



PCI ExpressTM Architecture

Enterprise Platform Design Considerations

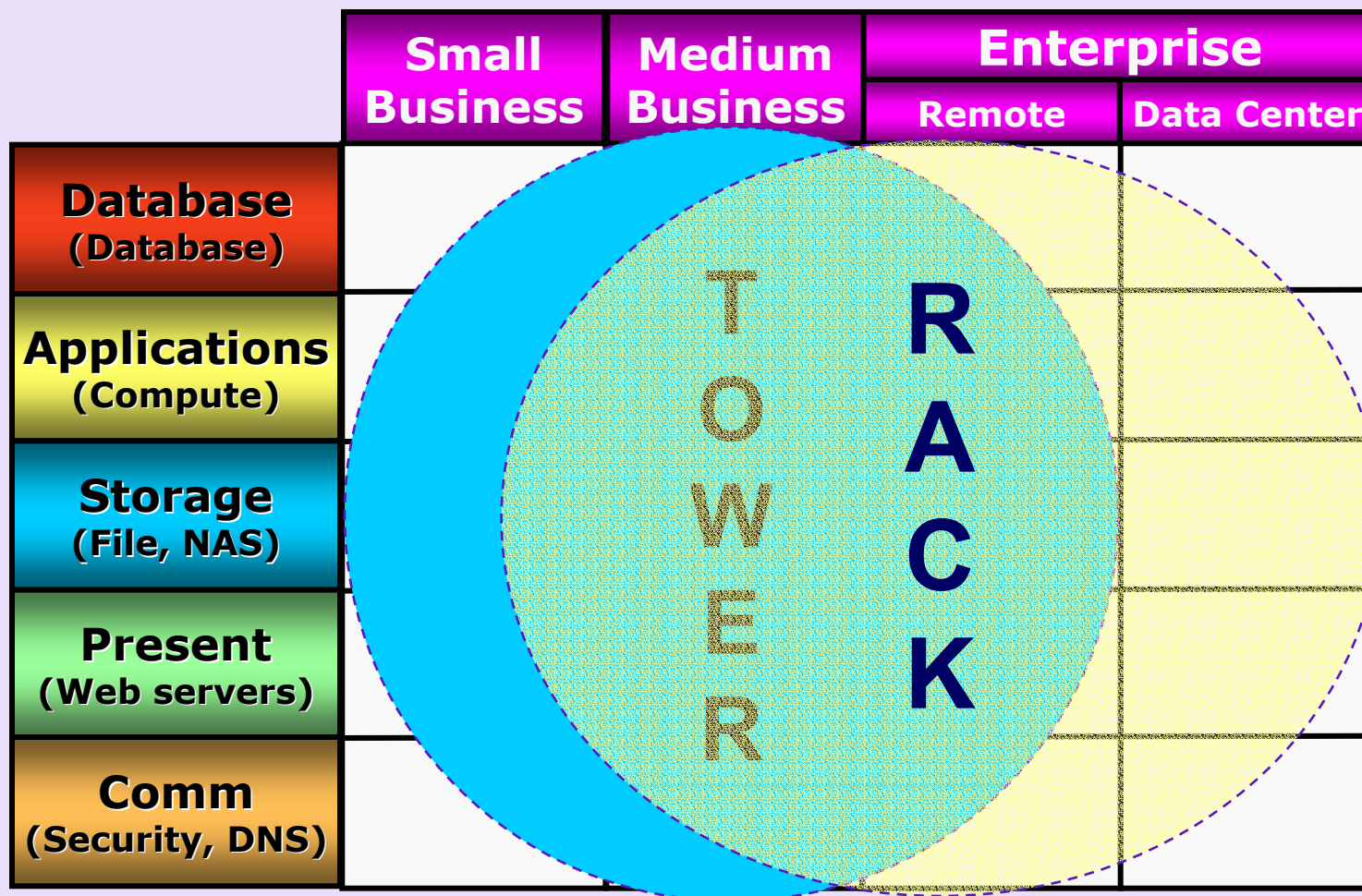
Jimmy Pike

Director, Server Architecture

Dell Computer Corporation



Enterprise Environment



Trends in Enterprise I/O

STORAGE



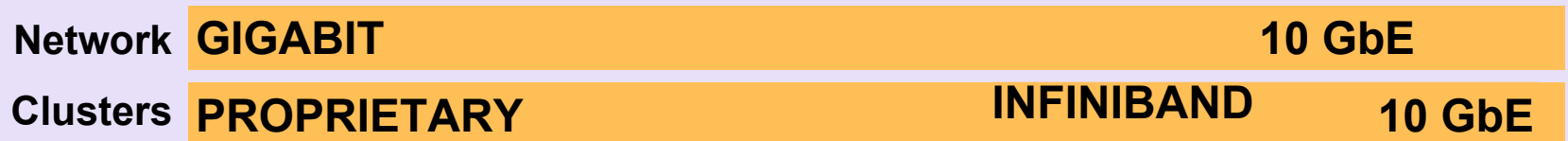
DISK



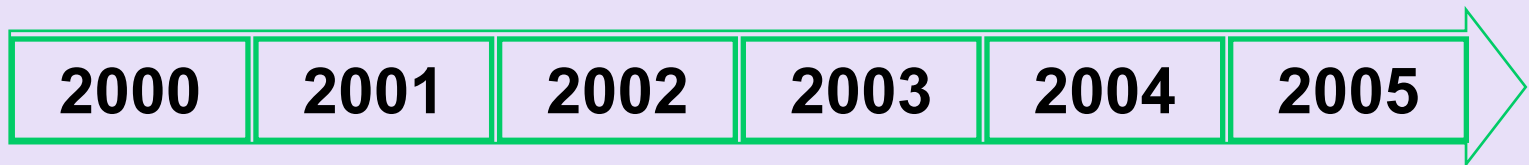
INTERNAL CONNECTIVITY



EXTERNAL CONNECTIVITY



PARALLEL
SERIAL

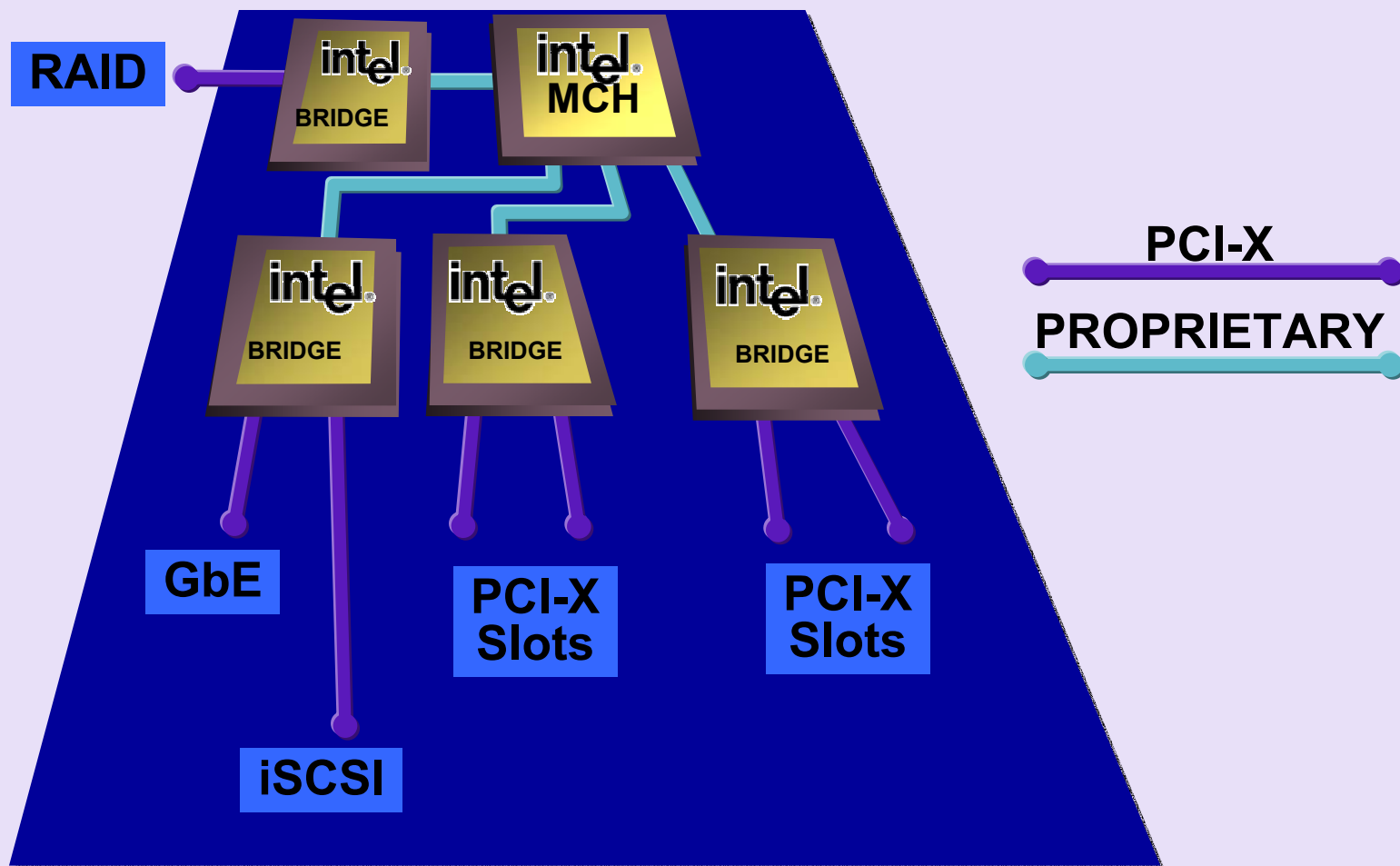


Challenges For Enterprise I/O

- Enterprise application I/O bandwidth requirements increasing
 - ✓ Transitioning to 10 Gb/sec and beyond
- Many new technologies emerging
 - ✓ Networking: 10Gb Ethernet, TOE
 - ✓ Clustering/IPC: InfiniBand* Architecture, RDMA enabled Ethernet
 - ✓ Storage: SAS, SATA, iSCSI, FC
- System ASPs continuing to decline
 - ✓ TCO is a key end-user priority
- Emerging modular/blade solutions driving new requirements
 - ✓ More technologies direct to motherboard vs. slots
 - ✓ Need for common fabric via direct attach to memory controller

PCI Express Architecture Meets All These Needs

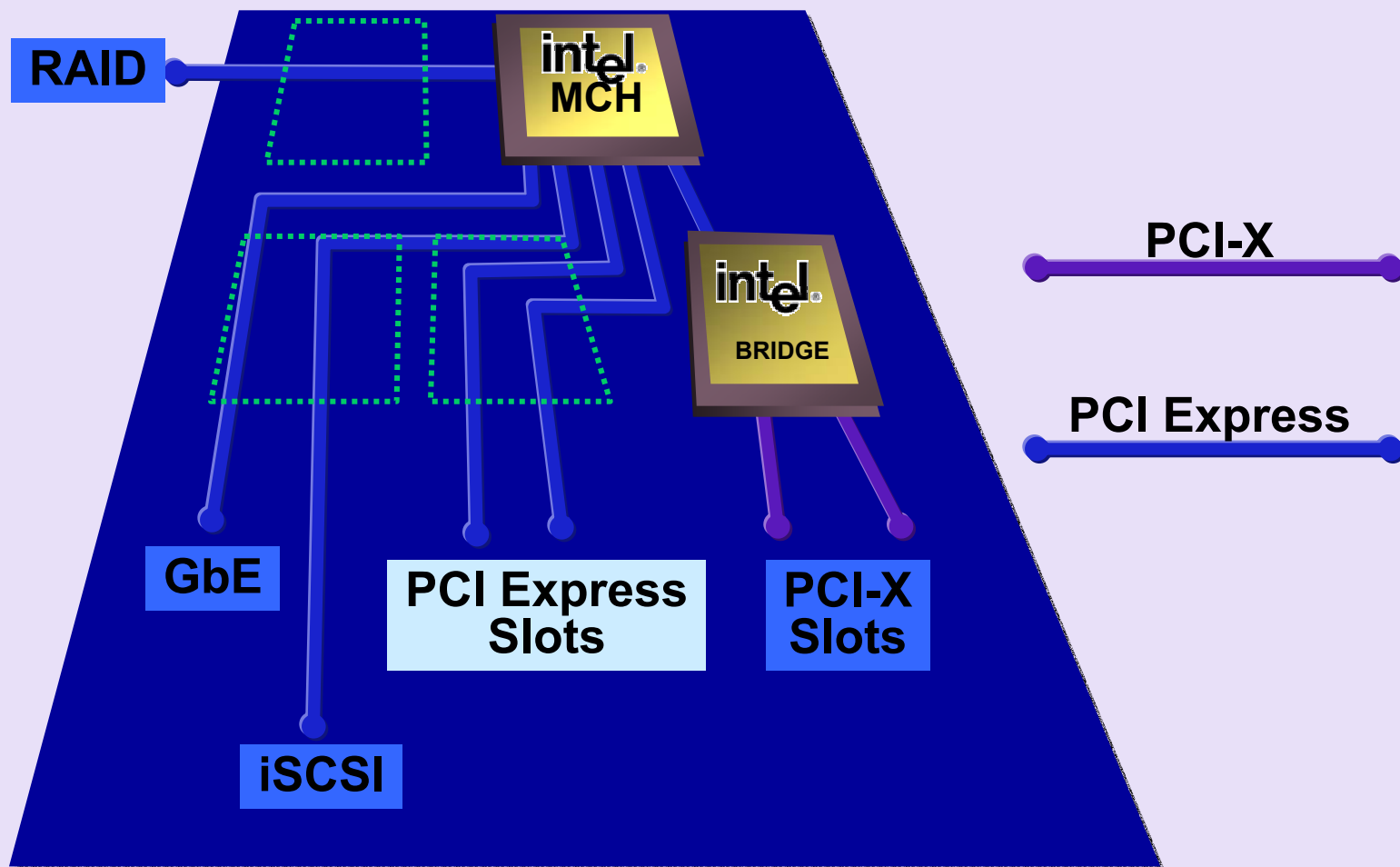
Enterprise Platforms Today



Key Server Requirements

- **Increased Performance**
- **High Connectivity**
- **Improved Scalability**
- **Cost and Power Reduction for Newer Form Factors**
- **Improved Reliability, Availability and Serviceability (RAS)**
- **Legacy Support**

Enterprise Platforms With PCI Express



PCI Express Lowers Cost, Complexity While Improving RAS, Performance

PCI Express for Enterprise

Increased ROI

Future for chip-to-chip & slot interconnect is PCI Express
“Future Proofing”: IT requires 3-5yrs server life cycle
Low Risk: Software compatibility, Proven electricals,
Industry Standard (owned by PCI-SIG)

Enhanced Performance

Bandwidth for 10Gb/s adapter technologies at PCI Express
launch (2004)
Extends high-end graphics capabilities beyond AGP8X
Roadmap for increased performance (4X) with the same
connectors

Advanced RAS

Standardized error reporting mechanisms
Standard hot-plug usage model
Better system reliability due to elimination of bridge
component

PCI Express is a strategic technology benefiting IT managers



PCI Express Architecture

Strongly Augments Data Center I/O

FC

iSCSI

Ethernet

IBA

PCI Express

- PCI Express offers “inside-the-server” technology required to support demanding InfiniBand*, Ethernet, and Fibre channel Data Center interconnects

“PCI Express is going to provide some additional flexibility in the way we design systems...higher availability server solutions...with a lot more performance headroom for the future...I/O subsystems that scale.”

- Randy Groves, Dell CTO & VP

Standardized RAS Advancements

- Excellent error reporting and diagnostic capability
 - ✓ Advanced Error Reporting Capability (*standardized*)
 - ✓ Logs/counts correctable and uncorrectable error types
 - ✓ Uncorrectable events may be programmed to non-fatal or fatal
 - ✓ Physical, data link, and transaction layer errors are individually logged
 - ✓ Hooks to identify agents involved in failures
- *Native* PCI Express Hot Plug support
 - ✓ Both slot and module form factor support
- Future modular form factor
 - ✓ More reliable connection and better protection for on-board circuitry
 - ✓ Usability and Serviceability improved

PCI Express targets server RAS requirements

PCI Express Error Registers

	Base Reporting Registers	Advanced Error Reporting Registers (in addition to Base)
Error Type Detected	Device Status	Uncorrectable Error Status Correctable Error Status Header Log Registers Advanced Error Capability and Control Register <i>Root Error Status*</i>
Error Mask	Device Control	Uncorrectable Error Mask Correctable Error Mask
System Alert Method	<i>Root Control*</i>	<i>Root Error Command*</i>
Severity	N/A	Uncorrectable Error Severity
Error Source		<i>Error Source ID Register*</i>

* Root Complex registers

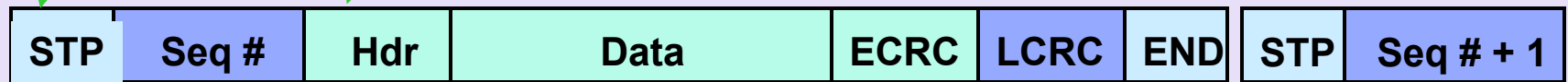
Advanced Error Reporting Provides Server-level Detail

Robust Error Protection

Unambiguous Framing with 8b/10b

Explicit error forwarding mechanism

End-to-end 32b CRC coverage



False/missed start correction

Transaction Layer Packet protected with 32b CRC

False/missed termination correction

Errors in all Packet Phases are Recoverable

PCI Express Addresses I/O Gap

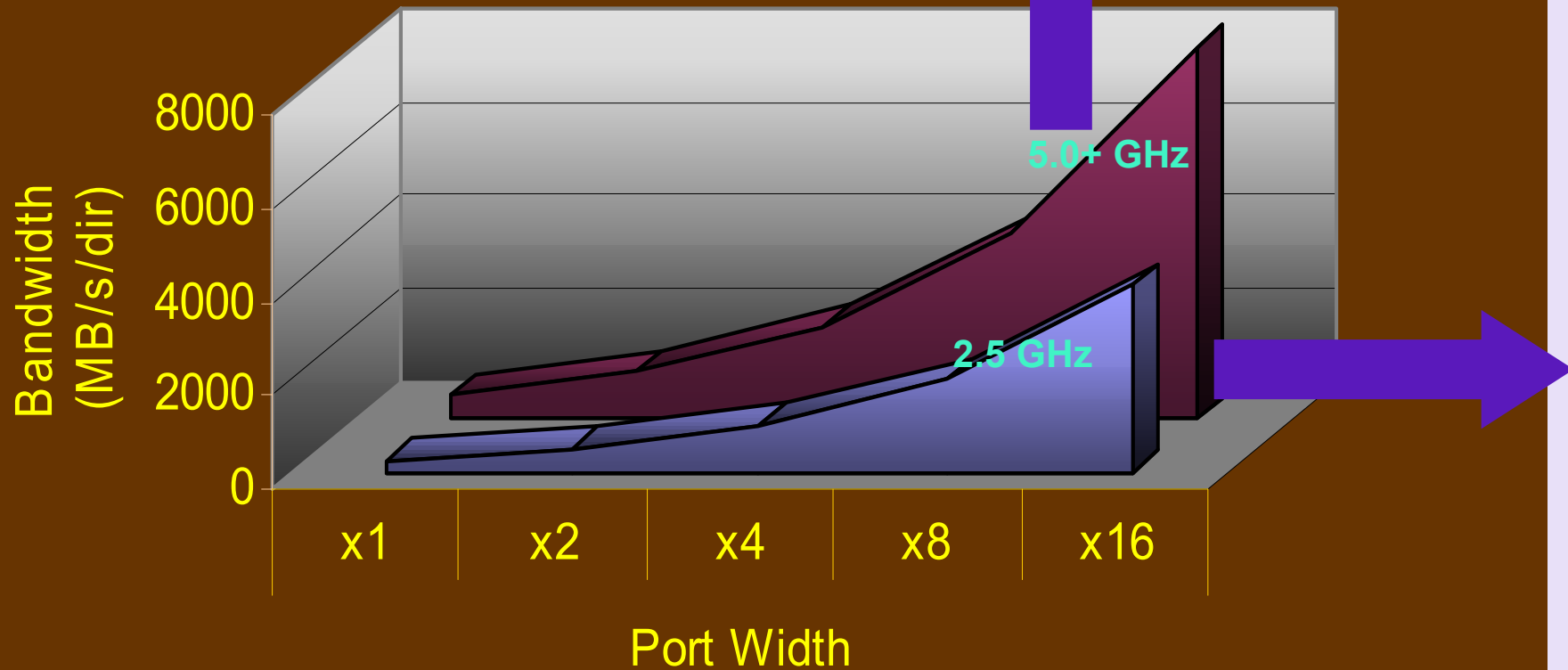
	2000		Q1'03	Q1'03 vs 2000		Q1'04	Q1'04 vs 2000
CPU	733 MHz		3.06 GHz	4X			
FSB	133 MHz		533 MHz	4X			
Memory	PC133 SDRAM		DDR266	4X			
LAN	100 MB		1 Gb	10X			
Local I/O	500 MB/s		1 GB/s	2X		4 GB/s	8X
Graphics	1 GB/s		2 GB/s	2X		2 GB/s	4X

PCI Express allows I/O to keep pace with platform

PCI Express Architecture

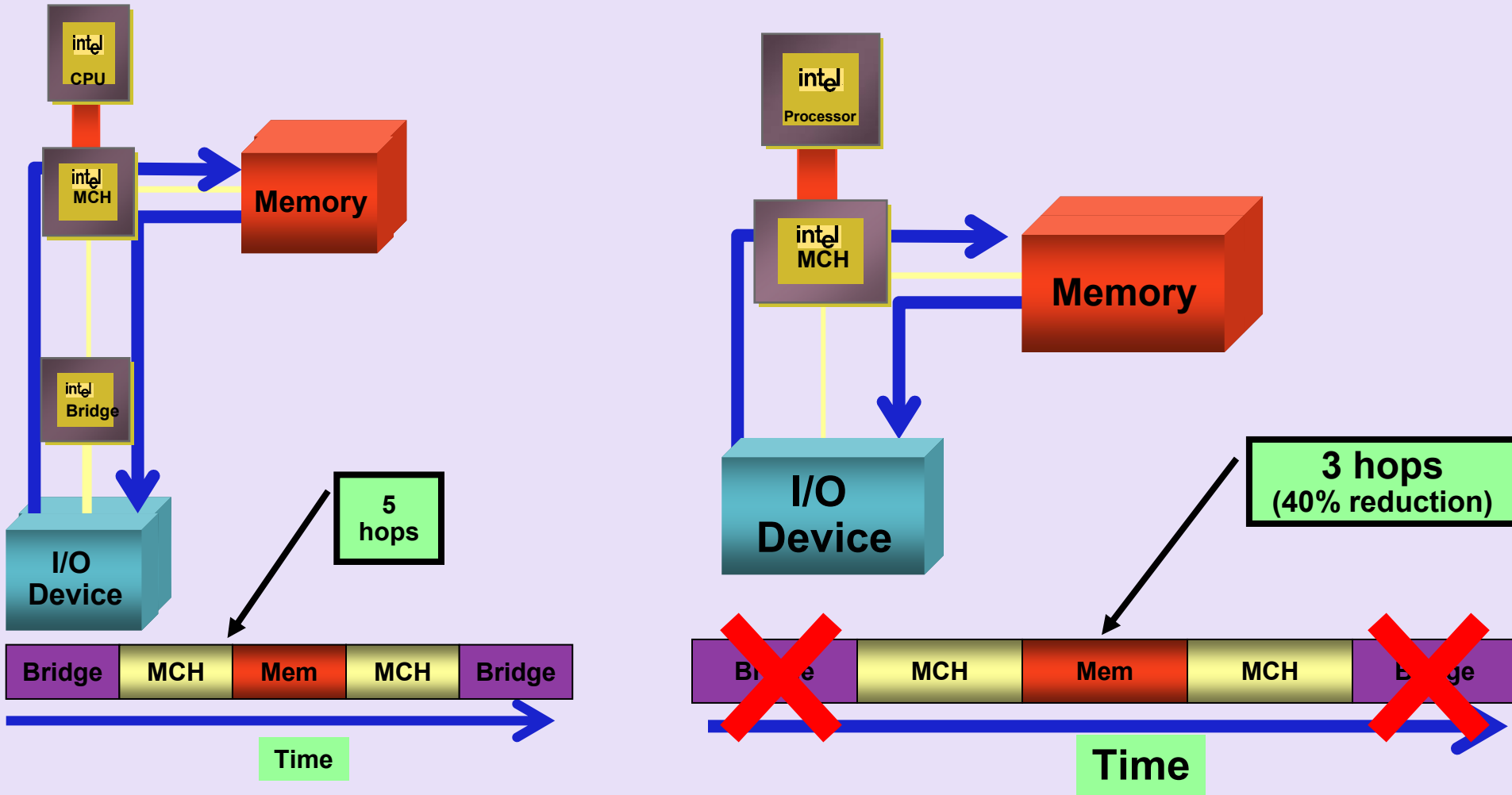
Performance Future Needed In Enterprise

To The Limits Of Copper



Highly Scalable *with* Low Signal Count

System Latency Savings



PCI Express Lowers Latency to Memory

Latency Reduction With PCI Express

Memory Latency

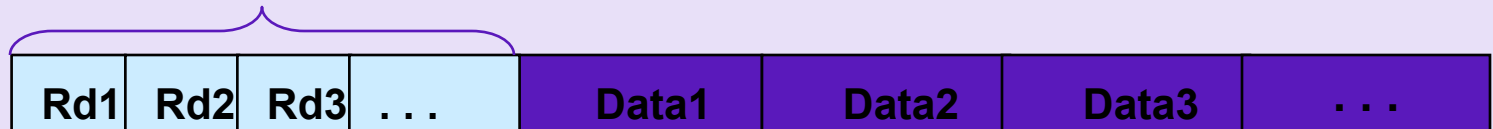
Memory Latency

Serialized:



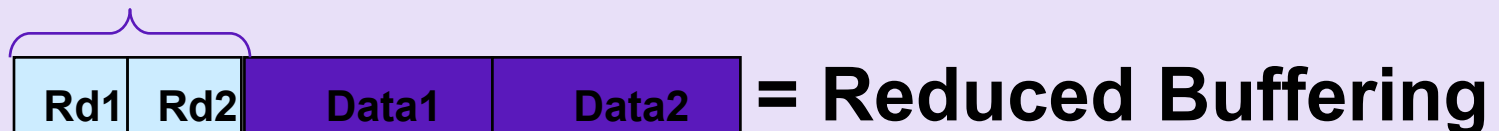
Pipelining extent dependent on latency and traffic characteristics

Pipelining:



Pipelining requirements reduced

PCI Express
Pipelining:

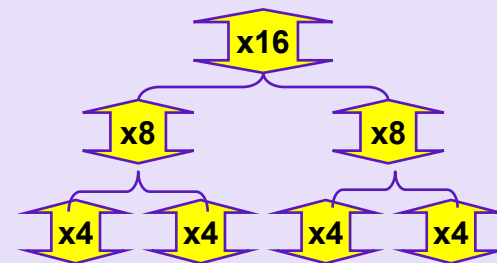
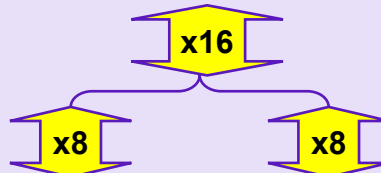


Simpler, Cheaper, & Higher Performing Adapters

PCI Express Architecture

Enables Easier Bandwidth Fan-Out

- Port bifurcation feature:
 - ✓ Fanout of bandwidth without switch
 - ✓ x8 to dual x4s in some chipsets



PCI Express Logic

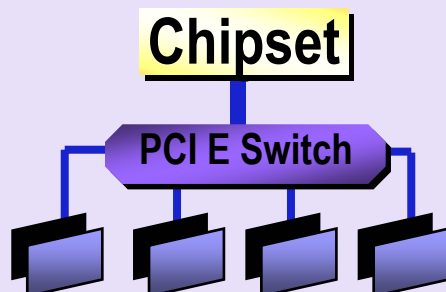
TXN
LNK
PHY

PCI Express Logic

TXN	TXN
LNK	LNK
PHY	PHY

PCI Express Logic

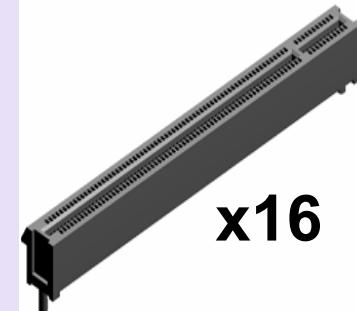
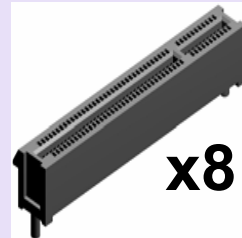
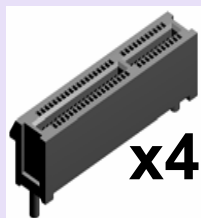
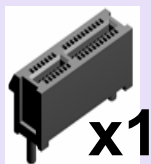
TXN	TXN	TXN	TXN
LNK	LNK	LNK	LNK
PHY	PHY	PHY	PHY



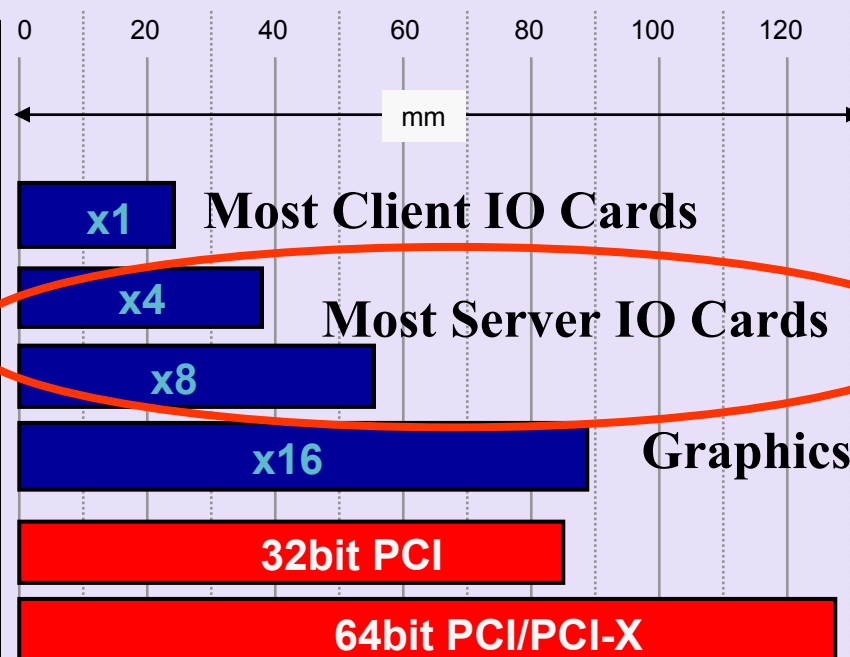
- Fanout via switches
 - ✓ Set of link widths provide optimized fanout per component hop
 - ✓ Enables switches for new topologies

PCI Express Architecture

Enables Smaller Connectors



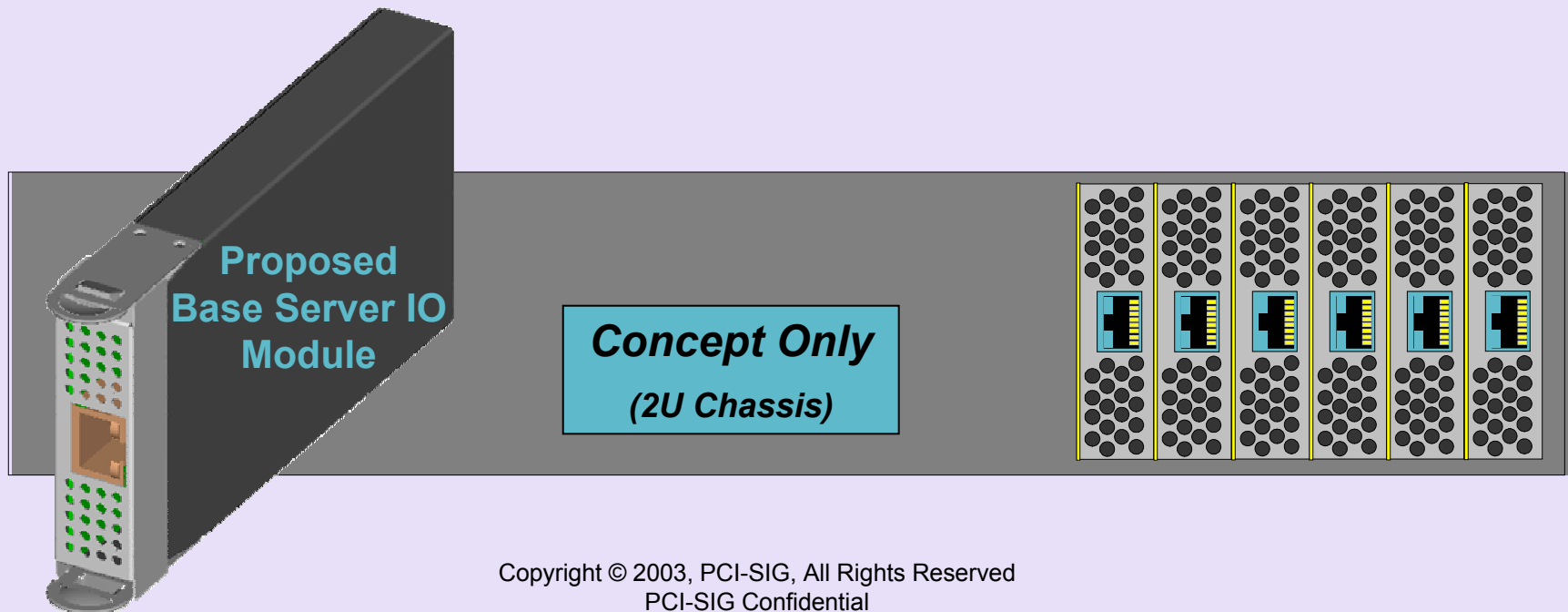
Performance Point		Connector Length: mm (inches)	
PCI Express	x1	25.00	(0.984)
	x4	39.00	(1.535)
	x8	56.00	(2.205)
	x16	89.00	(3.504)
32b PCI		84.84	(3.400)
64b PCI/PCI-X		128.02	(5.040)



PCI Express Architecture

Server I/O Modules Enable New System Flexibility

- Delivers on the promise of hot plug
 - ✓ End user does not have to open the chassis
 - ✓ Takes advantage of PCI Express native hot plug
- Improved OEM cost
 - ✓ Eliminates hot-plug power controller and interface switches



Thank you for attending the
2003 PCI-SIG Developers Conference.

For more information please go to
www.pcisig.com

PCI  **SIG**TM

PCI  **EXPRESS**TM