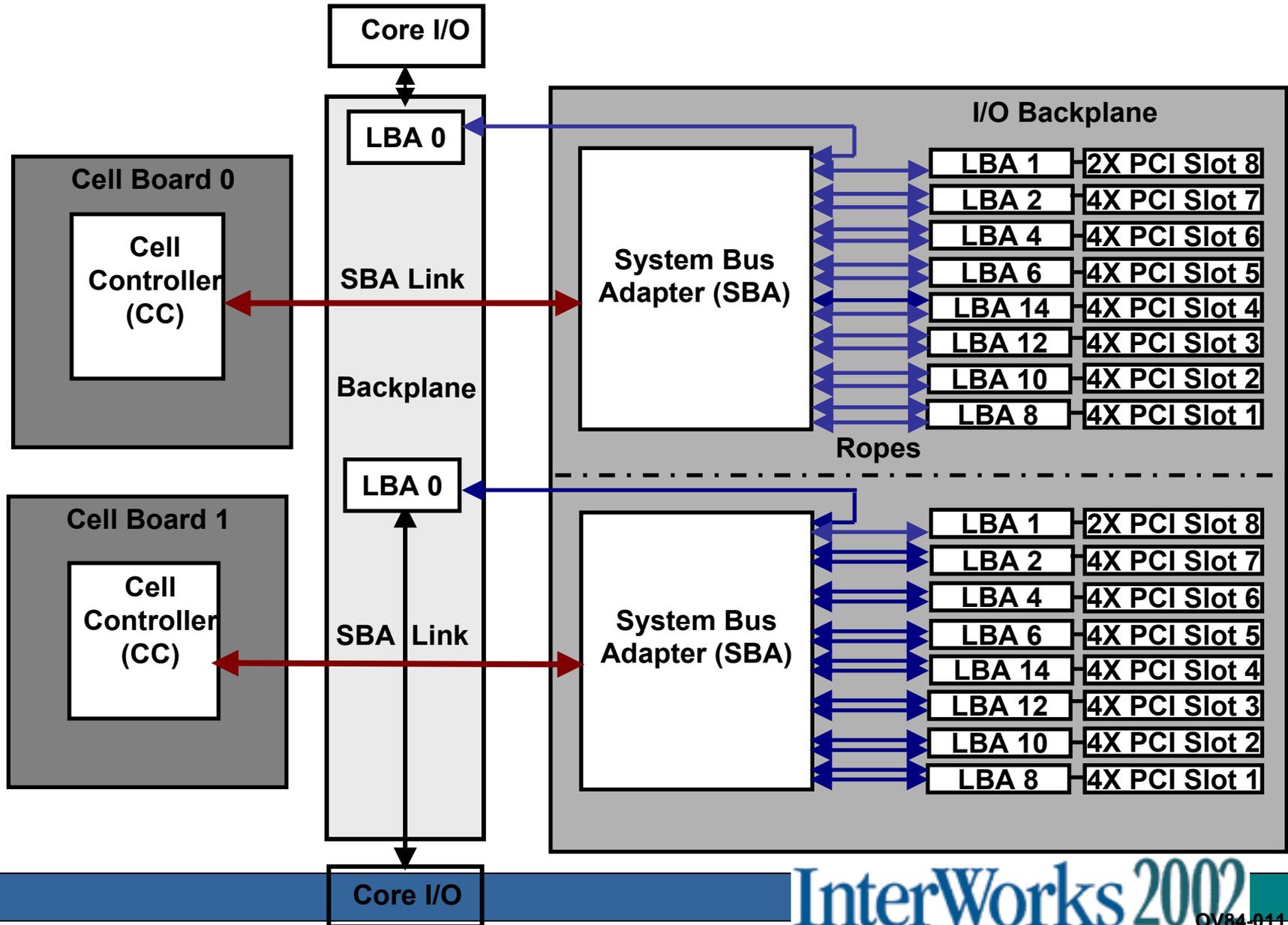
Partitioning Continuum

Hard Partitions with multiple nodes
(at least 1 OS image per node)
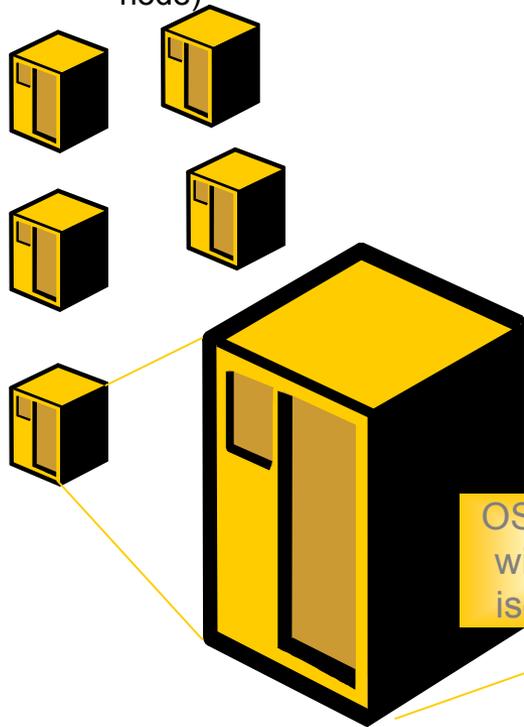
Hard Partitions within a node
(multiple OS images with HW isolation)

Virtual Partitions within a hard partition
(multiple OS images with SW isolation)
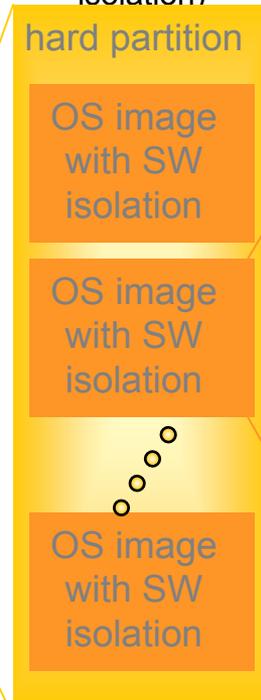
Resource Partitions within an OS image
(resource allocation by app)

OS image with HW isolation

OS image with HW isolation

OS image with HW isolation

hard partition

OS image with SW isolation

OS image with SW isolation

OS image with SW isolation

1 OS image

Application 1
with guaranteed compute resources

Application 2
with guaranteed compute resources

Application n
with guaranteed compute resources

Based on SLOs or percentages

HyperPlex

New!
nPartitions

New!
Virtual partitions

PRM
(Process Resource Manager)

New release!
HP-UX WLM
(Workload Manager)

Partitions: What are they?

- Partitions are physical entities defined around cell board boundaries
- 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.
- Dynamic addition of new Partitions
- Adding/removing cell boards to existing Partitions does not impact any other Partition
- Follow-on release will increase the dynamism.
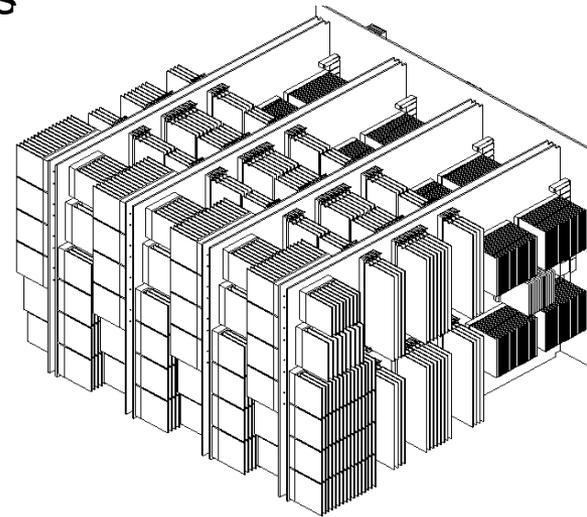
Partition Features

- Collection of any cells and I/O (1 - 16) within the system runs an OS
    - Plan wisely for performance and availability
    - Proper system configuration eliminates sharing hardware

- Changes made through Parmgr GUI or command line I/f

- Hardware Firewalls built into the system ASICs
    - Memory and I/O are private to the partition, not shared
    - CC screens out external traffic
    - XBC isolates link failures
    - Lock and key protects XBC critical data structures

Partitions on an eight cell Superdome SD32000
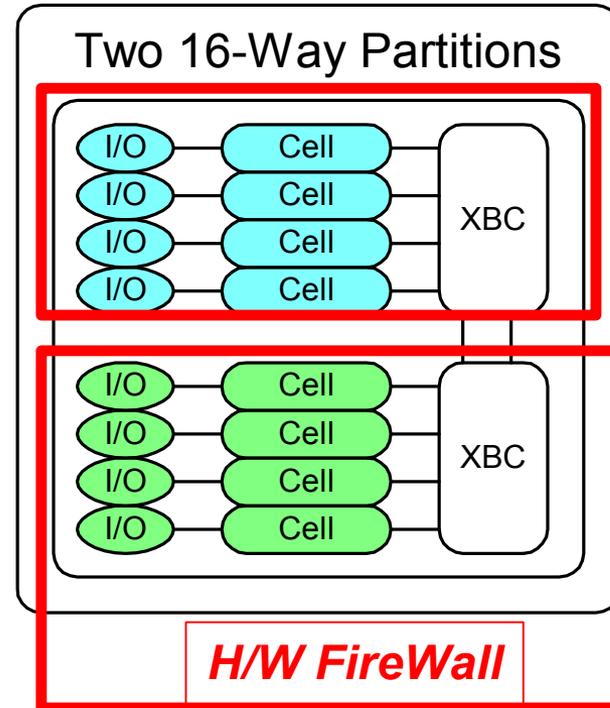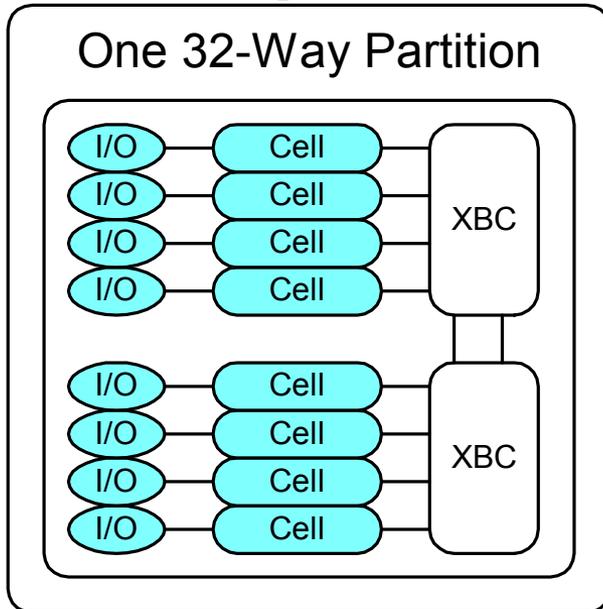
Complex profile concepts

- **What is the complex profile?**
  - **A set of data structures which define:**
    - **SCCD - Global parameters (affecting all partitions)**
      - **which cells are in which partitions**
      - **system name**
    - **PCD - Partition parameters (affecting one partition)**
      - **boot path**
      - **partition IP address**
      - **SWID**
- **Where is the complex profile stored?**
  - **Stored in GSP (master copy) and on every cell**
- **How is it modified?**
  - **Initially created by the GSP**
  - **Read and modified by thePartition Manager, the partition commands, and system firmware (aka PDC)**

Partition boot process

- **Cell reset is released after power up or RS command**
- **System FW begins booting cell (not part of a partition yet)**
- **System FW sets the "shared memory good" flag**
- **PDHC and GSP begin checking complex profile for consistency across cells. PDHC and GSP may push out a modified complex profile.**
- **Concurrently, System FW does memory and I/O discovery**
- **System FW on a cell checks the "Boot Is Blocked" (BIB) flag and spins waiting for it to be cleared**
- **PDHC and GSP clear the BIB flag after the complex profile has been checked**
- **Cell boots, reads the complex profile, and rendevous with the other cells in its partition.**

Boot-Related Availability Features

- **Each cell contains everything the cell needs to boot to FW rendezvous with other cells in the partition.**
  - **Cell-local complex profile storage**
  - **Cell-local timeout on GSP communications**
- **Partition can boot with a partial complement of cells**
- **Partition can be configured with redundant I/O including console I/O**

Platform Manageability
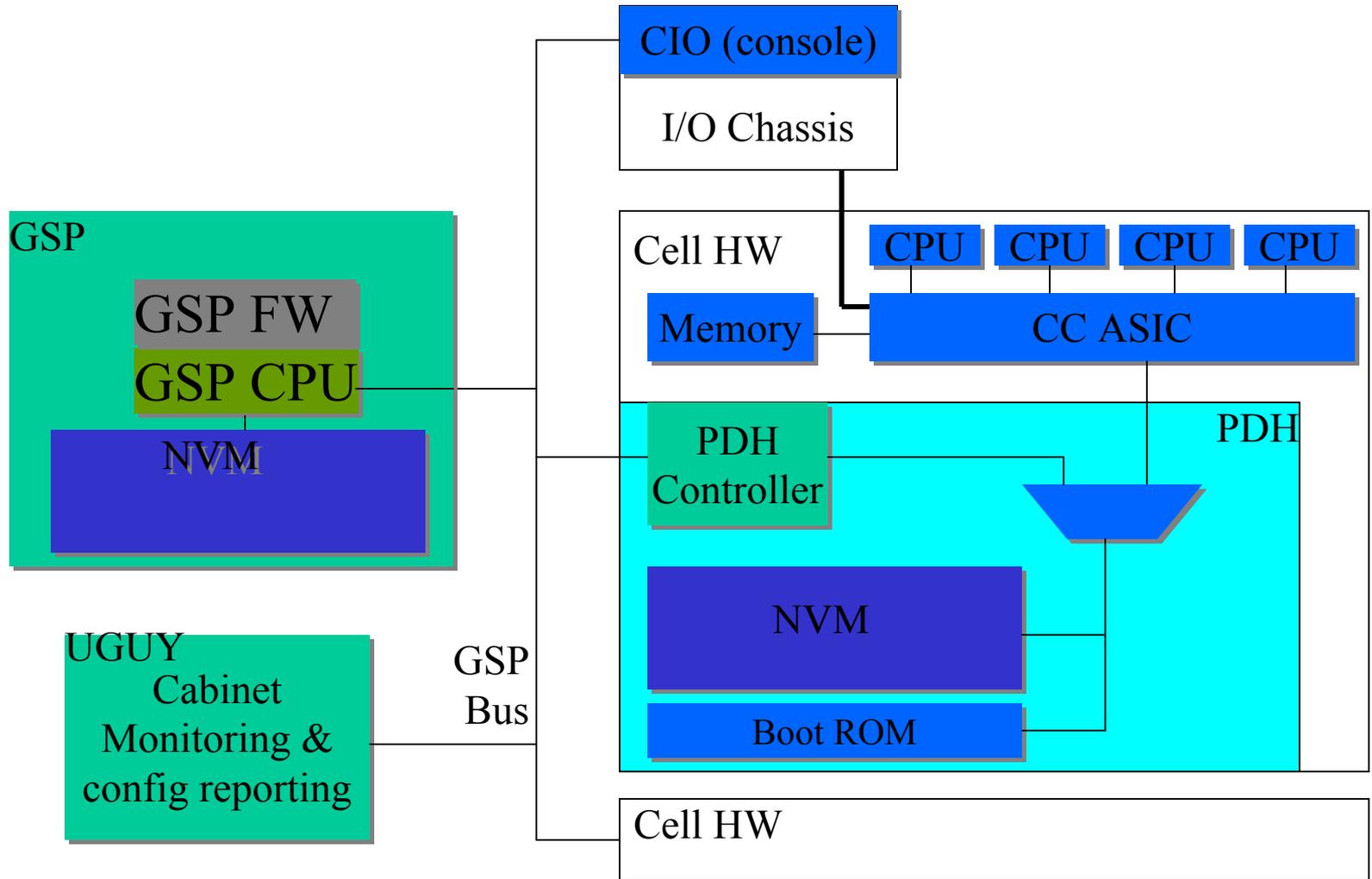
Partition Reconfiguration

① Parmgr sends new complex profile to GSP
② GSP distributes it to affected cells
③ PDC & OS use it next boot or during OLA

Mgmt Apps | Parmgr | Parstatus

OS | HP-UX

FW | PDC FW

Cell HW | CPU | CPU | CPU | CPU

Memory | CC ASIC

①

GSP FW

GSP CPU

NVM

Complex Profile

②

PDHC uC

NVM

③

① Command to GSP

Complex Profile

Cabinet Monitoring & config reporting (UGUY Board)

GSP Bus

Boot ROM

Cell HW
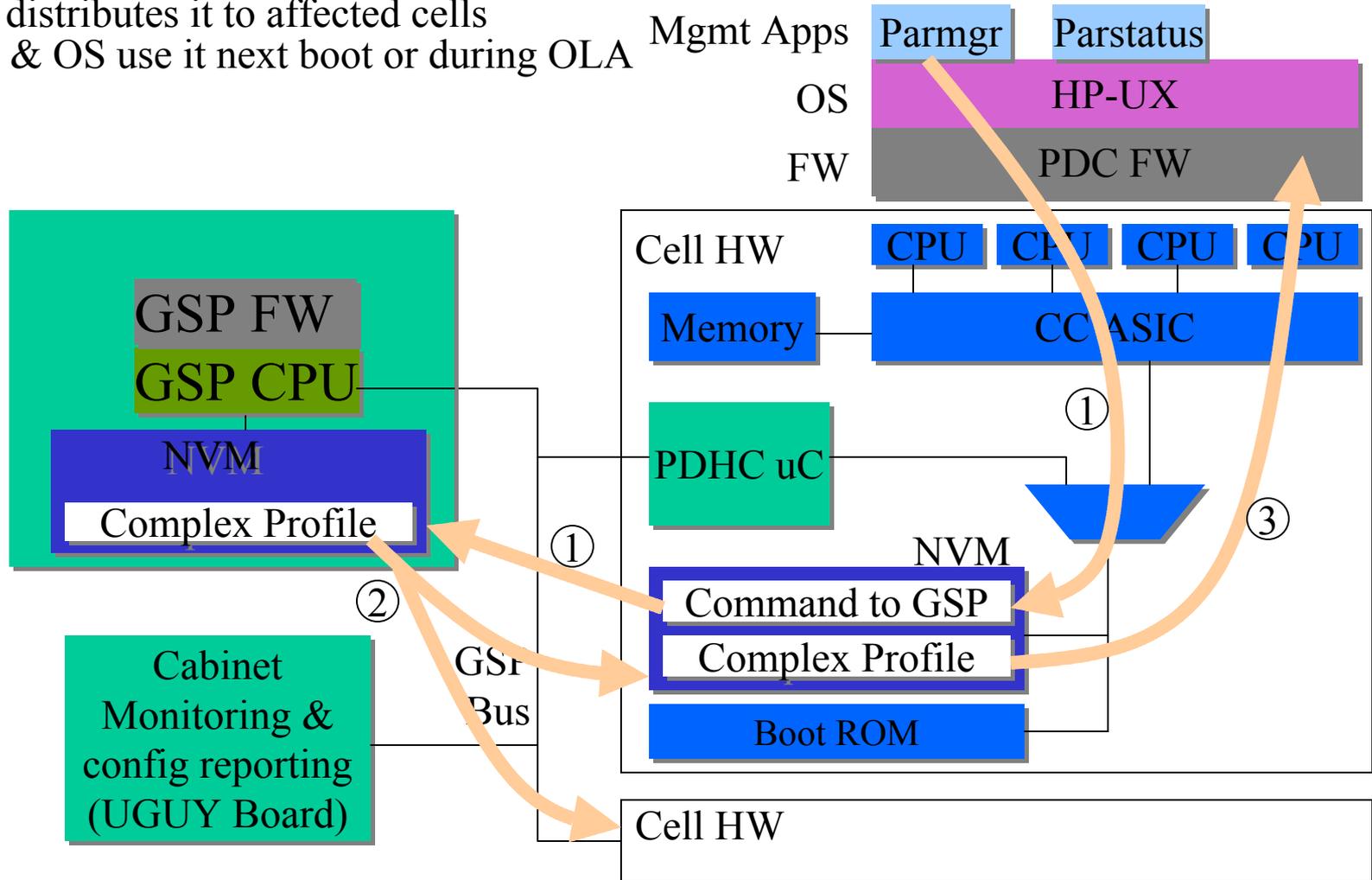
Event Management System

HAO

Mgmt App

① Events (Chassis logs) sent from UGUY, PDC, HP-UX, to GSP ③
② GSP reflects logs to all partitions
③ EMS notifies management apps

EMS

OS   HP-UX

FW   PDC FW

Cell HW   CPU CPU CPU CPU

GSP FW
GSP CPU

Memory   CC ASIC

NVM

PDHC uC   ①   ②

Log buffers   ①

NVM

①   ②   Chassis Logs to GSP

Cabinet Monitoring & config reporting (UGUY Board)   GSP Bus   Reflected Logs

Boot ROM

Cell HW
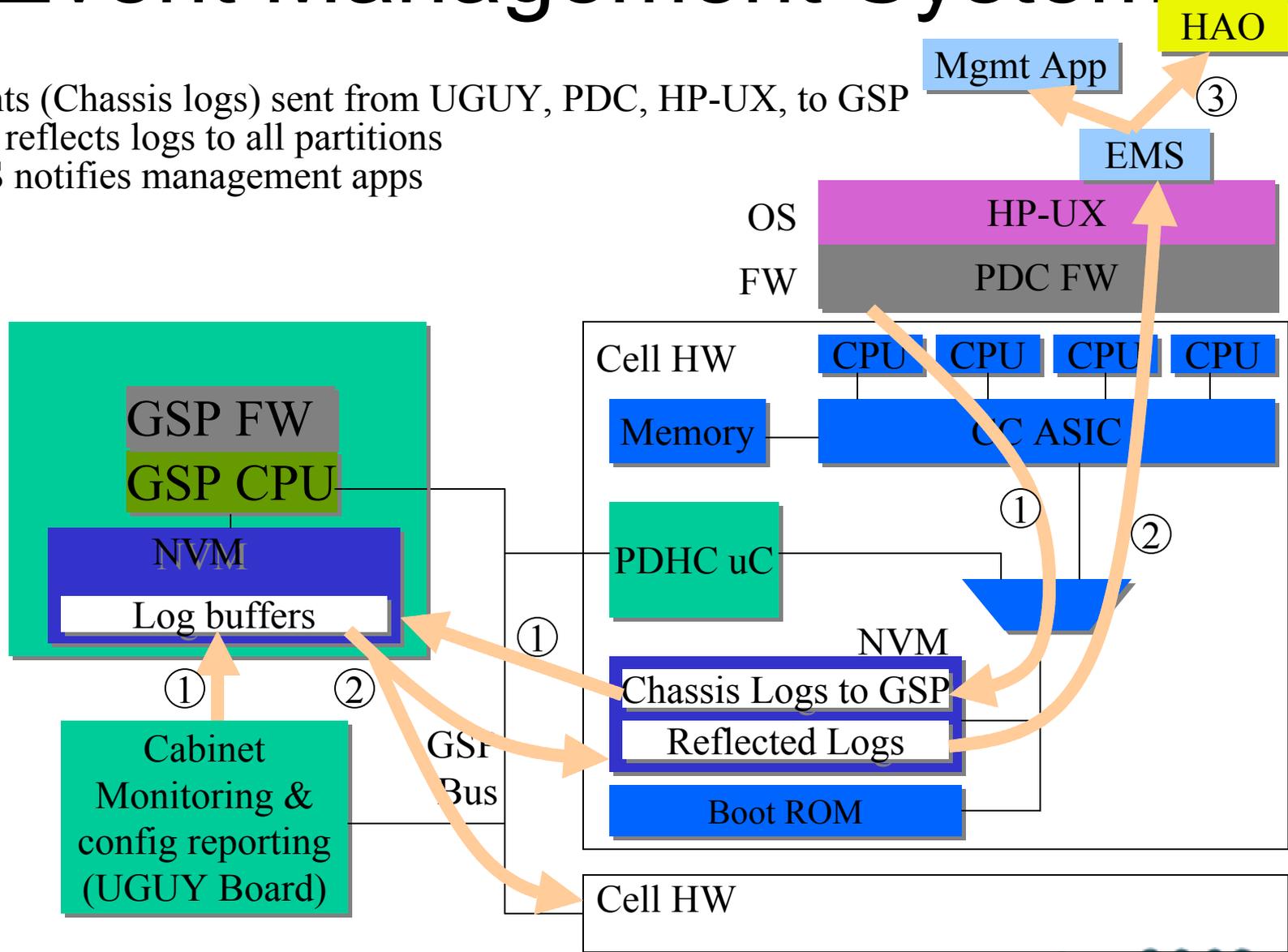
Event Management Software Stack

HP Response Center

HAO/HPEN Predictive

Service-Guard

Enterprise Mgmt. (OpenView, CA, Tivoli, Remedy, etc.)

User

Customer Tools

HP Local Event Browser

FN client | tcp/udp | opc | snmp | email | syslog | text log | console

EMS Framework

EMS

H/W Monitors

SW Monitors

Kernel Drivers | Firmware | Networking

APIs, system calls, etc.

Server Hardware

O/S

Applications

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

Console and SMS Local Subnet Configuration
Single Superdome

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

MODEM*

SUPPORT MANAGEMENT STATION

Customer Site LAN
(If Desired)

LAN Console Port

10/100 Base-T Ports

PC/WORKSTATION CONSOLE DEVICE

LAN HUB

Superdome Server

Guardian Service Processor

Unconnected ← **Private LAN Port**

Customer LAN Port

RS232 Local

RS232 Remote

CIO LAN — Partition 0

CIO LAN — Partition 1

CE TOOL (PC)

Dashed lines indicate temporary connections
Solid Lines indicate permanent connections
* Modem and HAO bundled with Service Agreements

For Initial Install of SMS
Serial Cable 24542G
needed

Console and SMS General Network Configuration
Single superdome

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

SUPPORT MANAGEMENT STATION

LAN Console Port

10/100 Base-T Ports

PC/WORKSTATION CONSOLE DEVICE

LAN HUB

Superdome Server

Guardian Service Processor

Private LAN Port

Customer LAN Port

RS232 Local

RS232 Remote

Customer Site LAN
(If Desired)

MODEM*

CIO LAN   Partition 0

CIO LAN   Partition 1

CE TOOL (PC)

For Initial Install of SMS Serial Cable 24542G needed

Dashed line indicates temporary connections
Solid Line indicates permanent connections
* Modem and HAO bundled with Service Agreements

Service Processor

➢ Main Console (local serial, modem or LAN)
➢ Provides access to partitions' consoles
➢ Per partition Virtual Front Panel
➢ View chassis logs and console logs
➢ Display SPU status
➢ Command Menu
➢ One SP per complex

GSP Main Menu
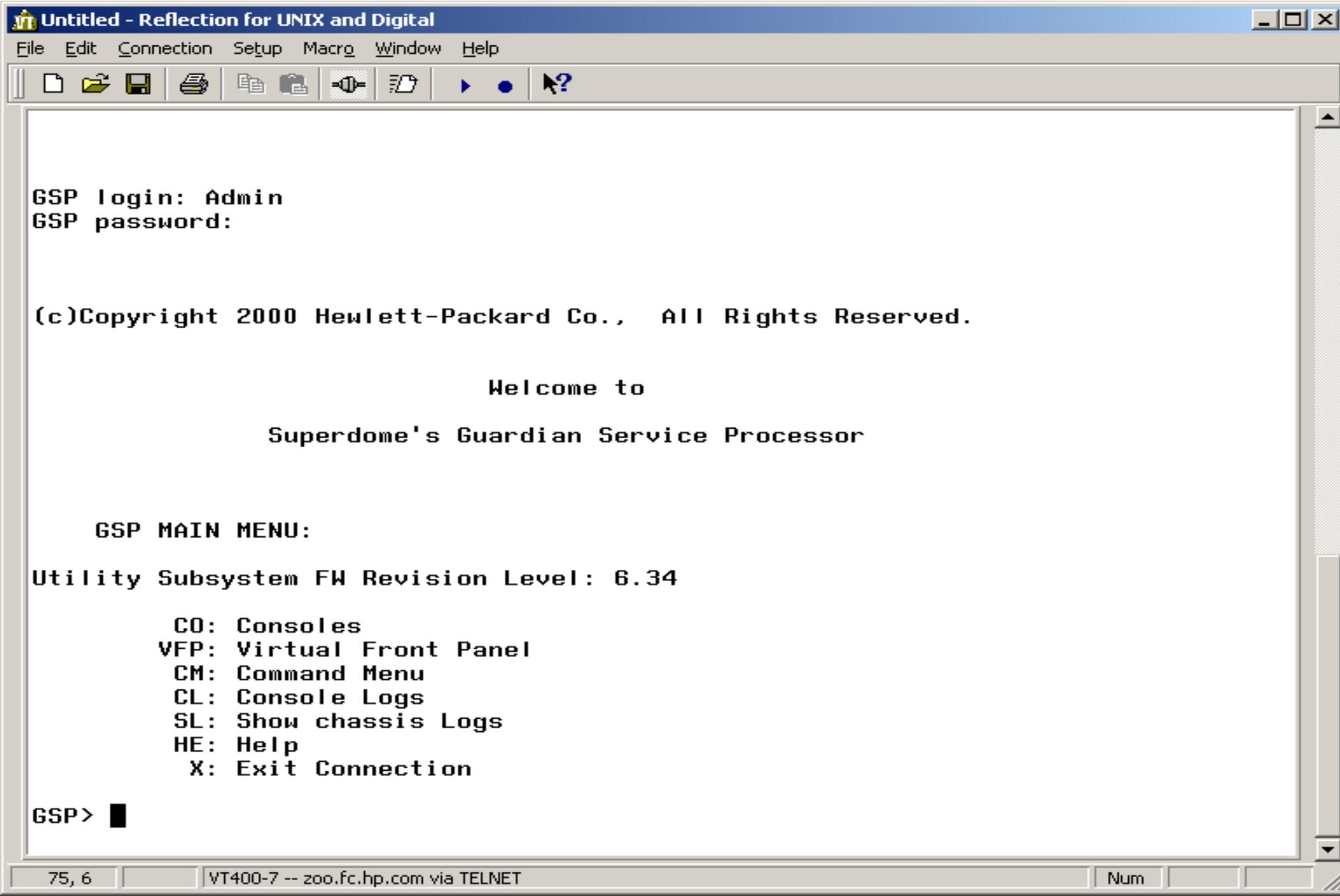

```
GSP login: Admin
GSP password:


(c)Copyright 2000 Hewlett-Packard Co.,  All Rights Reserved.


                          Welcome to

              Superdome's Guardian Service Processor


    GSP MAIN MENU:

Utility Subsystem FW Revision Level: 6.34

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

GSP> █
```

Console Interface

```
GSP> co
    Partitions available:
     #    Name
    ---   ----
     0)   Partition 0
     1)   Partition 1
     2)   Partition 2
     3)   Partition 3
     Q)   Quit
    Please select partition number: 1
    Connecting to Console: Partition 1
        (Use ^B to return to main menu.)
        [A few lines of context from the console log:]
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
    MFG menu                           Displays manufacturing commands
    DIsplay                            Redisplay the current menu
    HElp [<menu>|<command>]            Display help for menu or command
    REBOOT                             Restart PD
    RECONFIGRESET                      Reset to allow Reconfig Complex
Profile
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
- You are now connected to the console
- ^ECF to gain control of the console
- ^B to escape back to the main menu
```

Console Logs

```
GSP> cl


    Partition Console Logs available:
     #   Name
    ---  ----
     0)  Partition 0
     1)  Partition 1
     2)  Partition 2
     3)  Partition 3
     Q)  Quit
GSP:VW>
DeBug Menu: Enter command > mr 0x20000 16

(N)ext or <cr>, (P)revious, ^B to exit to menu
```

Virtual Front Panel

```
E indicates error since last boot
    Partition 0  state                  Activity
    ------------------                   --------
    Cell(s) Booting:  █ 857 Logs

 #  Cell state                          Activity
 -  ----------                          --------
 0  Early CPU selftest                  Processor firmware slave rendezvo  121  Logs
 2  Early CPU selftest                  Processor firmware slave rendezvo  121  Logs
 3  Early CPU selftest                  Processor firmware slave rendezvo  121  Logs
 4  Early CPU selftest                  Processor firmware slave rendezvo  121  Logs
 6  Early CPU selftest                  Cell firmware test                 120  Logs
12  Early CPU selftest                  Processor test                     118  Logs
14  Early CPU selftest                  Processor initialization           117  Logs




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

f1	f2	f3	f4	hpfcjd hpter	f5	f6	f7	f8

VFP – All Partitions
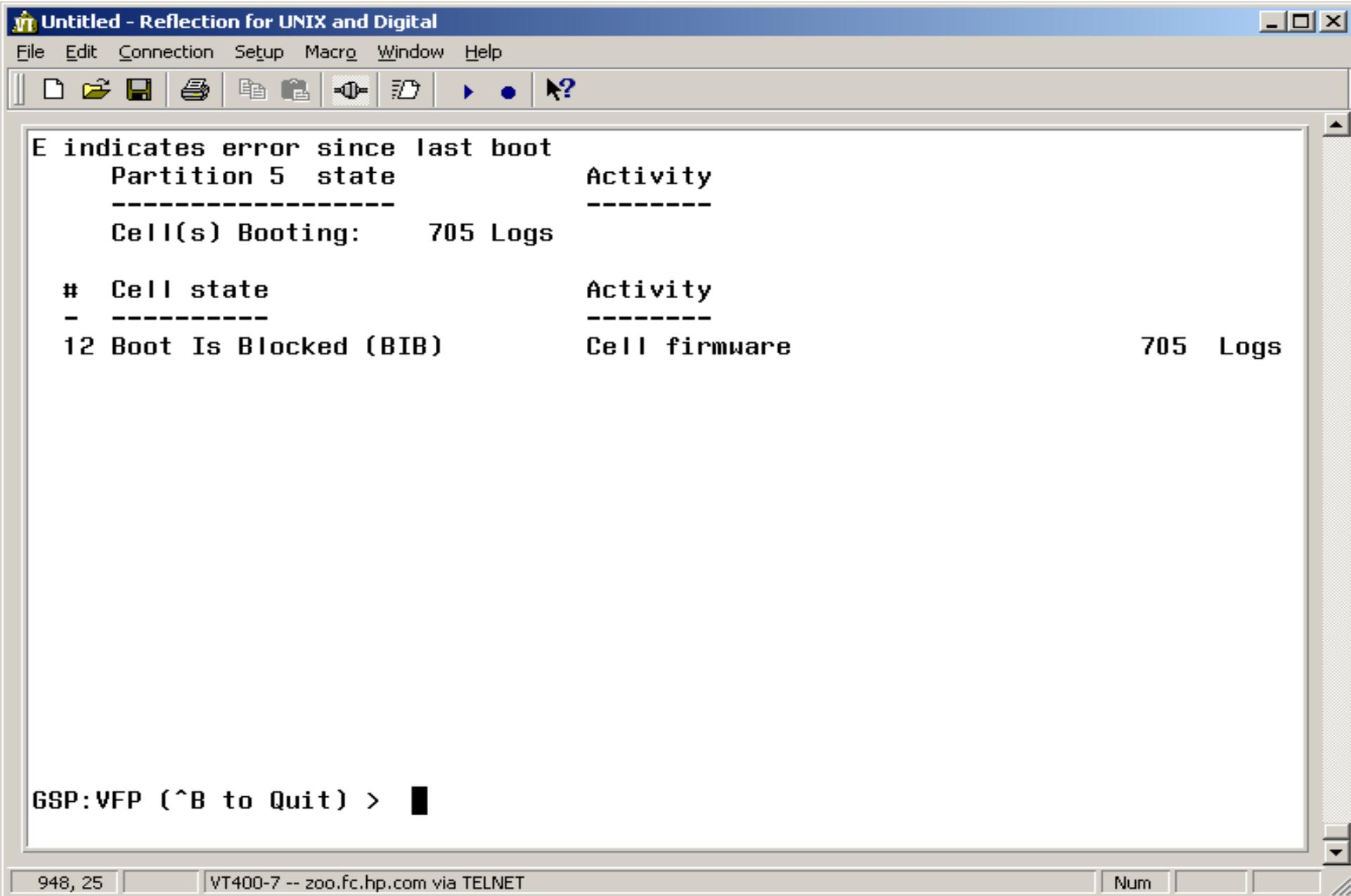
```
File  Edit  Connection  Setup  Macro  Window  Help
```

```
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
```

VFP – Partition at BIB

```
E indicates error since last boot
      Partition 5   state                Activity
      ------------------                  --------
      Cell(s) Booting:      705 Logs


  #  Cell state                          Activity
  -  ----------                          --------
  12 Boot Is Blocked (BIB)               Cell firmware                      705   Logs

















GSP:VFP (^B to Quit) >  █
```

Log Viewer (1 of 2)

```
GSP> sl
Chassis Logs available:
    (A)ctivity
    (E)rror
    (Q)uit

GSP:VW> e
To Select Entry:
    (<CR> or <space>) View next or previous block
    (+) View next block (forwards in time)
    (-) View previous block (backwards in time)
    (F)irst entry
    (L)ast entry
    (J)ump to entry number
    (V)iew Mode Select
    (H)elp to repeat this menu
    ^B to exit
Press any key to continue
```

Log Viewer (2 of 2)

```
GSP:VWR >
#      Location Alert  Keyword                        Timestamp
642  PDC   0.0.0 *3   A402000                         05/29/2000 11:32:19
641  SINC 0.0   *2   CCI_FULL                         05/29/2000 11:32:18
640  PDC   0.0.0 *10 A204071                          05/29/2000 11:32:17
434  PDC   0.7.4 *8   9F020D2                         05/29/2000 10:53:56
433  PDC   0.7.4 *8   9F020D2                         05/29/2000 10:53:56
432  SUB   0      *2   GSP_BUS_DEVICE_DETACH           05/29/2000 10:53:53
431  SUB   0      *2   GSP_BUS_DEVICE_DETACH           05/29/2000 10:53:48
430  SUB   0      *2   GSP_BUS_DEVICE_DETACH           05/29/2000 10:53:43
429  PDC   0.3.4 *2   ERR_DNA_SEC_HEALTH              05/29/2000 10:53:42
GSP:VWR (<CR>,<sp>,+,-,F,L,J,V,H,^B) >
```

GSP Command Menu

```
GSP:CM> HE


Guardian Service Processor (GSP) - FW Revision 6.34, Oct 26 2000 at 23:51:44

Utility Subsystem FW Revision Level: 6.34


The following are the service commands that are available

BO        : Boot a partition.
DF        : Display FRU information of an Entity.
MA        : Return to the Main menu.
MFG       : Exit manufacturing mode
MR        : Modem Reset
PE        : Power entities on or off.
RE        : Reset entity.
RR        : Reset a partition for reconfiguration.
RS        : Reset a partition.
TC        : Send a TOC signal to a partition.
TE        : Broadcast a message to all users of the GSP Command Handler.
VM        : Margin the voltage in a cabinet
WHO       : Display a list of GSP connected users


Press q to quit or <CR> to continue...
```

Command Menu

- **There are over 40 commands. The most frequently used ones are:**
  - **RS   reset partition**
  - **TC   TOC partition**
  - **RR   reset partition for reconfiguration**
  - **BO   boot partition**
  - **CP   display complex profile**
  - **CC   create genesis complex profile**
  - **LS   display LAN status**
  - **LC   set LAN configuration**
  - **ND   enable/disable network diagnostics**
  - **PE   power enable cabinet**
  - **PS   display power and configuration status**

```
GSP> cm

       Enter HE to get a list of available commands
GSP:CM>

       q to quit out of parameter entry for a command
       MA command goes back to the main menu
```

Command menu concepts

- **Shared environment**
    - **all users in the command menu share the environment,**
    - **i.e. all users see commands typed by other users**

- **Command ownership**
    - **once a user types a command name, e.g. RS, they have exclusive ownership of that command until the command completes or an inactivity timer expires.**

- **Command timeout**
    - **if a user types a command but doesn't complete the action before the timeout expires, the command will be aborted**

Command menu concepts

- **Three Capability levels**

  1. Single_Partition user

  Can only connect to assigned partition

  Can only execute commands that affect their assigned   partition

  2. Operator

  Has single partition user capabilities on all partitions

  Can't reconfigure the GSP

  3. Administrator

  Has operator level capabilities

  Can reconfigure the GSP

Restart Commands

RS (Reset) Command: Irrecoverably halts all the system processing and I/O activity and restarts the Partition chosen.

TC (TOC) Command: Irrecoverably halts all the system processing and I/O activity and dumps core then restarts the Partition chosen.

RR (Reset for Reconfiguration) Command: Irrecoverably halts all the system processing and I/O activity and restarts the chosen Partition in a way that it can be reconfigured. (Partition stops at BIB.)

BO (Boot) Command: BO releases a partition to boot that has been reset for reconfiguration, or that has failed to boot because of an invalid configuration. BO clears the "BIB" bit (Boot Is Blocked). RR and BO are natural pairs. However, it should not be necessary to use them because Parmgr is the preferred tool. But, like RS, they can get you unstuck if the system won't shut down or won't boot.

RS Command

```
GSP:CM> rs
Execution of this command irrecoverably halts all the system
processing and I/O activity and restarts the Partition
chosen.
    #    Name
   ---   ----
    0)   Partition 0
    1)   Partition 1
    2)   Partition 2
    3)   Partition 3
    4)   Partition 4
    5)   Partition 5
    6)   Partition 6
    7)   Partition 7

    Select a partition number : 0
    Do you want to reset Partition number 0,
    named Partition 0? (Y/[N]) y

    -> Partition selected will be reset.
```

CP Command

- **Displays the assignment of cells to partitions in the current complex profile**

```
GSP:CM> cp

--------------------------------------------------------------------
Cabinet |   0    |   1    |   2    |   3    |   4    |   5    |   6    |
--------+--------+--------+--------+--------+--------+--------+--------+
 Slot   |01234567|01234567|01234567|01234567|01234567|01234567|01234567
--------+--------+--------+--------+--------+--------+--------+--------+
 PD  0  |XX......|........|........|........|........|........|........|
 PD  1  |..XX....|........|........|........|........|........|........|
 PD  2  |....XX..|........|........|........|........|........|........|
 PD  3  |......XX|........|........|........|........|........|........|
```

CC Command

- **Allows creation of a "Genesis complex profile" comprising one cell in the partition. Used at initial configuration to create a valid complex profile.**

```
GSP:CM> cc

WARNING: You must shut down all Partitions before executing
         this command.


G - Genesis Complex Profile
L - Last Complex Profile
   Select Profile: g
   Enter Cabinet number: 0
   Enter Slot number: 0
   Do you want to modify the complex profile? (Y/[N]) n
```

PS Command - (1 of 3)

- **This command displays detailed power and hardware configuration status.**

```
GSP:CM> ps


The following devices were found on the GSP bus:
                                            |          Core IOs
              |    UGUY     |               | IO Bay | IO Bay | IO Bay | IO Bay
              |             |     Cells     |   0    |   1    |   2    |   3
|Cab.|        |      |      |               |IO Chas.|IO Chas.|IO Chas.|IO
Chas.
| #  | GSP | CLU | PM  |0 1 2 3 4 5 6 7 |0 1 2 3 |0 1 2 3 |0 1 2 3 |0 1 2 3
+----+-----+-----+-----+---------------+-------+-------+-------+-------
-
| 0 | * | * | * |* * * * * * * * | * * | * * |         |


You may display detailed power and hardware status for the following items:
     1 - Cabinet
     2 - Cell
     3 - Core I/O
     4 - GSP
     Select Device: 1

Enter Cabinet number: 0
```

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

PS Command - (2 of 3)

```
HW status for SD32000 compute cabinet #0: NO FAILURE DETECTED
Power switch: on;   Power: enabled, good;   Door: closed
Fan speed: normal;   Temperature state: normal
Redundancy state: blowers redundant, BPSs redundant


                       | Main BP|                    |   IO Backplanes    |
                       | Power  |                    | IO Bay 0 | IO Bay 1 |
             | Main | Boards |     Cells       | Chassis  | Chassis  |
             | BP   | 0 1 2  | 0 1 2 3 4 5 6 7 | 0 1 2 3  | 0 1 2 3  |
+-------------+------+--------+-----------------+----------+----------+
  Populated    |  *   | * * *  | * * * * * * * * |  *    *  |  *    *  |
  Power Enabled |  *   | * * *  | * * * * * * * * |  *    *  |  *    *  |
  Powered On   |  *   | * * *  | * * * * * * * * |  *    *  |  *    *  |
  Power Fault  |      |        |                 |          |          |
  Attention LED |      |        | *               |  *       |          |


                    | Cabinet |   I/O   |
             |    BPS      | Blowers |  Fans   |
             | 0 1 2 3 4 5 | 0 1 2 3 | 0 1 2 3 |
+-----------+-------------+---------+---------+
  Populated | * * * * * * | * * * * | * * * * |
  Failed    |             |         |         |
```

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE

PS Command - (3 of 3)

```
Voltage margin: nominal;   Clock margin: nominal

PM firmware rev 5.0;    PM firmware time stamp:  THU MAY 18 19:50:51  2000
CLU firmware rev 5.0;  CLU firmware time stamp: WED MAY 24 15:38:15  2000


               CLU Status      PM Status      CLU POST
UGUY LEDs:       ***_*___        ***_            ____


                                     | Parity    | Connected  | Location
Inter-cabinet connections | Connected | error     | to cabinet |
(Upper/Lower) |
+------------------------+----------+---------+-----------+----------
--
  XBC [7-0]                | _*_*_*_* | _____ | 00000000   | N/A
  RC [7-0]                 | _____  | _____ | 00000000   | LLLLLLLL
```

InterWorks 2002
THE HP TECHNICAL TRAINING CONFERENCE