



HP Superdome  
Technologie;  
Partitionierung mit  
HPUX 11i.  
Konsolidierungs  
Plattform

Martin Beuse  
HP CSPS  
Technical Consultant BCS

Agenda 10.04.2003  
DECUS 3G03 Raum Schumann



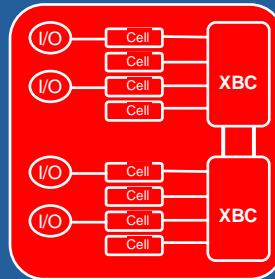
**13.00 – 15:00**

- HP Superdome Technologie
- Partitionierung unter HPUX 11i
- Beispiele eine Konsolidierung

*Martin Beuse*



## Die Superdome im Überblick



Überblick der Superdome-Modelle  
Prinzip der Ccnuma Technologie

## hp Superdome Überblick



### SD16000

- 1 – 4 Cell Boards
- 2 – 16 CPU's
- 4 – 64 Gbyte RAM
- 1 – 4 I/O Chassis
- 12 – 48 PCI Slots
- 1 – 4 nPartitionen

### SD32000

- 1 – 8 Cell Boards
- 2 – 32 CPU's
- 4 – 128 Gbyte RAM
- 1 – 4 I/O Chassis
- 12 – 48 PCI Slots
- 1 – 4 nPartitionen

### SD64000

- 1 – 16 Cell Boards
- 2 – 64 CPU's
- 4 – 256 Gbyte RAM
- 1 – 8 I/O Chassis
- 12 – 96 PCI Slots
- 1 – 8 nPartitionen

### SD64000 + IOX

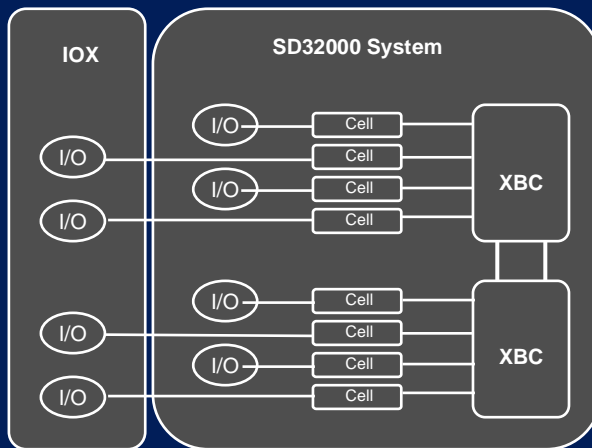
- 1 – 16 Cell Boards
- 2 – 64 CPU's
- 4 – 256 Gbyte RAM
- 1 – 8 I/O Chassis
- 12 – 192 PCI Slots
- 1 – 16 nPartitionen

# Superdome System Architektur im Überblick an Beispiel einer SD32000



Die drei Grundbausteine der Ccnuma Technologie:

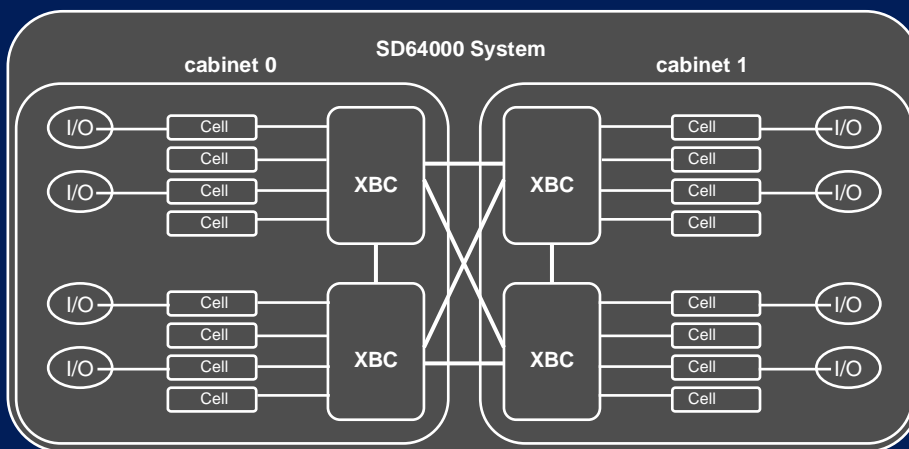
- 1. IO
- 2. Cellboard
- 3. Crossbar



# Superdome System Architektur im Überblick an Beispiel einer SD64000



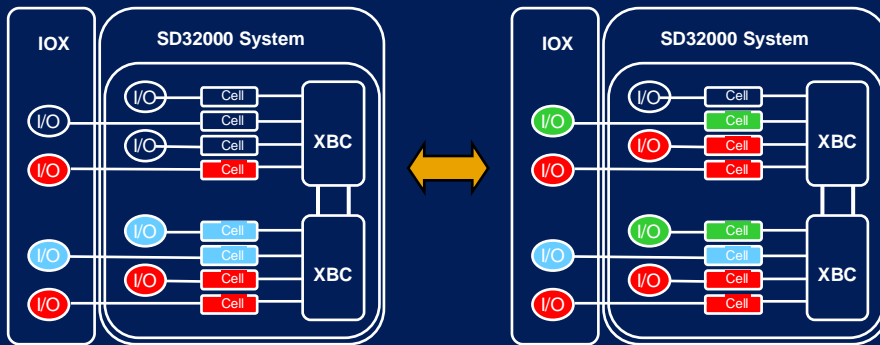
Superdome dual cabinet 64-way



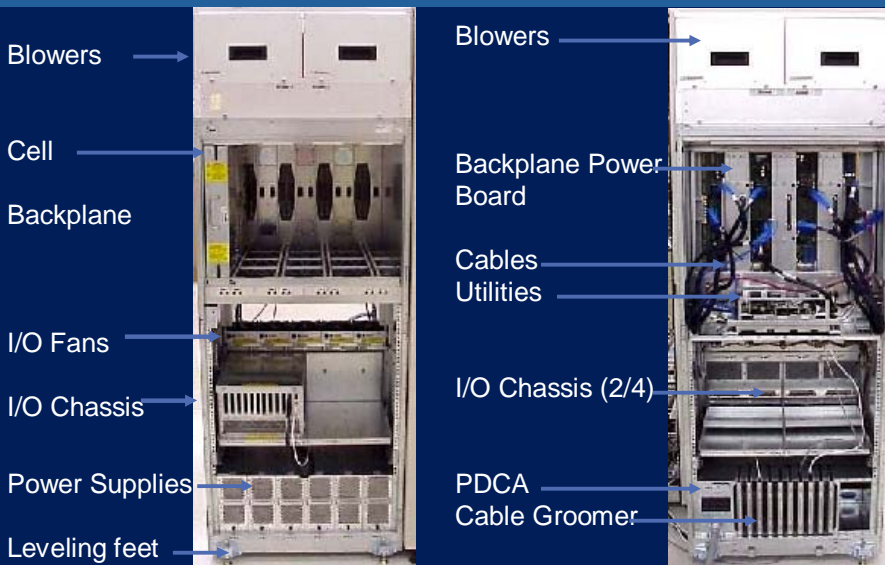
# Superdome flexibel durch nPartitionen



Eine nPartition besteht aus einem oder mehreren Cellboards!  
 Die Abgrenzung wird auf den Level der Crossbar vollzogen  
 Änderungen der nPars z.Zt. durch deren Reboot

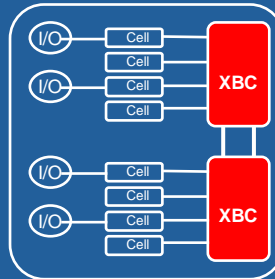


# superdome system view





## Die Superdome Crossbar (XBC)

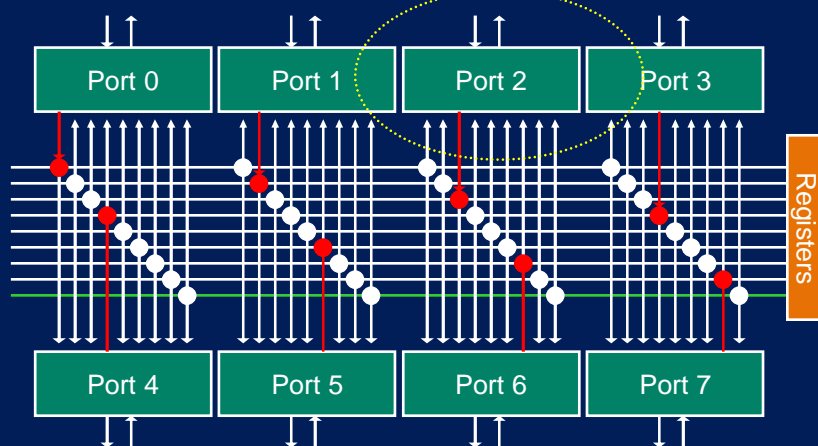


Die XBC ASIC:

- hohe Bandbreite
- niedrige Latenzzeit

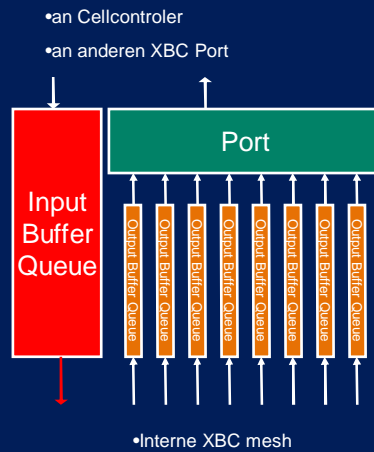
application specific integrated circuit (ASIC) chips in a 0.18µm, low-voltage CMOS process

## XBC Detail



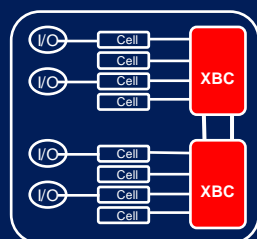
- 8 identische ports, an cell controller oder andere XBC
- 1 register block
- 2 Physikalische bilden einen logischen XBC

## XBC Port Detail



Queues driving queues

## XBC am Beispiel SD32000



Das XBC ASIC :

•2 logische XBC

•in Summe 16 logische Ports

davon sind:

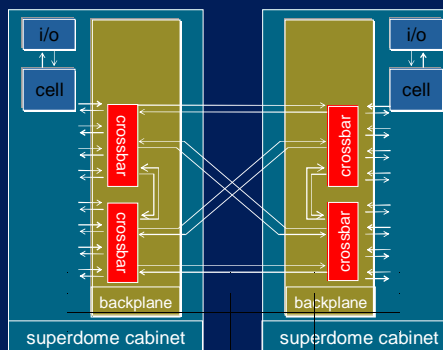
- 8 für Cellcontroller
- 4 permanent Verbunden
- 2 konfigurierbar
- 2 not usable

- Die XBC Links weisen nahezu keine Fehlerbitrate auf
- eine 2 Level SEC/DED Schema verhindert eine Beeinflussung untereinander
- ECC über die komplette Einheit
- keine Beeinflussung der Latency durch doppelte Parity
- Jeder Port hat individuelle Queues und Datenpfade

## crossbar mesh in einer SD64



- Maximalausbau der crossbar mesh
  - 4 crossbars
  - 4 cells pro crossbar
- alle Links haben gleiche Bandbreite und Latenzzeit
  - minimierte Latenzzeit
  - maximale Bandbreite
- Eingebaute point-to-point Paket Filterung und routebares Netzwerk
  - ermöglicht Hardware Isolation von Fehlern
- Verbindung von 16 Zellen mit 3 Latenzbereichen
  - lokale Zelle
  - lokale Crossbar
  - remote Crossbar



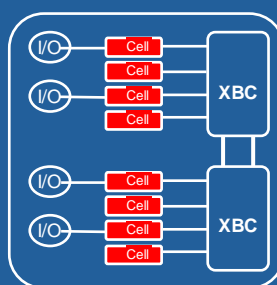
Durchschnittliche Latenzzeiten

processes	ns
4	200
8	250
16	275
32	315
64	335

### speeds & feeds

Operating Frequency	500 MHz
Max. Mesh, GB/s	64 GB/s
Max. I/O, GB/s	28.8 GB/s
TPC-C	423,414

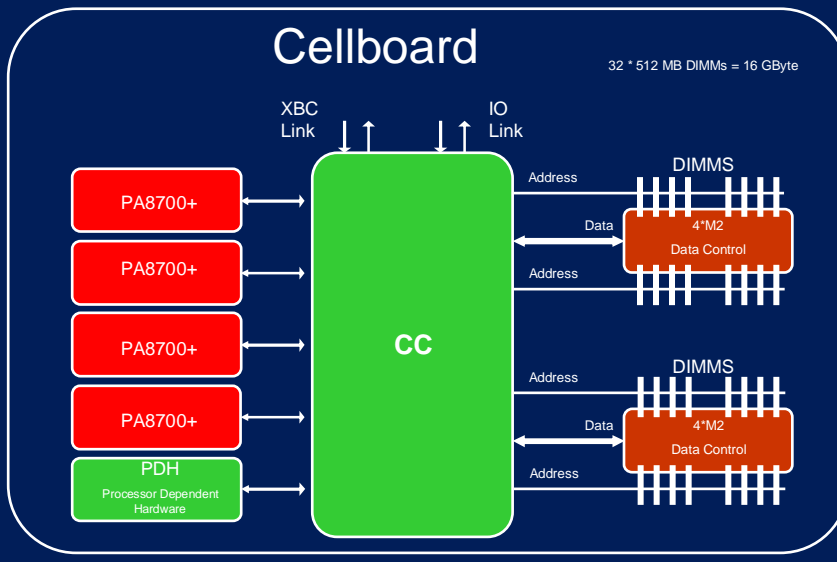
## Das Superdome Cellboard



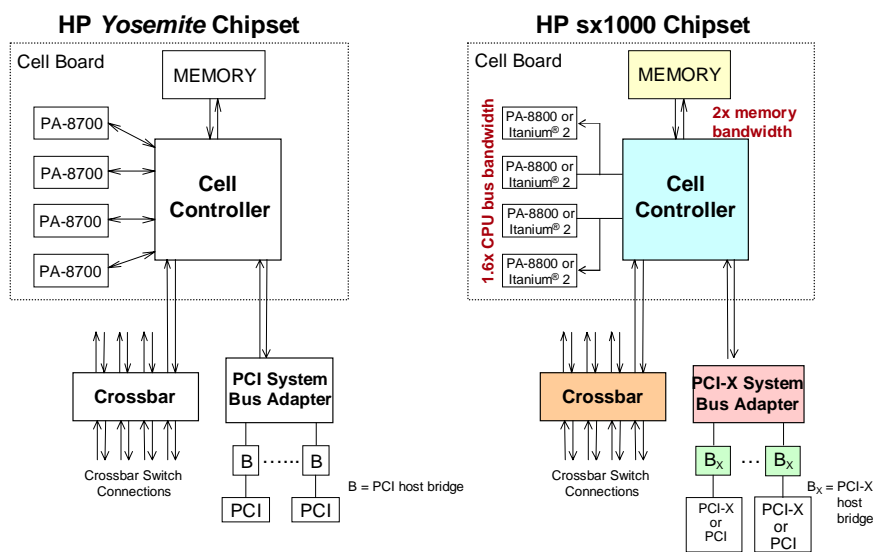
Das Cellboard besteht aus

- dem Cellcontroller (ASIC)
- Prozessoren
- Memory
- Verbindung zum IO System
- Verbindung zur Crossbar XBC

# Cellboard - Detail



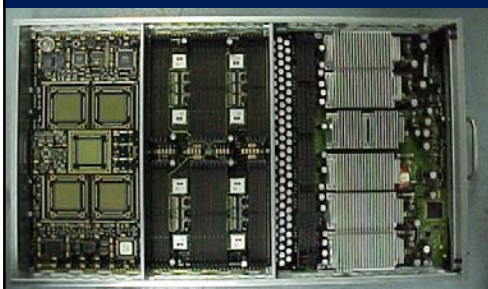
# sx1000 boosts memory & CPU bus bandwidth over Yosemite chipset







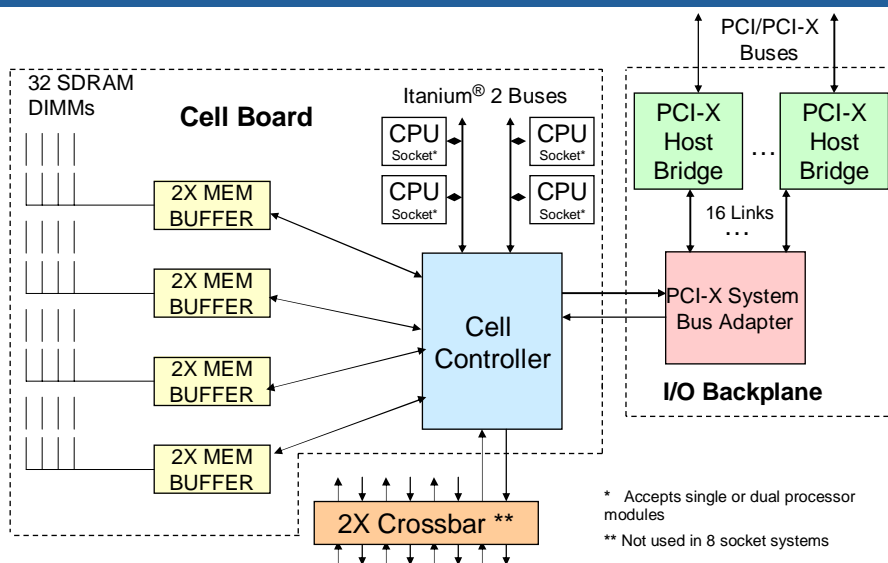
## Superdome PA-RISC cell board



## Itanium cell Board



## sx1000 chipset includes five distinct chips

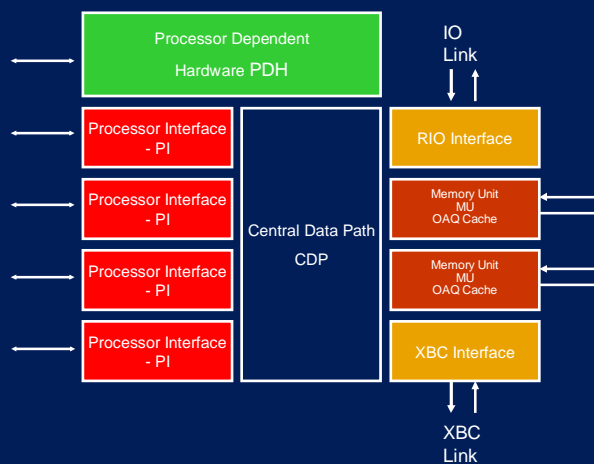


## Comparison: HP Yosemite and sx1000 chipsets



Feature	Yosemite	sx1000
<b>Processors</b>	PA-8600, PA-8700	<b>PA-8800, Itanium 2</b>
Core frequency	250 MHz	250 MHz
<b>Processor bus</b>	4 buses, 8 bytes, 250 MT/s	2 buses, 16 bytes, <b>400 MT/s</b>
<b>Physical address</b>	40 bits	<b>44 bits</b>
<b>CPUs per cell</b>	4	<b>4-8</b>
<b>Peak memory bandwidth per</b>	8 GB/s	<b>16 GB/s</b>
<b>I/O card support</b>	PCI	PCI and <b>PCI-X</b>
<b>High availability</b>	Hard partitions; ECC on memory, I/O, crossbar; memory chip-kill	Hard partitions; <b>ECC on CPU bus</b> , memory, I/O, and crossbar; memory chip-kill; <b>cell &amp; PCI online add/delete</b>

## Cellboard – Cellcontroller (CC)

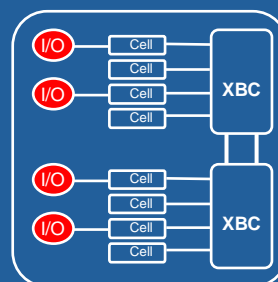


# Zellboard Chipsets und Prozessoren

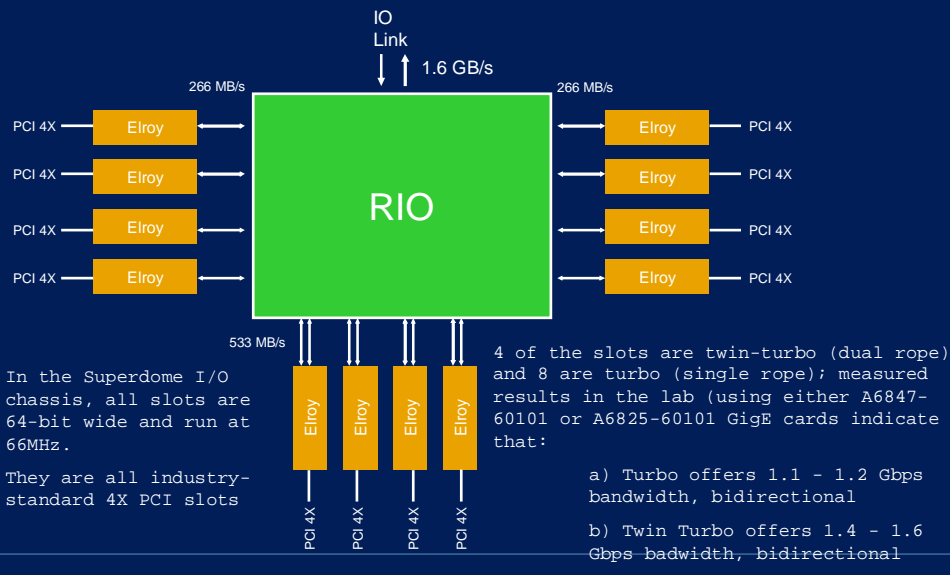


Name (ASIC)	Yosemite (Yosemite)	Sx10000 (Pinnacles)	madison	PA8800 - Mako	mx2 (sherpa)	hondo
Beschreibung	Today's chipset for PA8600, PA8700, PA8700+ Processors	A new chipset (Pinnacles) that supports the next generation Itanium® 2 and PA-RISC processors – Madison and PA-8800			two processor (madison) + 32 Mbyte L4 cache on a daughter board that fits into madison socket	
processor family	PA-Risc	IA64 + PA-Risc	IA64	PA-Risc	IA64	IA64
Betriebssystem	HP-UX	HP-UX, Windows Server 2003, Linux	HP-UX, Windows Server 2003, Linux	HP-UX	HP-UX, Windows Server 2003, Linux	HP-UX, Windows Server 2003, Linux

## Das Superdome IO-System



## 12 Slot PCI chassis



## hp Superdome



### Superdome SD16

- 4 Cell Boards
- 2-16 CPU
- PA8700/552,750,875
- 64 GB RAM
- 48 PCI I/O Slots

### Superdome SD32

- 8 Cell Boards
- 4-32 CPU
- PA8700/552,750,875
- 128 GB RAM
- 96 PCI I/O Slots

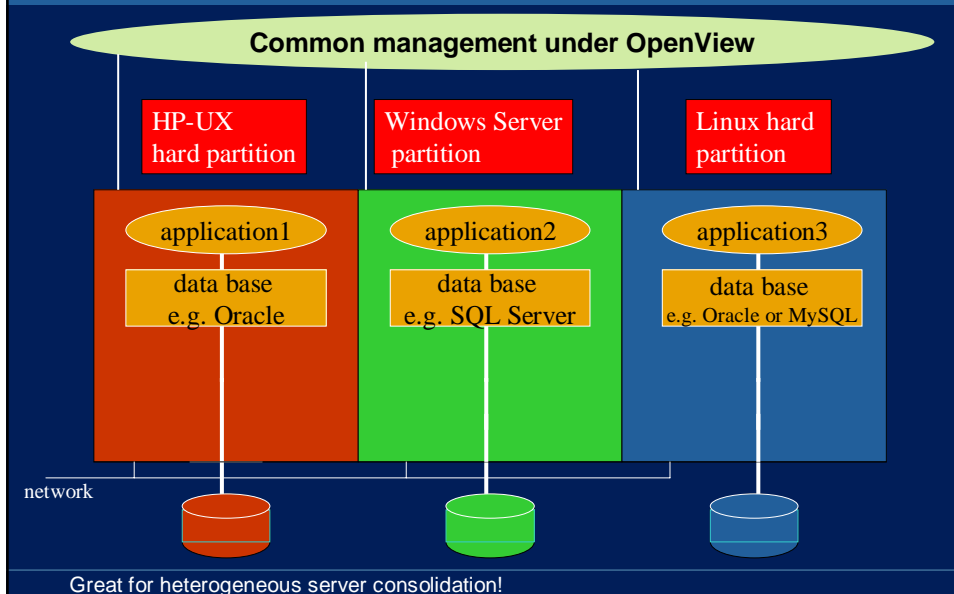
### Superdome SD64

- 16 Cell Boards
- 8-64 CPU
- PA8700/552,750,875
- 256 GB RAM
- 192 PCI I/O Slots

### Betriebssystem

- HP-UX 11i

## Example of multiple OSs running simultaneously in an Orca Server with Itanium2 processors



## hp Superdome



### HP-UX 11i v2

-V2: third release of HP-UX to support Itanium (does not support PA-RISC)

-Supports all OEs and adds 6 more dynamically tuneable kernel parameters

-Only supports Itanium CPUs

### Windows Server 2003

- Scales to 64 processors

-for breakthrough performance

-for enterprise use and enterprise applications

- 64 bit windows for increased performance and large memory capability

### Linux IA64

- Red Hat 3.0

- modified to support Pinnacles chip set

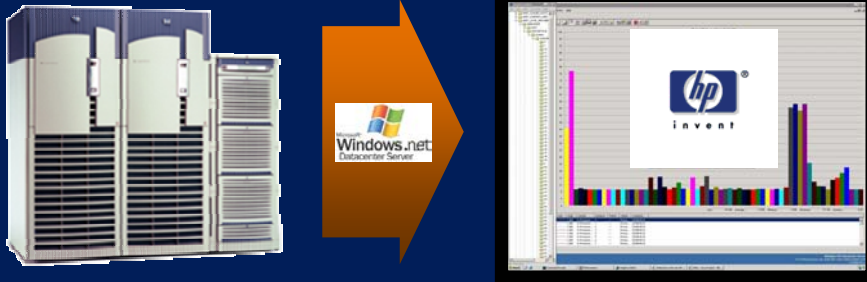
- scales to 4 processors/partition, targeted at technical, "network edge", and a wide range of commercial applications

- current plan:

-Developers kit in 2H03

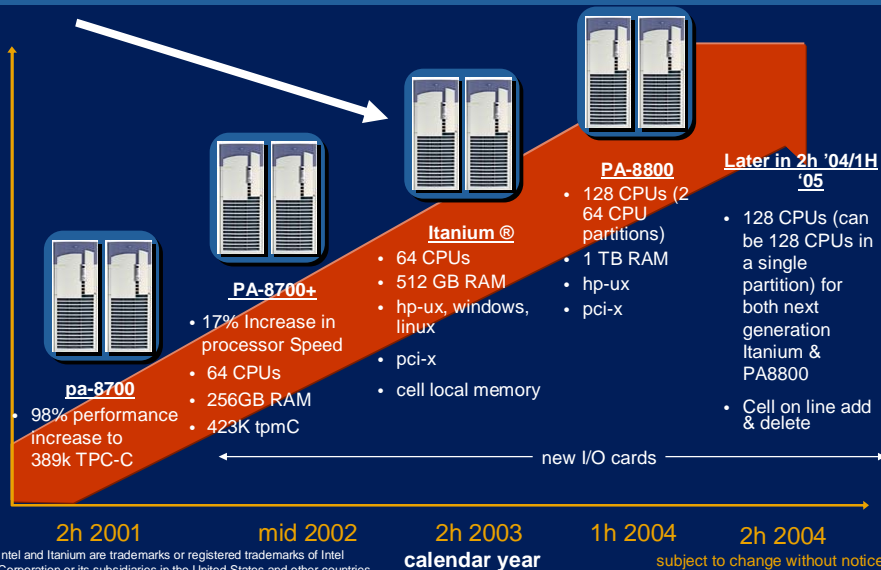
-Productized in '04

# World's first boot of 64-way Windows



HP is successfully running Windows on HP's Superdome server configured with 64 Itanium 2 processors and 512GB memory

# superdome: built for the future with investment protection today



Intel and Itanium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

## In-box upgrade path from Superdome to Orca



- Pull out existing cell board
- Swap existing memory into new Itanium cell board (protects investment in current memory)
- Plug in new Itanium cell board
- Update HP-UX and firmware
- Some I/O cards may need to be changed (windows does not support the identical set of I/O cards that HP-UX supports)
- NO change to backplane

## Orca summary



- Hot new Itanium2 (Madison) processor
- multi OS
- Double memory 256 GB => 512 GB
- Industry-unique in-box upgrade PA-RISC to Itanium2
- HP affirmation of Itanium by introducing Itanium in the premier high end server
- new Pinnacles chip set makes it all possible

all in-box upgrades!



Screaming performance!!



HP Confidential NDA required for customer presentation

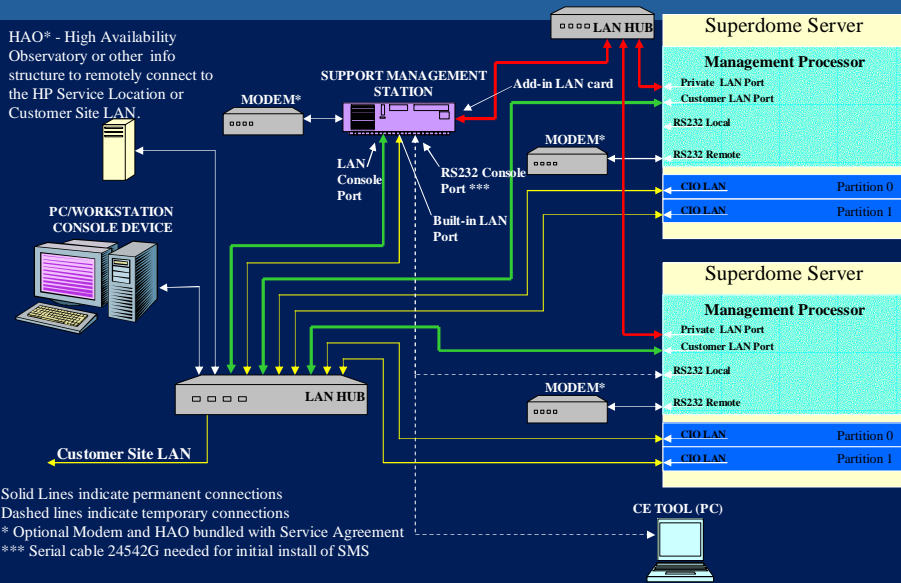
Presentation is HP restricted – NDA required

# Maintaining via Service Processor Console Architecture



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

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

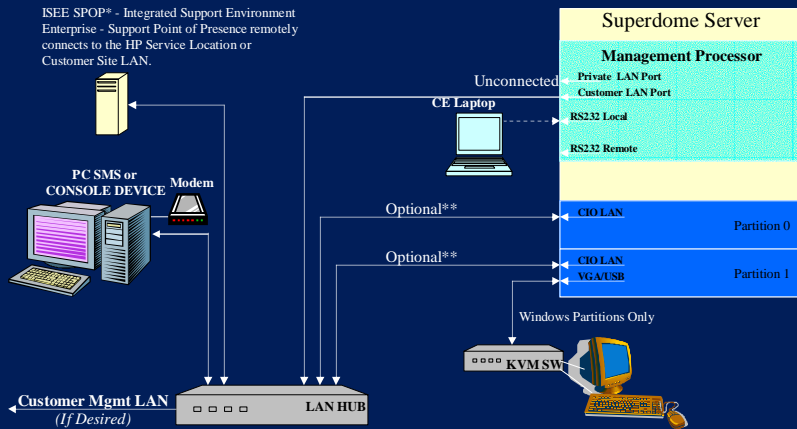




## Console and PC SMS Network Configuration Single IPF-Enabled Superdome



ISEE SPOP\* - Integrated Support Environment Enterprise - Support Point of Presence remotely connects to the HP Service Location or Customer Site LAN.



Solid Lines indicate permanent connections  
Dashed lines indicate temporary connections

\* Optional ISEE Support Point of Presence (SPOP) bundled with Service Agreement

\*\* Security concerns may dictate that a partition LAN not be connected to the management LAN. Access from a management station to a partition LAN through a secure router is an alternative.

\*\*\* Serial cable 24542G needed for initial install of SMS

## 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

