



IBM eServer® X3 Architecture: Third-Generation EXA

## IBM xSeries Technology

Scale-up / Scale-out  
*High-End offering*



## Executive Briefing

Yann de Visme – I/T Specialist – RHCE  
ydevisme@fr.ibm.com

© 2005 IBM Corporation

# Agenda

- ▶ IBM Offering - Scale-Up / Scale-Out positioning
- ▶ *Detailed answer to yesterday's question*
- ▶ IBM Initiative: EXA Architecture – Scale-up
- ▶ Products: x260 / x366 / x460
- ▶ X3 Chipset details
- ▶ Intel Xeon vs AMD Opteron: fundamental differences

# The Power of server



## xSeries

X-Architecture™  
Intel-based Servers

## pSeries

Performance UNIX®  
Servers

## iSeries

Integrated Solutions  
Servers

## zSeries

Zero Downtime  
Servers

Thriving in the unpredictable e-business market requires a robust and adaptable IT infrastructure, designed for your business.

No other company in the world offers the enterprise expertise applied to such a breadth of product as IBM eServer.


## IBM Offering - Scale-Up / Scale-Out positioning

# IBM eServer Industry-Standard Portfolio

## eServer xSeries

### Enterprise Scale Up Servers


2-way to 32-way SMP




**xSeries 460**

### Scale Out Rack Optimized Servers


Uni to 4-way SMP




**xSeries 336**



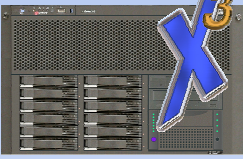
**xSeries 306**



**xSeries 346**




**xSeries 343  
NEBs Solution**




**xSeries 260**

### Distributed Tower Servers


Uni to 4-way SMP



**xSeries 206**



**xSeries 226**




**xSeries 236**


## eServer

### Enterprise Scale Out Offerings


Uni to 4-way nodes, scalable clustering



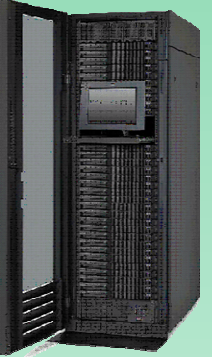
**eServer  
BladeCenter**



**eServer  
BladeCenter-Telco**

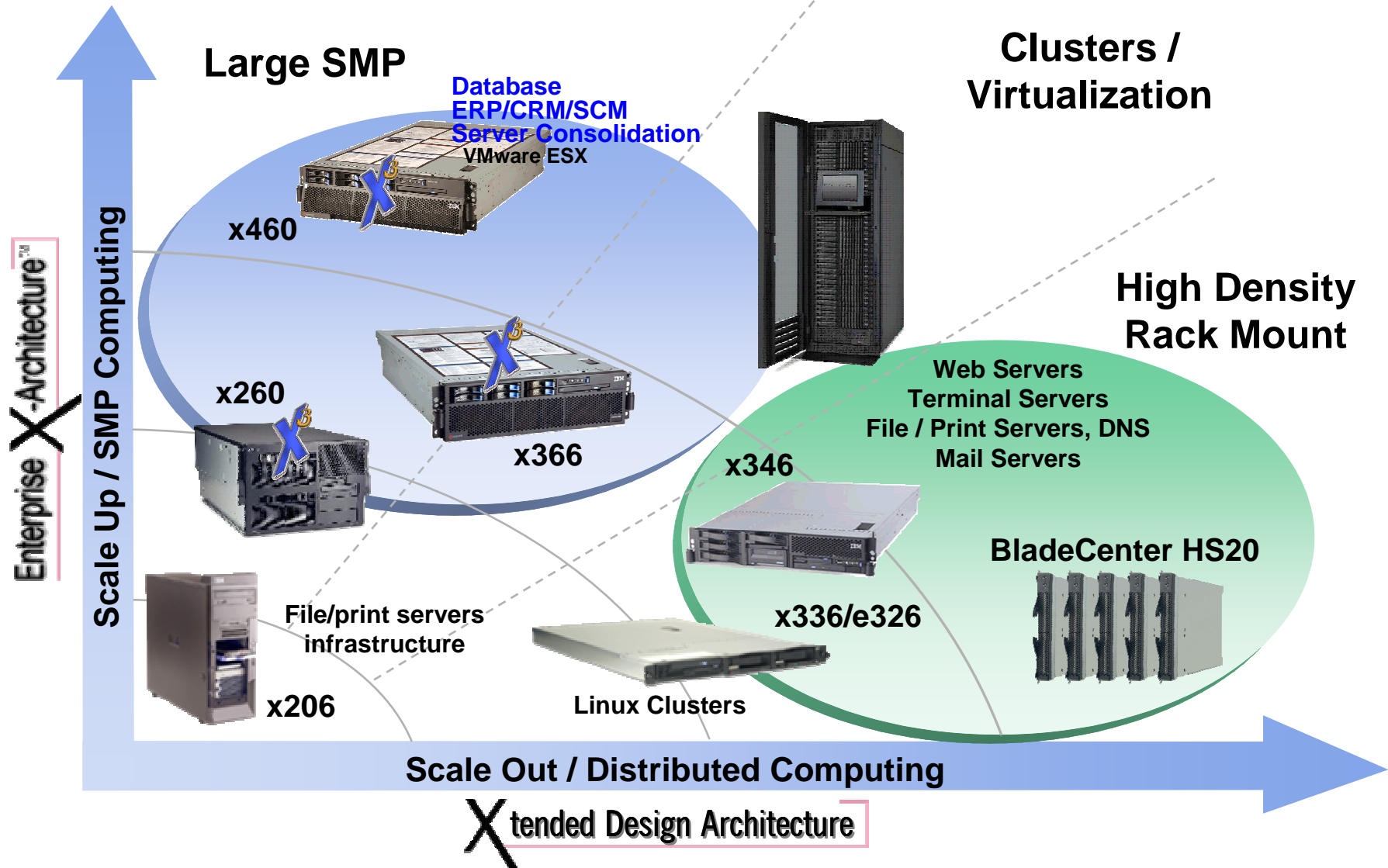


**eServer 326**



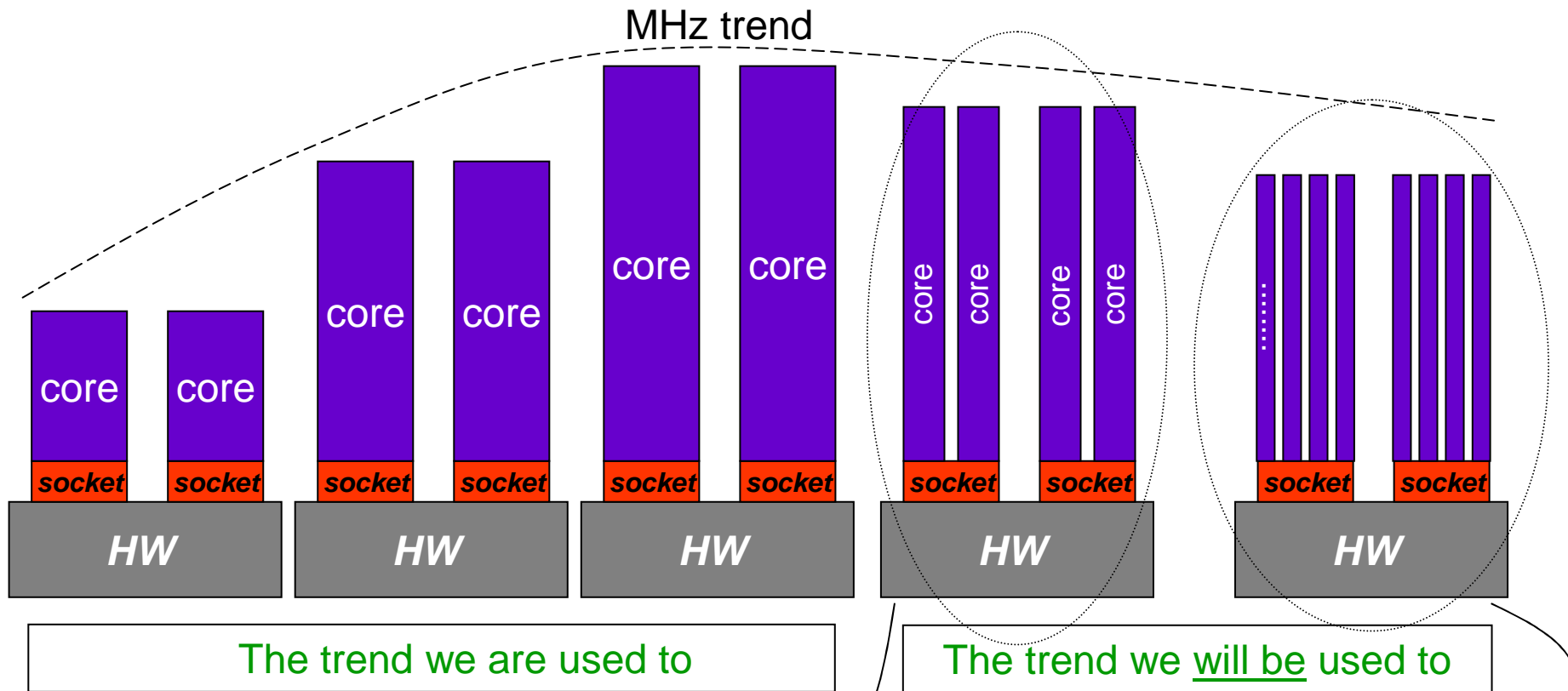
**IBM eServer Cluster 1350**

# Advancing X-Architecture



*Detailed answer to yesterday's question*

## Dual Core and Multi Core ?



(4-way) dual-core is what's happening today with:

- Paxville (soon Tulsa) from Intel
- Opteron from AMD

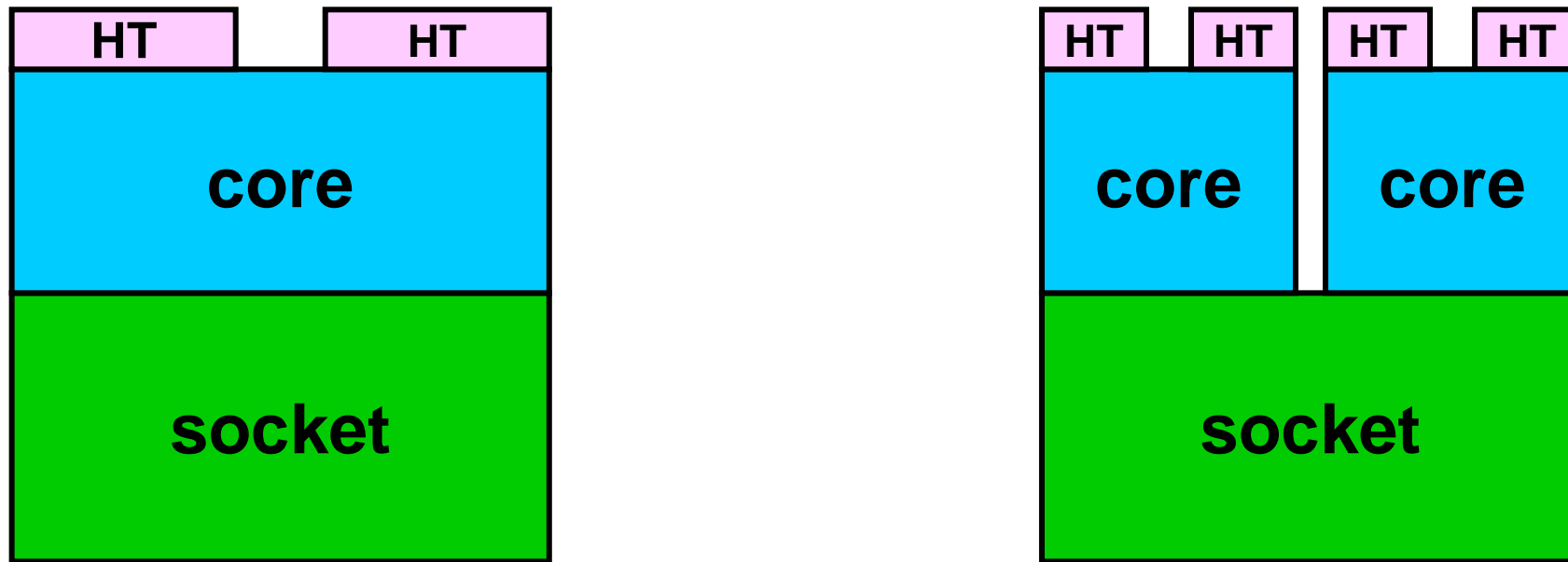
(4-way) quad-core is what will happen in 2007(?) with:

- Tigerton (used to be Whitefield) from Intel
- Opteron (same current micro-architecture) from AMD

**Power (and cooling) are becoming THE challenge and priority #1 from now on !**



## Dual core and HyperThreading ? .....



- HT builds on top of the core whether it is single or dual !

- HT and dual core has no correlation with each other .... (for example Opterons are dual core but do not have any HT concept/technology)

- Two HT logical CPU's share much more things of a core than what two cores share on a socket

	Architectural state	Execution resources	On-board caches	System Bus Interface
"Legacy" CPU	exclusive	exclusive	exclusive	exclusive
"HyperThreaded" CPU	exclusive	shared	shared	shared
Dual/Multi Core CPU	exclusive	exclusive	exclusive/shared	shared

## IBM Initiative: EXA Architecture – Scale-up

# IBM defines High-end Industry-Standard Servers

## 1st Generation: 2001

- x360: 6-month time to market advantage, Most rack dense 4w (3U) ever introduced
- x440: 12-month TTM, Most rack dense 8w (4U), Most successfully benchmarked server in history (35 #1's)
- XpandOnDemand Scalability up to 16-way plus Remote I/O
- Industry-first High Availability Technologies: Active Memory & Memory ProteXion
- Leadership Virtualization for Server Consolidation

## 2nd Generation: 2003

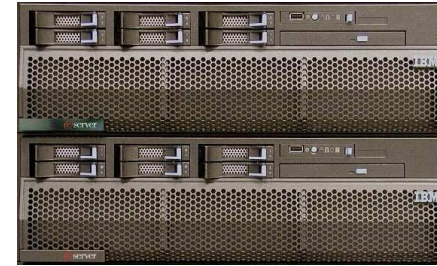
- x365: Leadership density (3U) with 4X storage capacity & advanced EXA features
- x445: the fastest industry-standard server in history, 20 more #1 benchmarks (little competition to compare)
- x455: Unleashing EXA on Itanium2 for pure 64-bit
- XpandOnDemand Scalability up to 32-way plus Remote I/O
- 10 Consecutive Quarters (3Q02) as #1 8-way database server in the Industry

## 3rd Generation: 2005

- x366: Leadership 4-socket performance, First-to-market with 64-bit Xeon MP
- x460: xSeries 32-way flagship optimized for scalability & virtualization with 100%+ higher performance
- x260: Extending EXA to the 4-way Tower space with maximum storage for SMB
- 64-bit Extensions for higher performance, application flexibility (32-bit & 64-bit) and investment protection

Products: x260 / x366 / x460

# High-Performance xSeries Positioning



## x260: 4-way Tower

Target: SMB, Remote/  
Branch Office

**Apps: SMB, Collaboration,  
Departmental database**

eServer X3: 3rd Gen EXA  
64-bit Intel Xeon MP  
Competitive Price-perf  
Latest technology  
SAS, PCI-X2, DDR2  
High Availability  
7U Tower or Rack  
Maximum Int. Storage  
Tape Backup Support  
Dual-core Capable

## x366: 4-way Rack

Target: Enterprise & Mid-  
market

**Apps: Collaboration,  
Database, ERP, SCON**

eServer X3: 3rd Gen EXA  
64-bit Intel Xeon MP  
Leadership 4-way perf  
**38% better than x365**  
Latest technology  
SAS, PCI-X2, DDR2  
High Availability  
3U Rack-optimized  
Maximum Int. Storage  
Dual-core Capable

## x460: 4-way+ Rack

Target: Enterprise & Mid-market  
**Database**

**MS SQL, DB2, Oracle  
ERP/CRM/SCM  
SAP, Siebel, i2**

**Server Consolidation**

**VMware ESX Server**  
X3: 3rd Gen EXA  
64-bit Intel Xeon MP  
Leadership 8-way perf  
**50% better than x445**  
Leadership 16-way+ perf  
**80% better than x445**  
Latest technology  
SAS, PCI-X2, DDR2  
High Availability  
XpandOnDemand to 32way  
Dual-core Capable



x255



x260

- ServerWorks GC-HE Chipset
- 1-way to 4-way, 32-bit with 400MHz FSB
- Intel Xeon MP: 2.0/1M, 2.2/2M, 2.7/2M, 3.0/4M
- 12 DIMMs Total: All standard
- 24GB Max Memory (12 x 2GB DIMM)
- PC1600 DDR SDRAM, 2-way Interleaving
- Ultra160 SCSI, Optional RAID
- Max Storage = 12 HDDs x 146 GB = 1.7TB
- Active PCI-X: 6 slots @ 100MHz
- Broadcom 5703 Single Port Gigabit Ethernet
- Chipkill + Active Memory: Online Spare
- 4 x 370W Hot-swap Power Supplies, N+N, 110V/220V
- Remote Supervisor Adapter 2 Optional
- 3-year Next Business Day 9x5 Warranty
- 7U Rack or Tower with Internal Tape Support

- **XA-64e 3rd Generation Chipset**
- **1-way to 4-way SMP, Dual-core Capable**
- **Intel Xeon MP: 3.16/1M, 3.66/1M, 667MHz FSB**
- **16 DIMMs Total: 4 Standard, 12 Optional**
- **64GB Max Memory (16 x 4GB DIMM)**
- **DDR2 SDRAM PC2-3200, 2-way Interleaving**
- **Adaptec Serial Attached SCSI (SAS), Int. RAID5**
- **Max Storage = 12 3.5" HDDs x 300 GB = 3.6TB**
- **Active PCI-X 2.0: 6 slots @ 266MHz**
- **Broadcom 5704 Dual Port Gigabit Ethernet**
- **Chipkill + Memory ProteXion + Memory Mirroring**
- **4 x 775W Hot-swap Power Supplies, N+N, 110V/220V**
- **Remote Supervisor Adapter 2 Slimline optional**
- **3-year Next Business Day 9x5 Warranty**
- **7U Rack or Tower with Internal Tape Support**



## x365

- XA-32 2nd Generation Chipset
- 1-way to 4-way, 32-bit with 400 MHz FSB
- Intel Xeon MP: 2.0/1M, 2.2/2M, 2.7/2M, 3.0/4M
- 16 DIMMs Total: 8 Standard, 8 Optional
- 32GB Max Memory (16 x 2GB DIMM)
- PC2100 DDR SDRAM, 2-way Interleaving
- LSI 53C1030 Ultra320 SCSI, Integrated RAID-1
- Max Storage = 6 HDDs x 146 GB = 876 GB
- Active PCI-X: 4@133MHz, 1@100MHz, 1@33MHz
- Remote I/O + RIO Sharing between x365's
- 24X CD-ROM
- Broadcom 5704 Dual Port Gigabit Ethernet
- Chipkill + Memory ProteXion + Memory Mirroring
- 2 x 950W Hot-swap Power Supplies, N+N, 110V/220V
- Remote Supervisor Adapter 2 Standard
- 1-year or 3-year Next Business Day 9x5 Warranty
- 3U: 17.46"(444mm) x 5.07"(129mm) x 28.1"(715mm)

## x366



- XA-64e 3rd Generation Chipset
- 1-way to 4-way SMP, Dual-core Capable
- Dual-bus x86-64 Architecture, 667MHz FSB
- Intel Xeon MP: 3.16/1M, 3.66/1M
- 16 DIMMs Total: 4 Standard, 12 Optional
- 64GB Max Memory (16 x 4GB DIMM)
- DDR2 SDRAM PC2-3200, 2-way Interleaving
- Adaptec Serial Attached SCSI (SAS), opt. RAID5
- Max Storage = 6 2.5" HDDs x 73 GB = 438 GB
- Active PCI-X 2.0: 6 slots @ 266MHz, No Remote I/O
- 8X DVD-ROM
- Broadcom 5704 Dual Port Gigabit Ethernet
- Chipkill + Memory ProteXion + Memory Mirroring
- 2 x 1300W Hot-swap Power Supplies, N+N, 220V
- Remote Supervisor Adapter 2 Slimline opt.
- 3-year Next Business Day 9x5 Warranty
- 3U: 17.46"(444mm) x 5.07"(129mm) x 28.1"(715mm)

## x445



- XA-32 2nd Generation Chipset
- 2-way to 32-way SMP, 32-bit with 400MHz FSB
- Intel Xeon MP: 2.0/1M, 2.2/2M, 2.7/2M, 3.0/4M
- Intel Xeon DP 3.0GHz up to 4-way
- 64MB XcelL4 per CEC, 512MB Max
- 64GB max addressable memory supported
- DDR SDRAM PC2100, 2-way Interleaving
- LSI Ultra320 SCSI, Integrated RAID-1
- Max Storage = 2 HDDs x 146 GB = 292 GB
- Active PCI-X: 2@133MHz, 2@100MHz, 2@66MHz
- Remote I/O + RIO Sharing
- 2 x 1200 Watt Power Supplies, Hot-swappable
- Broadcom 5704 dual port GbE
- Remote Supervisor Adapter II for EXA (std)
- Active Memory + Hot-swap and Hot-add Memory
- 3-year Next Business Day 9x5 Warranty
- 4U: 19"(483mm) x 7"(178mm) x 28.1"(714mm)
- 20 #1 Benchmarks...and counting!!!

## x460



- XA-64e 3rd Generation chipset
- 2-way to 32-way SMP, Dual-core Capable
- Dual-bus x86-64 Architecture, 667MHz FSB
- Intel Xeon MP: 2.83/4M, 3.16/8M, 3.33/8M
- 256MB XcelL4v per chassis, 2GB Max
- 64GB Max Memory per chassis, 512GB Max Total
- DDR2 SDRAM PC2-3200, 2-way Interleaving
- Adaptec Serial Attached SCSI (SAS), Opt. RAID5i
- Max Storage = 6 2.5" HDDs x 73 GB = 438 GB
- Active PCI-X 2.0: 6 available slots, all 266MHz
- MXE-460 Modular Xpansion Enclosure (>4-way)
- 2 x 1300 Watt Power Supplies, Hot-swappable
- Broadcom 5704 dual port GbE
- Remote Supervisor Adapter II Slimline (std)
- Active Memory + Hot-swap (All DIMMs accessible)
- 3-year Next Business Day 9x5 Warranty
- 3U: 19"(483mm) x 5.25"(133mm) x 27.5"(698mm)
- Up to 125% performance improvement over x445



# x460 & MXE Supported Configurations

## *XpandOnDemand™ Scalability*

Modular Building-block Scalability eliminates the need for fork-life upgrades and provides an easier growth path to larger, scale-up high-performance SMP configurations

Perfect for:



x460 2w-4w Single Chassis  
Up to 64GB Memory



x460 + (1) MXE-460  
Two Chassis 8-way  
Up to 128GB Memory



x460 + (3) MXE-460  
Four Chassis 16-way  
Up to 256GB Memory

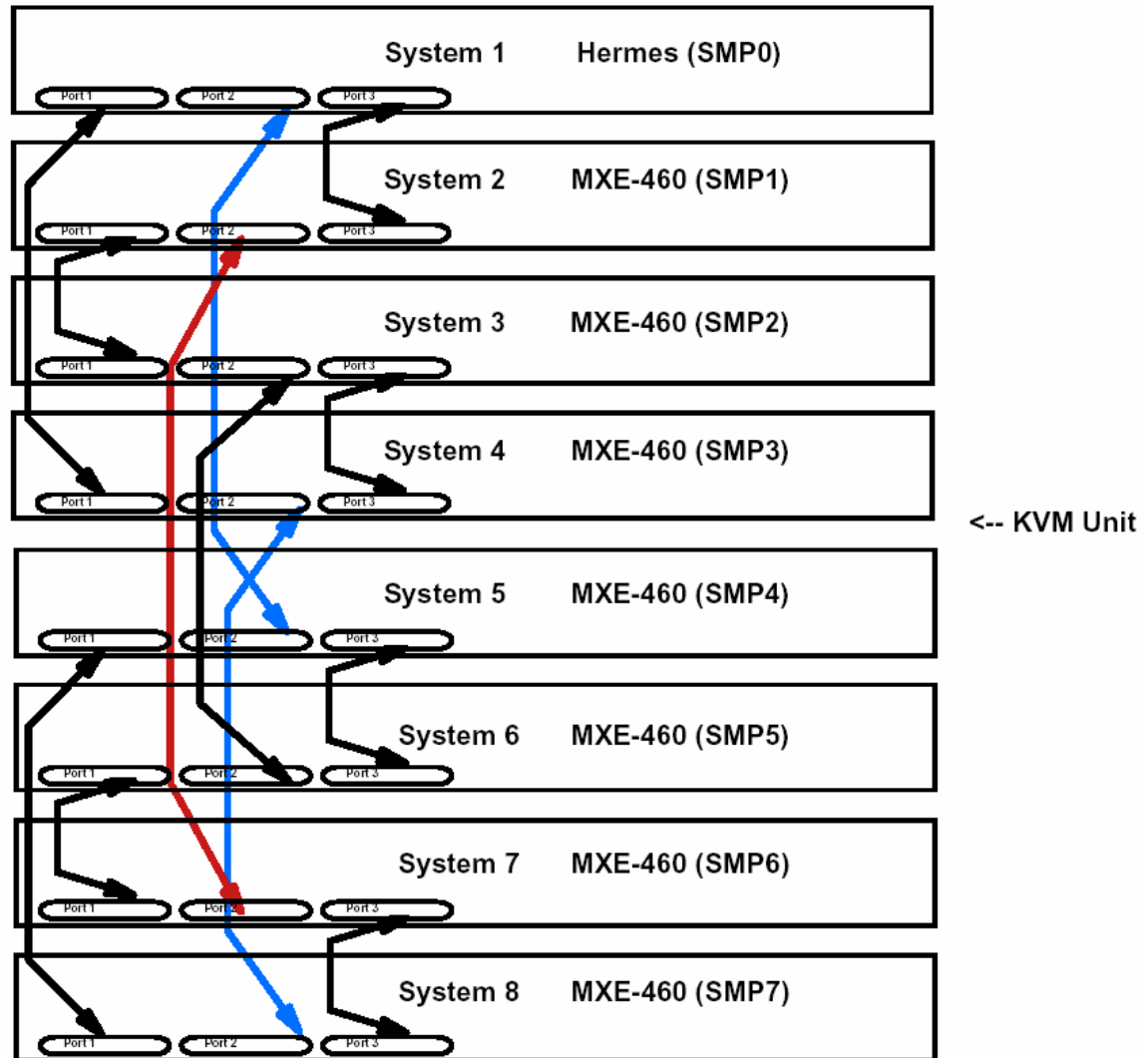


x460 + (7) MXE-460  
Eight Chassis 32-way  
Up to 512GB Memory



# x460 32-way Cabling Diagram

- All 8-way and 16-way Configurations only require the 2.3m cable
- Only the 32-way requires the 2.9m cable
  - ▶ Red & blue cables: 2.9m
  - ▶ Black cables: 2.3m
- This cable diagram applies to both MXE and partitioning with the x460



## IBM X3 Chipset details

# eServer X3: Third-generation Enterprise X-Architecture

Mainframe-inspired innovation that delivers break-through performance, mission-critical availability, and unmatched modular scalability to become *the leading 64-bit solution architecture for commercial enterprise applications*, virtualization, and web services.

## Performance

- #1 x86 4w, 8w, 16w, 32w Performance
- 32-bit/64-bit x86 compatibility
- Reduced latencies of 3G Chipset
- Xcel4v™ Server Accelerator Cache
- PCI-X2, SAS, DDR2 Memory
- Optimized for Windows & Linux and the **application-serving tier**



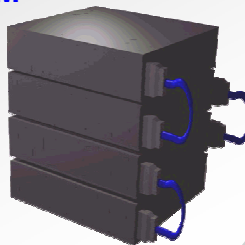
## High Availability

- 64GB Active Memory
- OS-independent Mirroring
- Chipkill & Mem ProteXion
- Hot-swap & Hot-add in all major subsystems
- Reliability of Intel Xeon MP **Front-side bus architecture**



## Scalability

- Improved pay-as-you-grow with more granularity in CPU, I/O, RAM
- 2-32-sockets, Up to 512GB Memory, Dual-core Capable
- Flexibility with MXE scalability or x460 partitioning
- Optimized for Windows & Linux and the **database-serving tier**



## Manageability

- Integrated hardware & remote mgmt software
- Integrated hardware-based security (TPM)
- Comprehensive alerting with PFA and Light Path Diagnostics
- **Multi-chassis partitioning**



# Enterprise X-Architecture Chipset Changes

## XA-32 1st generation

2002: x360, x440

Memory Controller:

- ▶ Cyclone Jr. 2.2
- Memory Mirroring

Processor Controller:

- ▶ Twister 2.1.1
- 32MB XceL4

PCI Bridge I/O Controller:

- ▶ Winnipeg 4 Ceramic
- Remote I/O Support

## XA-32 2nd generation

2003: x365, x445

Memory Controller:

- ▶ Cyclone 3
- Hot-swap memory
- 40% more aggregate I/O

Processor Controller:

- ▶ Twister 3
- 12% lower latency
- 64MB XceL4

PCI Bridge I/O Controller:

- ▶ Winnipeg 4 Plastic
- \$50 cost savings

## XA-64e 3rd generation

2005: x366, x460, x260

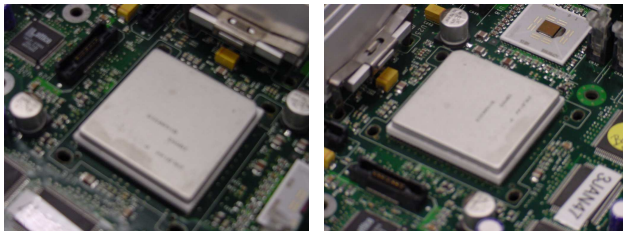
Integrated Processor & Memory Controller:

- ▶ Hurricane
- Dual-bus architecture
- Support for EM64T
  - x86 64-bit Extensions
- 256MB XceL4v Cache
  - Virtual L4 Cache

PCI Bridge I/O Controller:

- ▶ Calgary
- PCI-X 2.0 Support
- All slots 266MHz

–Used by x, i, p, zSeries

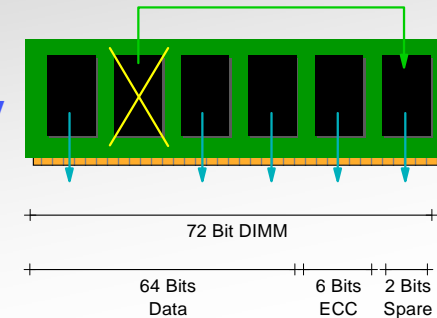


Normalized for CPUs, the EXA 3G chipset represents a 25% to 40% performance improvement over 2nd Generation

## X3: Leadership Technology – Memory

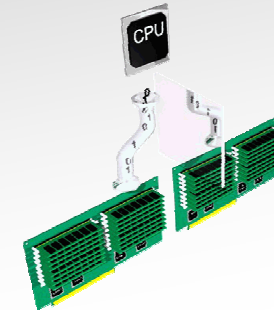
### Memory ProteXion™ - Redundant Bit Steering

- ▶ Redundant Bit Steering is similar to the "hot-spare" of a DASD Array
- ▶ Utilizes unused bits in each memory DIMM (hot spare bits)
- ▶ Double the number of Chipkills sustainable per server
- ▶ Included at no additional cost, requires no additional hardware, and works independently of operating system



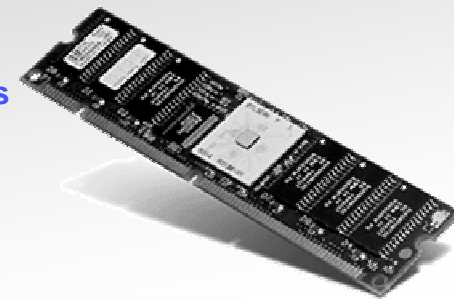
### Memory Mirroring

- ▶ Propels Intel-based servers towards continuous operations
- ▶ Dramatically increases uptime and allows scheduled maintenance
- ▶ Mainframe capability and reliability
- ▶ Operating System Independent: Does not require drivers or OS Support
- ▶ First introduced on the x440 and x445

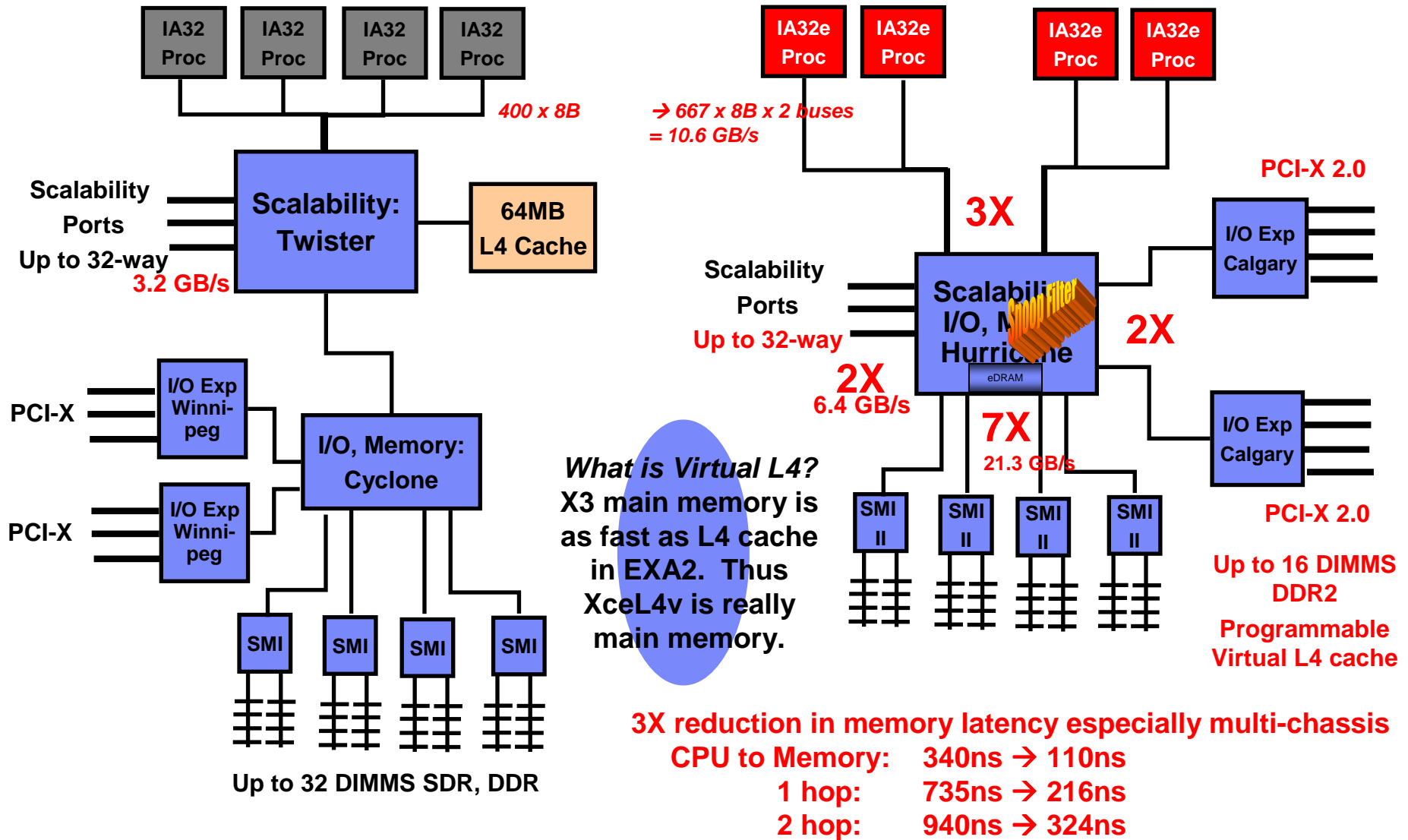


### Chipkill™ Memory

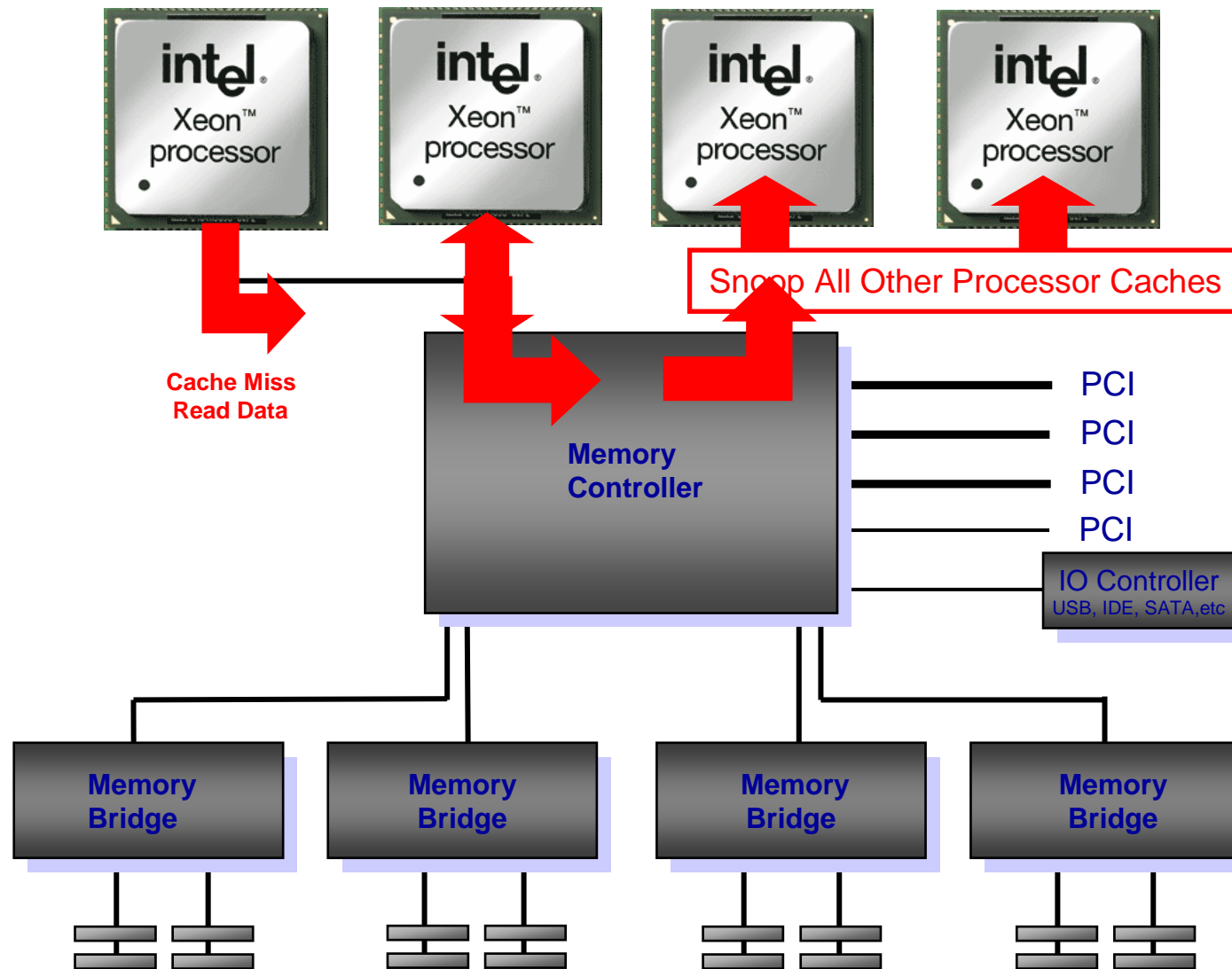
- ▶ Integrated into XA-32 and XA-64 chipsets for using off-the-shelf DIMMs
- ▶ Better memory reliability to support In-Memory Databases
- ▶ Chipkill Memory enables increased availability by detecting and correcting multiple-bit memory DIMM errors
- ▶ Third-Generation Chipkill design (1: 7000 M10, 2: 7600/6000)



# EXA2 vs. EXA3: Fatter pipes & lower latencies

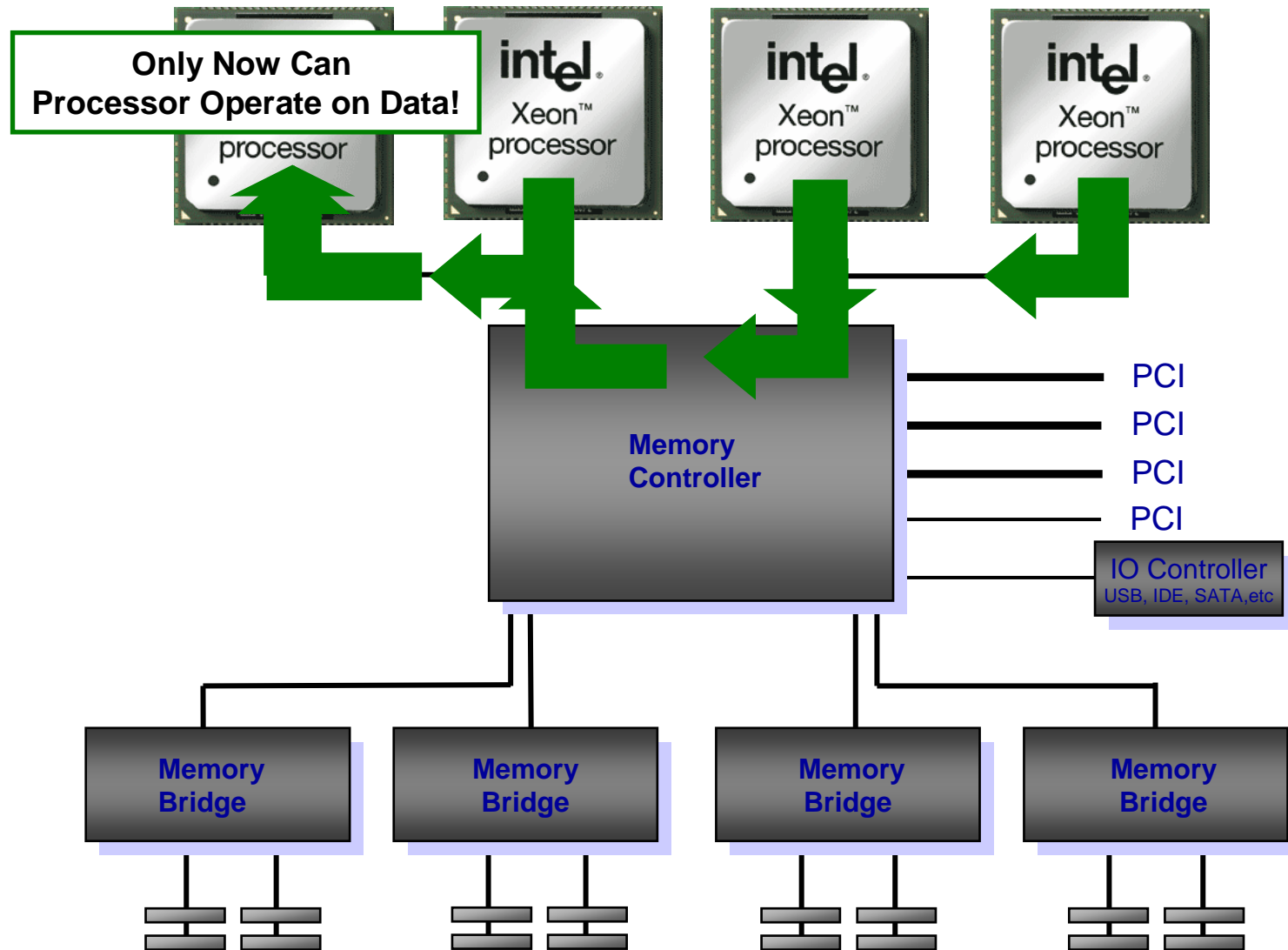


# Xeon Coherency Protocol – CPU Snoop Request

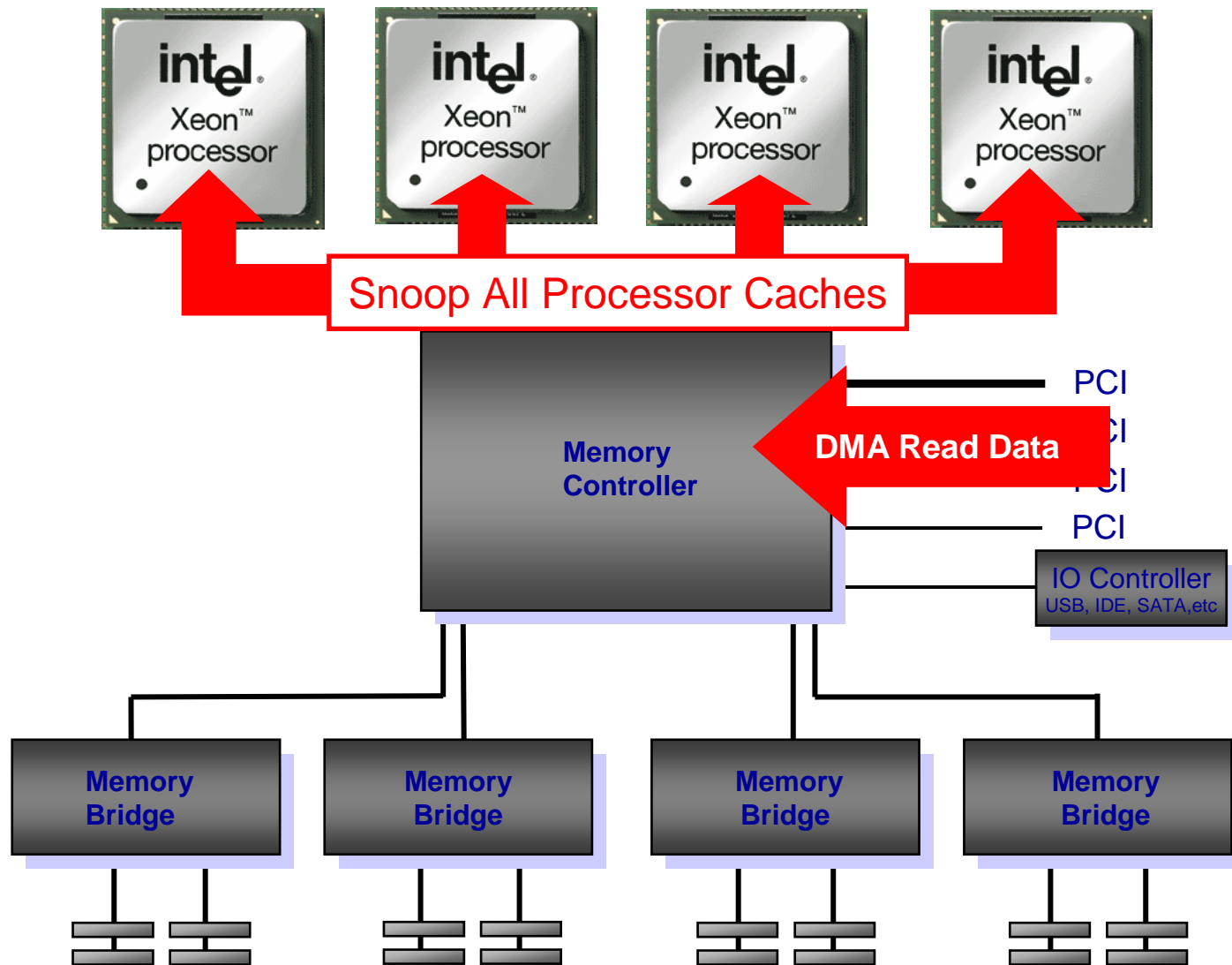




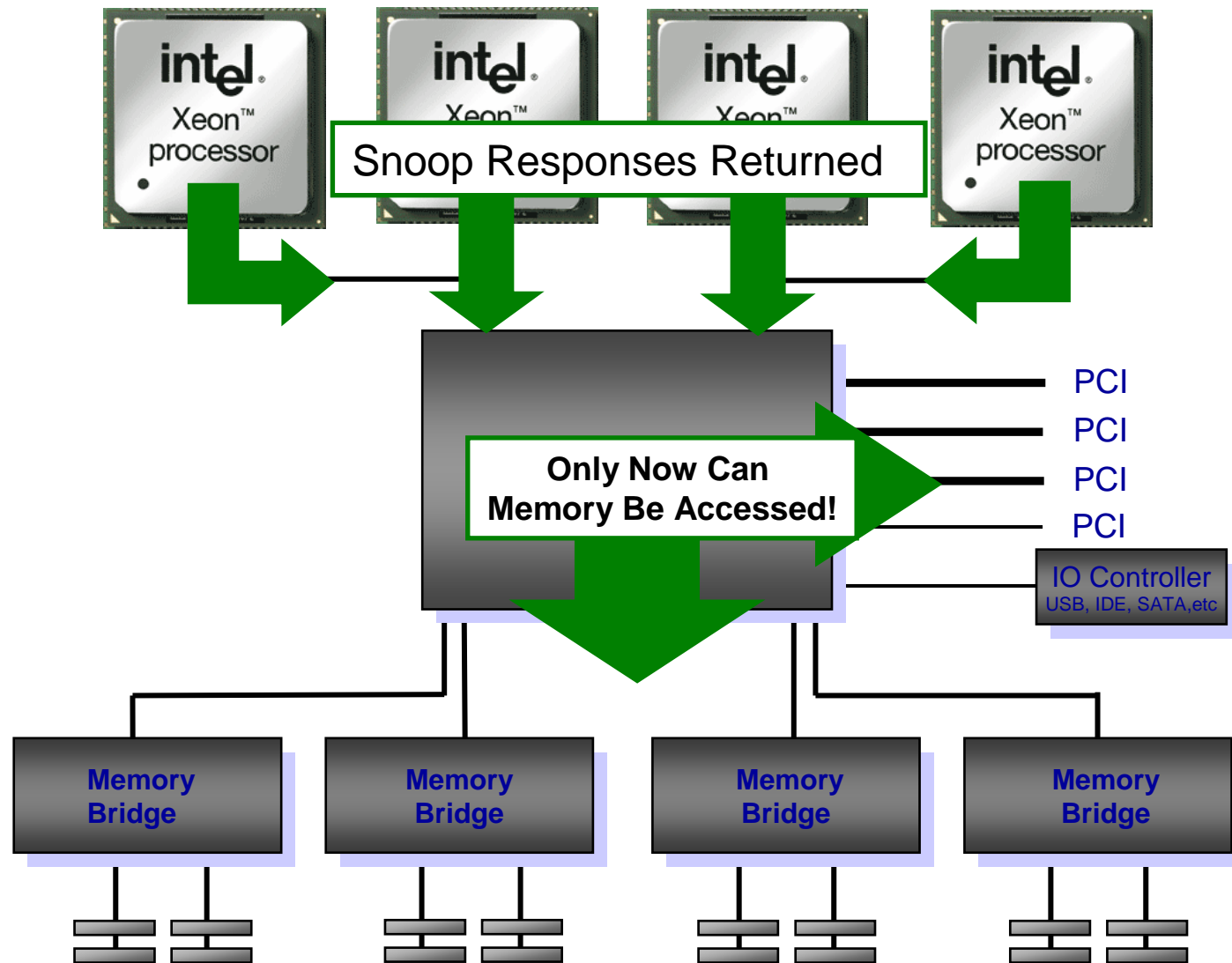
# Xeon Coherency Protocol – CPU Snoop Response



# Xeon Coherency Protocol – DMA Snoop Request

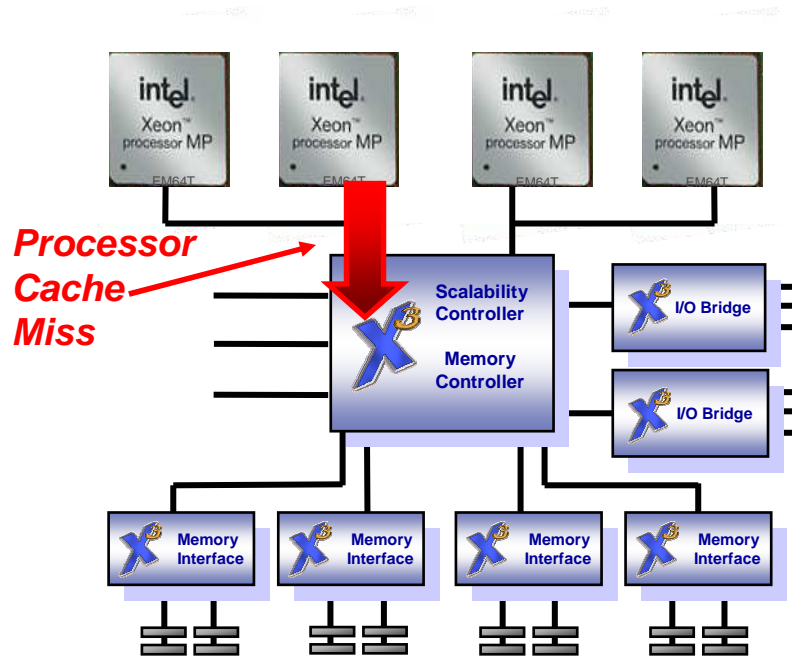


# Xeon Coherency Protocol – DMA Snoop Response



# X Chipset – Snoop Filter

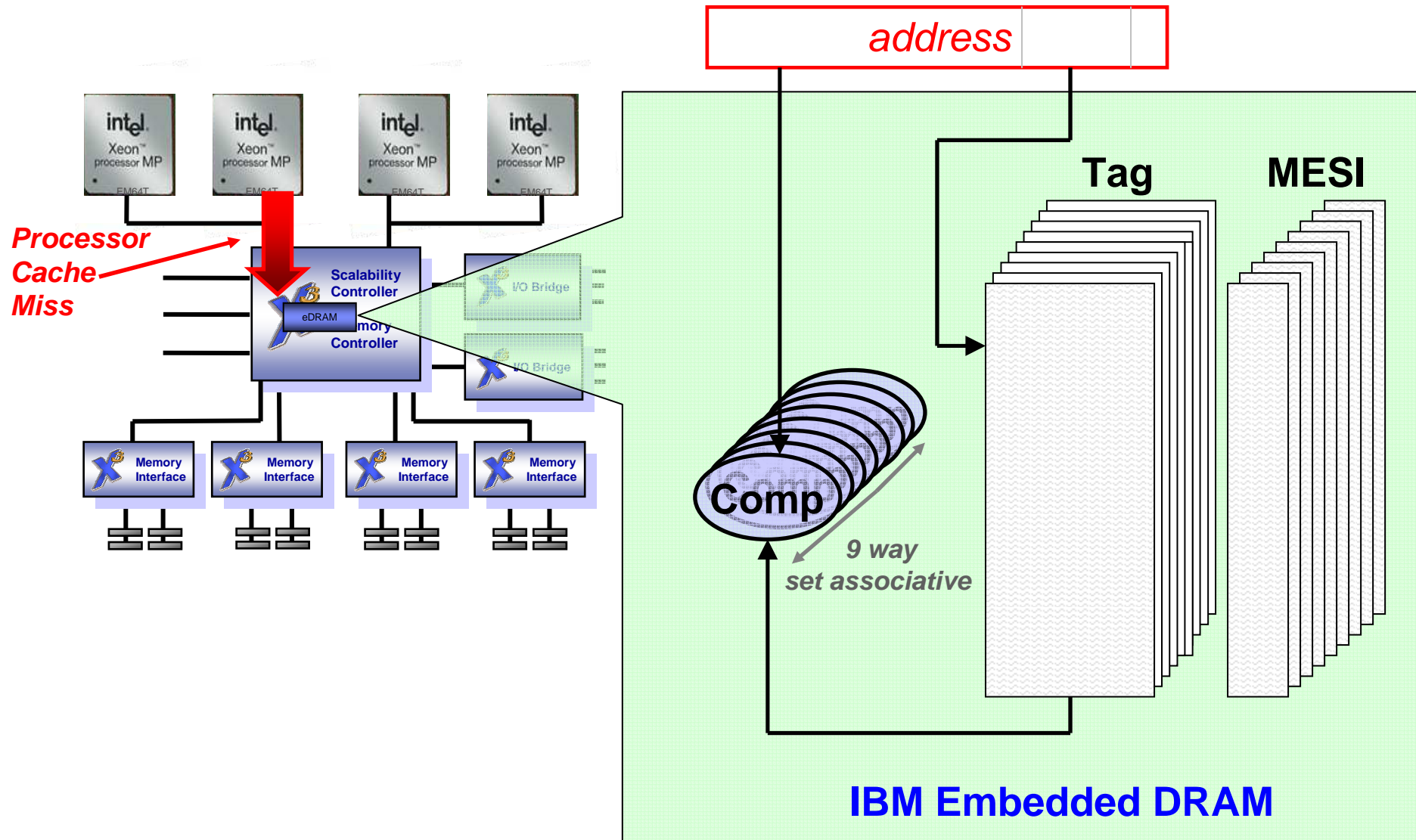
## Cache Miss – Snoop Filter Miss Example



# SNOOP FILTER

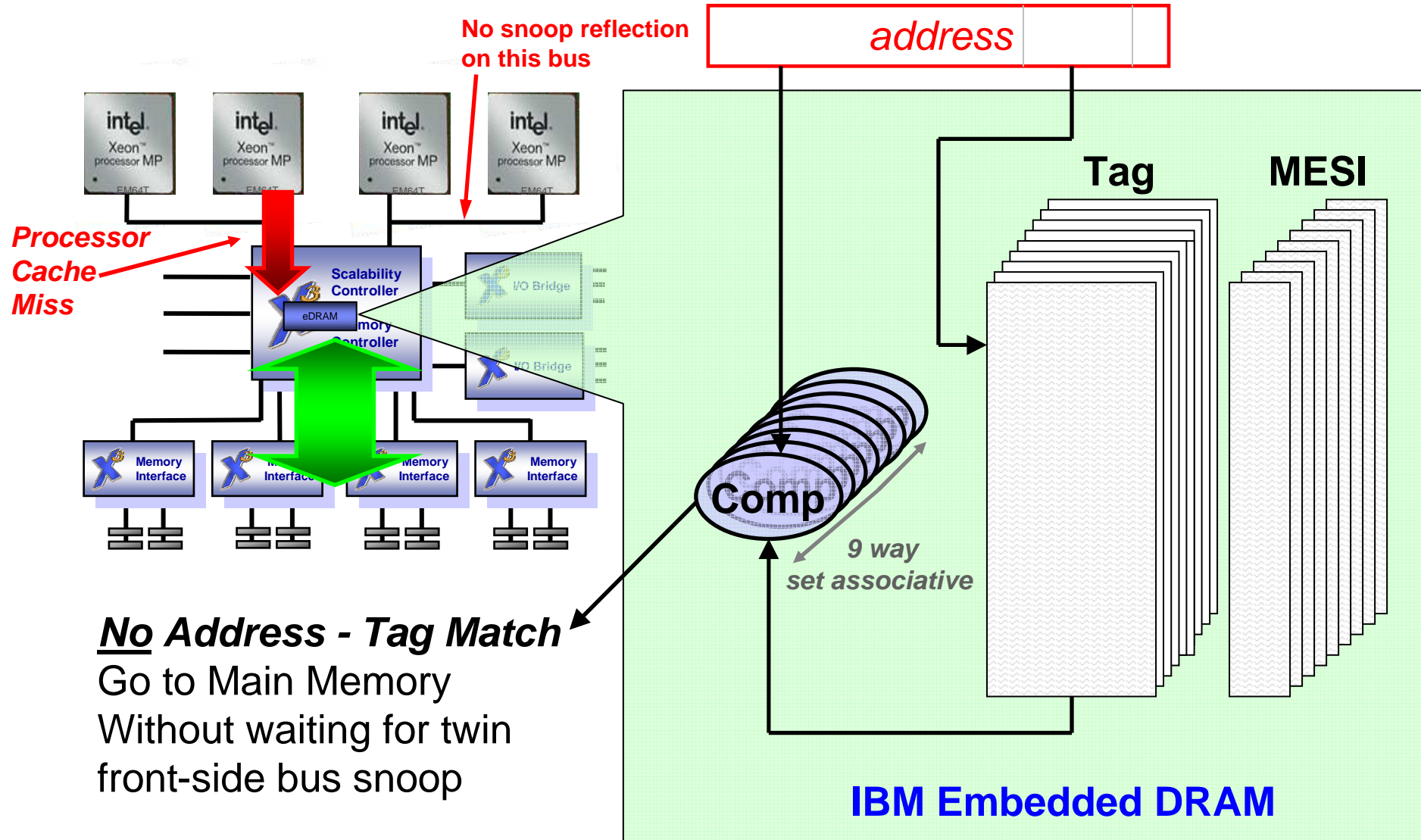
# X Chipset – Snoop Filter

## Cache Miss – Snoop Filter Miss Example

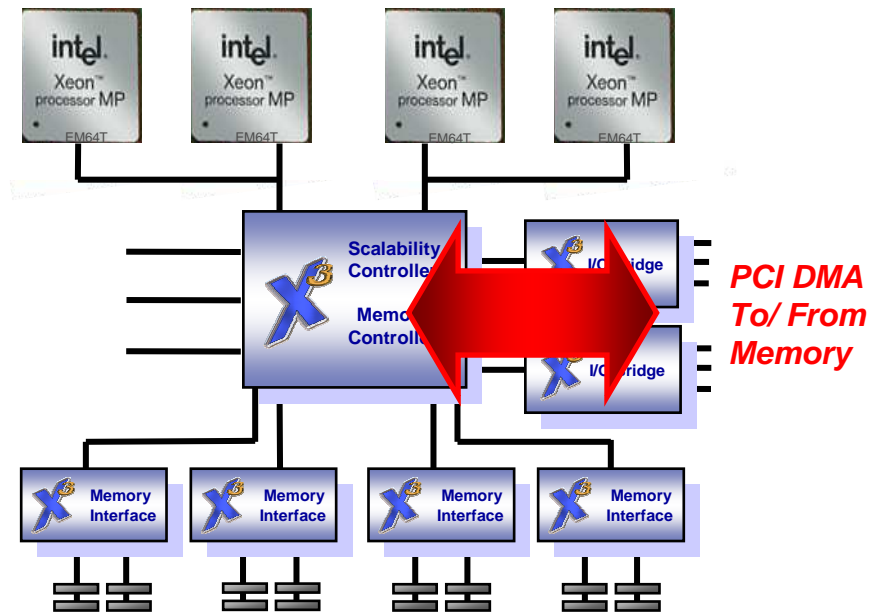


# X Chipset – Snoop Filter

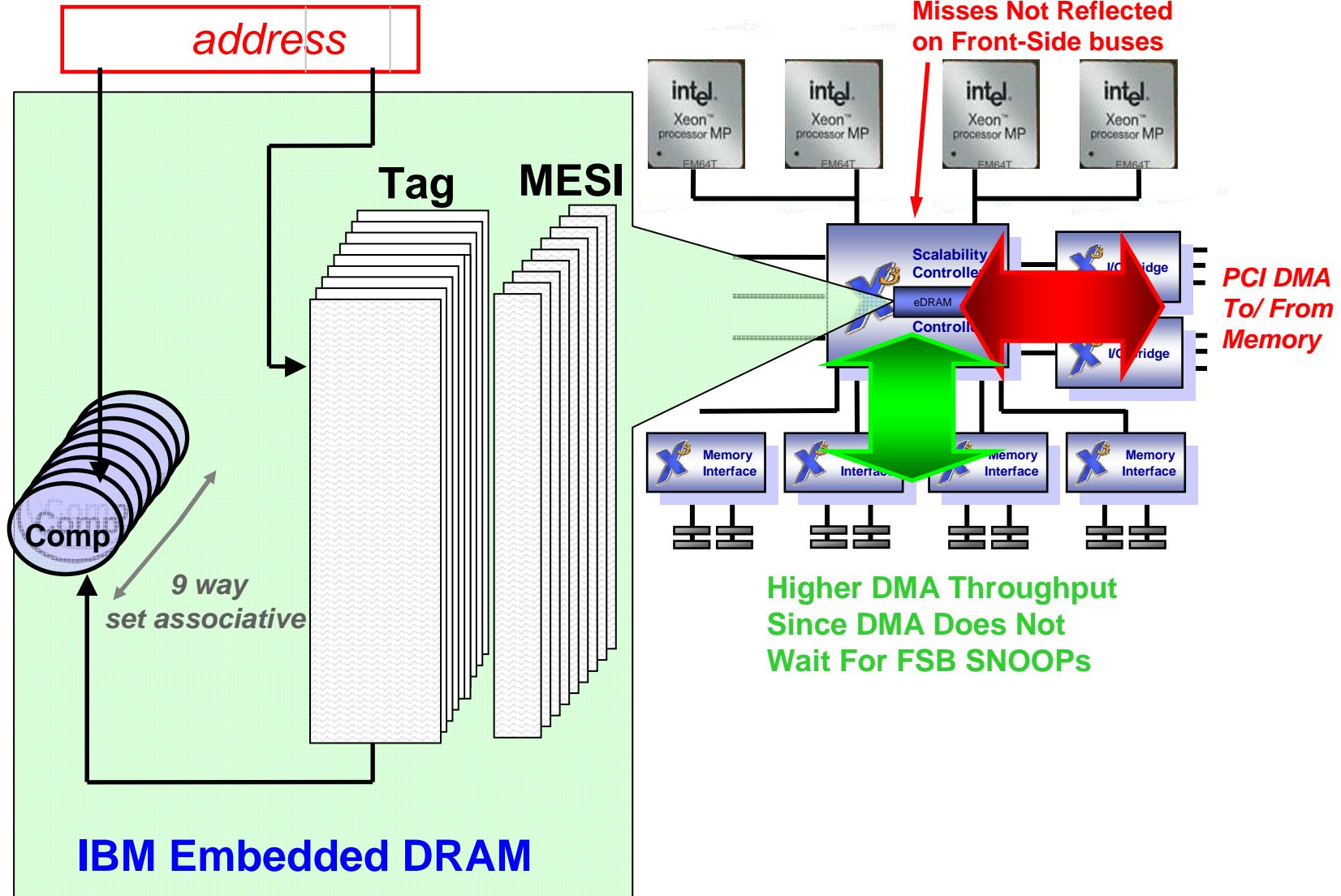
## Cache Miss – Snoop Filter Miss Example



# X Chipset – Snoop Filter DMA I/O – Snoop Filter Miss Example



# X Chipset – Snoop Filter DMA I/O – Snoop Filter Miss Example



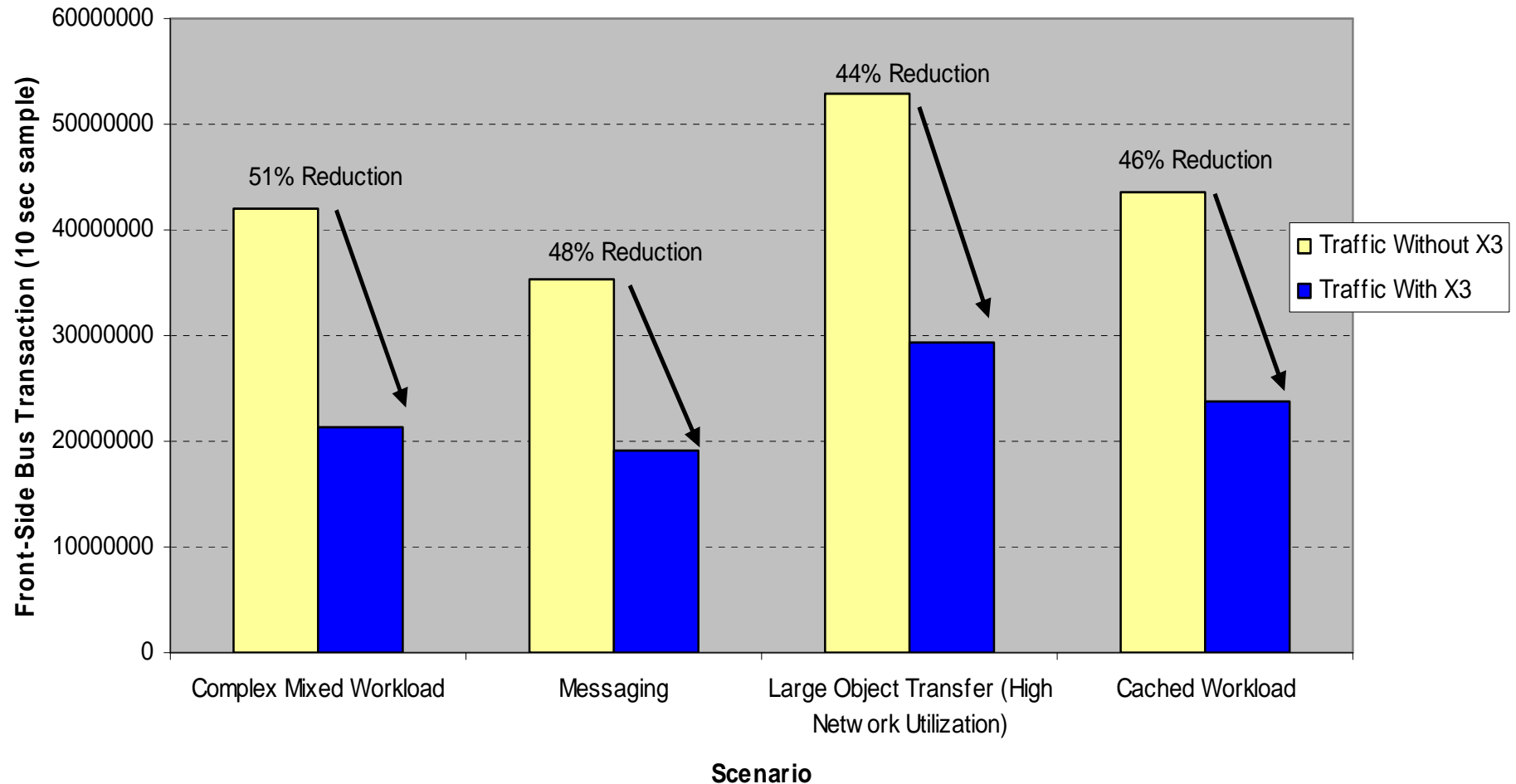


# Reduction in FSB Traffic = Higher Performance

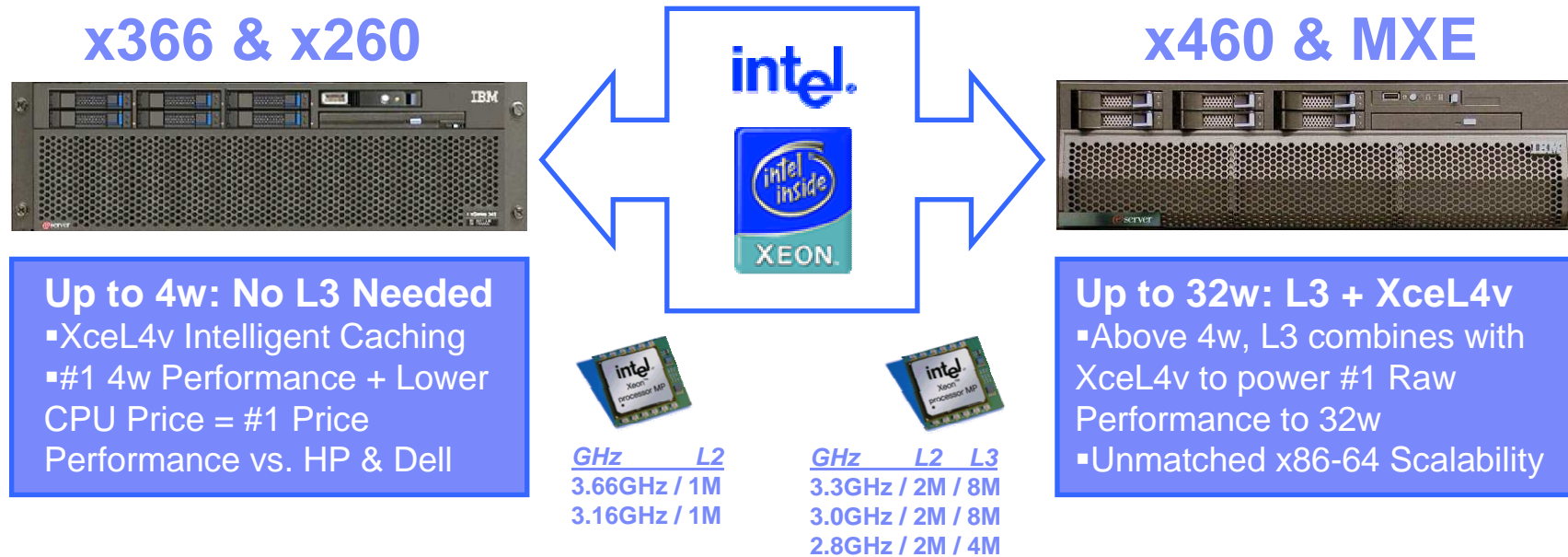
## IBM X3 Architecture

System Traffic Reduction = Performance Increase

Performance increases achieved by reduction of traffic between system resources (such as processors, memory, and I/O)



# Intelligent Caching for Ultimate Performance

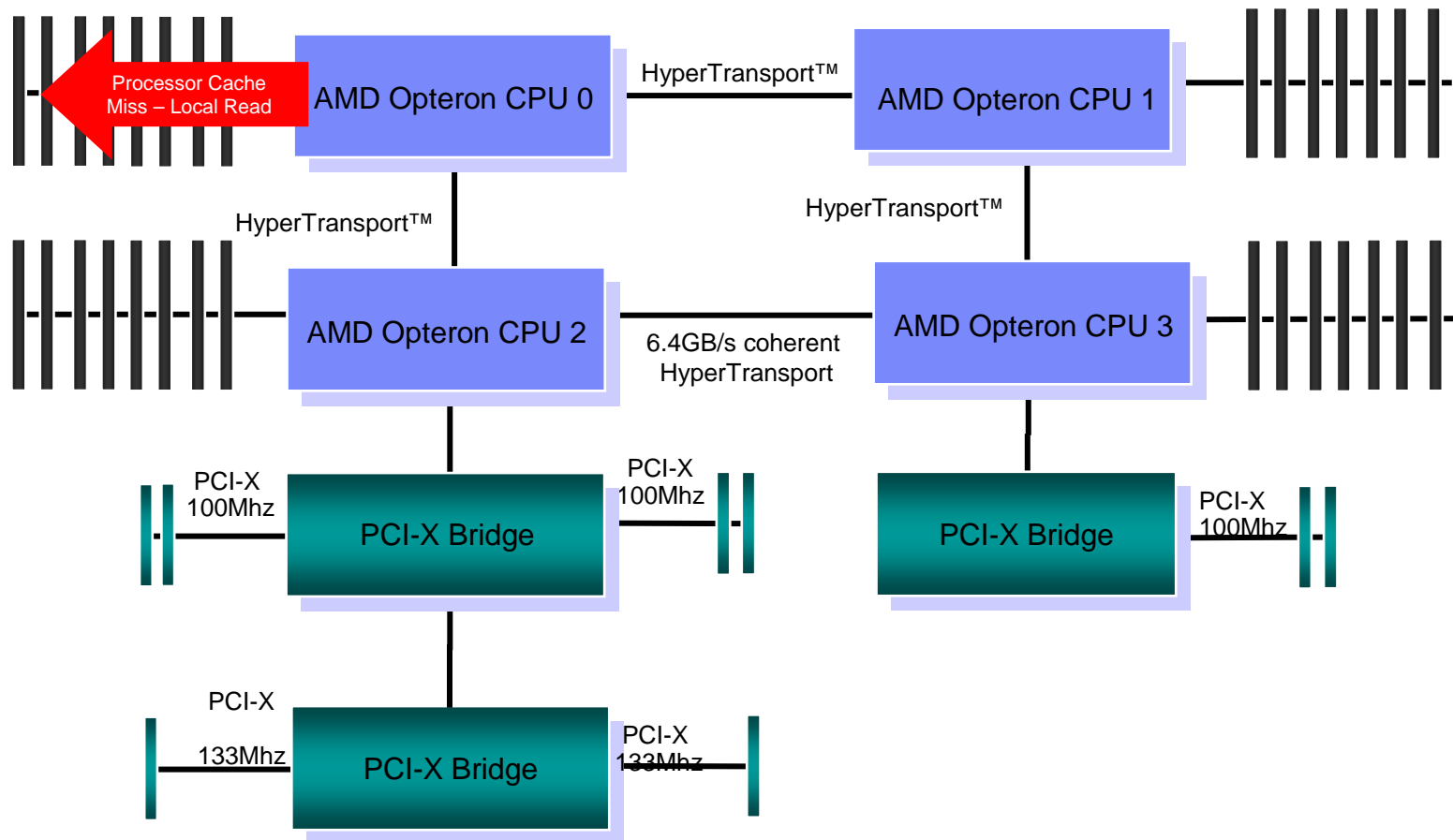


- Intel has introduced two variations of its next generation 64-bit Xeon MP
  - Cranford: L2 only, Potomac: up to 8M L3
- XceL4v: 256MB of DD2-based virtual L4 cache per 4 CPUs (up to 2GB max)
  - Compare to 32MB per 4 CPUs in x440, 64MB in x445
- XceL4v™ provides a performance boost to scalability powering #1 performance with 64-bit Intel Xeon MP with up to 8M L3 cache.
  - L3 is required only for scalability greater than 4-way

## Intel Xeon vs AMD Opteron: fundamental differences

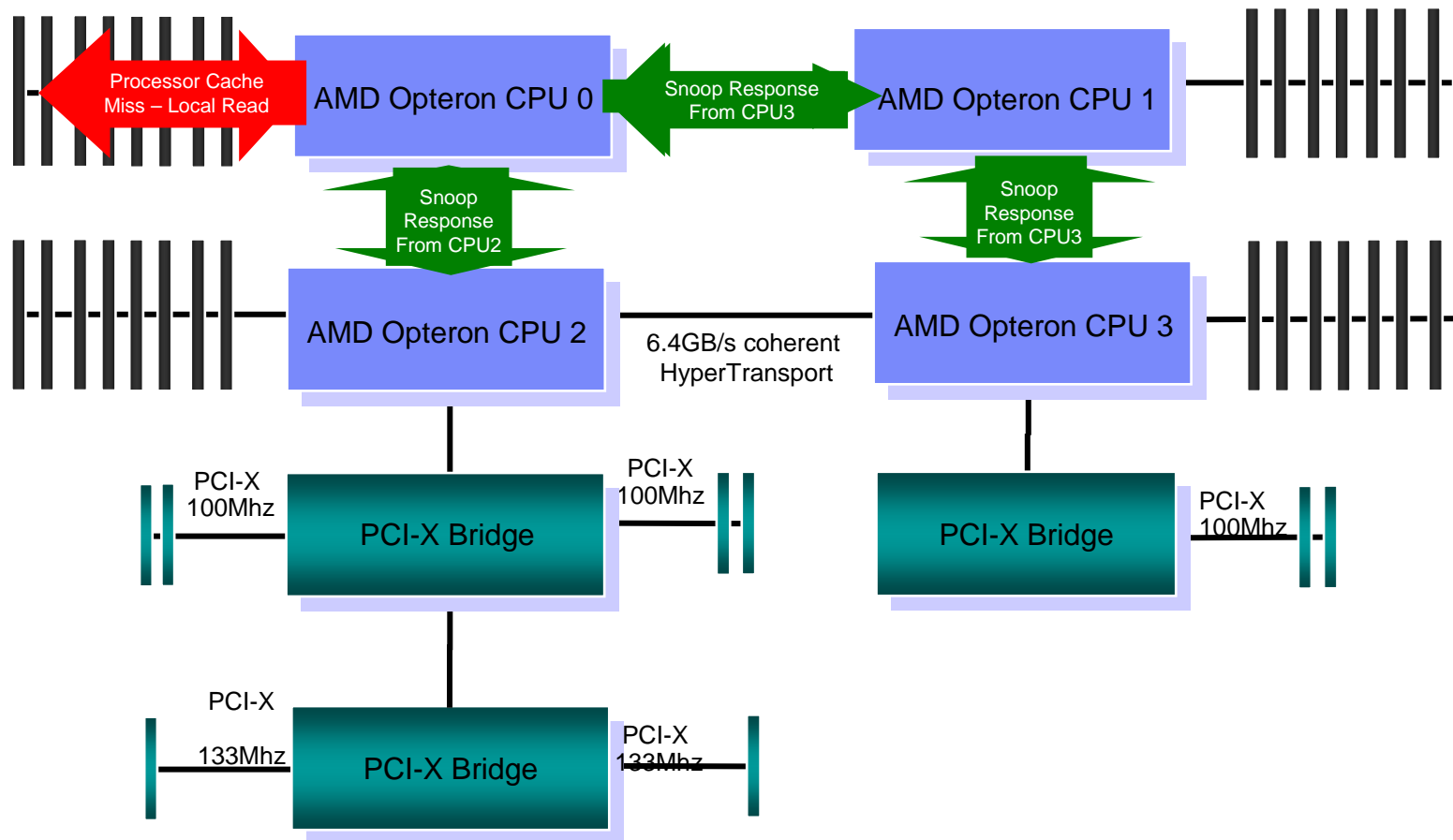
# 4-way Opteron Architecture – Local Memory Access

1. Local memory read happens fast – This low latency is well publicized



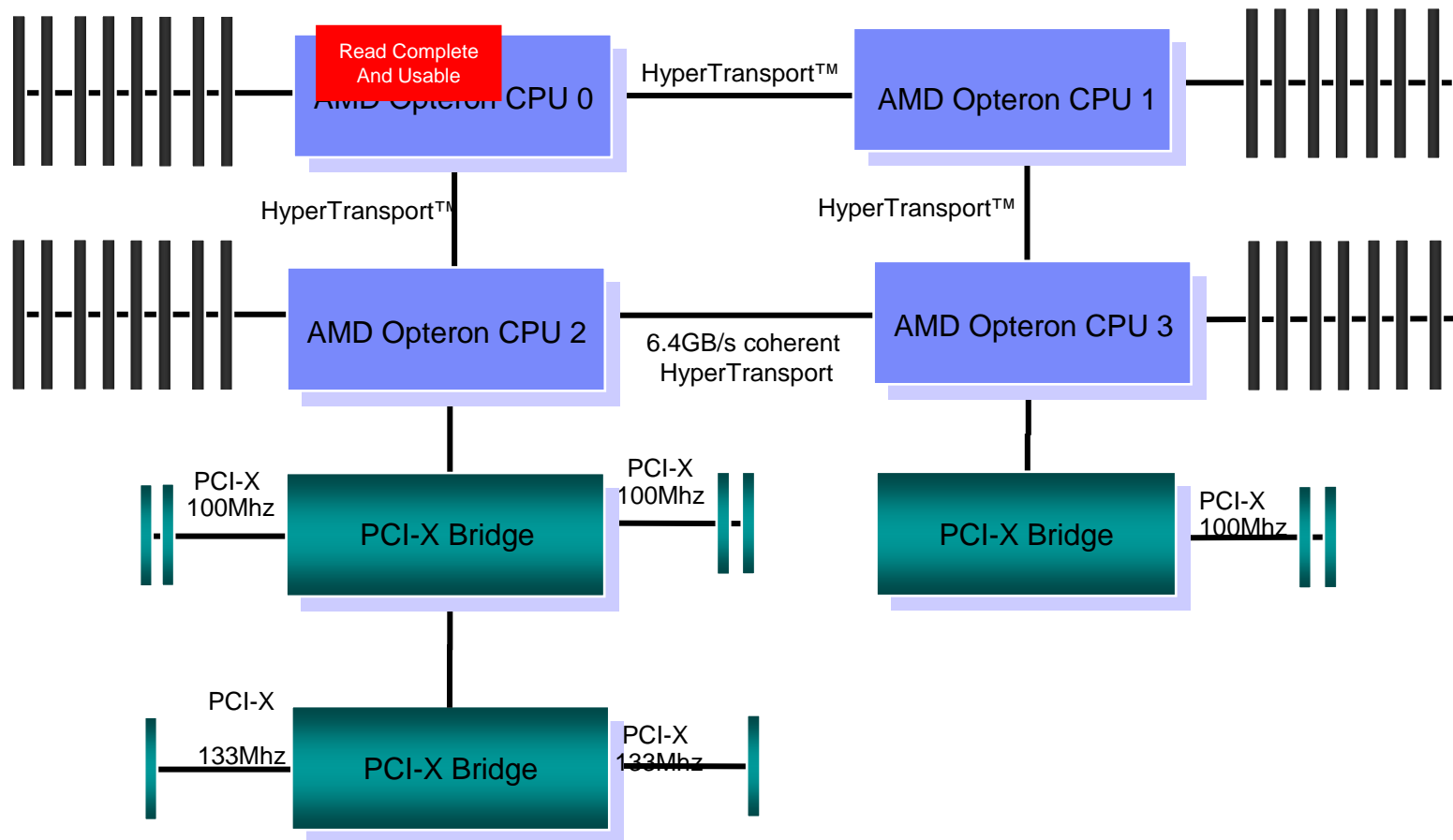
## 4-way Opteron Architecture – Local Memory Access

1. Local memory read happens fast – This low latency is well publicized
2. But processor cannot use data until ALL snoops complete
3. In 4-way there are always two hops for snoops



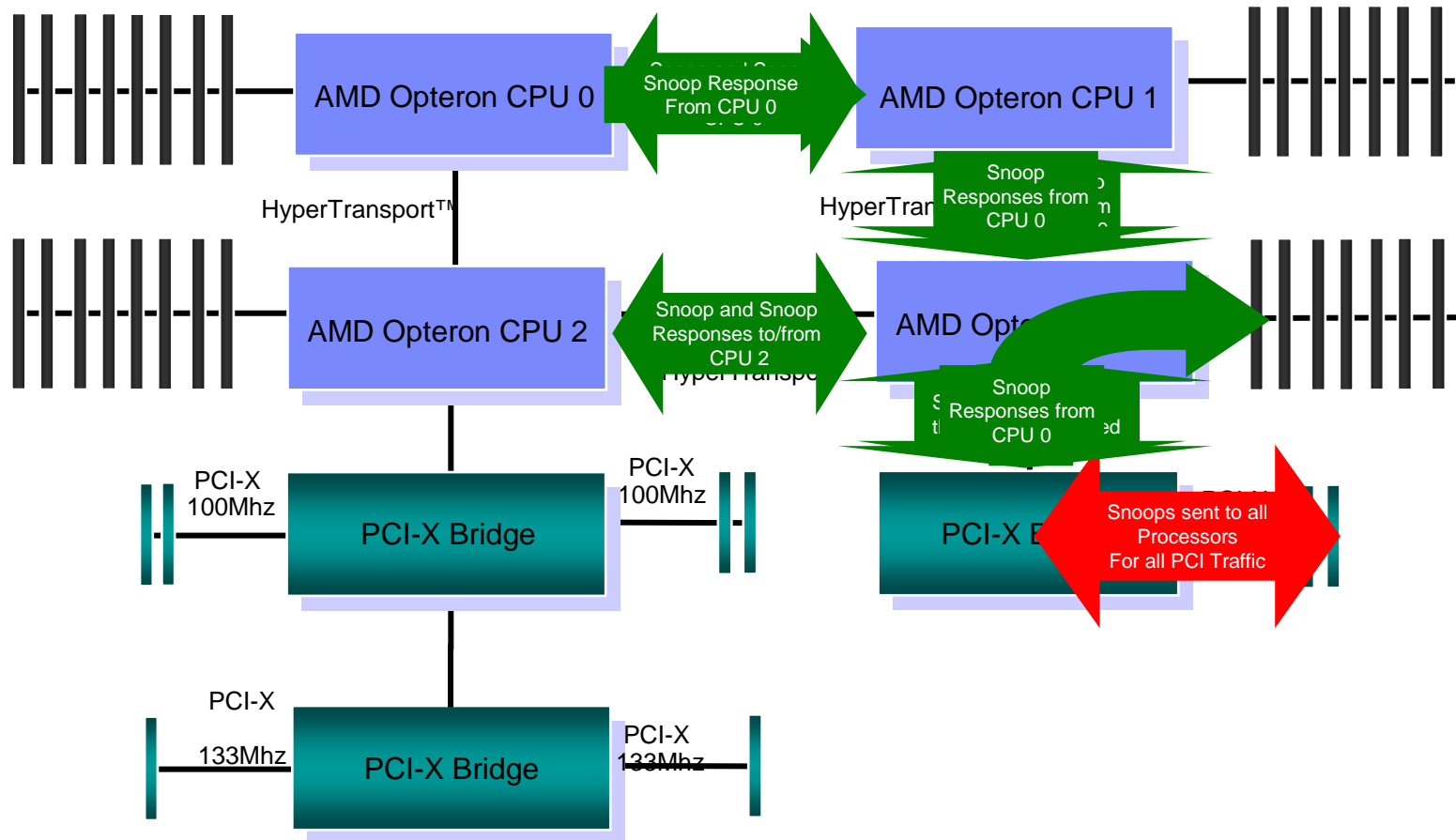
## 4-way Opteron Architecture – Local Memory Access

1. Local memory read happens fast – This low latency is well publicized
2. But processor cannot use data until snoop completes
3. In 4-way there are always two hops for snoops
4. CPU 3 is always the farthest away and determines minimum load to use time
5. Only now can data be used by processor

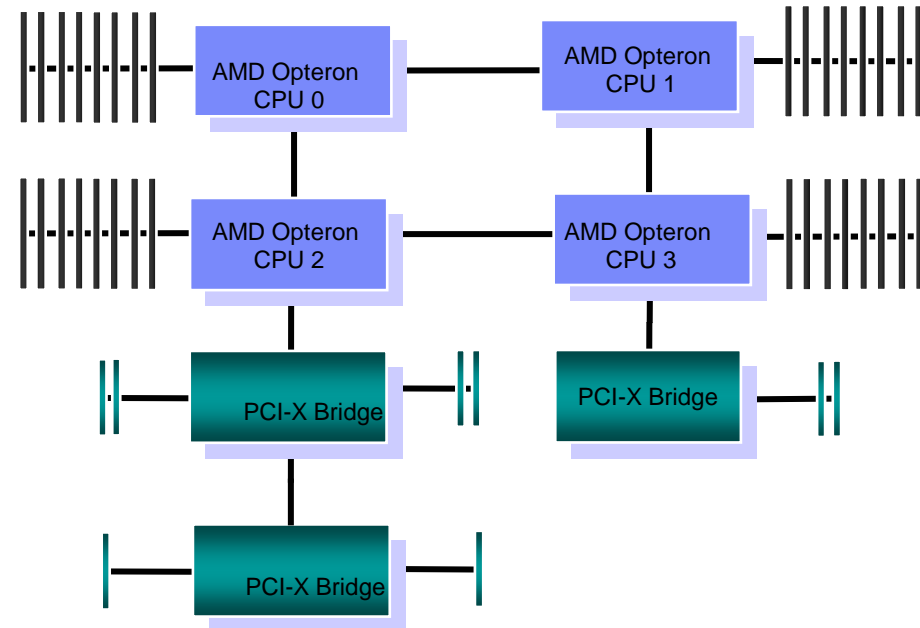
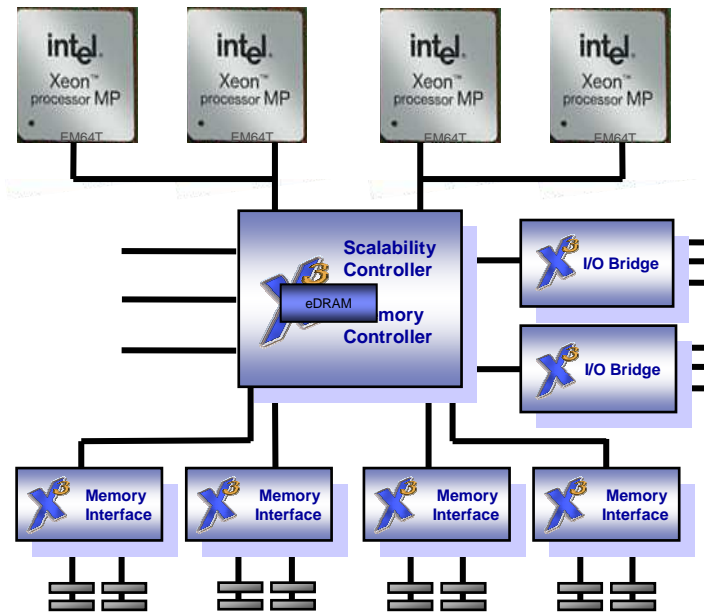


## 4-way Opteron Architecture – Local Memory Access

1. ALL PCI I/O also have to send snoops across the HT links to EVERY processor
2. The snoops must complete before PCI I/O can complete
3. As long as memory accesses are local and PCI traffic is local
  1. The HT Links are mostly used for snoop traffic
4. But for remote memory or I/O, snoops queue with remote traffic and performance suffers



# Xeon vs. Opteron – Fundamental Differences



- **Memory Bandwidth Performance**
  - ▶ Significant difference between memory architectures of Xeon and Opteron processors.
  - ▶ Xeon processors uses a Front-Side Bus (FSB) that is shared among all processors connected to external memory controller.
  - ▶ Memory controller employs two channels to the actual DDR-II 400 (PC3200 DDR-II) memory DIMMs
  - ▶ Opteron processors use embedded memory controller; all processors connected by HyperTransport Bus
  - ▶ Each processor has two channels to DDR-I memory (going EOL)
- **Memory Latency Performance**
  - ▶ Similar for both processors with advantage depending upon NUMA awareness and access characteristics of workload
  - ▶ Multi-threaded, multi-user workloads cause highly random memory address activity causing frequent page faults
  - ▶ Latency, not memory bandwidth, dominates performance because page-hits rarely occur

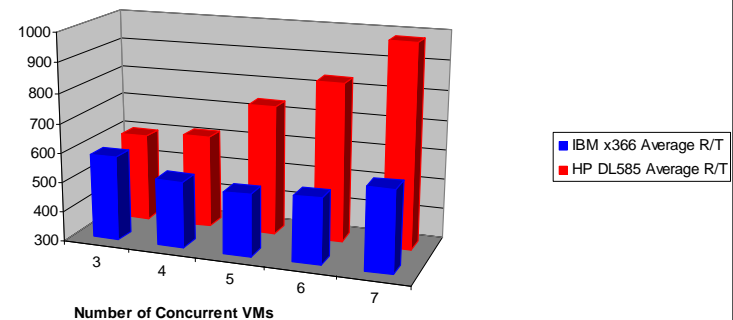




## New Test Results IBM vs. HP

- New report July, 2005
  - ▶ IBM commissioned study
  - ▶ x366 vs. DL 585
  - ▶ Similarly priced systems
- Simulates real-world virtualized workloads driving 90% processor utilization (as opposed to AMD sponsored study)
- 41% faster response times for Exchange and SQL Server applications

Exchange Average Response Time (ms) - Lower is Better



### Key findings

- In a 'real world' mixed application virtualized environment (including Terminal Services, two file and print, static and dynamic WEB, Collaboration and Database applications, all running simultaneously), VeriTest found that the IBM x366 Intel® Xeon Processor MP based server, using X3 Architecture, out-performed a similarly configured and priced HP DL585 AMD Opteron based server on Exchange and SQL workloads, delivering response times that were over 41% faster.

[http://www.veritest.com/clients/reports/ibm/IBM\\_eServer\\_X3\\_0705.pdf](http://www.veritest.com/clients/reports/ibm/IBM_eServer_X3_0705.pdf)

# Thank you !