

PCI

SIGTM

**PCI ExpressTM
Architectural Overview**

Ajay Bhatt

Chairman, PCI Express Working Group

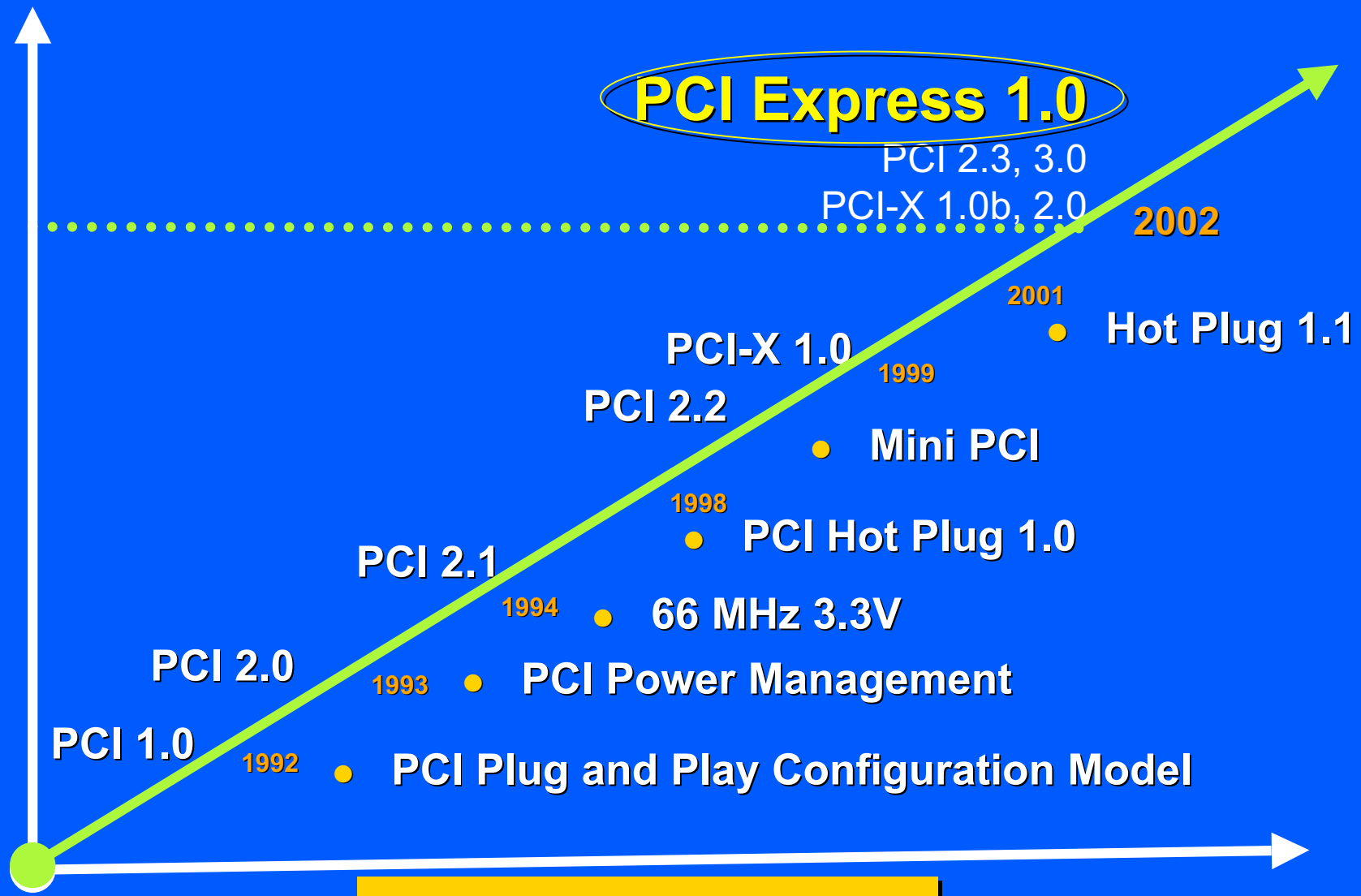
PCI-SIG

PCI 
EXPRESS

Agenda

- ➔ ■ **Requirements**
 - **PCI Express Architectural Overview**
 - ✓ **Physical Layer**
 - ✓ **Transaction and Data Link Layers**
 - ✓ **Mechanicals and Form Factors**
 - **PCI Express based Platforms**
 - **Call to Action**

PCI History



PCI is Everywhere!!!

PCI Express Requirements

- **Cost Effective**
 - ✓ Cost at or below PCI Cost structure at system level
 - ✓ Utilize high volume system ingredients: Si, boards, connectors

- **Support multiple market segments and applications**
 - ✓ Mobile, Desktops, Servers and Communications Devices
 - ✓ Balance performance and feature set by segments

- **Stable, Scalable and Extensible**
 - ✓ Suitable for future applications and technologies over ~10 years
 - ✓ Support advanced features and performance

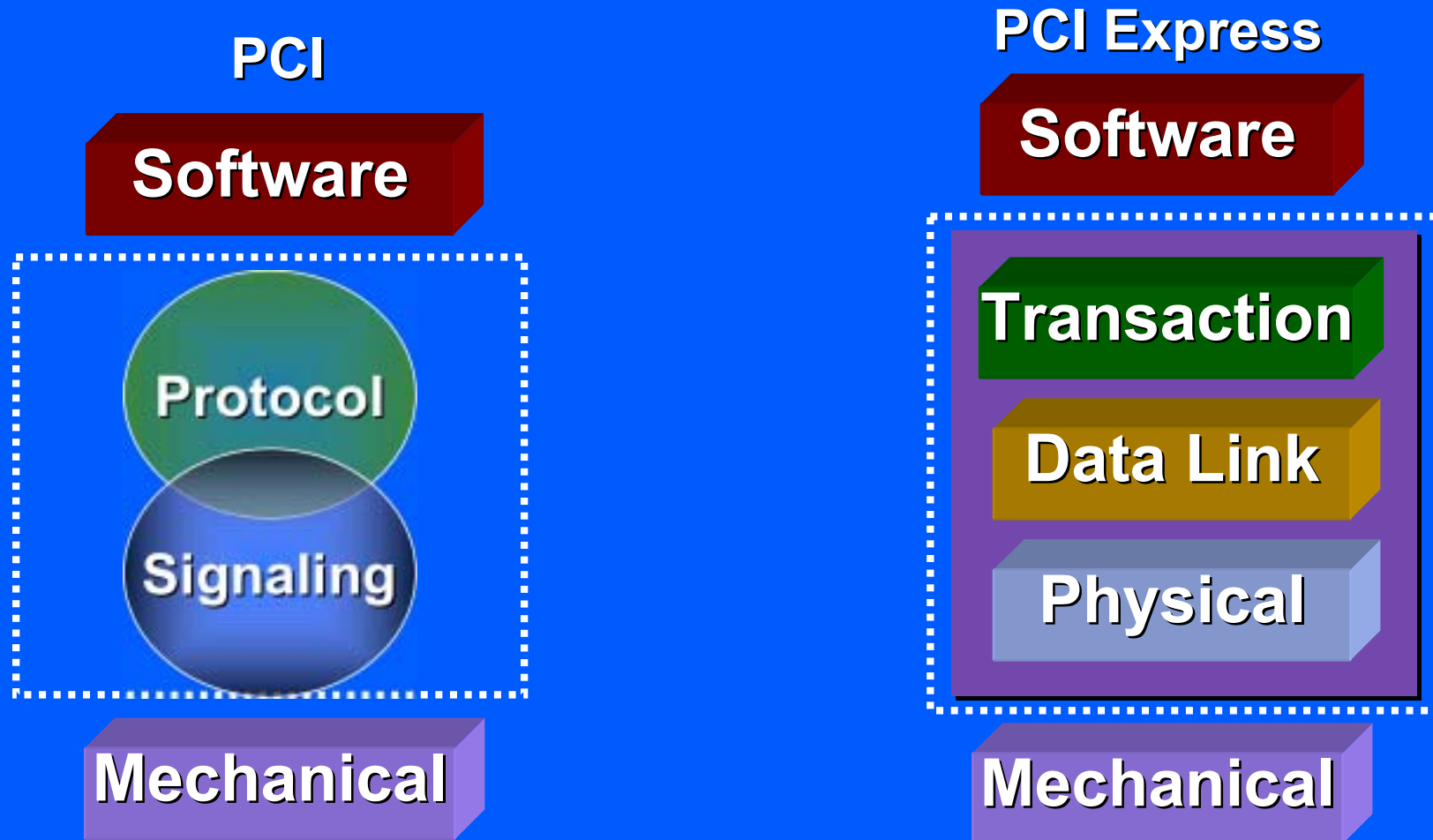
- **Compatible with PCI Architecture and Infrastructure**
 - ✓ Boot existing OSs without any changes
 - ✓ Utilize existing system infrastructure without any changes

Drive high volume Industry implementations starting 2H '03

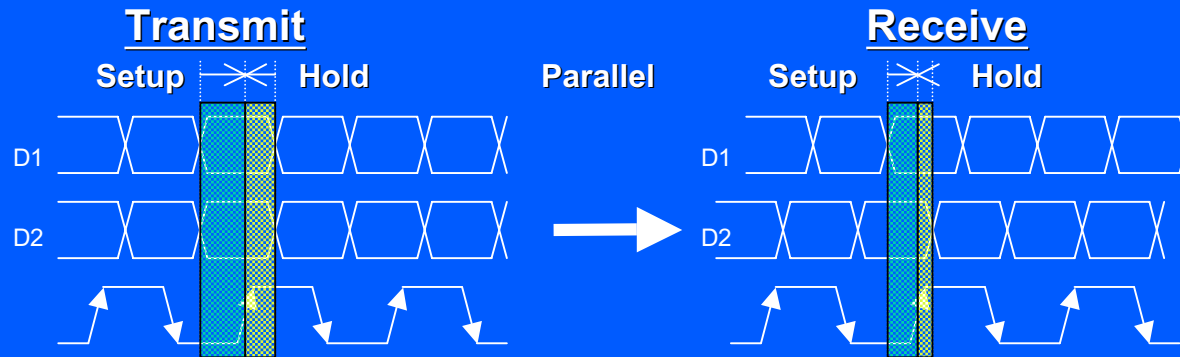
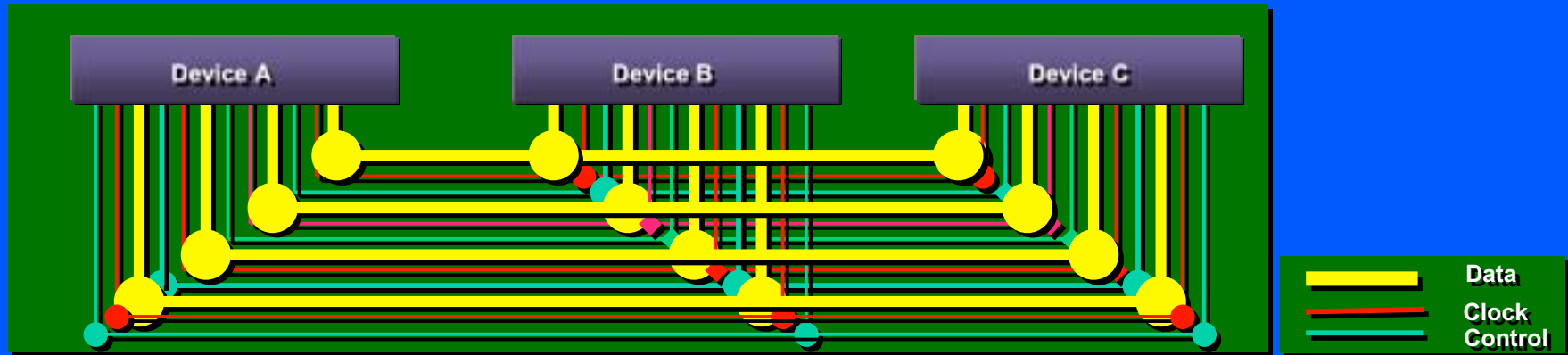
Agenda

- Requirements
- ■ **PCI Express Architectural Overview**
 - ✓ Physical Layer
 - ✓ Transaction and Data Link Layers
 - ✓ Mechanicals and Form Factors
- PCI Express based Platforms
- Call to Action

Layered Architecture

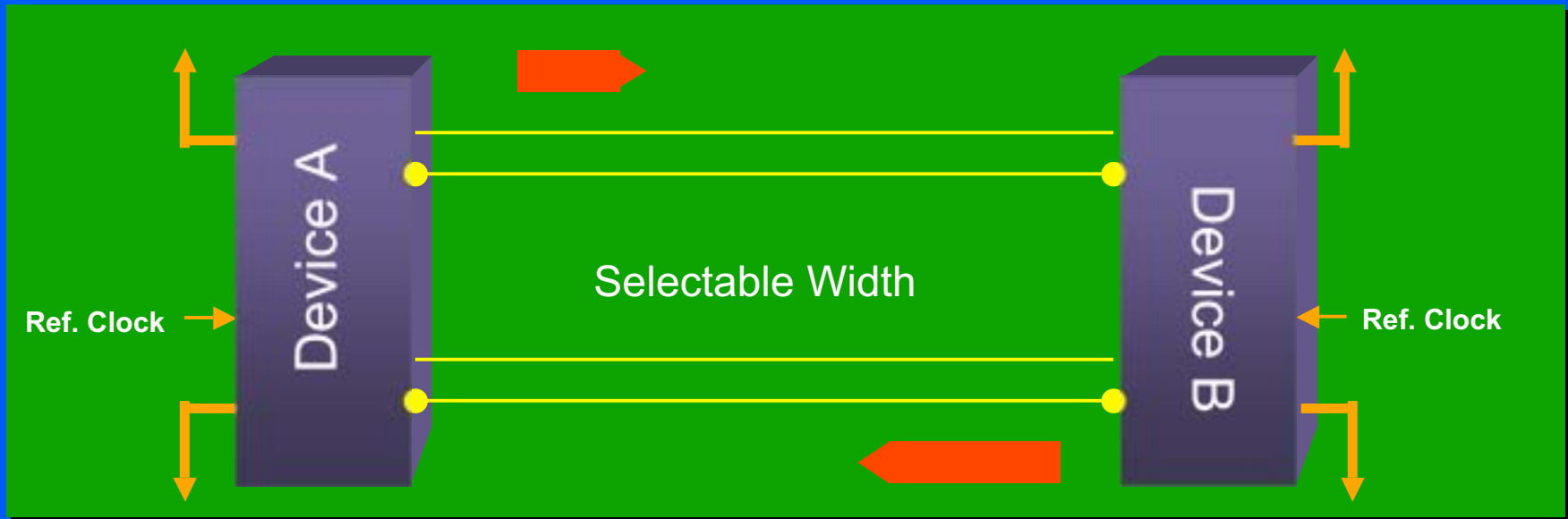


PCI Express Layering Enables Scalability, Modularity and Reuse



- Stringent routing rules
- Limited scaling with frequency and voltage
- Lack of advanced features

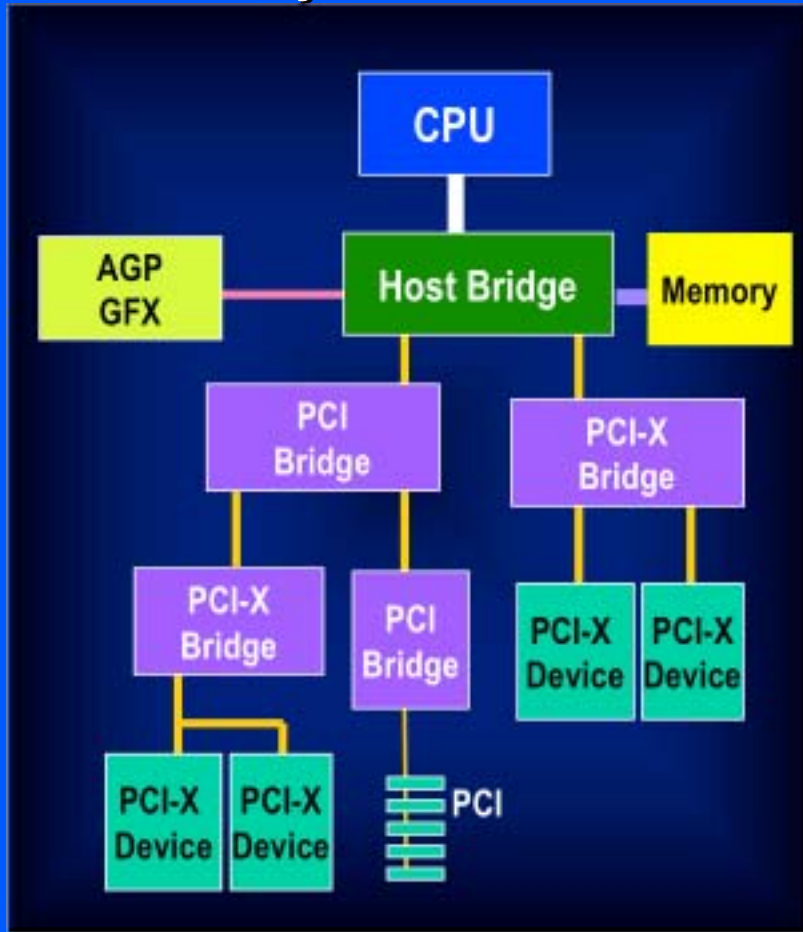
Parallel I/O interconnects have reached practical limit



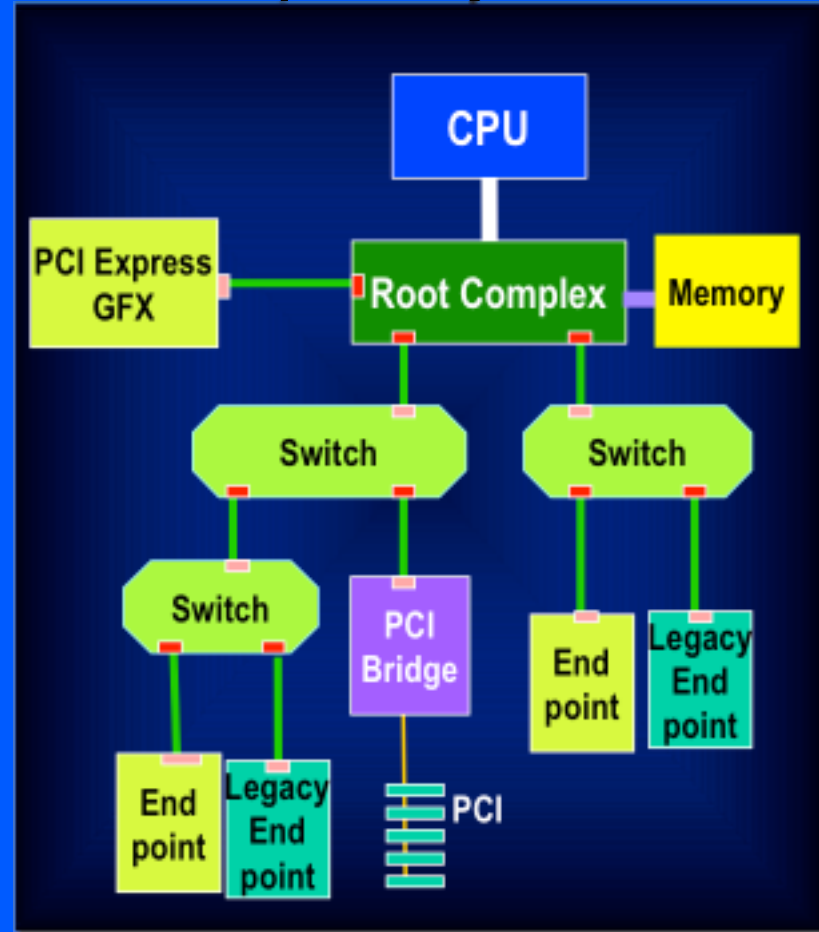
- Dual Simplex Point to point topology
- Differential low voltage interconnect
- Bit rate: > 2.5Gb/sec/lane/direction and beyond
- Selectable lane width:
 - ✓ x1, x2, x4, x8, x12, x16, x32

Functional Elements

PCI System

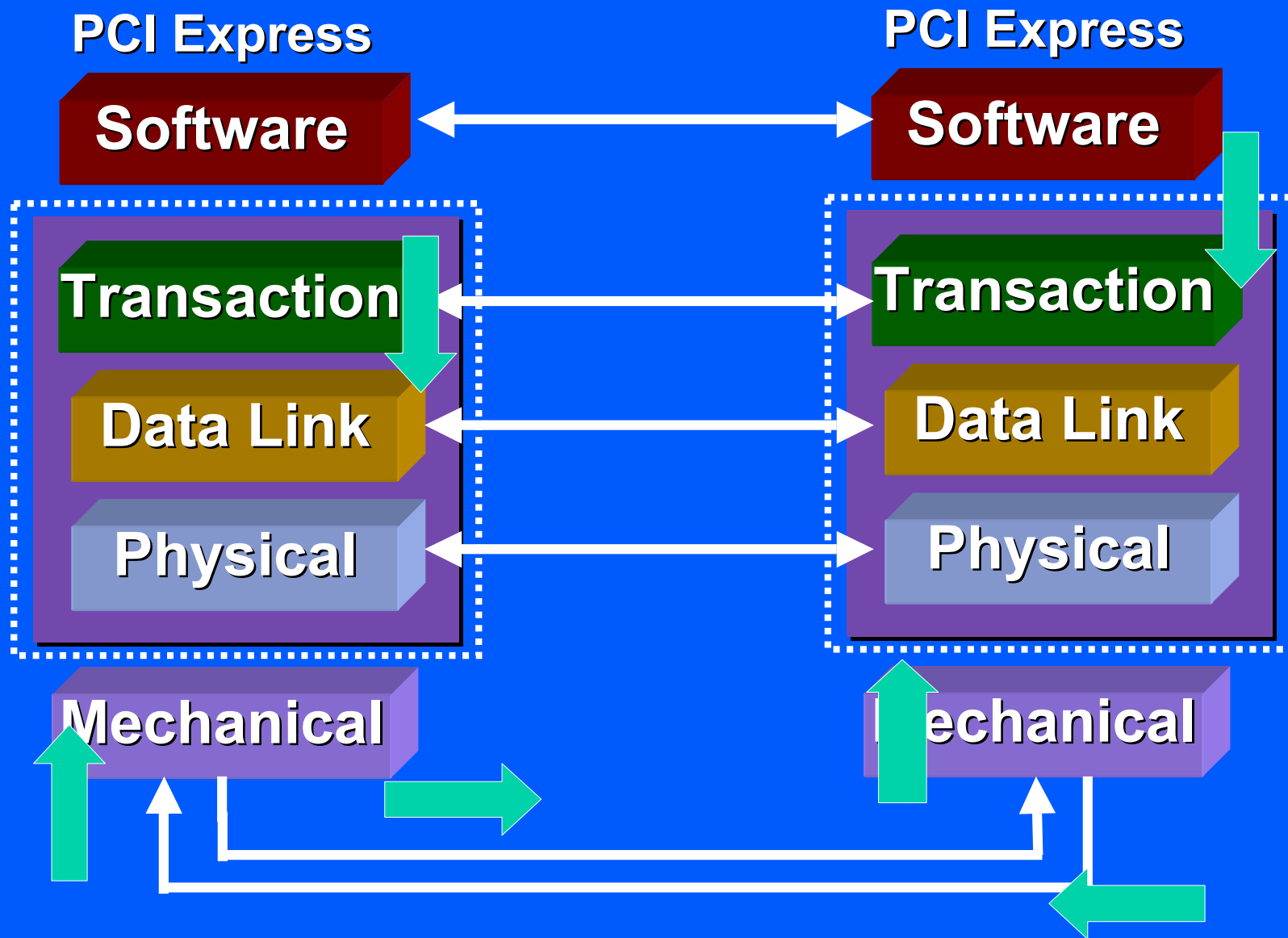


PCI Express System

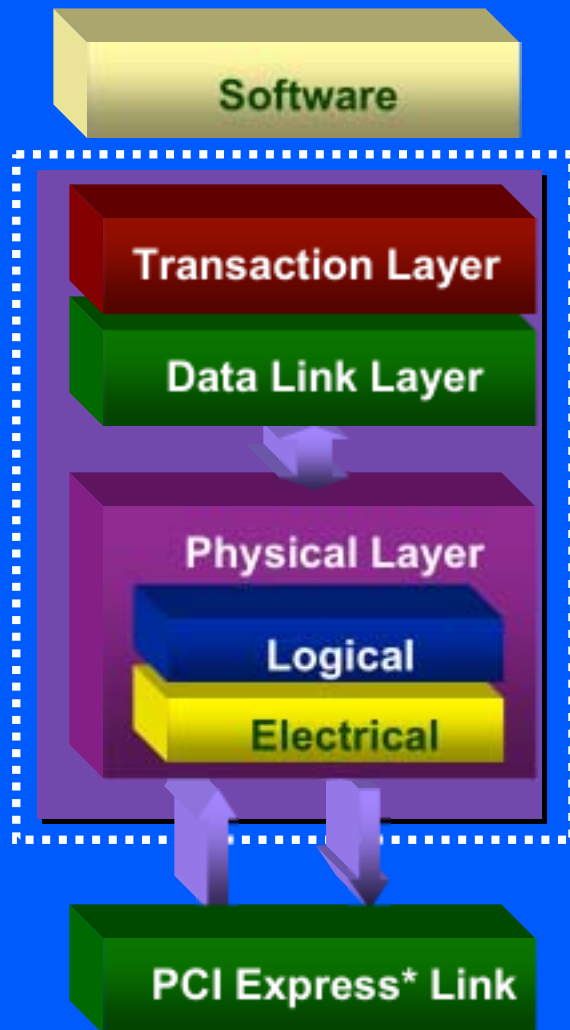


PCI Express elements emulate PCI configuration environment

Data flow thru PCI Express Layers



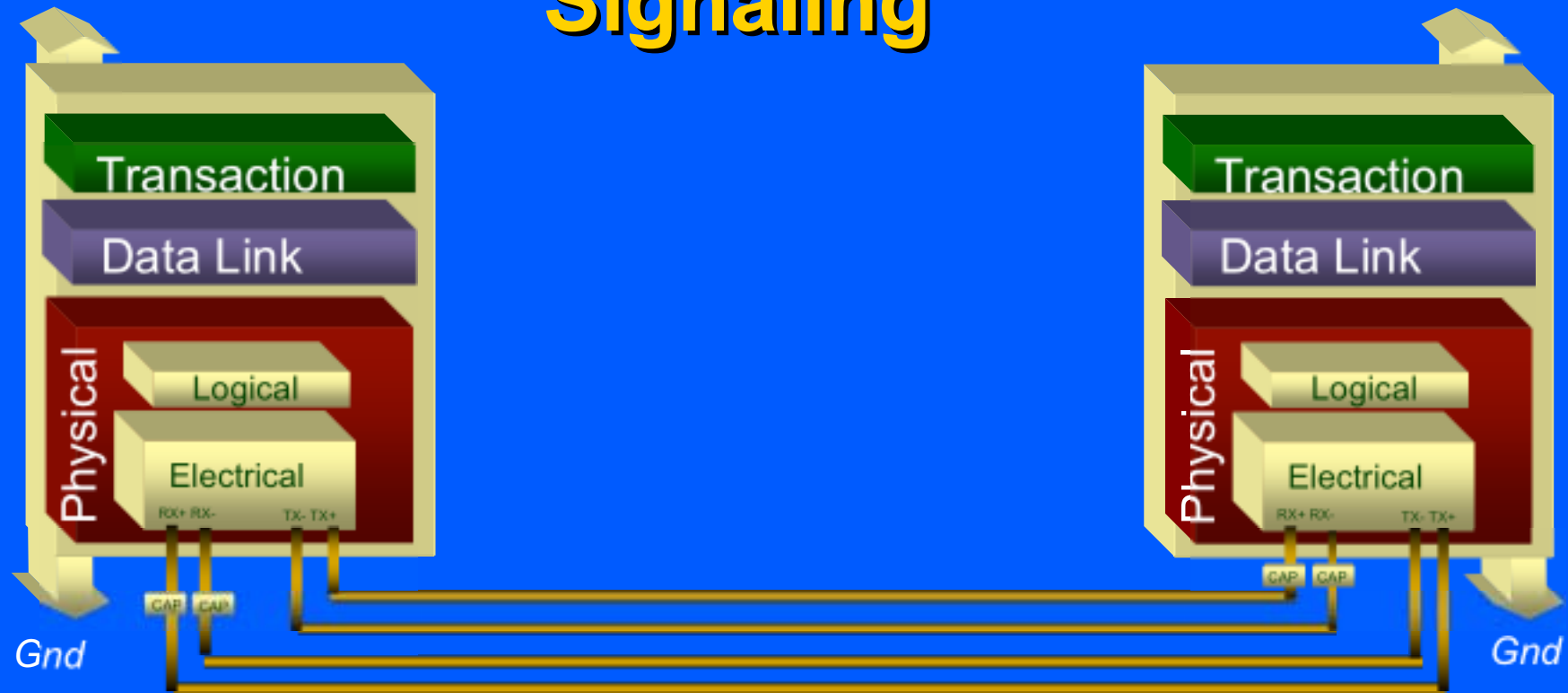
Physical Layer Basics



- Provides isolation from higher layers
- Logical Functions
 - ✓ Encoding/decoding/scrambling
 - ✓ Reset, initialization, de-skew
 - ✓ Configuration: speed, link width, lane mapping
- Electrical Functions
 - ✓ Transmitter/receiver
 - ✓ Packet exchange
 - ✓ Link power management

PHY upgrades do not affect Data Link layer and above

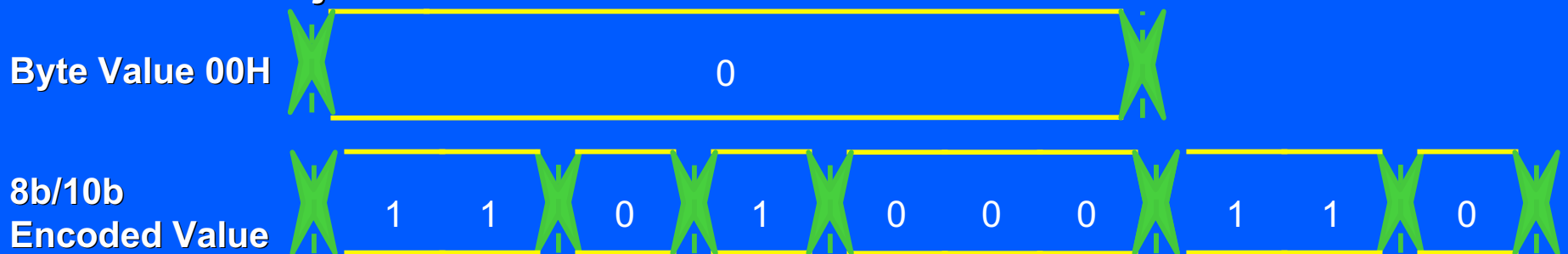
Signaling



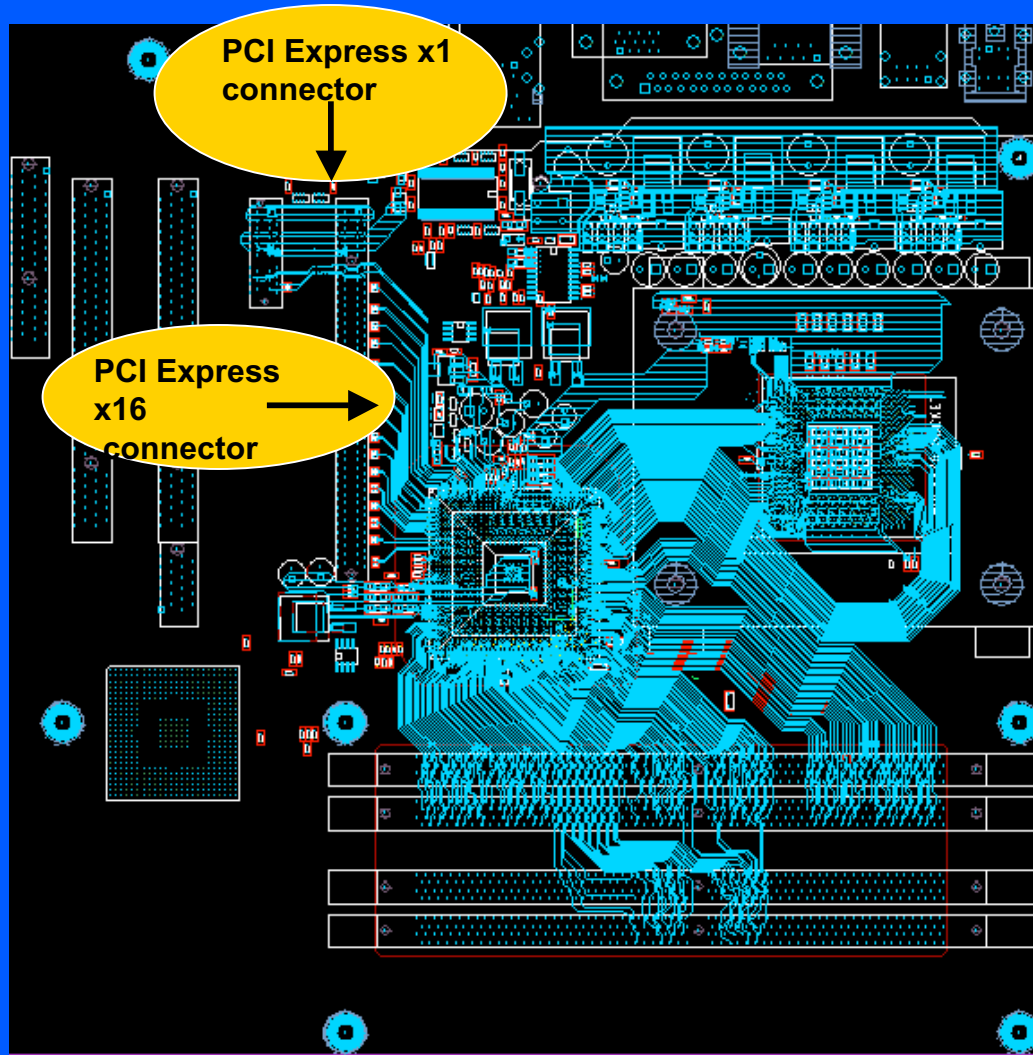
- **Differential**
 - ✓ Superior voltage margins to single-ended
 - ✓ Voltage independence; VCC can vary
 - ✓ Reduces EMI problems
- **Serial Lanes**
 - ✓ Removes skew requirements between parallel lanes
 - ✓ Relaxed routing rules within a link
- **Routing length target**
 - ✓ 4 layer FR-4 boards; ~20 inches, 2 connectors

Embedded Clocking

- **8b/10b Encoding**
 - ✓ Limits the run length of 1's and 0's
 - Data transitions enable bit level clock recovery at the receiver
- **Routing flexibility**
 - ✓ Removes skew relationships between clock and data signals
- **Higher performance**
 - ✓ High frequencies -> lossy transmission lines
 - ✓ Skew/jitter between bits



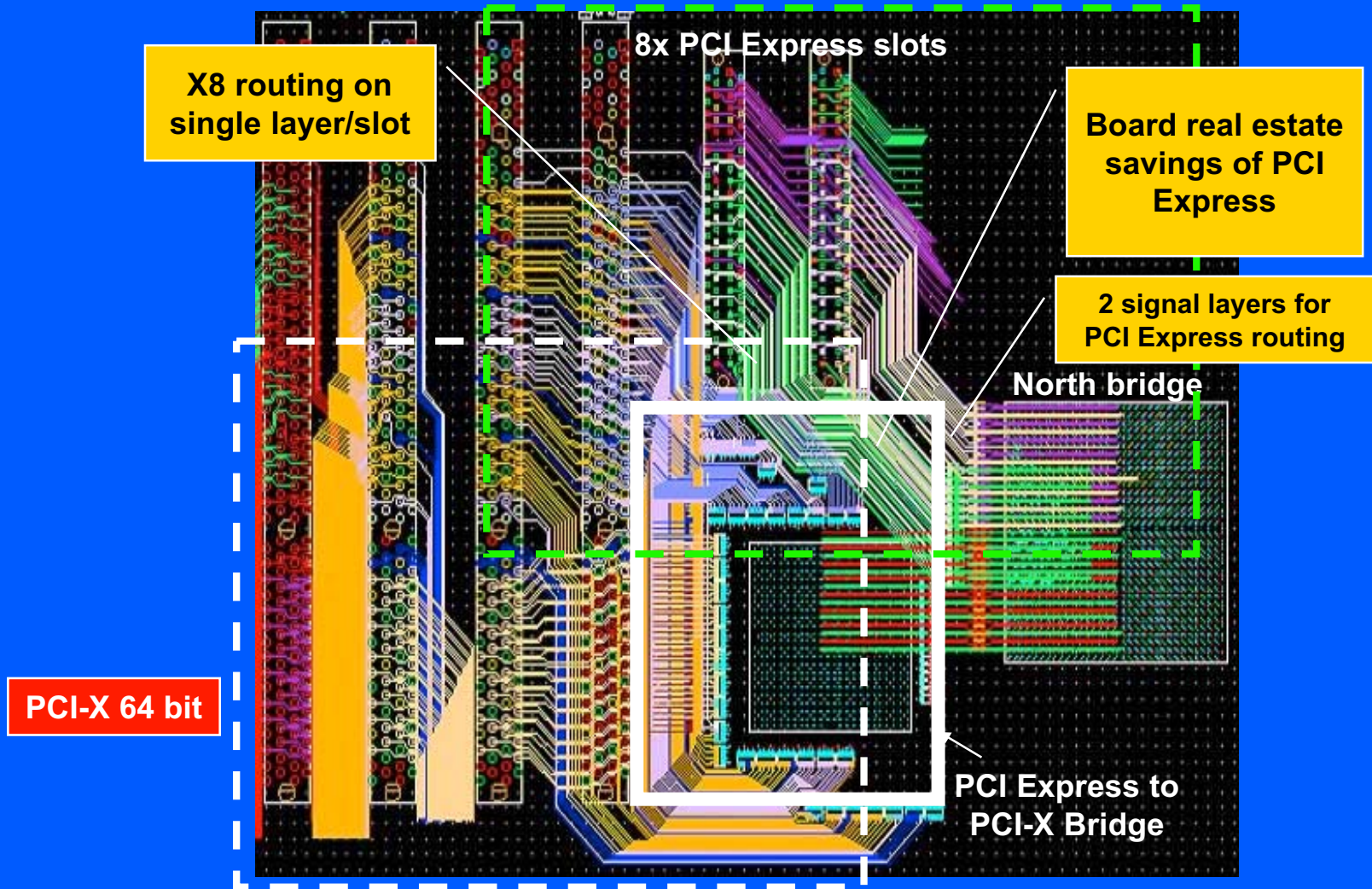
Motherboard Routing



_ATX 4 Layer, P4 Motherboard

- PCI Express* layout & connectors in 4 layers
- Graphics – x16 PCI Express
- Gen/O – x1 PCI Express
- Maintain flexibility with PCI Express & PCI

PCI Express Routing



Fewer pins and simpler routing -> Lower cost and complexity

Physical Layer Summary

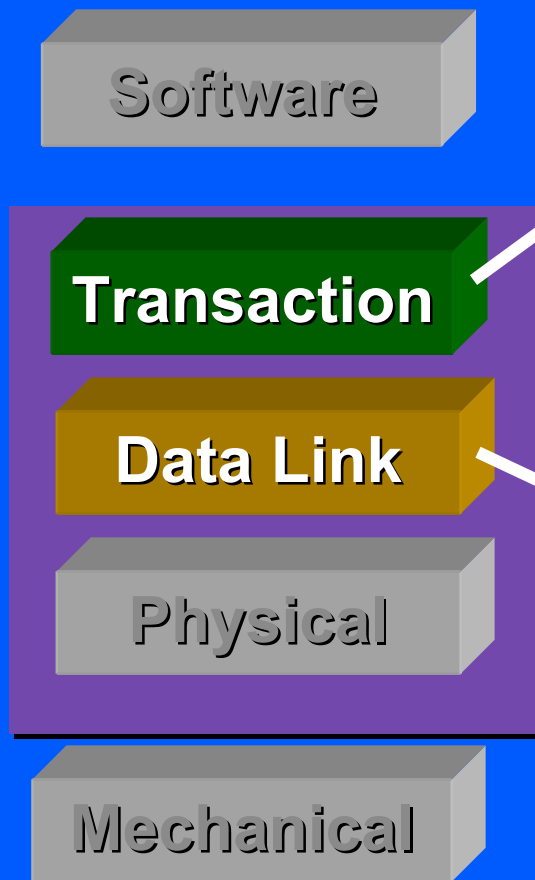
- **The PCI Express PHY technology addresses interconnect bottlenecks**
 - ✓ Balances low cost with high performance

- **Investment in PCI Express PHY technology applies today**
 - ✓ Friendly to future low voltage processes

- **Advanced features beyond transfer speed**
 - ✓ Auto-configuration, simple clock solutions, EMI friendly, choice of power management mechanisms

- **The technology is mature and silicon ready**

Packet Sources and Types



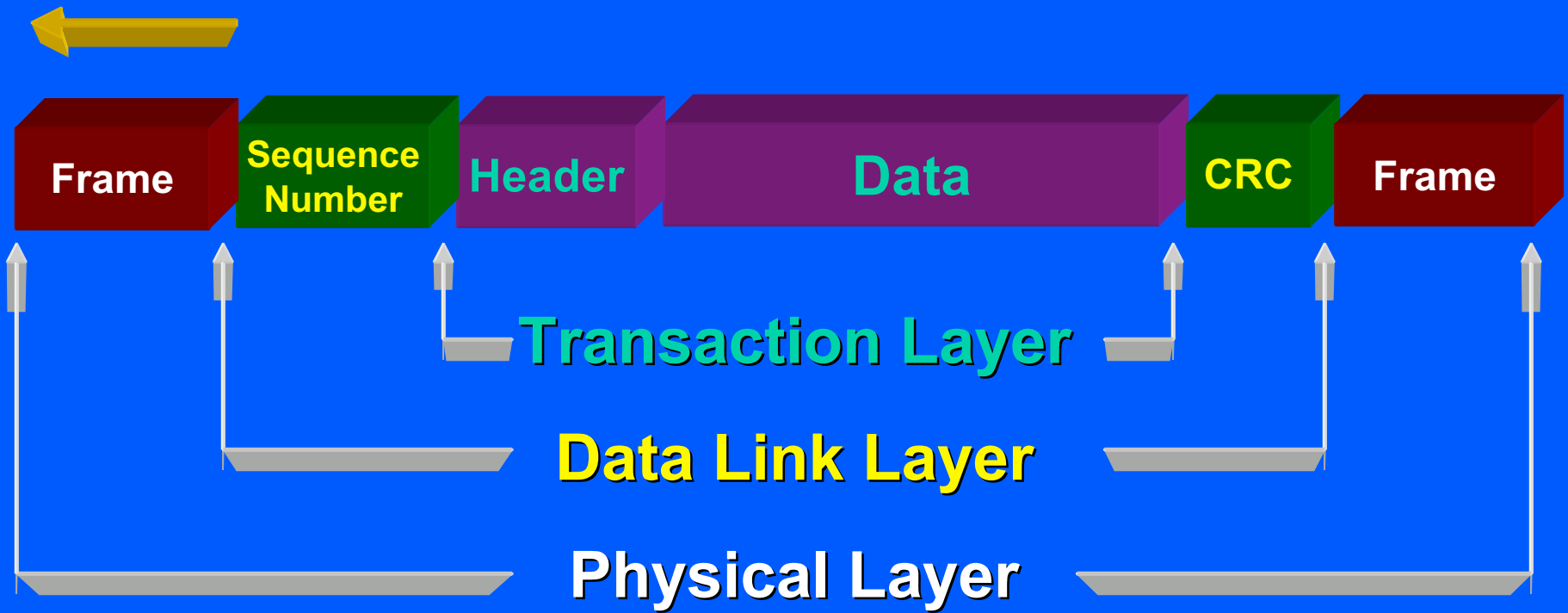
- Transaction Layer Packet (TLP)

- ✓ Memory, IO, Configuration and Message transactions
- ✓ Request and Completion packets

- Data Link Layer Packet (DLLP)

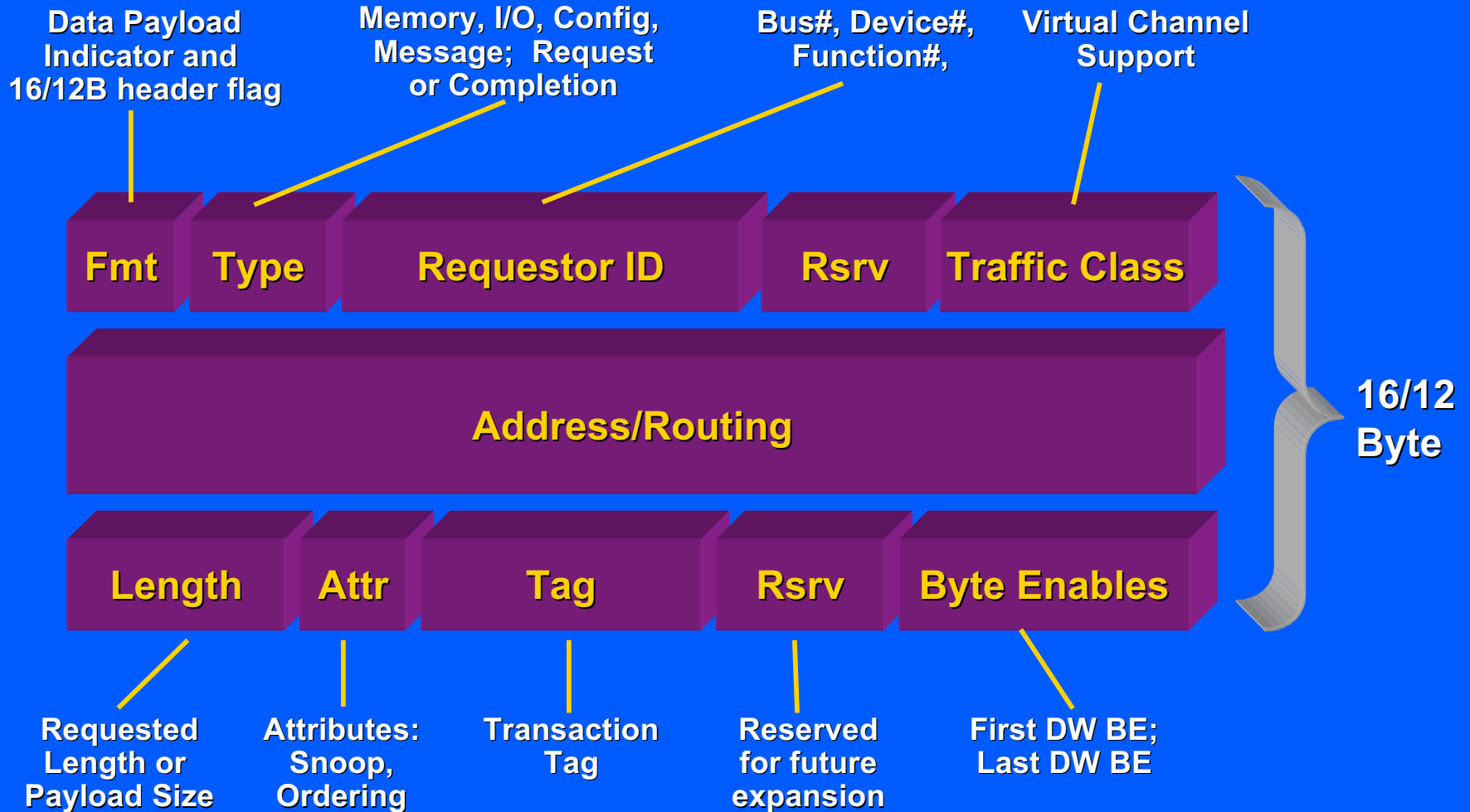
- ✓ Link Data Integrity Ack/Nack Support
- ✓ Transport mechanism for Flow Control information exchange
- ✓ Low-level Power Management

Packet Formation



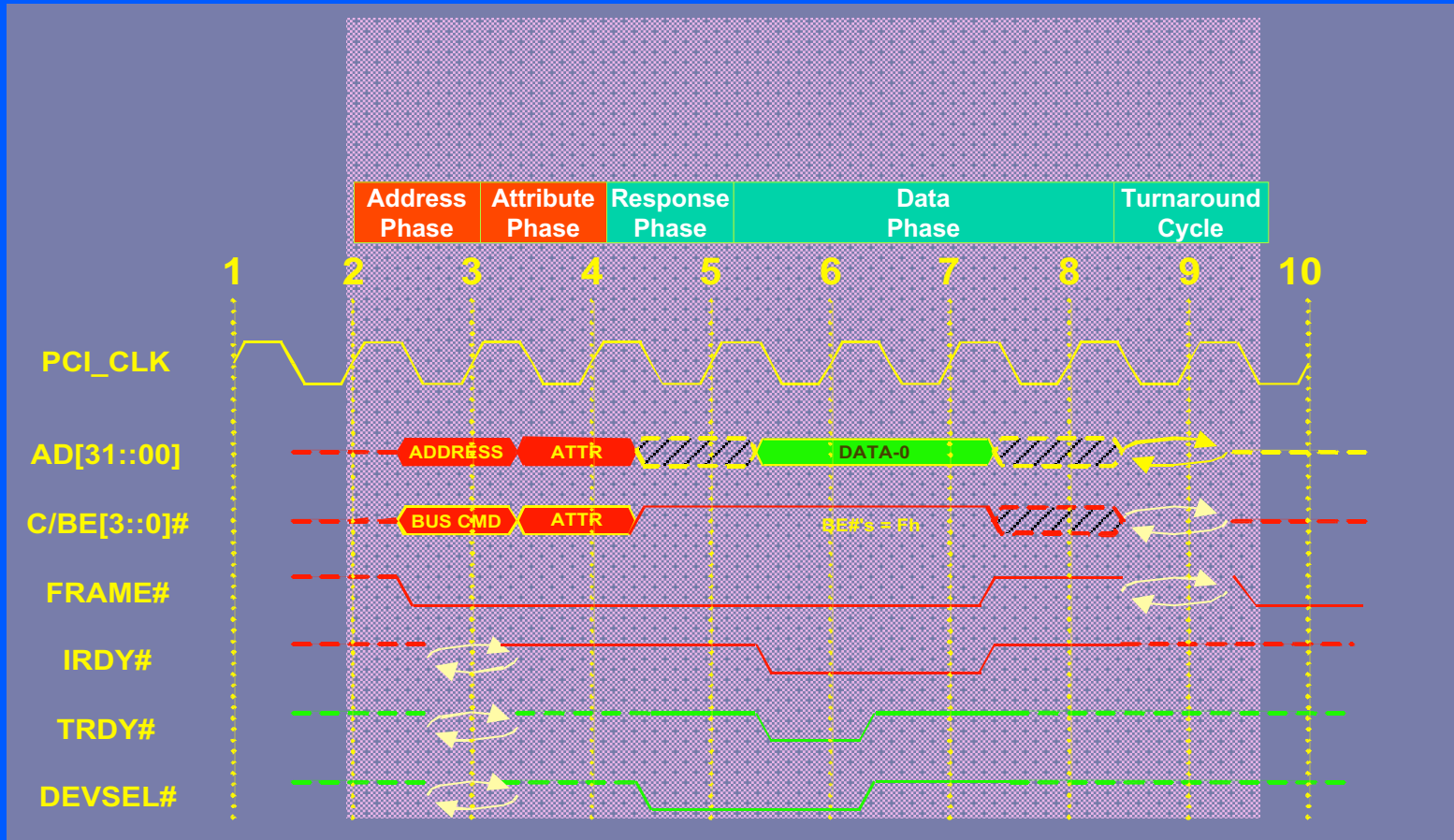
Formation of packets reflects layered architecture

Packet Headers



Packet Header Contains Address, Control and Attribute Fields

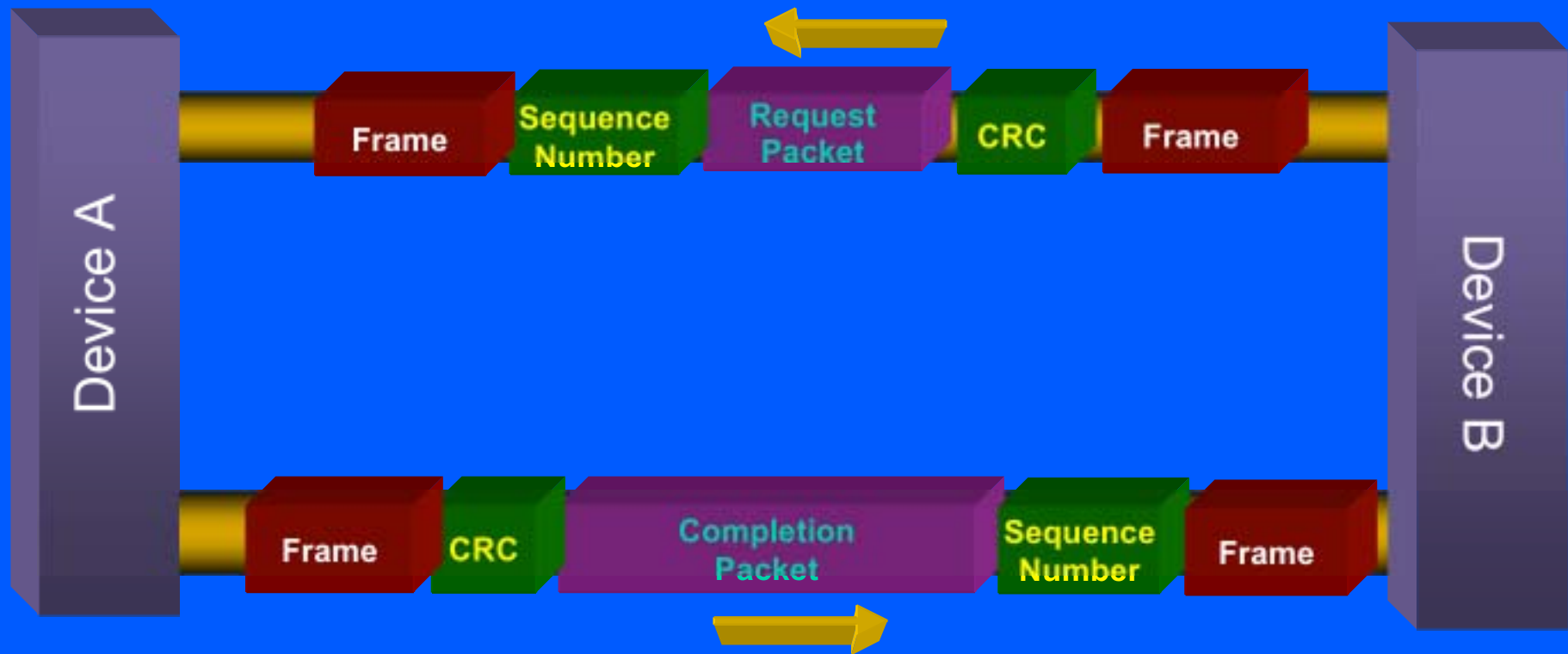
Basic PCI-X Transaction



PCI/PCI-X Transactions

- Memory transactions
- I/O transactions
- Configuration transactions

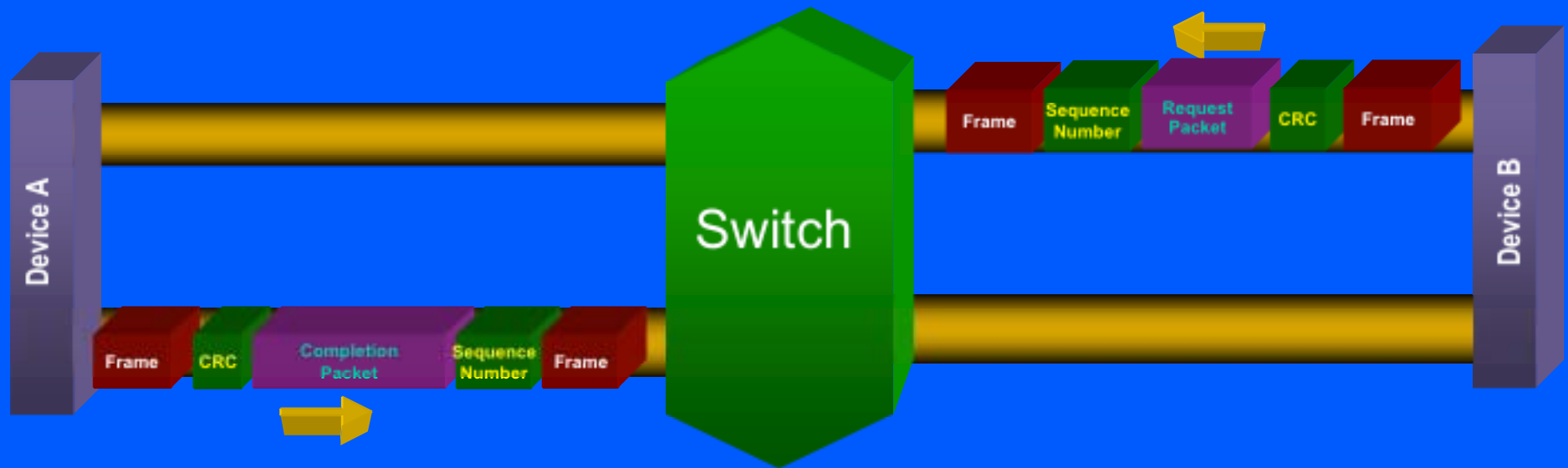
Basic PCI Express Transaction



PCI Express Transactions

- Memory transactions
- I/O transactions
- Configuration transactions
- Message transactions

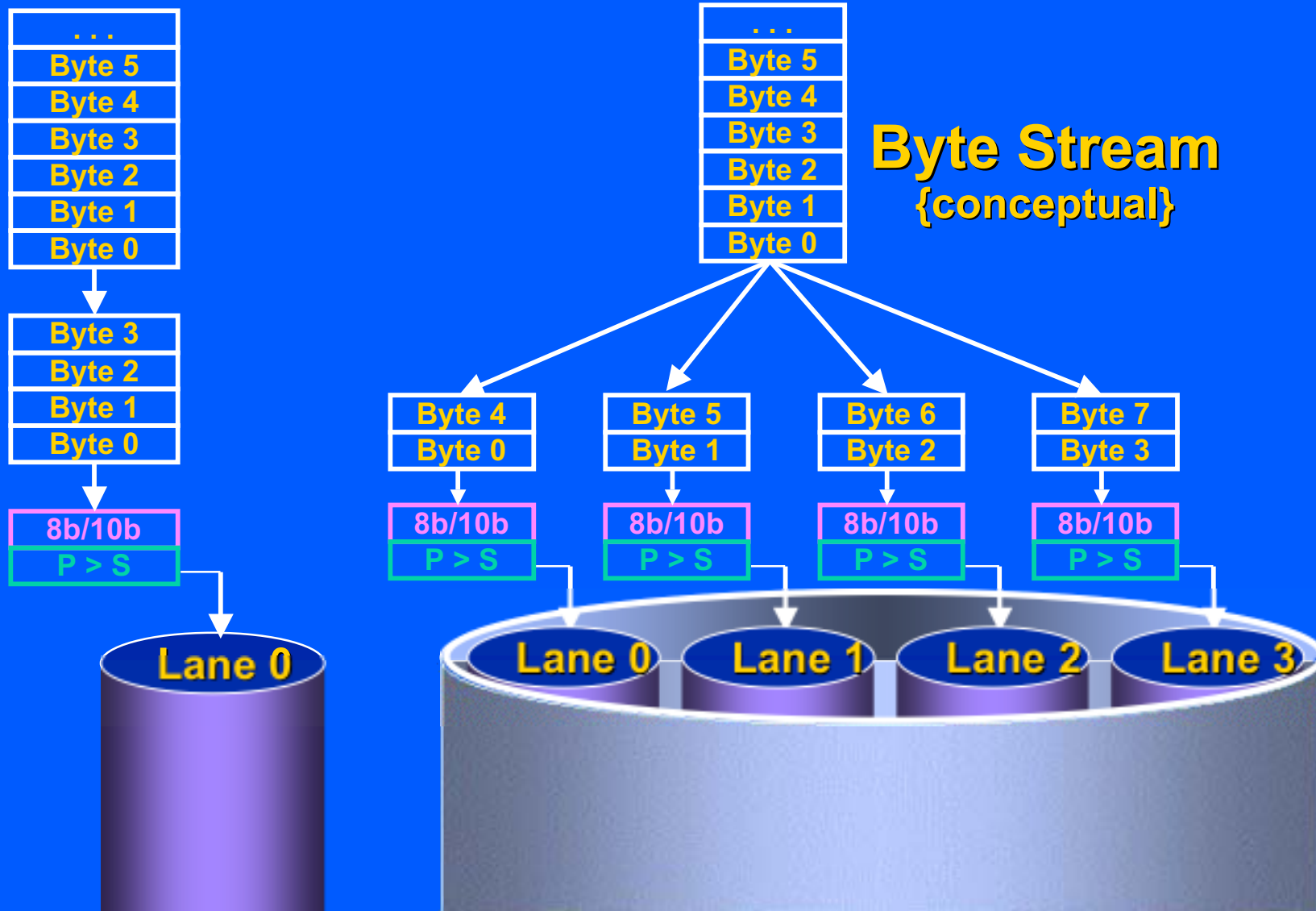
Basic PCI Express Transaction



PCI Express utilizes credit based flow control

- Flow control credits applicable per link
- Requester must accept the returning data
- ACK/NACK used only for the purpose of data integrity

PCI Express Transfers



Bandwidth is selectable using multiple lanes

PCI Software Compatibility

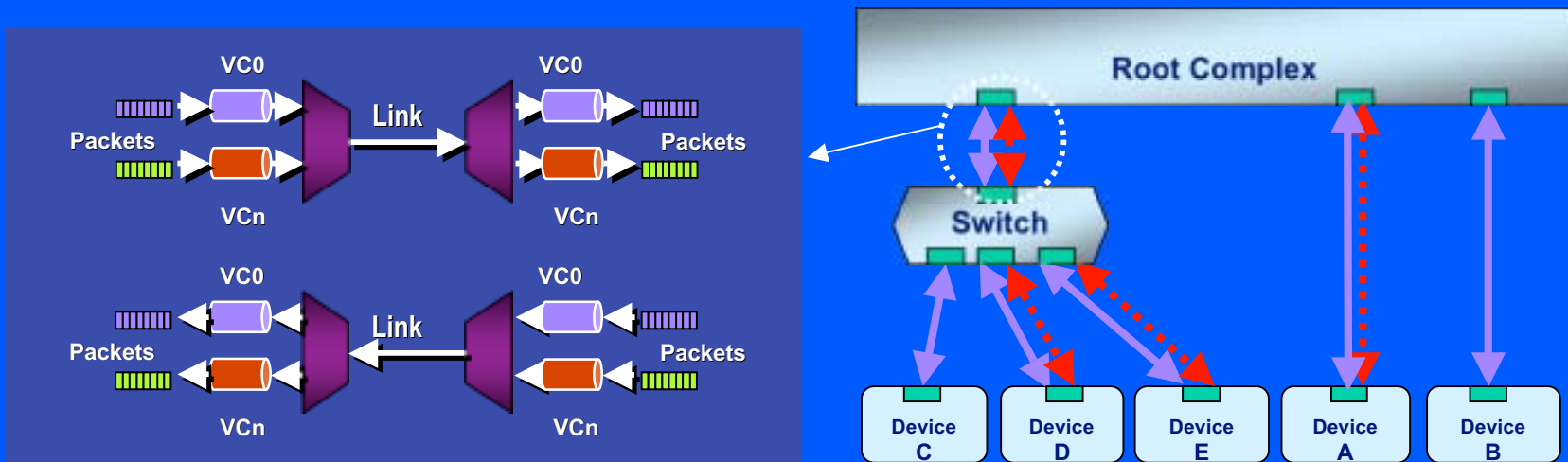
- **PCI-compatible configuration mech.**
 - ✓ Defined to boot PCI 2.2 compliant OSs unmodified
 - ✓ Compatibility with PCI device driver model and existing software stacks for optional PCI capabilities
 - PCI Power Management, SHPC Hot-plug, MSI support

- **PCI compatible in-band signaling**
 - ✓ PME supported using in-band messaging
 - ✓ MSI is the native interrupt mechanism in PCI Express

- **Enhanced Configuration Mechanism**
 - ✓ Flat memory mapped access, removes bottlenecks associated with PCI mechanism
 - ✓ Extends device configuration space
 - ✓ Improves configuration transaction performance

PCI Express does not require new OS and BIOS

Virtual Channels



- VCs are independent (ordering and flow control) paths
- Traffic Class (TC) labeling for differentiation of traffic
- Up to 8 Virtual Channels with associated servicing priorities
- Mapping of TCs to Virtual Channels for platform flexibility
- Configuration of TC/VC mapping and VC arbitration by software

Virtual Channels = Support of Qos

Power Management

- Builds on PCI Power Management (PM)
 - ✓ Compatible with existing PCI PM software stacks
- Device PM States: D0-D3_{hot/cold}
- Link PM States: L0, L0s, L1-L3
- Enhanced PM capabilities
 - ✓ Aggressive power reduction through Active State PM (L0s, L1)
 - ✓ Improved PME using in-band messaging
 - ✓ Improved definition and SW control of Vaux

PCI Express Advances Platform PM While Preserving Software Investment

Data Integrity Support

- Requirements for Robust Data Integrity
- Data Link Layer Mechanisms (Link/local):
 - ✓ TLPs protected using 32bit CRC
 - ✓ DLLPs protected using 16bit CRC
 - ✓ TLP error recovery through Link-level retry
 - ✓ Supplemental coverage through 8b/10b
 - ✓ Loss of packets detected using Sequence Numbers
- Transaction Layer Mechanisms (End-to-End):
 - ✓ Optional coverage using 32bit CRC
 - ✓ Data Poisoning capability

**Robust Data Integrity Allows for Signaling
Frequency Headroom**

RASUM Capabilities

- RASUM = Reliability, Availability, Serviceability, Usability and Manageability
- Server, Communication and Client platforms
- PCI Express Portfolio
 - ✓ Data Integrity
 - ✓ Error Signaling and Logging
 - ✓ Hot-Plug/Swap and Surprise Removal
 - ✓ Fabric Resource Management

**Standard Scaleable RASUM Capabilities from
Server to Client**

PCI Express Performance



Best Bandwidth/Pin -> Best Value

Protocol/Software Summary

- **PCI Express advances overall platform while preserving PCI investments**
 - ✓ PCI Express Layering provides architecture longevity
 - ✓ Defined to boot PCI 2.2 compliant OSs unmodified

- **Preserves PCI Architecture while addressing basic limitations by defining new features**
 - ✓ Traffic Classes, Isochrony & QoS
 - ✓ Native Hot-plug/Surprise removal
 - ✓ Advanced RAS/Error Reporting
 - ✓ In Band messages
 - ✓ “Hooks” to enable advanced switching

PCI Express is the PCI you know and Love

PCI Express Mechanicals Scope

Card & Connectors

- Covers today's Desktop and Server applications
- Require no change to the system form factors

Mini PCI Express

- Covers Mobile Comm. applications
- Build to Order/Configure to Order

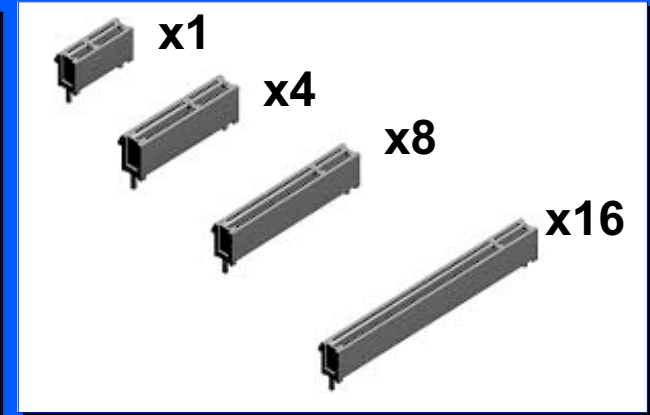
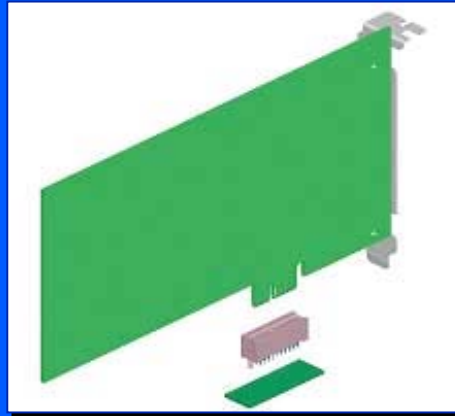
Modules

- Break away from 20 years of legacy form factors
- Suitable for Desktops, Mobile and Servers
- Hot Pluggable/Swappable

Comm. Backplanes Cables

- Opportunity to extend to back-plane and cabled applications

PCI Express DT Mechanicals



Card

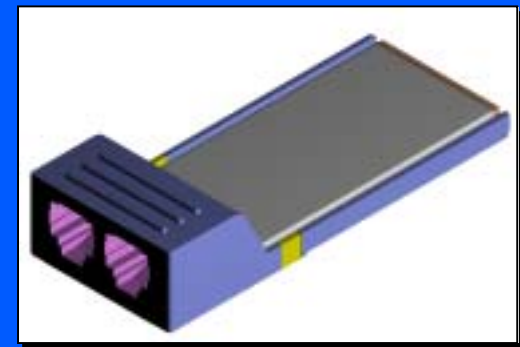
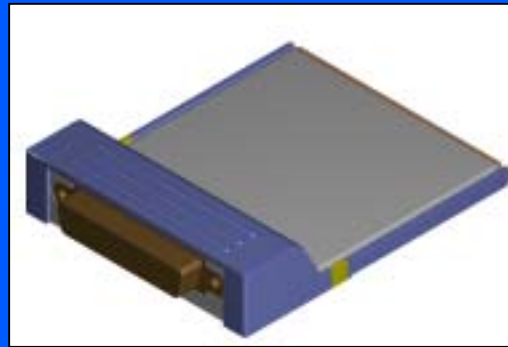
Connectors



Modules



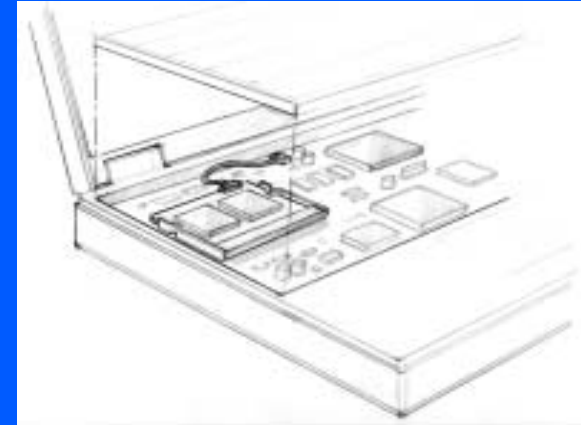
Cables



Mini-PCI-Express Module Overview

■ The Problem:

- ✓ Mini-PCI is too slow for Gigabit Ethernet.
- ✓ Many mini-PCI variations



■ The Solution is Mini PCI Express:

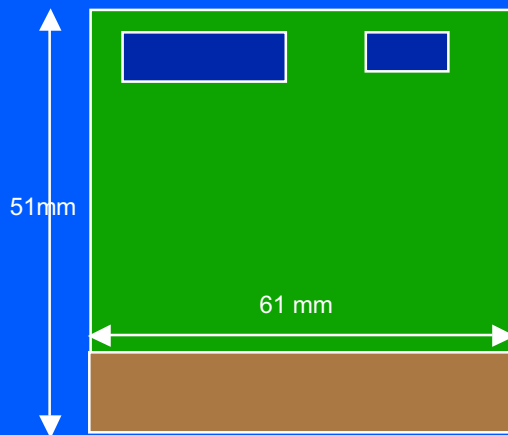
- ✓ An open standard BTO/CTO solution
- ✓ Higher bandwidth interface
- ✓ Balanced between power needs and form factor limitations
- ✓ Wireless apps need more power
- ✓ PCI-Express has better Power Management

Opens the door to new applications

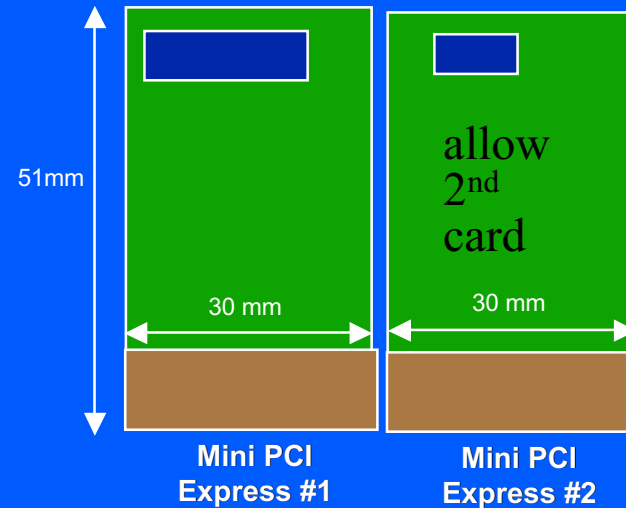
Card Sizes

- Size optimized for single function card
 - ✓ Half size of today's Mini-PCI Type III card
- Defined area for I/O connectors
 - ✓ LAN, modem, & antenna signals for wireless applications

Mini-PCI Type III Card

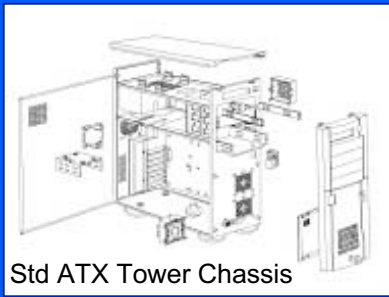


Mini PCI Express Cards

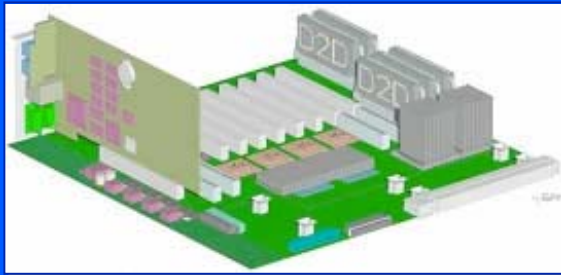
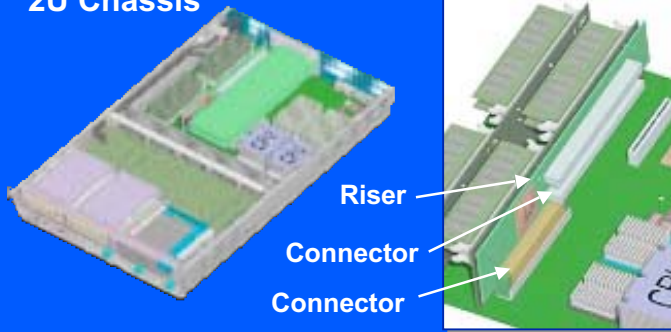


Removes need for dual functions cards

PCI Express Server Mechanicals



2U Chassis



- **Uses Existing Server Chassis**
 - ✓ Does not require chassis changes

- **Can Be Implemented in Rack Mount Server Chassis**
 - ✓ Can be routed through two connectors and riser card

- **Capable of Transition with PCI-X**

PCI Express does not require chassis changes

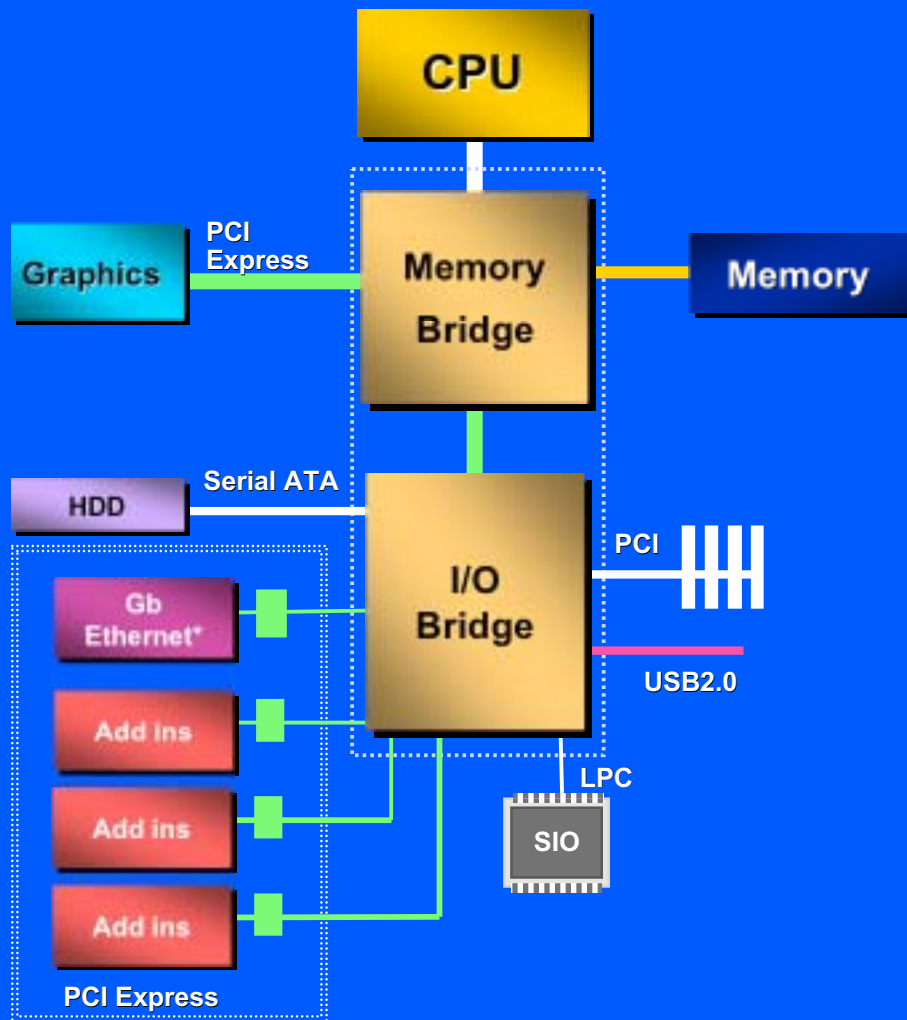
Mechanical Summary

- **PCI Express is Optimized for Cost**
 - ✓ Cost effective for migration into commodity infrastructure
 - ✓ Replaces PCI over time with 10+ years of life
- **PCI Express is Easy to Implement**
 - ✓ Does not require chassis changes in Desktops and Servers
 - ✓ Transition with Existing PCI form factors
- **PCI Express Provides Opportunities for Innovative Design**
 - ✓ Benefits for Industry and end users

Agenda

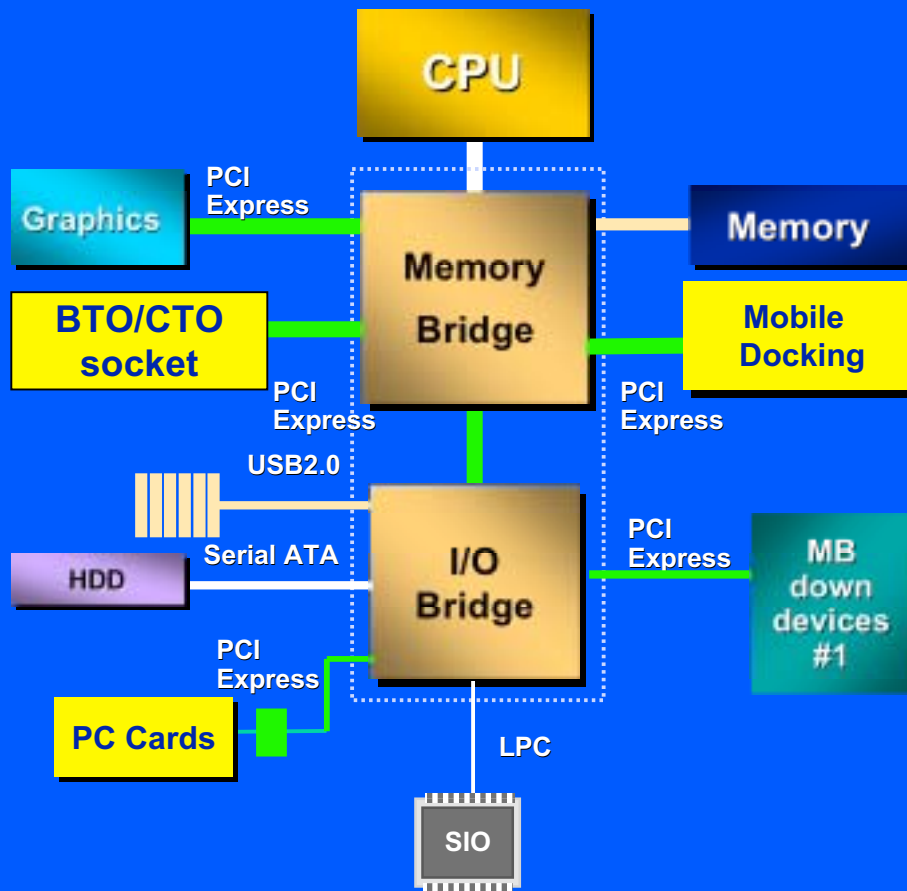
- Motivation and Requirements
- PCI Express Architectural Overview
 - ✓ Physical Layer
 - ✓ Transaction and Data Link Layers
 - ✓ Mechanicals and Form Factors
- ■ **PCI Express based Platforms**
- Call to Action

Desktop Platforms



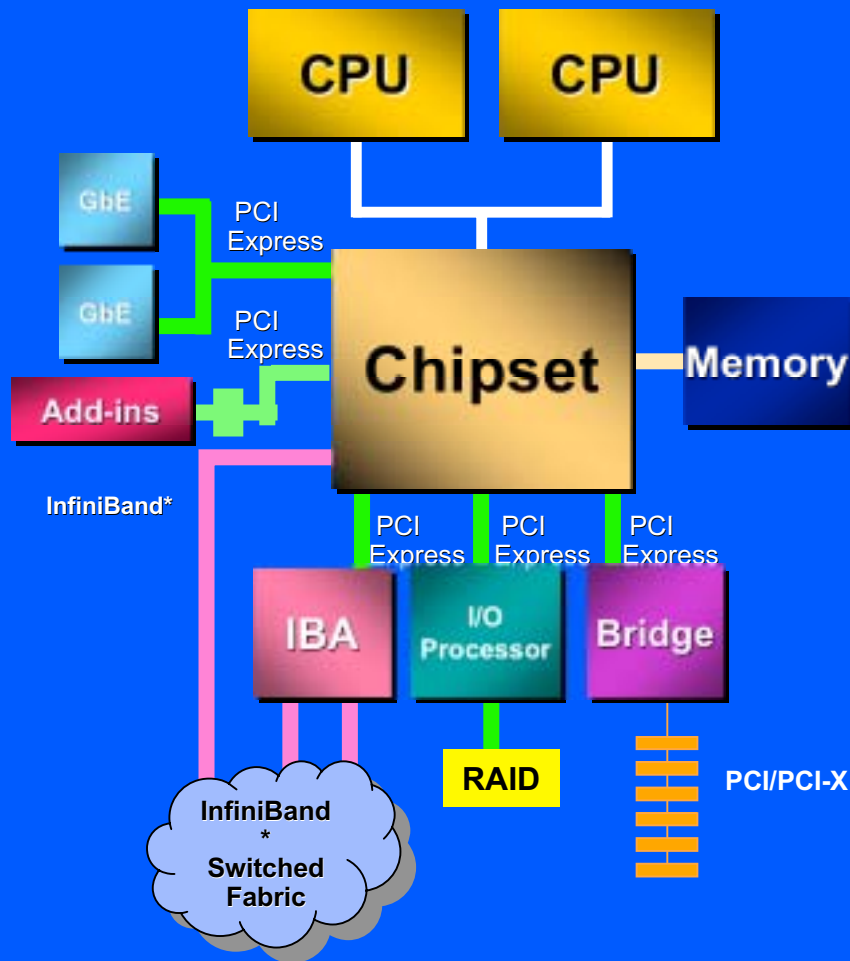
- Low Cost
 - ✓ Maximize reuse: design, validation
 - ✓ 4 layer motherboard
 - ✓ Lane reversal; flexible routing
- Software:
 - ✓ PCI compatible software model
 - ✓ ACPI and PCIPM compliant PM
- Performance:
 - ✓ Width: x1 and x16 interface
 - ✓ Isochrony done right
- Physical:
 - ✓ Card and Module form factors
 - ✓ Hot attach/Detach
 - ✓ Power Budgeting

Enables Next Generation Graphics, Multimedia and LAN



- Power Management:
 - ✓ PCIPM and ACPI compliant PM
 - ✓ Quick entry /exit from PM states via physical layer
- Physical
 - ✓ Differential low voltage signaling
 - ✓ Flexible routing
 - ✓ BTO/CTO form factors
 - ✓ Cables and docking
 - ✓ Hot attach/surprise removal

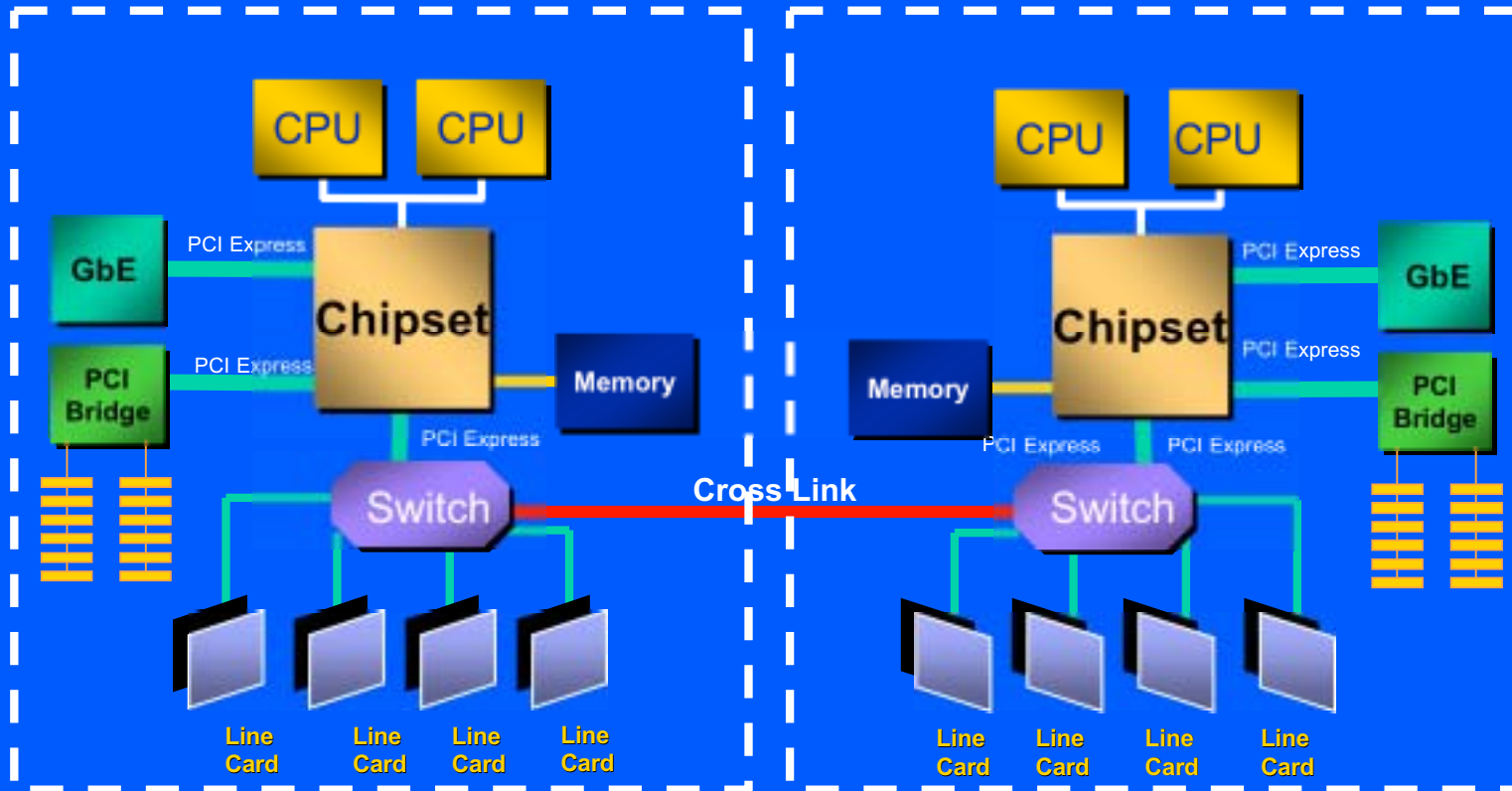
Featuring APM/ Form Factors/Docking



- **RAS Capabilities:**
 - ✓ 32bit CRC protection & link-level retry
 - ✓ End to End CRC Protection
 - ✓ Rich set of error logging and reporting
 - ✓ PCI PM and SHPC SW compatible
 - ✓ Power Budgeting
- **Performance and scalability**
 - ✓ Lower system latency and best bandwidth
 - ✓ Better connectivity and fan out
 - ✓ True QoS with Virtual Channels
- **Physical**
 - ✓ Max. Trace lengths >2X PCI-X 2.0 in FR4
 - ✓ Lane reversal to assist routing short distances (e.g. add-in card)
 - ✓ New signaling medium possibilities (e.g. cables)

Higher Performance, RAS Capability

PCI Express in Communications



- Advanced Switching Features:

- ✓ Multi-hierarchy support, Peer-to-peer switching, QoS features
- ✓ Advanced switching and message passing primitives
- ✓ Enables development of sophisticated upper layer protocols

Agenda

- Motivation and Requirements
- PCI Express Architectural Overview
 - ✓ Physical Layer
 - ✓ Transaction and Data Link Layers
 - ✓ Mechanicals and Form Factors
- PCI Express based Platforms
- ■ Call to Action

Summary

- **PCI Express is the PCI architecture for next decade**
 - ✓ Addresses limitations of the PCI and provides necessary extensions
 - ✓ PCI Express architecture to provides smooth migration to future
 - ✓ Provides Stable/scalable/extensible architectural foundation

- **Broad range of industry experts have contributed to the robust draft specification**
 - ✓ Features applicable to multiple market segments

- **High volume Industry implementations in progress for product introductions starting 2H '03**
 - ✓ Utilize existing system and software infrastructure without any changes

PCI Express* Specifications

- PCI Express Base Specification 1.0
- PCI Express Card Electro-mechanical Specification 1.0

Both
Available
NOW!!!

Visit www.pcisig.com for additional
information
or
contact PCI-SIG at **503.291.2569**

Call to Action

- Stay engaged with PCI-SIG for Specs, engineering collaterals and compliance
- Include PCI Express technology in your product roadmap
- Start developing designs for 2H'03 product introduction

Visit www.pcisig.com for additional information
or
contact PCI-SIG at 503.291.2569