

Itanium® 2-based Servers: Superdome and mid-range

Mark Shaw

Sr. Systems Architect, HP
mark.shaw@hp.com

Paul Bouchier

Management Hardware Architect, HP
paul.bouchier@hp.com



Agenda

- Cellular systems
 - Chipset overview
 - System Architecture
 - High-Availability and reliability features
 - Security
- System Management
 - Networked Service Processor/Console
 - Partition Configuration
 - Platform manageability HW/FW architecture
 - New security features

Related Presentations

- High Performance Systems Itanium Strategy #2594
 - Thurs 1:30
- HP-UX Partitioning hands-on lab #2473
 - Fri 8:00
- HP-UX Workload Manager hands-on lab #2474
 - Fri 10:10

HP Cellular Systems Family

Superdome (PA-RISC & Itanium®)



16 cells, 64 cpu's
512 DIMM slots
192 PCI slots

PA-RISC: rp8420 Itanium®



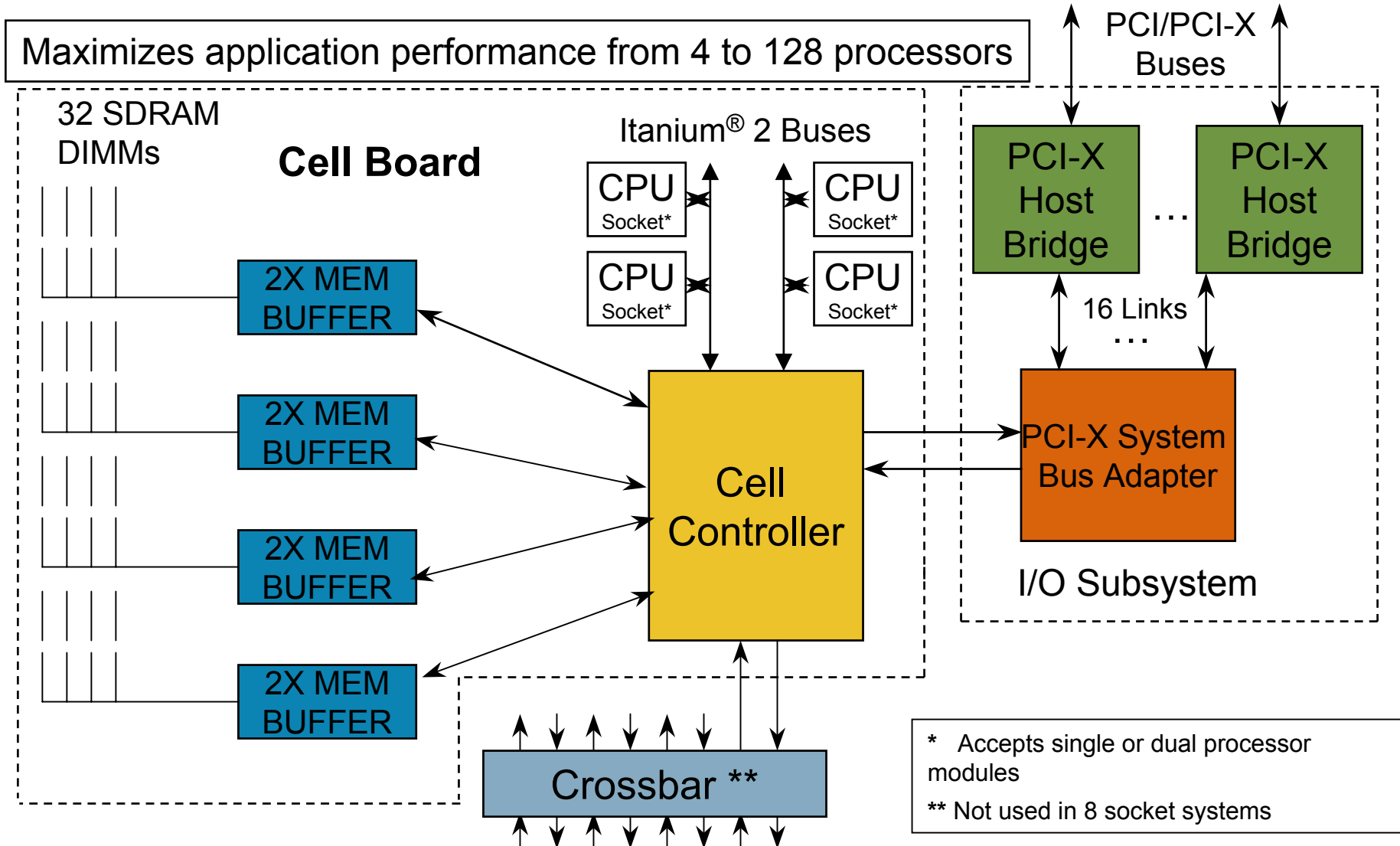
4 cells, 16 cpu's
64 DIMM slots
32 PCI slots,
17U high, 19" rack

PA-RISC: rp7420 Itanium®



2 cells, 8 cpu's
32 DIMM slots
16 PCI slots,
10U high, 19" rack

sx1000 chipset includes five distinct chips



System Performance and Scalability Features

- Low-latency, high-bandwidth backplane
- Cell-based Design
 - Highly integrated cell controller with internal crossbar to reduce latency and remove bottlenecks
 - Two processors (vs. four) attached to a single CPU front-side bus to provide more bandwidth per processor and reduce queuing delays
 - Large capacity, low-latency high-bandwidth memory
 - Memory directory coherency filters snoop traffic and dramatically reduces load on the system backplane
- Dedicated Host Bridge ASIC and bus PCI-X slot
 - Performance and error isolation for each I/O card

Cell and I/O assemblies

- Integrated CPU / memory cell controller with internal crossbar
 - Low latency, high-bandwidth through a single ASIC
 - 2500 pins among the largest in the world
 - 32GB memory on 32 double-height DIMMs, 64GB coming
 - Four memory sub-systems
 - Directory coherency
 - Four CPUs on two busses
 - Backplane and I/O links

- I/O Controller and dedicated PCI-X host bridge ASICs
 - 12 PCI-X Superdome slots
 - 8 PCI-X rx8620 & rx7620
 - Single PCI-slot per bus



System Capacities

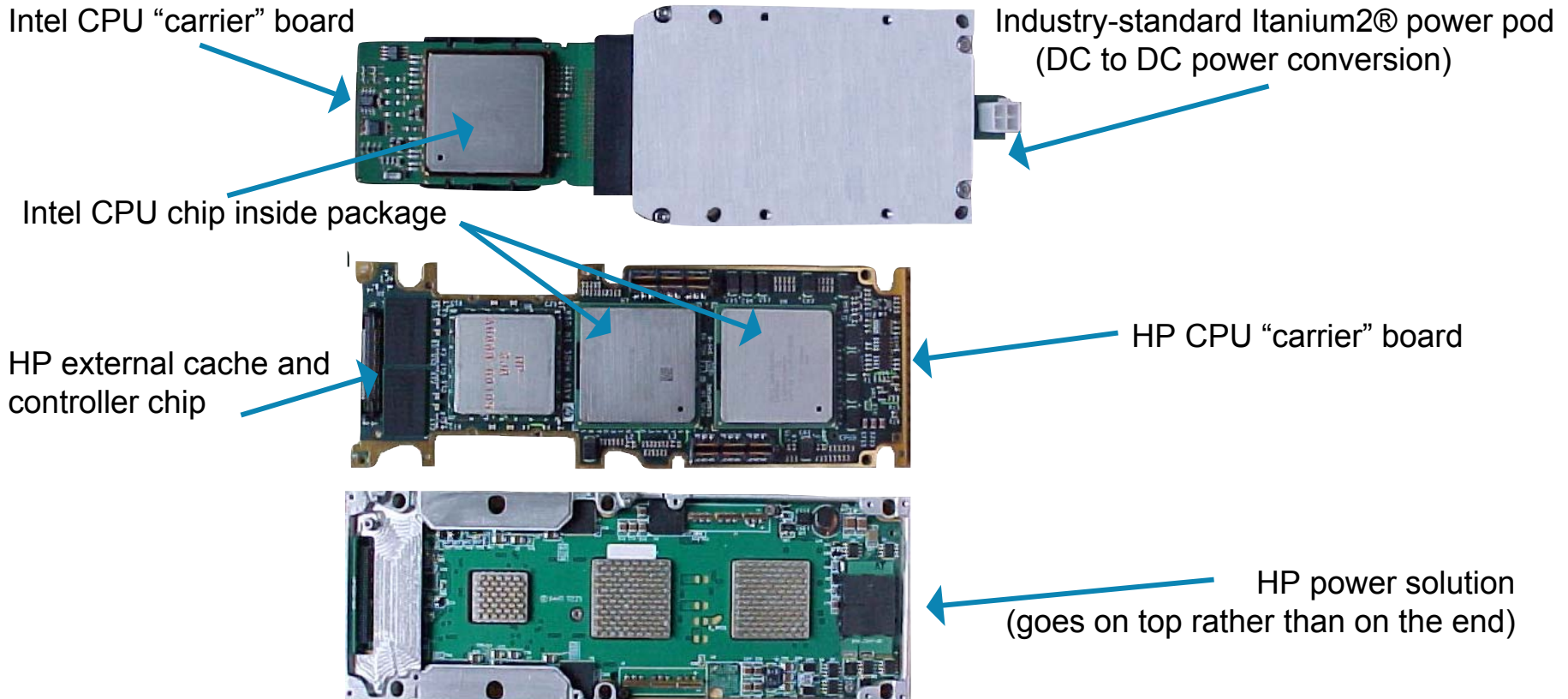
	CPUs	Memory	I/O slots	nPartitions
Superdome SD64000	64	512 GB ¹	192 ²	16 ²
Superdome SD32000	32	256 GB ¹	96 ²	8 ²
Superdome SD16000	16	128 GB ¹	48	4
rx8xx0	16	256 GB	32 ²	4 ²
rx7xx0	8	128 GB	16	2

¹ Memory capacities double next year

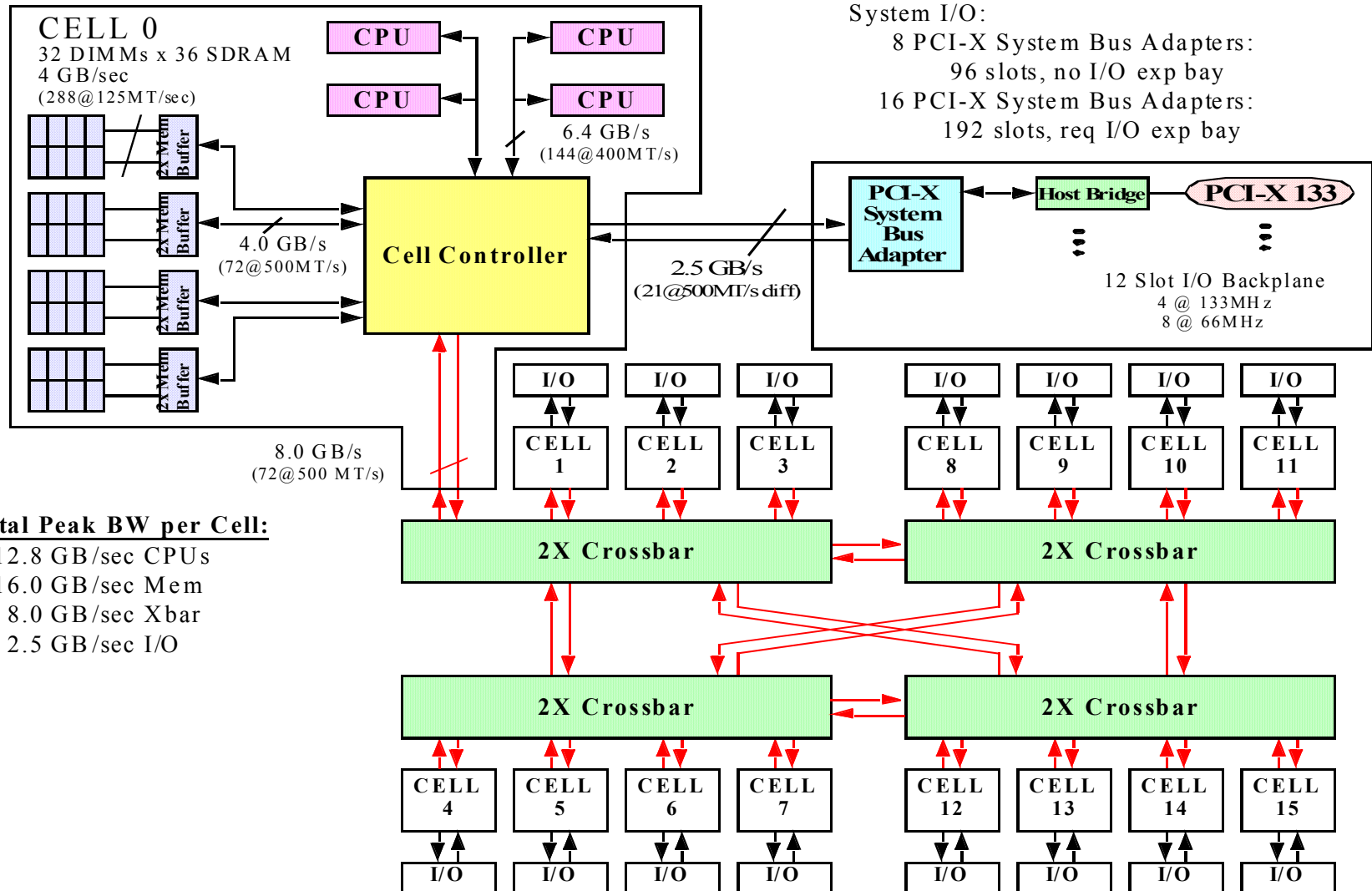
² I/O expander required

“Inventing” a dual core Itanium® CPU

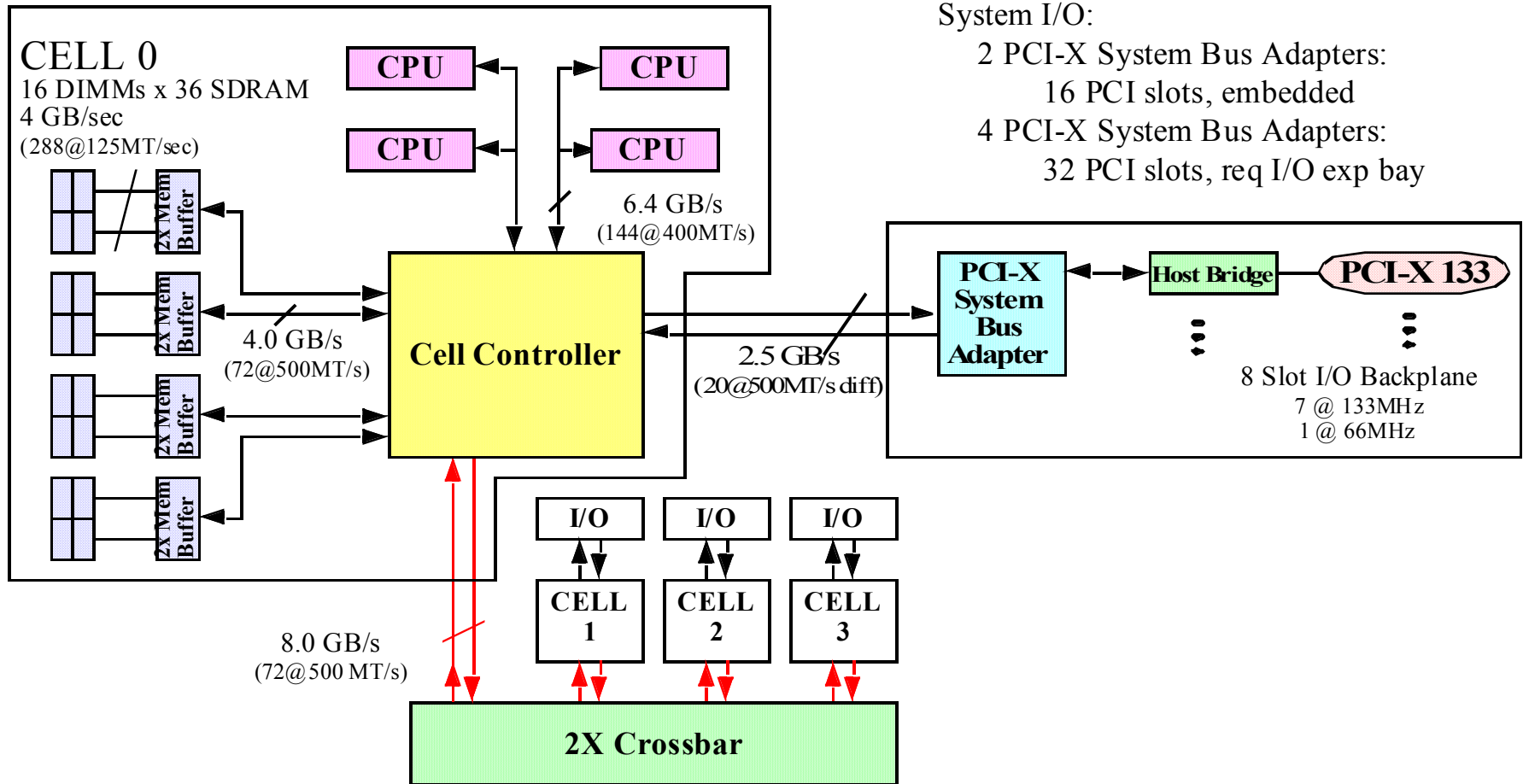
- The standard Intel cartridge packaging is not at maximum density
 - CPU silicon chip is in a package, on a carrier board, with power on the end
 - The basic chip and package could be packed much more densely



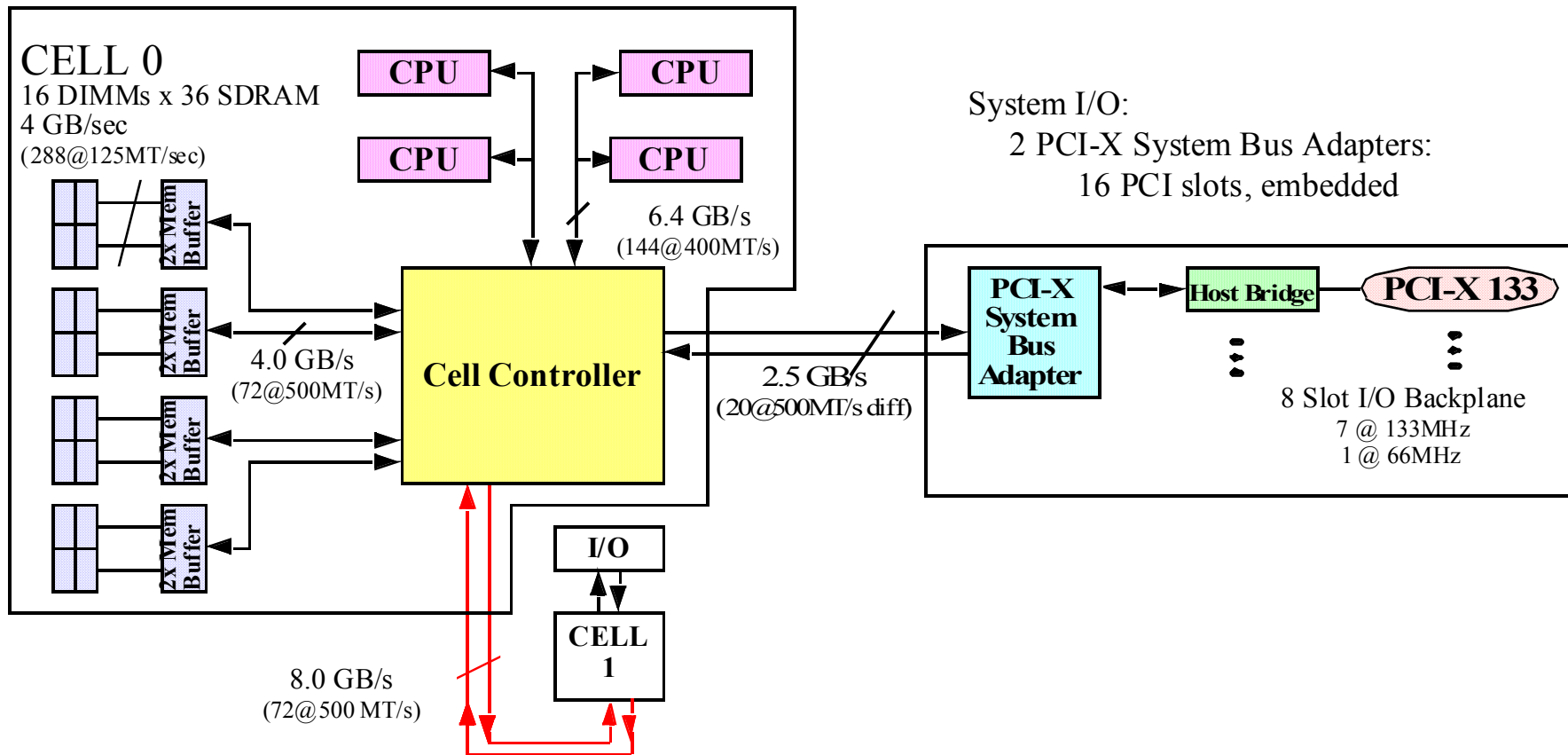
Superdome architecture



mid-range Itanium® architecture



mid-range Itanium® architecture



High availability features available across the family

Keep it running

n+1 features (hot swappable)

1. cabinet blowers
2. i/o fans
3. dc power supplies
4. cell backplane dc power supplies

error correction

1. ecc on cpu cache
2. parity protected cpu & i/o links
3. single wire correction on fabric & i/o
4. ecc on all fabric and memory paths
5. Chip-kill memory

redundant ac input power (optional)



Fix it fast

diagnostic features:

1. test station
 - asic level scan tools
 - remotely accessible via lan
2. enhanced predictive support
3. high availability observatory
4. ems monitoring system
5. dynamic processor resilience
6. dynamic memory resilience

fault isolation technologies

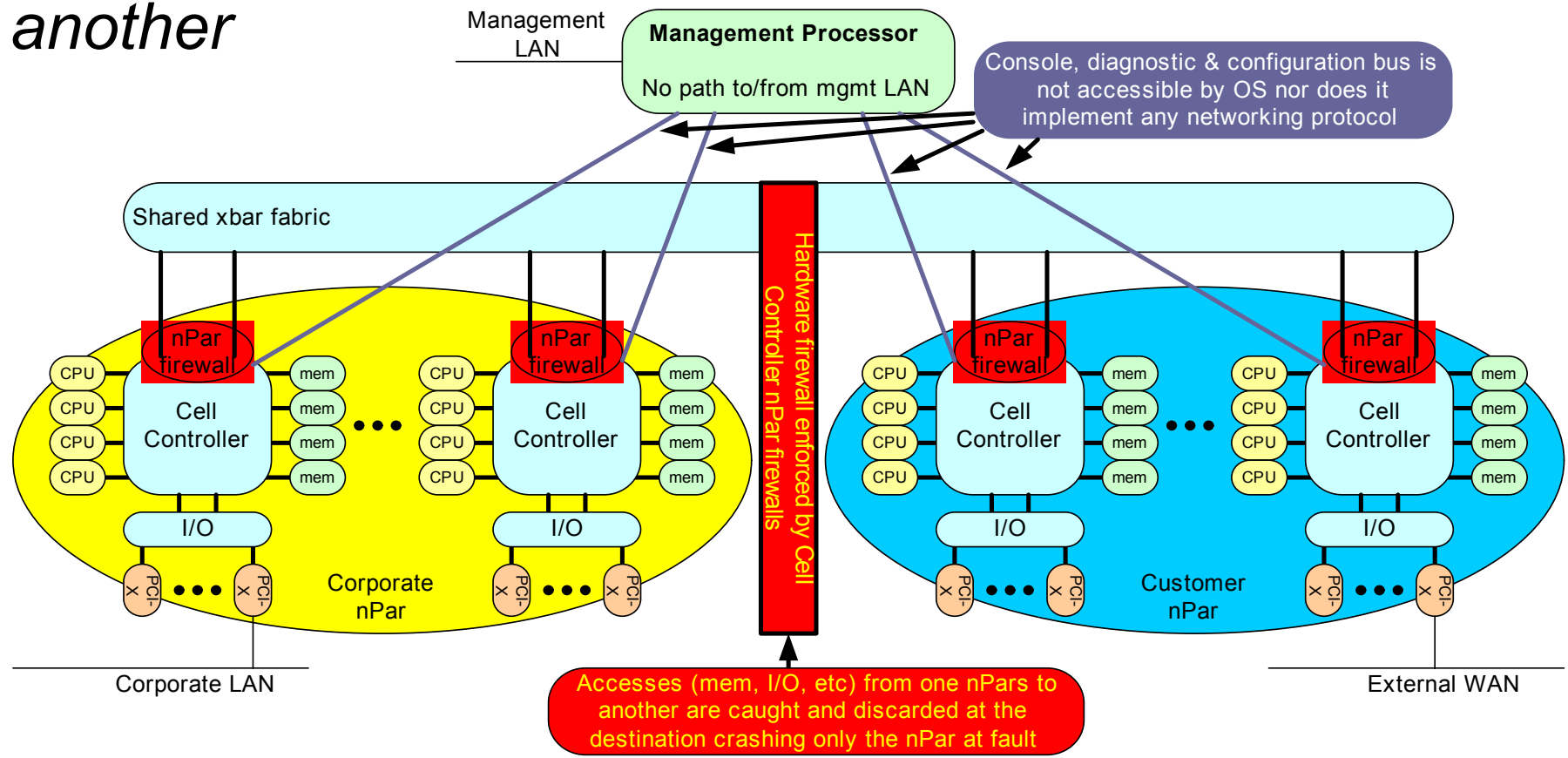
online removal, replacement:

1. cell assemblies*
2. i/o cards

* note: os version dependent

Superdome Security Hardware

No path exists for an entity on one nPar to invade another



Agenda

■ Cellular systems

- Chipset overview
- System Architecture
- High-Availability and reliability features
- Security

■ System Management

- Networked Service Processor/Console
- Partition Configuration
- Platform manageability HW/FW architecture
- New security features

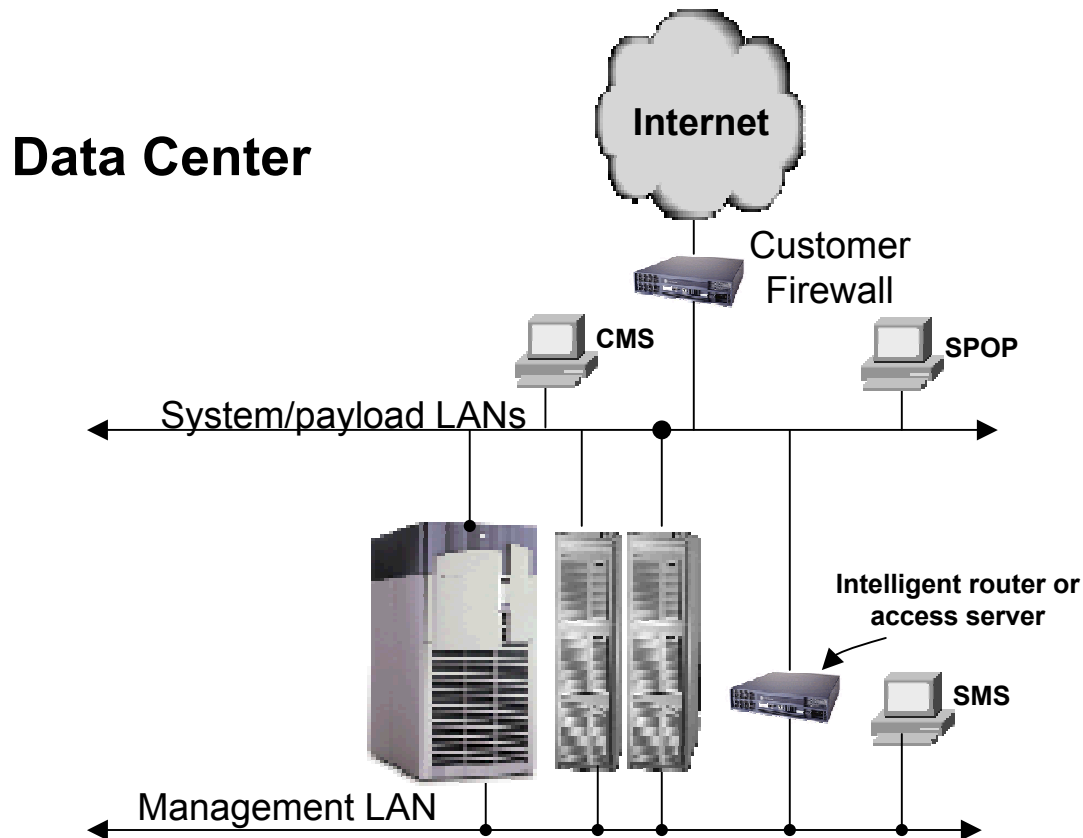
Platform Management

Definitions

- MP = GSP = Service Processor
- Complex: The entire box(es) including all partitions

Demo

SMS & SPOP configuration

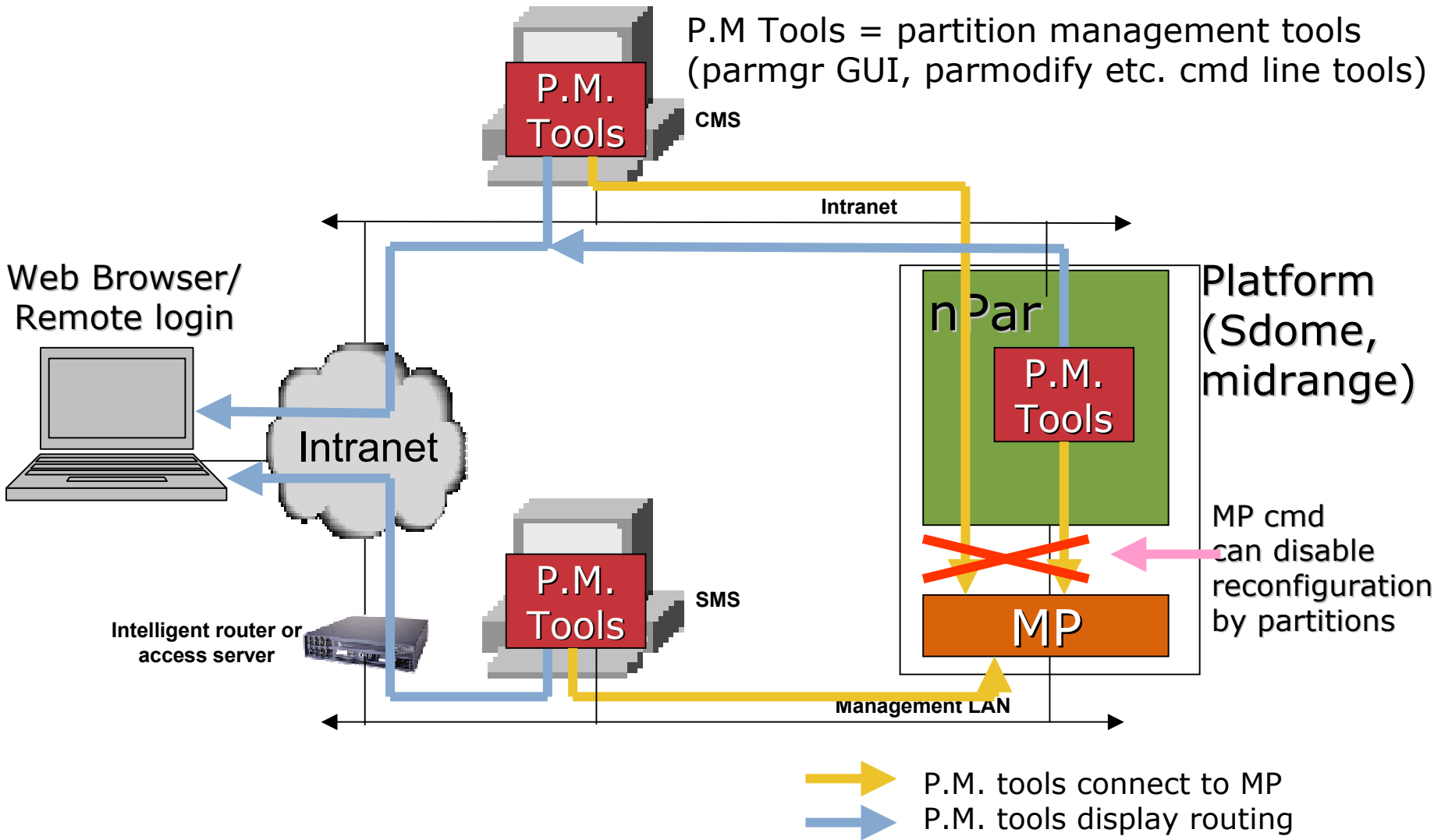


CMS: Central Management Station (Service Control Manager, Insight Manager)

SPOP: ISEE Support Point of Presence

SMS: Support Management Station

Partition Management



Partition Management

P.M Tools run on nPAR:

- **Benefit:** browser session can manage other aspects of OS (e.g. sam over web). Parmgr can launch other apps (e.g. PCI OL*).
- **Security consideration:** requires npar to be enabled to reconfigure the complex.

P.M. Tools run on CMS

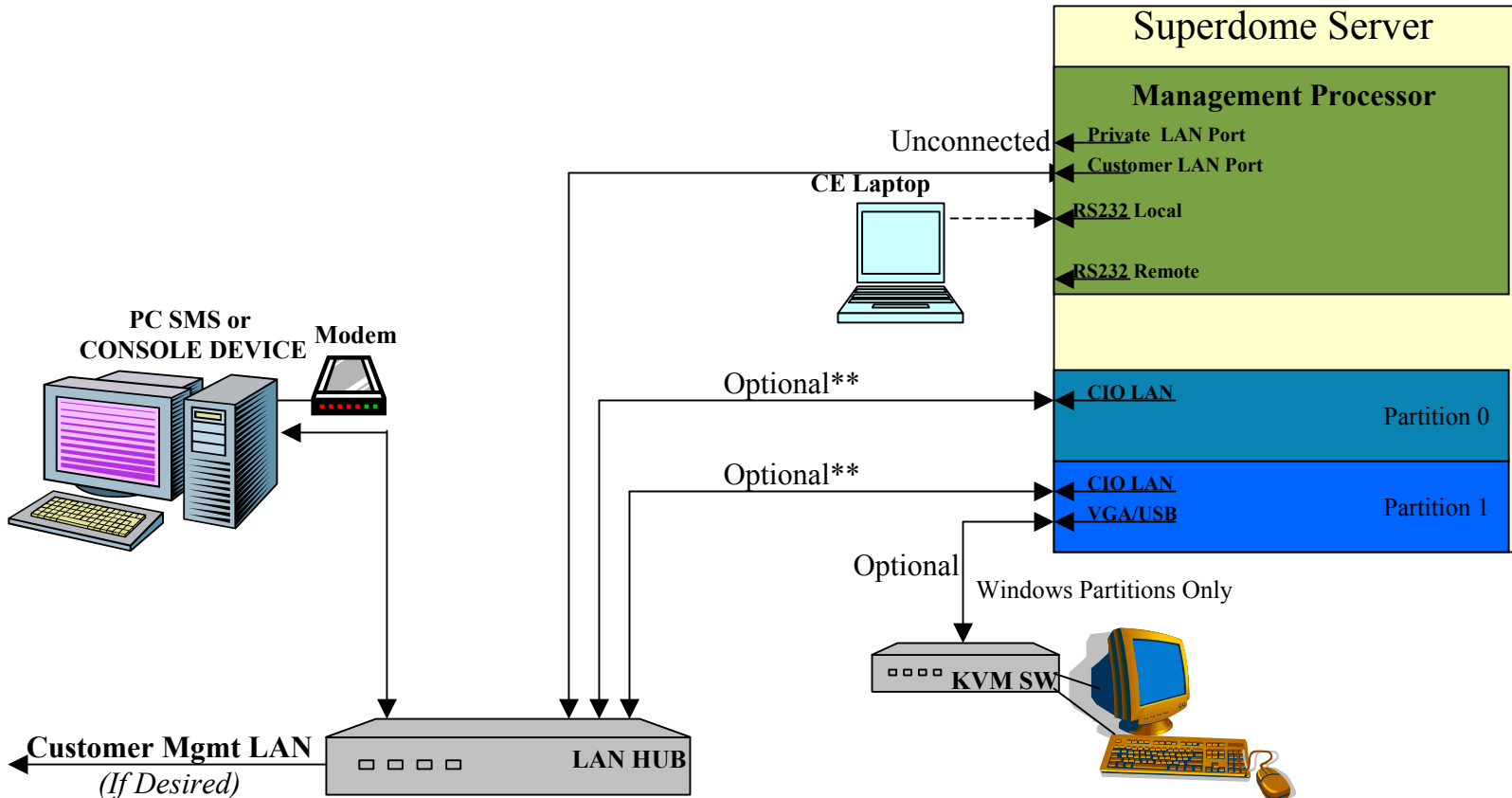
- **Benefit:** Tools can be launched from data-center level admin tool. Browser session can manage other aspects of OS (e.g. sam over web).
- **Security consideration:** requires npar to be enabled to reconfigure the complex.

P.M. Tools run on SMS

- **Benefits:** Partitions can be reconfigured in the absence of any running OS. Partitions can be disabled from reconfiguring partitions (security benefit).
- **Usability Consideration:** Partition management system may not have easy network access to the system for OS management. This may or may not be important to the customer.

Console and PC SMS Network Configuration

Single IPF or PA-8800 Superdome

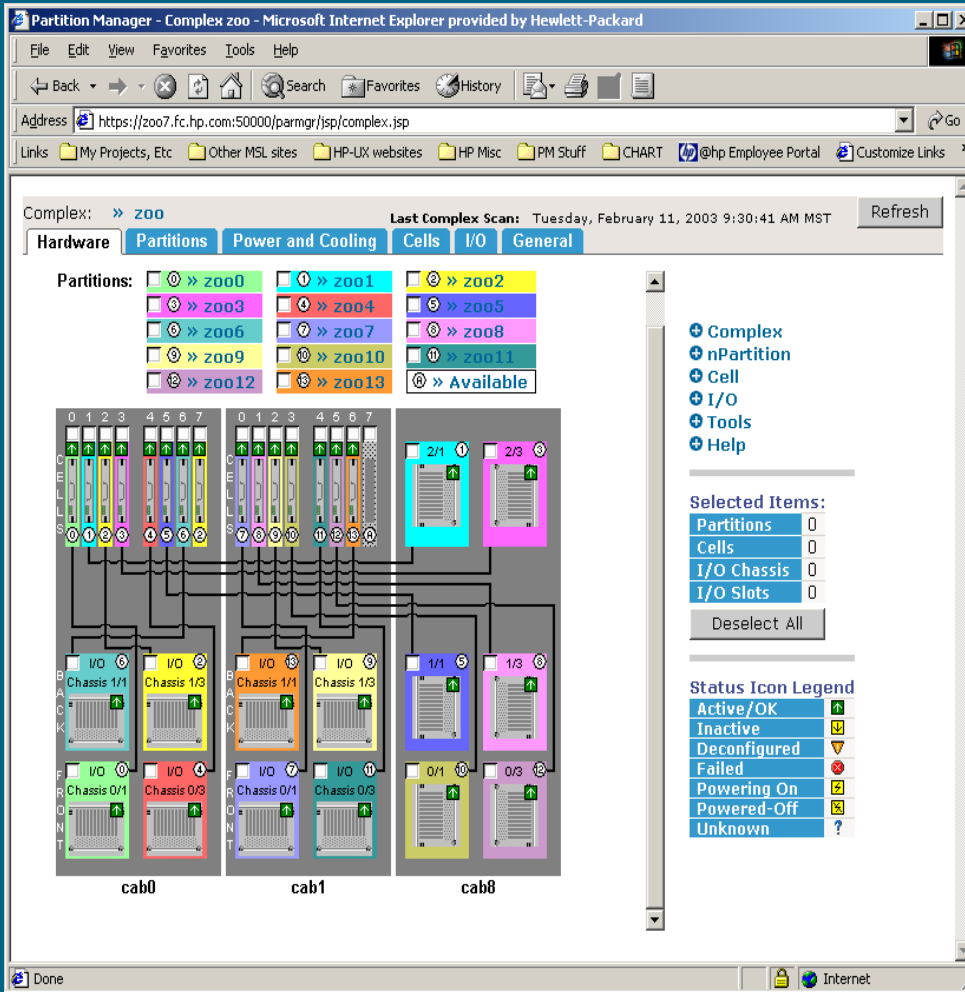


** Security concerns may dictate that a partition LAN not be connected to the management LAN.

Alternatives: 1) Access from a management station to a partition LAN through a secure router. 2) Text mode access to commands via the console.

*** Serial cable 24542G needed for initial install

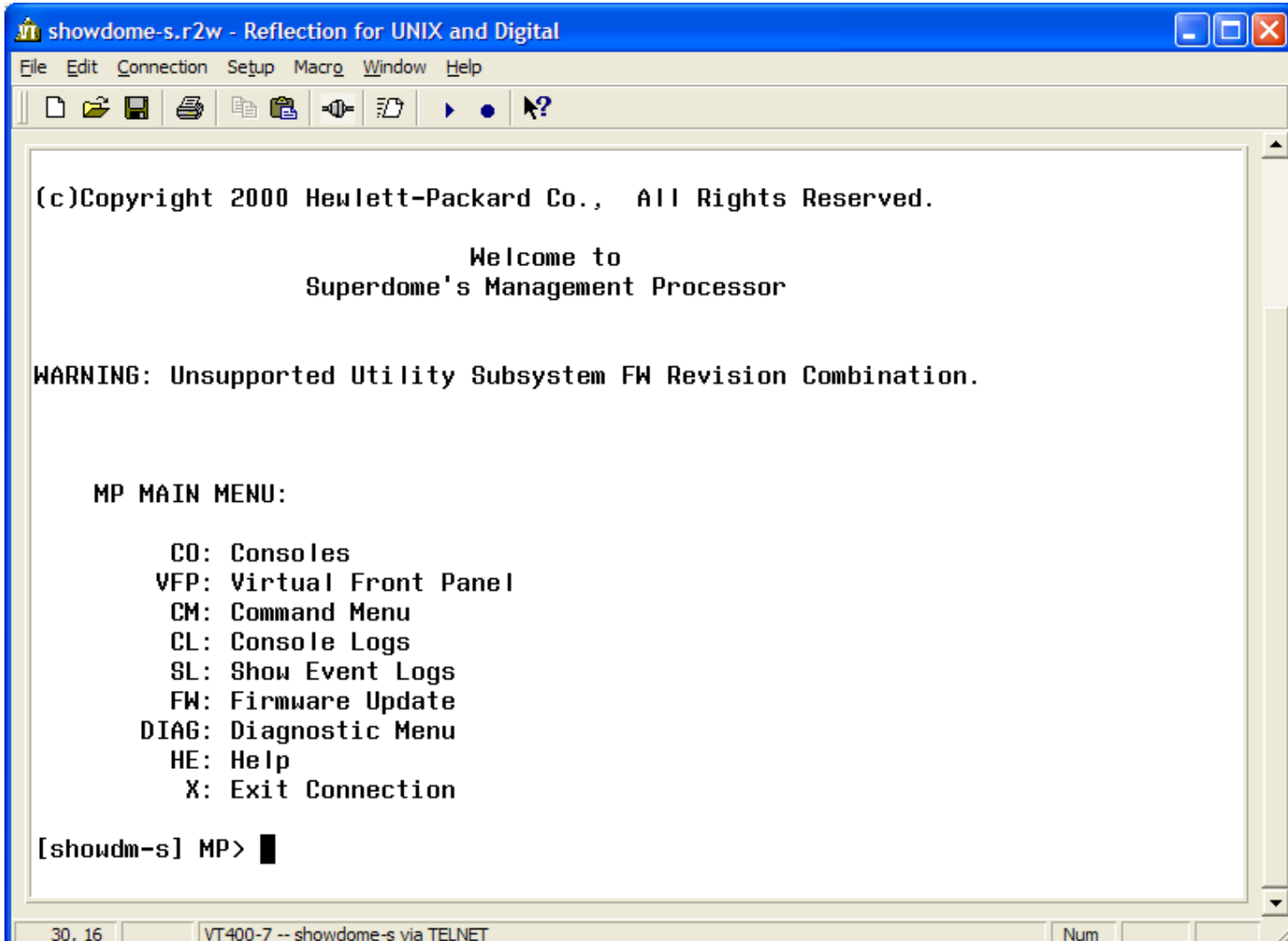
Partition Manager New Features



- ✓ New web interface
- ✓ Graphical “big picture” views of
 - nPars
 - Hardware in Complex
- ✓ Supports new OS/HW features
 - Cell local memory for HP-UX 11i v.2 partitions
 - Inter-partition security
- ✓ Remote admin of Superdome Madison complex
- ✓ Compatible with iCOD/Pay-Per-Use
- ✓ Increased integration with SCM 3.0
- ✓ Native on Windows (2H03)
- ✓ J2EE app runs in tomcat web server

MP operation

Always remember – HE from the main menu gets HELP!



The screenshot shows a terminal window titled "showdome-s.r2w - Reflection for UNIX and Digital". The window contains the following text:

```
(c)Copyright 2000 Hewlett-Packard Co., All Rights Reserved.  
  
Welcome to  
Superdome's Management Processor  
  
WARNING: Unsupported Utility Subsystem FW Revision Combination.  
  
MP MAIN MENU:  
  
CO: Consoles  
VFP: Virtual Front Panel  
CM: Command Menu  
CL: Console Logs  
SL: Show Event Logs  
FW: Firmware Update  
DIAG: Diagnostic Menu  
HE: Help  
X: Exit Connection  
  
[showdm-s] MP> █
```

The terminal window also shows a status bar at the bottom with the text "VT400-7 -- showdome-s via TELNET" and a numeric keypad icon.

MP operation – Console menu

```

showdome-s.r2w - Reflection for UNIX and Digital
File Edit Connection Setup Macro Window Help
[Icons: File, Folder, Disk, Print, Copy, Paste, Undo, Redo, Home, Help]

MP MAIN MENU:

    CO: Consoles
    VFP: Virtual Front Panel
    CM: Command Menu
    CL: Console Logs
    SL: Show Event Logs
    FW: Firmware Update
    DIAG: Diagnostic Menu
    HE: Help
    X: Exit Connection

[showdm-s] MP> co


Partitions available:

#   Name
---  ---
0)  showdome0 Windows
1)  showdome1 Linux
2)  showdome2 HPUX
Q)  Quit

Please select partition number:

186, 37 | VT400-7 -- showdome-s via TELNET | Num
  
```

Telnet Clients for Console Access

<u>Console</u>	Reflection1 for HP or hpterm (X)	Hyperterm or telnet from cmd.exe window	Other ANSI or VT100
HPUX	Best <input checked="" type="checkbox"/>	Limited 	<input checked="" type="checkbox"/>
Windows	<input checked="" type="checkbox"/>	Best <input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Linux	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
IPF FW	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
MP	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

MP operation - VFP

```

showdome-s.r2w - Reflection for UNIX and Digital
File Edit Connection Setup Macro Window Help
[Icons]
# Partition state          Activity
- -----
0 At System Firmware console  SAL_INFO_REC_CLEAR          55 Logs

# Cell state              Activity
- -----
0 Cell has joined partition  337 Logs
E 2 Cell has joined partition 247 Logs

E indicates error since last boot
MP:VFP (Use '?' to display help or ^B to Quit) >
281, 50 | VT400-7 -- showdome-s via TELNET | Num

```

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

- **3 User Capability Levels, set by SO command**
 - 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 MP**
 - 3. Administrator**
 - Has operator level capabilities**
 - Can reconfigure the MP**

Boot Startup / Shutdown / Reset



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

```
MP:CM> rs
```

Execution of this command irrecoverably halts all the system processing and I/O activity and restarts the Partition chosen.

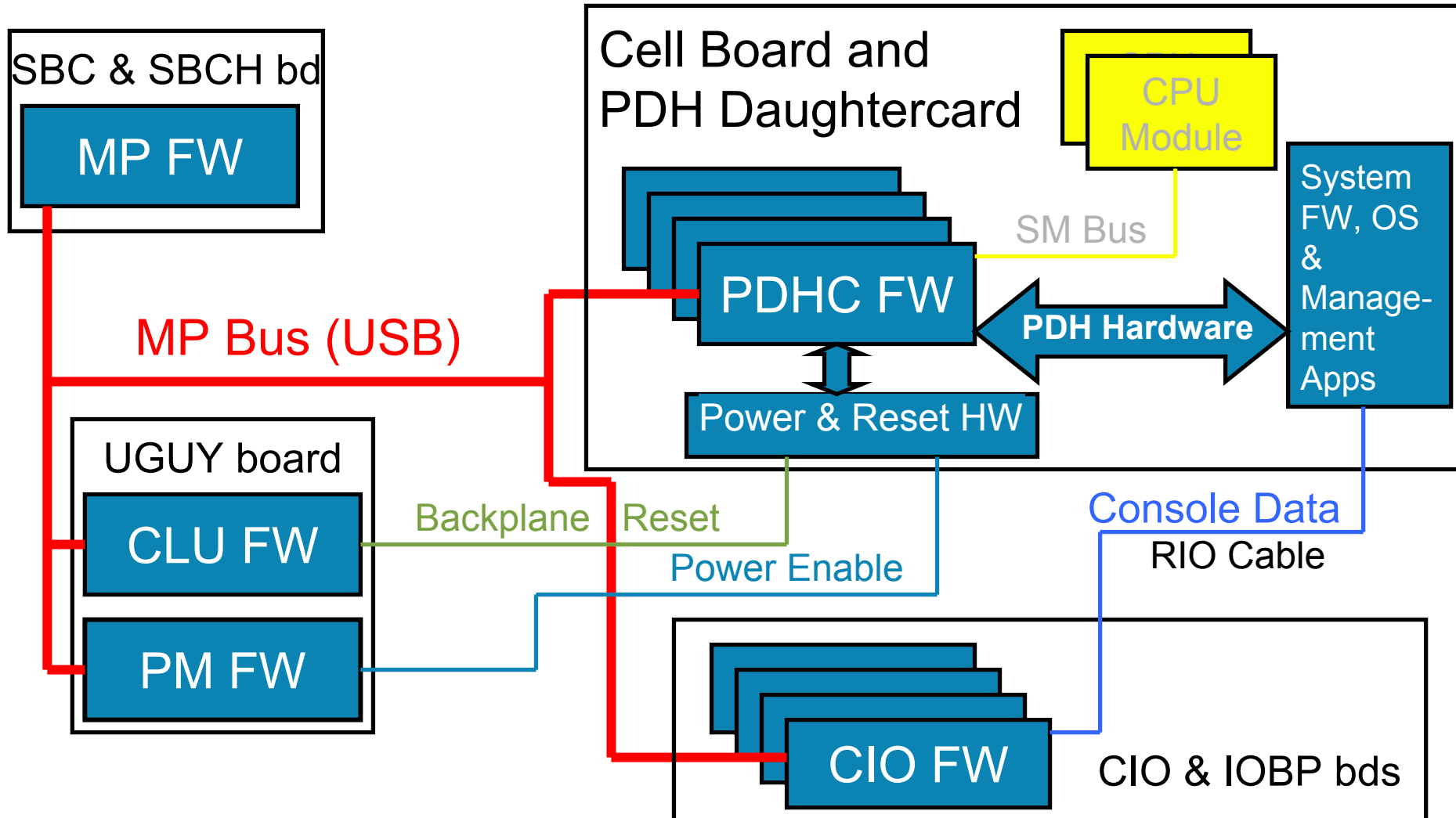
```
  #      Name
  ---    -
  0)    Partition 0
  1)    Partition 1
```

```
Select a partition number : 0
```

```
Do you want to reset Partition number 0,
named Partition 0? (Y/[N]) y
```

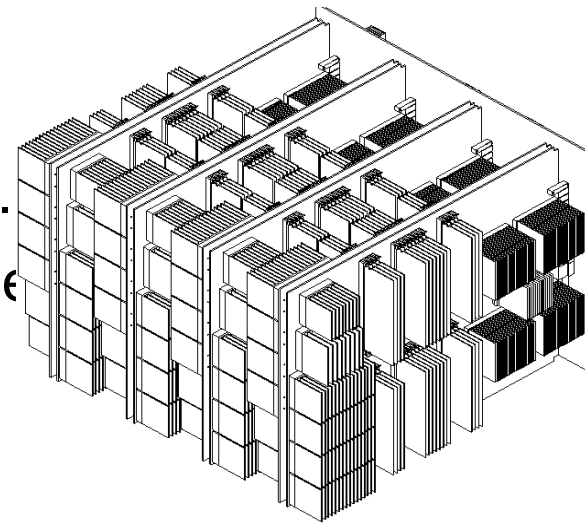
```
-> Partition selected will be reset.
```

Manageability Subsystem HW Architecture

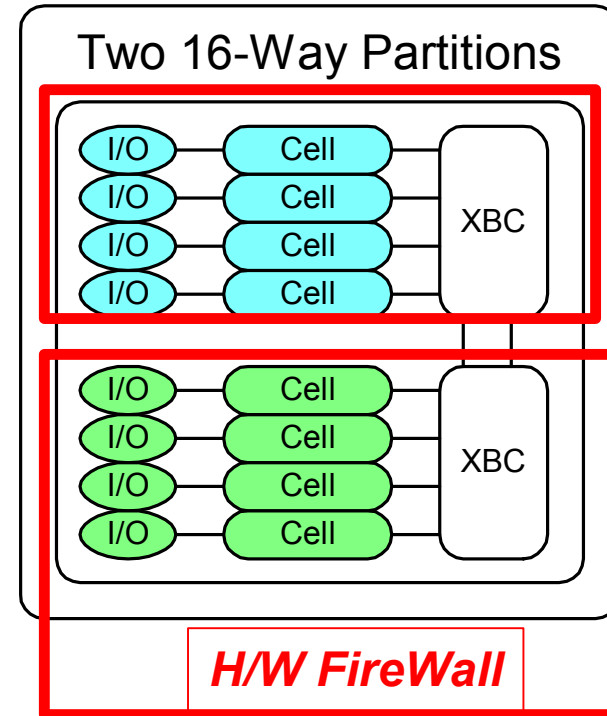
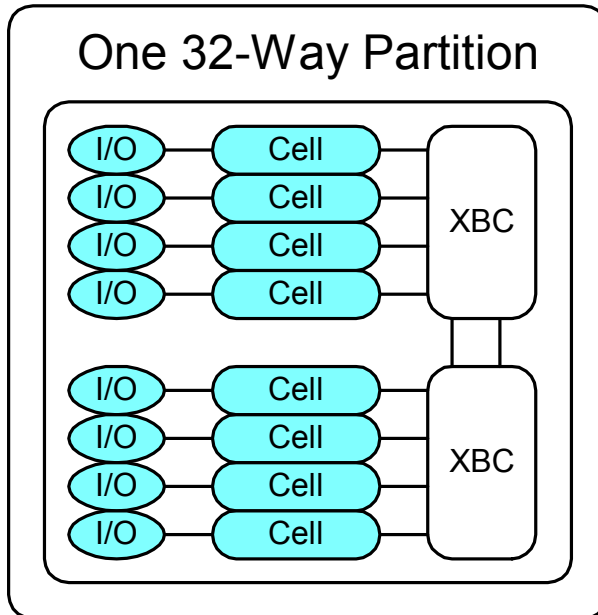


Superdome Partitions: What are they?

- 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.



Partitions on an eight cell SD32000



Complex profile concepts

Complex Profile is group of data structures which define:

- **SCCD - Stable Complex Config Data**
 - **Global parameters (affecting all partitions):**
 - **Cell \leftarrow \rightarrow partition association**
 - **system name, serial #, etc**
- **PCD – Partition Config Data**
 - **Partition parameters (affecting one partition)**
 - **Partition name**
 - **Active Cells**

Complex Profile stored in MP (master copy) and on every cell

Complex Profile initially created by the MP (CC command)

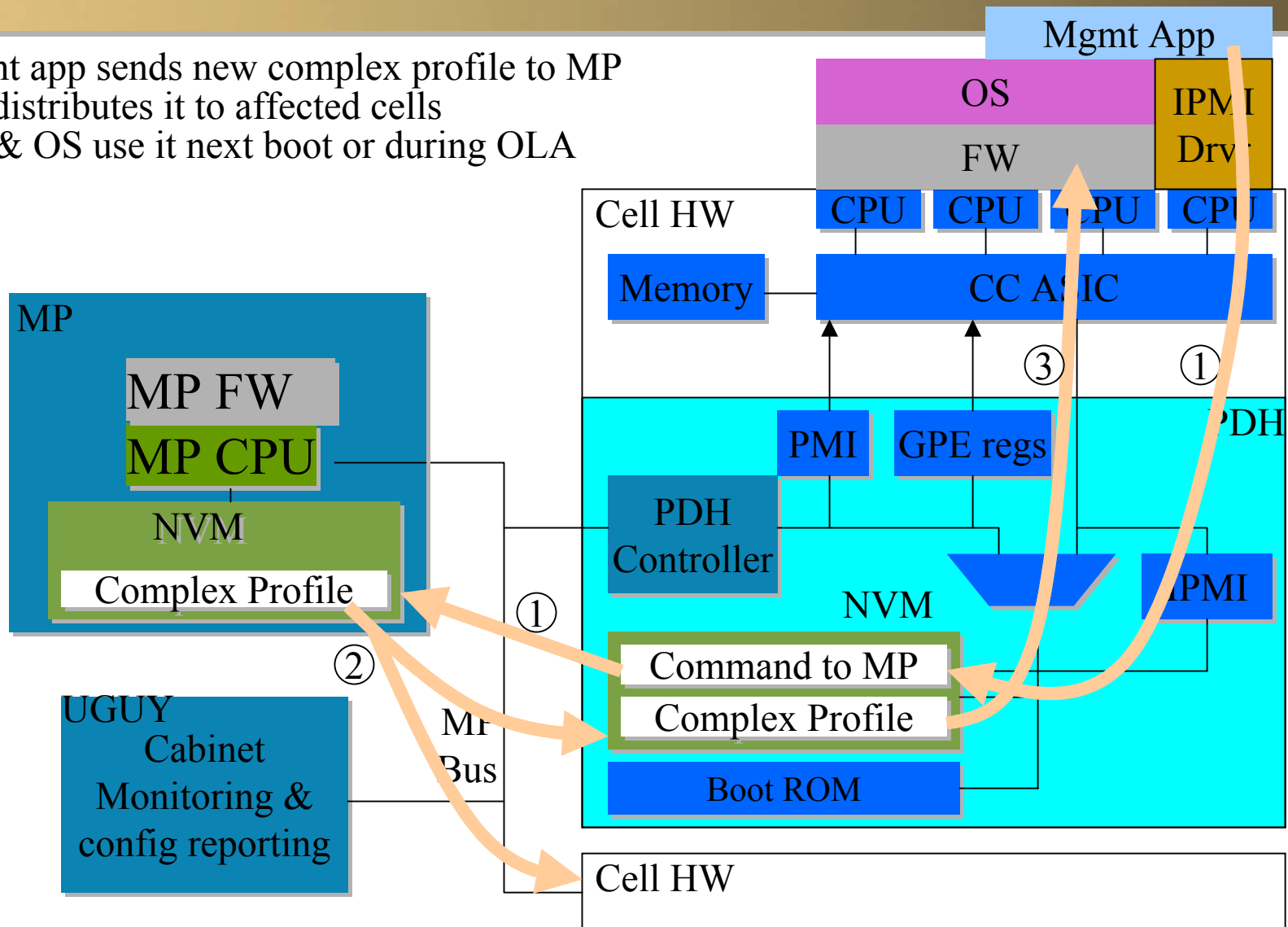
- **Read and modified by the Partition Manager, the partition commands, and system firmware**

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 MP begin checking complex profile for consistency across cells. PDHC and MP 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 MP clear the BIB flag after the complex profile has been checked
- Cell boots, reads the complex profile, and rendezvous with the other cells in its partition.

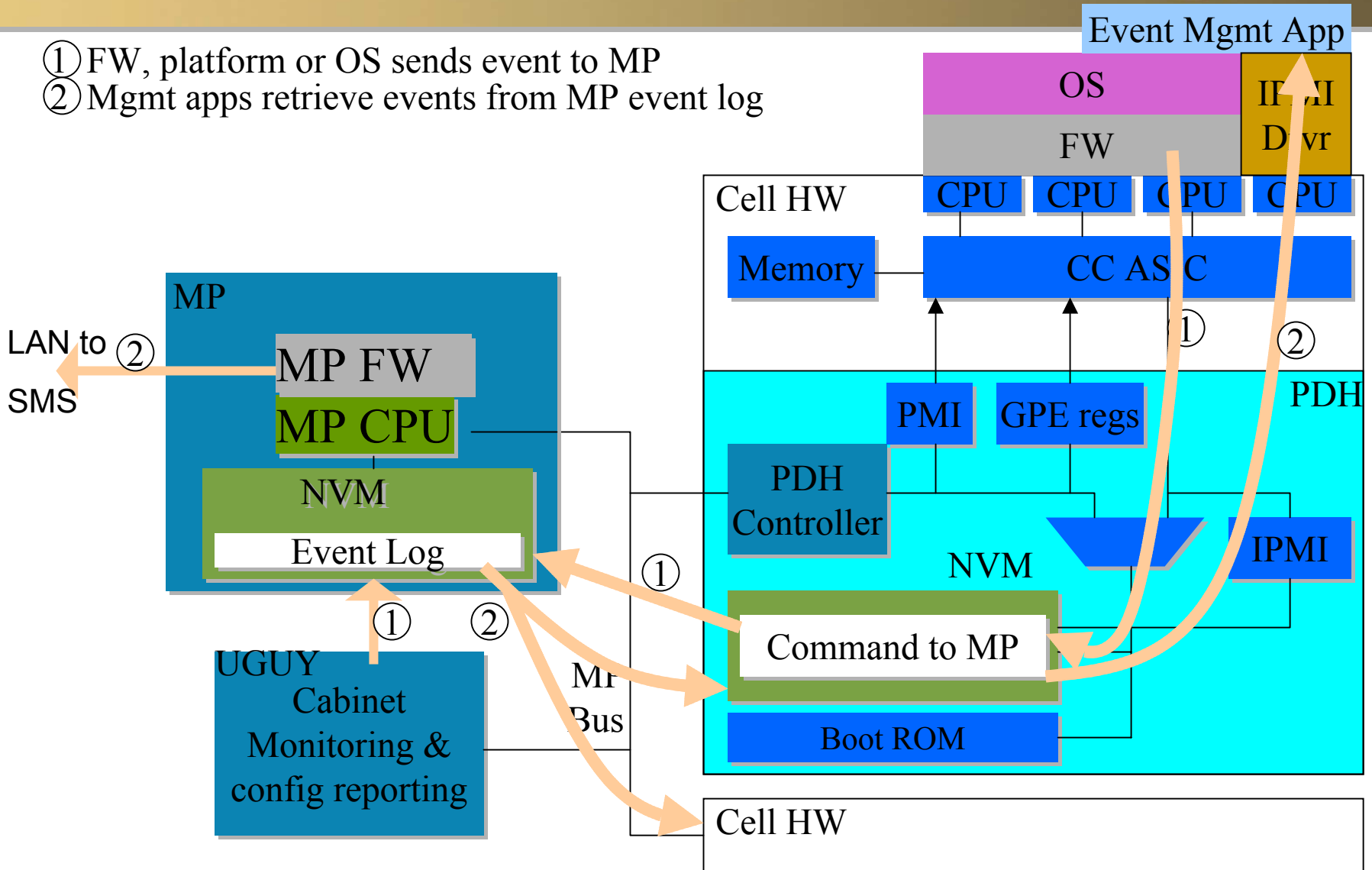
Configuration Management System

- ① Mgmt app sends new complex profile to MP
- ② MP distributes it to affected cells
- ③ FW & OS use it next boot or during OLA



Event Management System

- ① FW, platform or OS sends event to MP
- ② Mgmt apps retrieve events from MP event log



New & Forthcoming Security Features

- Parperm command, reconfigure disable, reflash disable
 - Enables admins to limit the capabilities of root on partitions
- IPMI password (SO command)
 - Enables admins to limit LAN user access to reconfiguration capabilities
- SA command
 - Enables admins to disable services
- Embedded Web Console
 - Provides privacy through encryption for MP/console access
- SNMP identification integrates with forthcoming Insight Mgr version



HP WORLD 2003

Solutions and Technology Conference & Expo

Interex, Encompass and HP bring you a powerful new HP World.

