

# Linux User Group of Davis

Marc J. Miller

Strategic Alliance Manager, AMD

October 7, 2003

## - *The Six System Challenges*

**#6: Watt density:**

**#5: The I/O infrastructure:**

**#4: Addressable memory:**

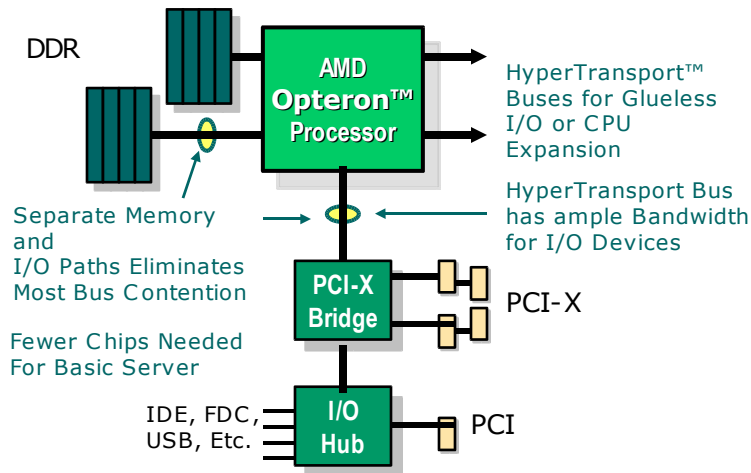
**#3: Memory bandwidth:**

**#2: Cost per processing node:**

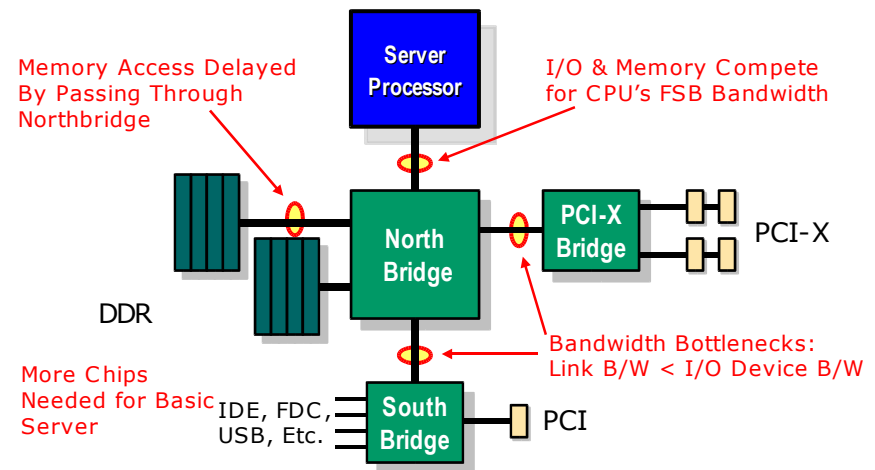
**#1: Backward compatible to x86-32:**

- ❑ There is a enormous investment is IA32 for all market segments. In many applications, porting code is not an option.
  - *Provide a solution that is not only 100% **backward-compatible**, but designed to run IA32 code faster than any existing 32-bit architecture available.*
  - *Provide a gradual and controlled migration path for porting to AMD64*
  - *Make the total cost of ownership minimal.*

## AMD64 System



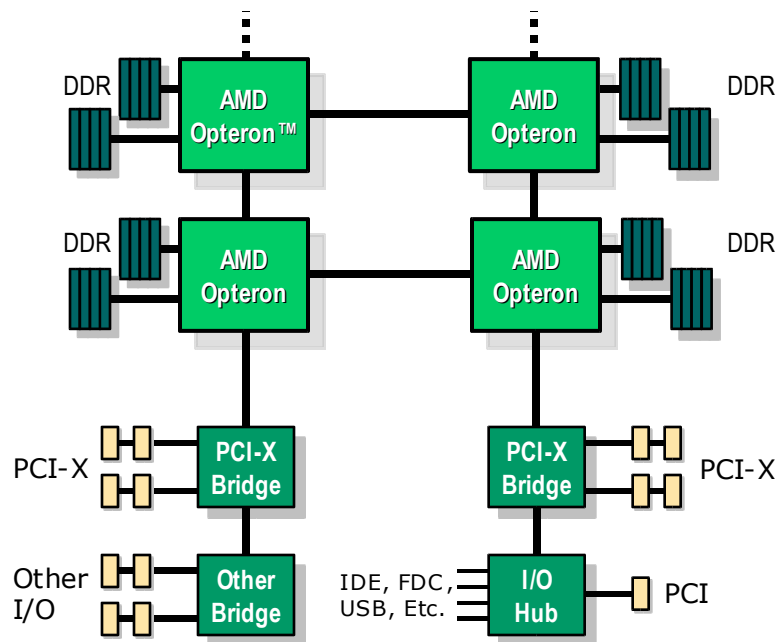
## Typical System



- Integrated memory controller
  - Low latency memory access speeds processing
- Separate Memory and I/O pathways
  - Eliminates I/O and memory bus competition
- Each processor has more memory & I/O paths
  - Memory and I/O bandwidth scales well
- Modular glueless logic using HyperTransport™ technology bus
  - Fewer chips and lower cost implementation

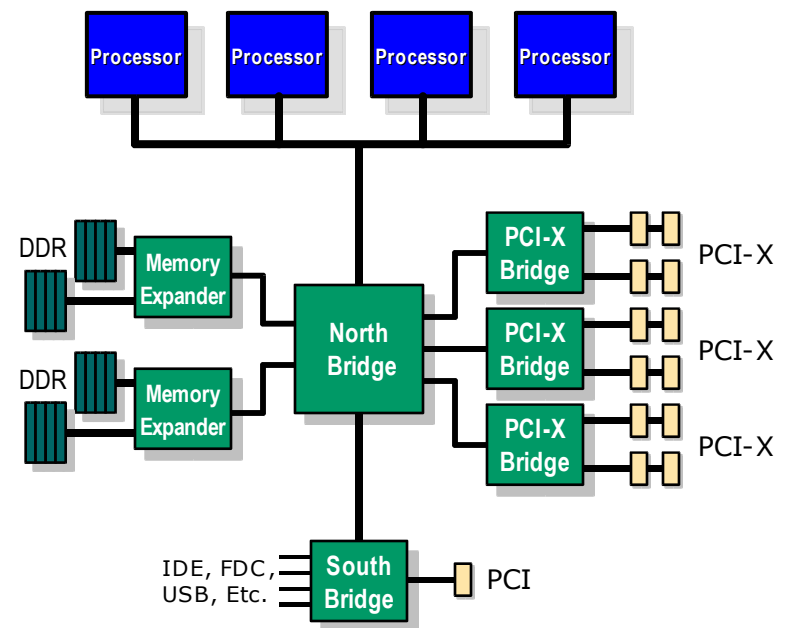
- Must access memory through Northbridge
  - Longer latency memory access
- Memory and I/O access on the same bus
  - I/O and Memory compete for bandwidth
- Memory or I/O paths originate from Northbridge
  - Bandwidth does not scale well with more CPUs
- System logic uses more chips and many buses
  - Systems cost more to design, build and test

## AMD Opteron™ System



- Scalable memory and I/O bandwidth
  - Up to 8 processors without glue logic
  - Each processor adds more memory
  - Each processor adds additional HyperTransport™ technology buses for more PCI-X and other I/O bridges
  - Fewer chips required

## Typical MP System



- System scalability limited by Northbridge
  - Maximum of 4 processors
    - Processors compete for FSB bandwidth
  - Memory size and bandwidth are limited
  - Maximum of 3 PCI-X bridges
  - Many more chips required

- System scalability limited by **Northbridge**

-Max of 4 processors

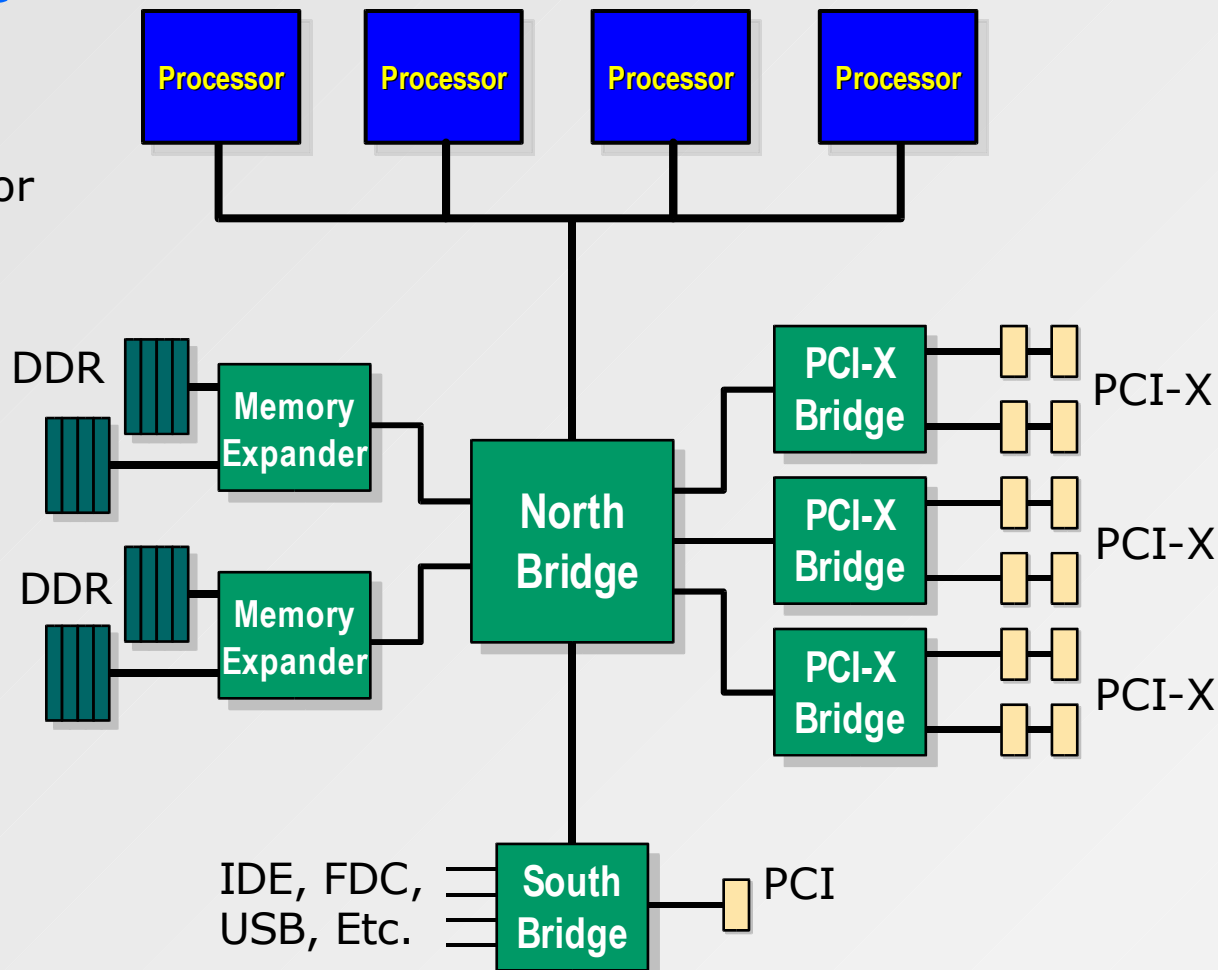
-Processors compete for FSB bandwidth

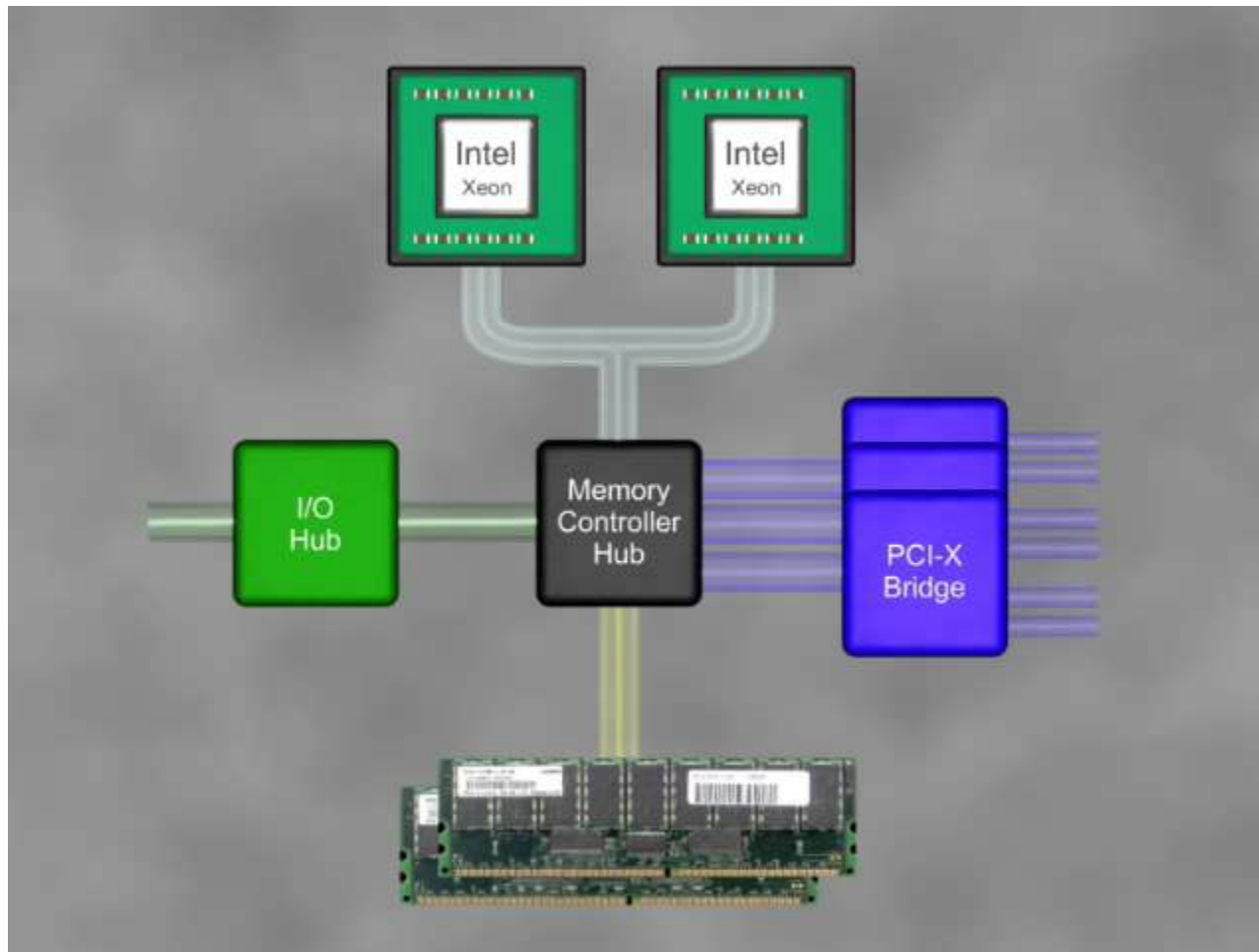
-Memory size and bandwidth are limited

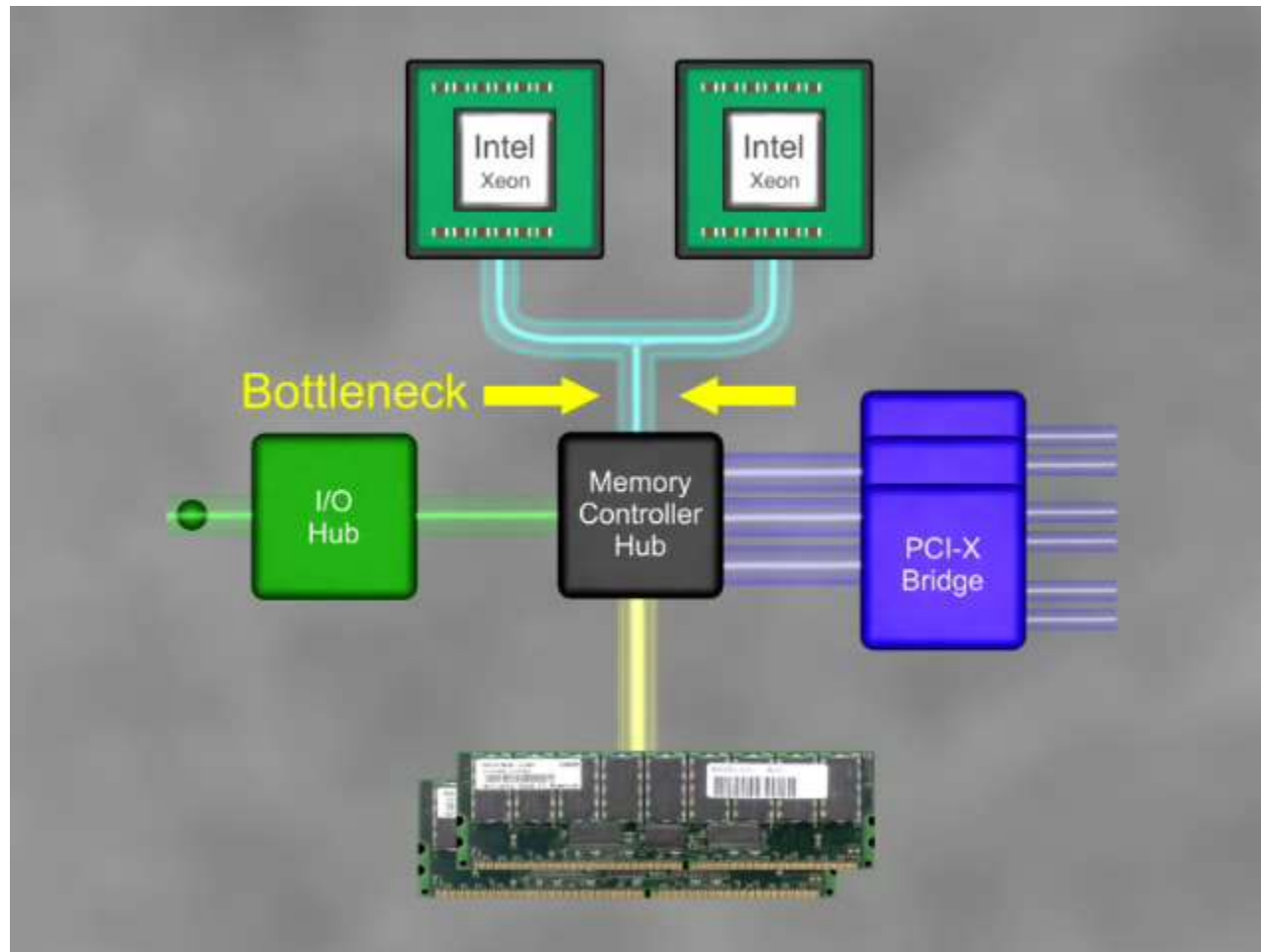
-Max of 3 PCI-X bridges

-Many more chips required

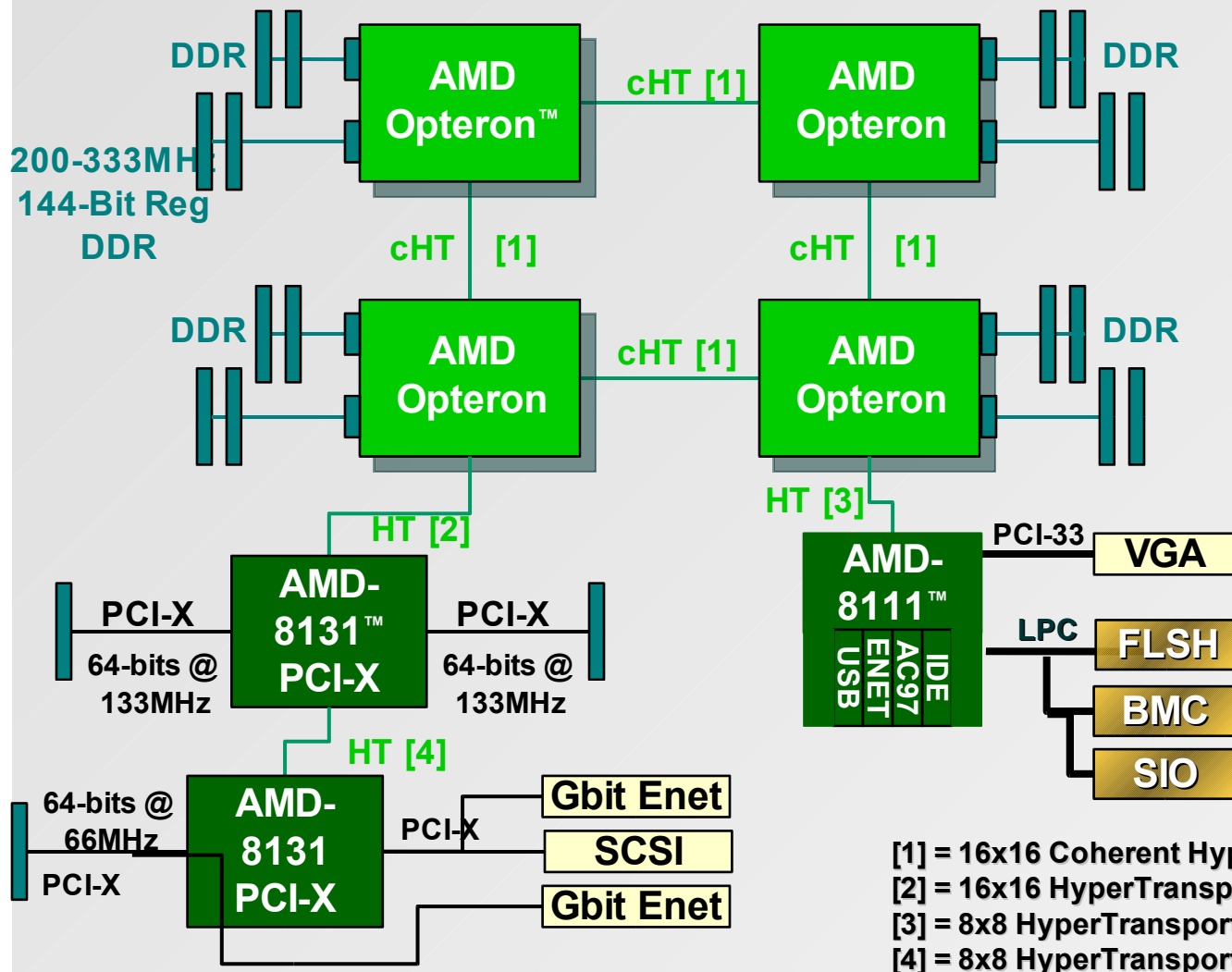
## Typical MP System





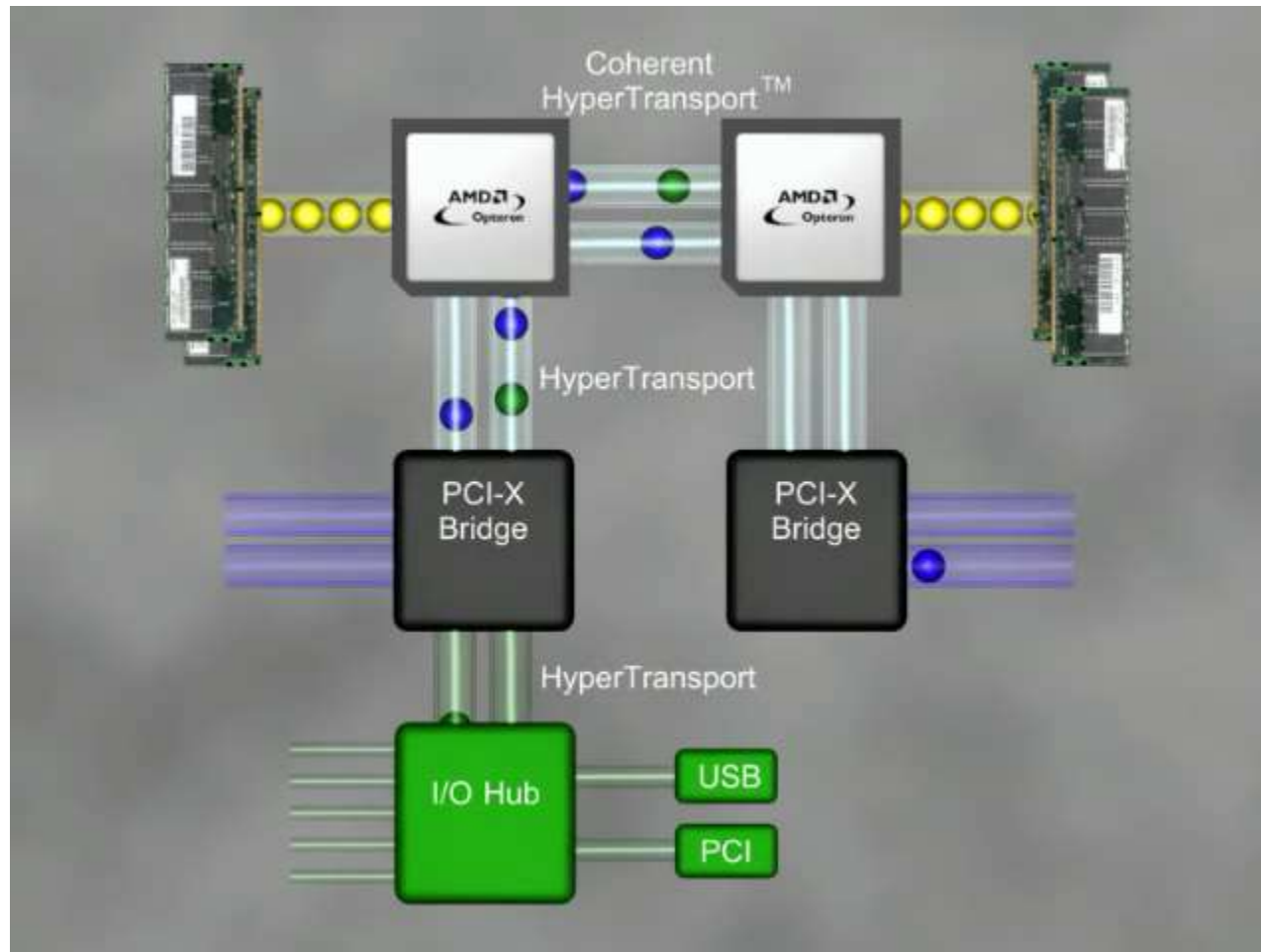


## Processor-based Server



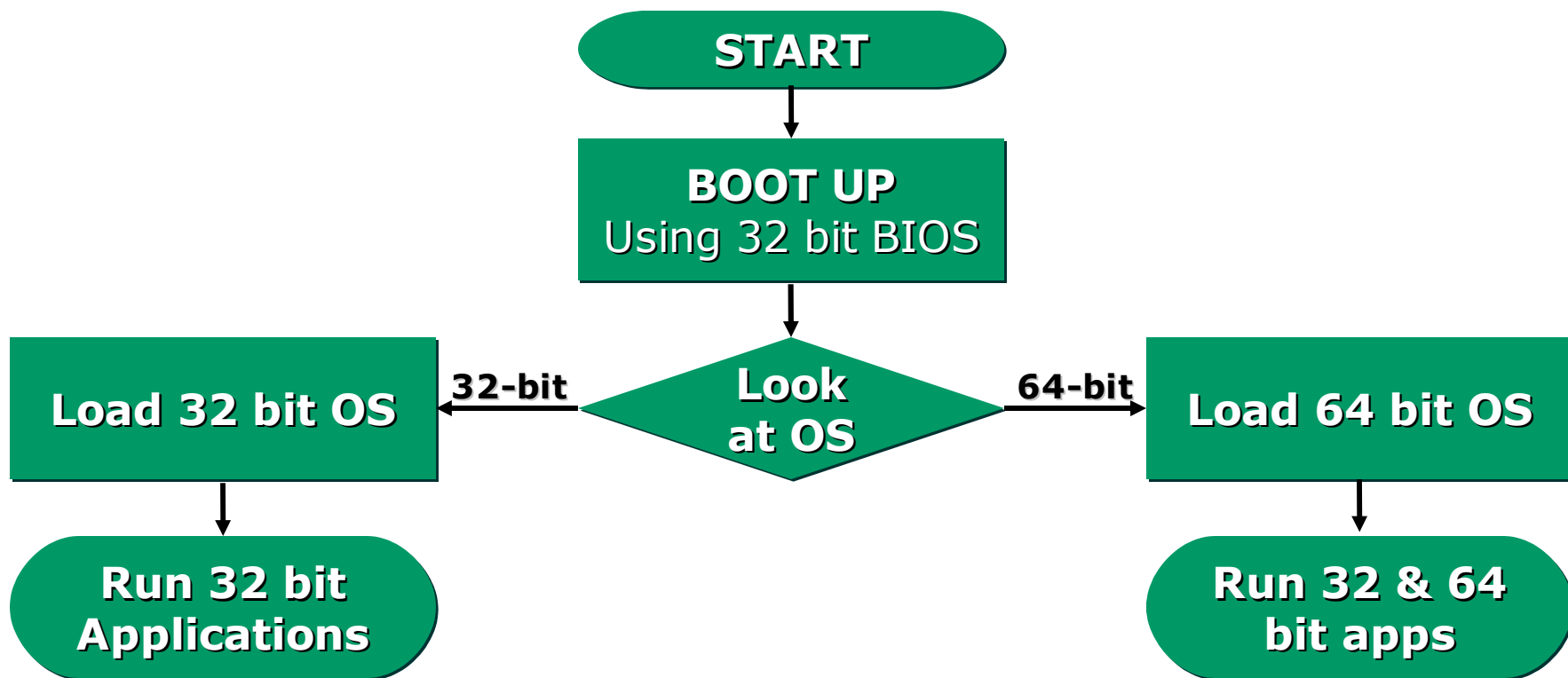
- Idle Latencies to First Data
- 1P System: <80ns
- 0-Hop in DP System: <80ns
- 0-Hop in 4P System: ~100ns
- 1-Hop in MP System: <115ns
- 2-Hop in MP System: <150ns
- 3-Hop in MP System: <190ns



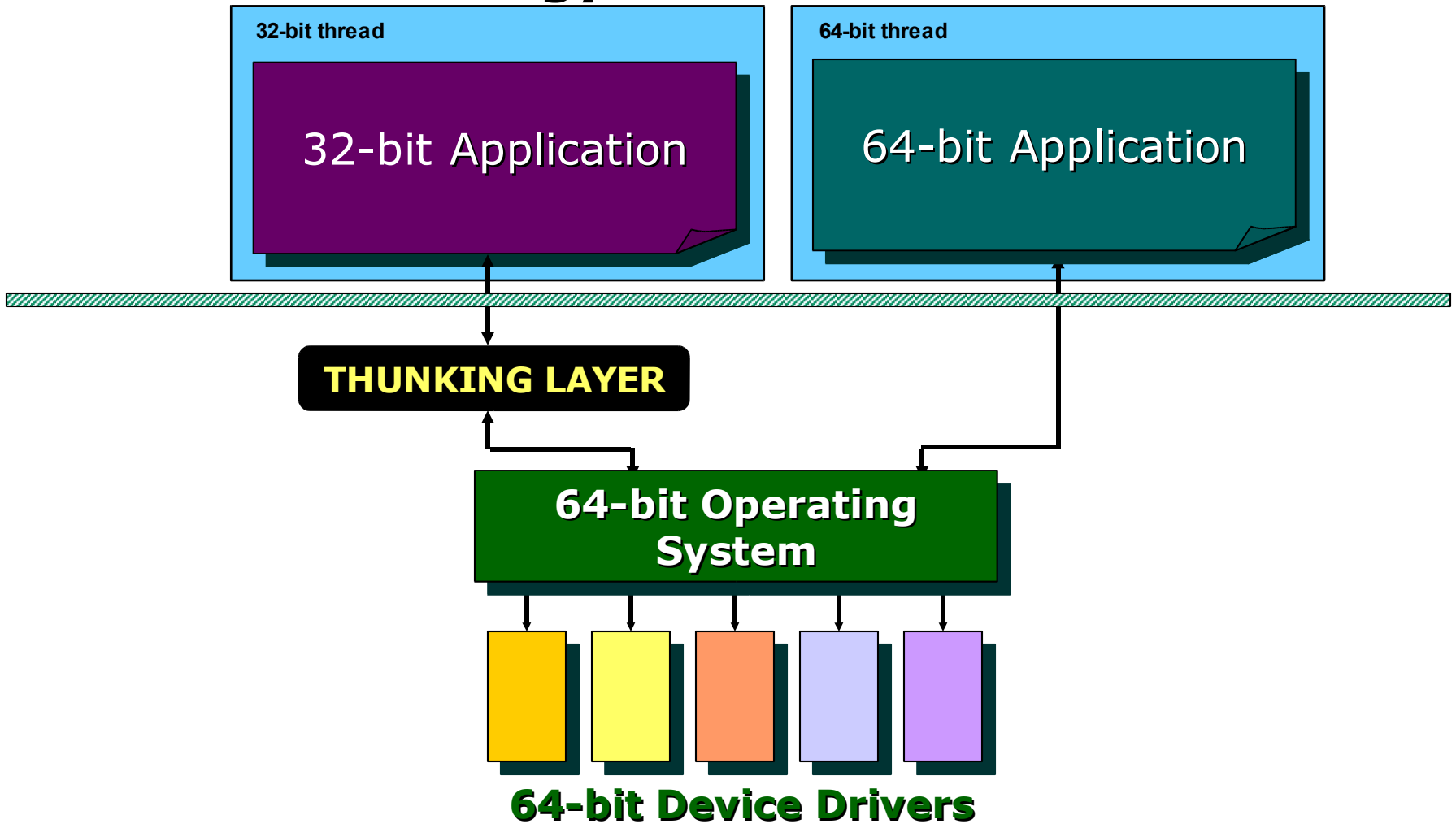


## AMD64 Technology

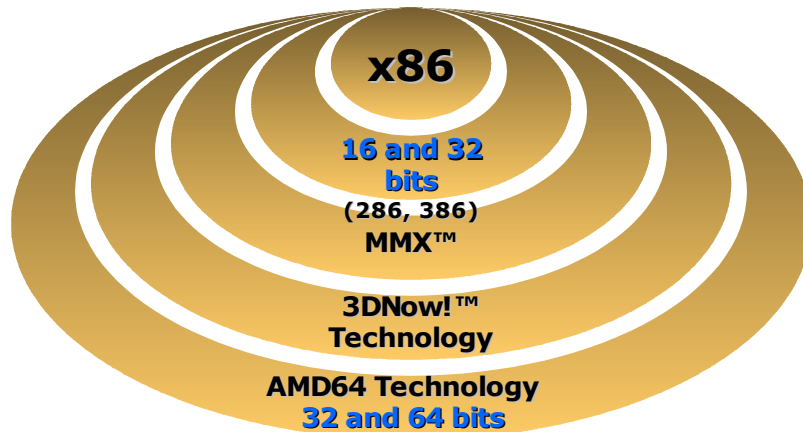
*An AMD64 PC can run  
both 32- and 64-bit operating systems*



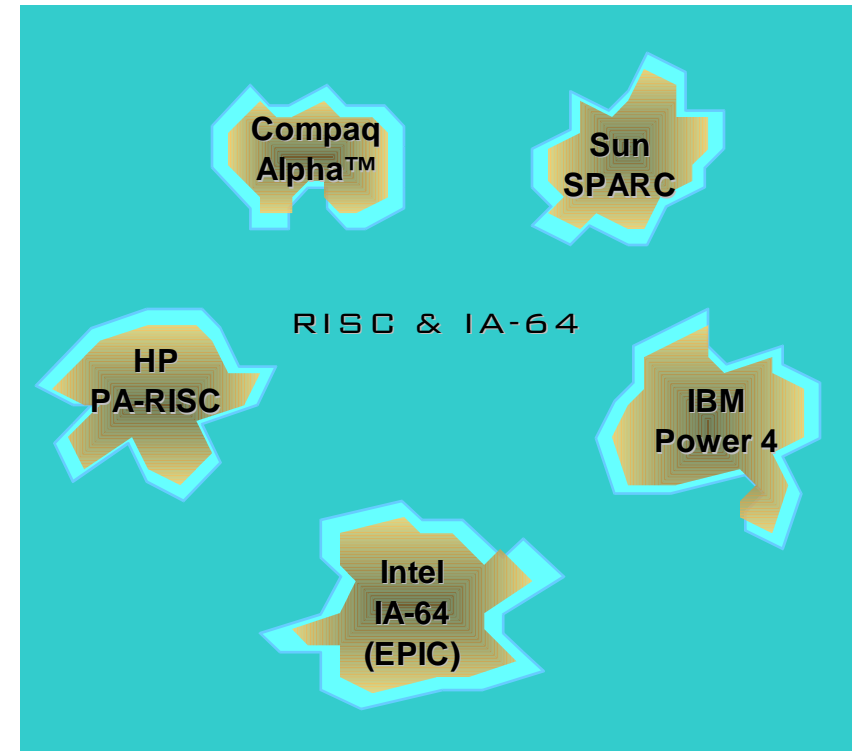
## AMD64 Technology



## AMD's x86 Technology Family of Processors



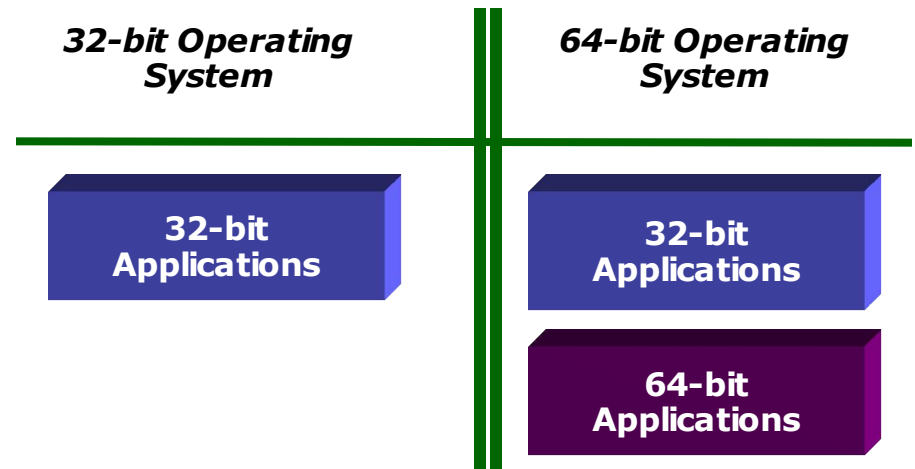
- A natural evolution of the current 32-bit architecture
- Similar to the 16- to 32-bit conversion of the 386
- Designed to retain compatibility with the current installed base of x86 operating systems and applications
- **Low-risk** and low-cost path to high-performance computing



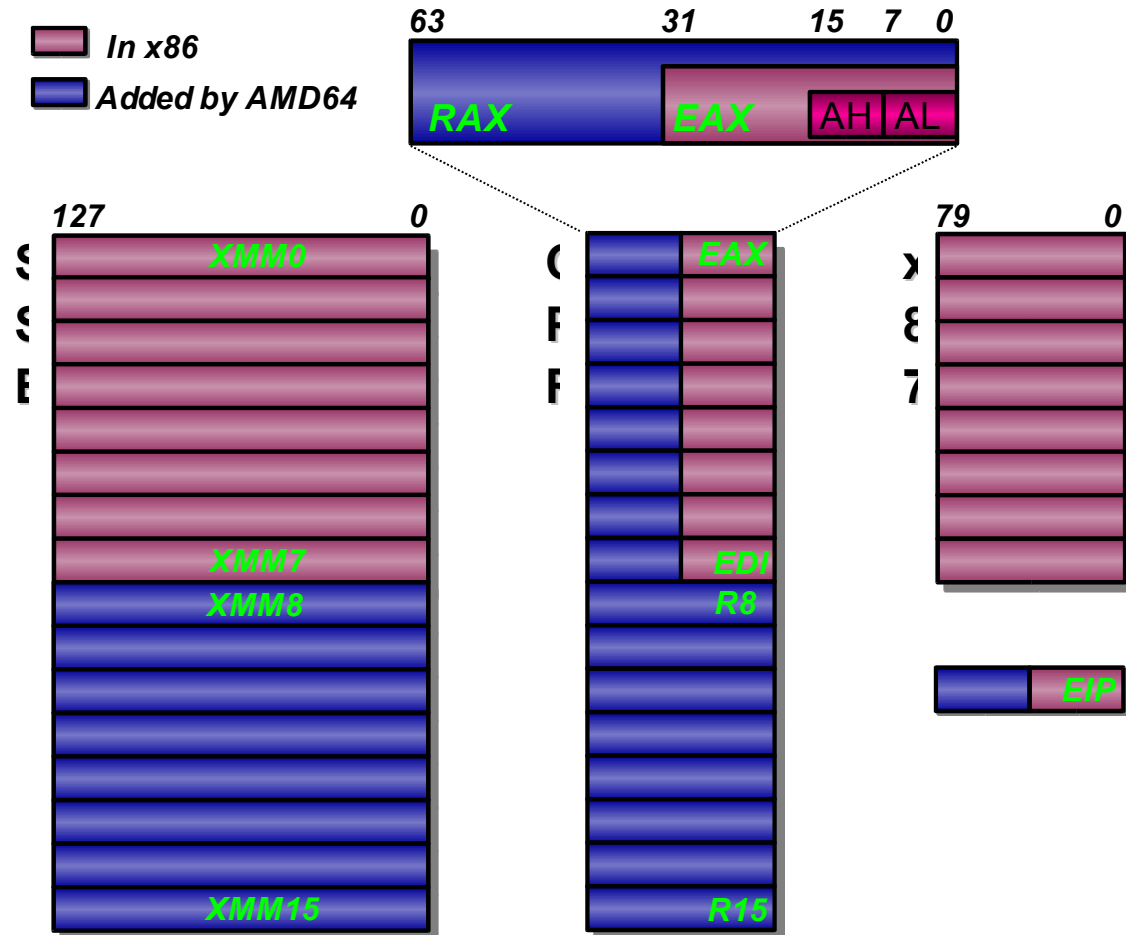
- There are many other 64-bit RISC solutions
- Each is a unique instruction set, all of which are incompatible with today's 32-bit code
- All require unique OS and applications

## *Building a Bridge from the 32- to the 64-bit World*

- Leverages the initial success of AMD Athlon™ MP processor
- Adds 64-bit capabilities to the world's highest performing 32-bit core for 2P and 4P servers
- Current 32-bit applications will work on both 32-bit and 64-bit operating systems
- Doesn't require special hardware or investment in a proprietary infrastructure
- Developing a solid ecosystem of motherboards, operating systems, development tools, and device drivers



- AMD64 Architecture:
  - 64-bit integer registers
  - 64-bit Virtual Address
  - 52-bit Physical Address
  - Sixteen 64-bit integer regs
  - Sixteen 128-bit SSE regs
  - SSE2 Instruction Set
  - Double precision scalar and vector operations
  - 16x8-, 8x16-way vector packed integer operations
  - SSE1 already added with AMD Athlon™ MP Processor



## ➤ **Performance**

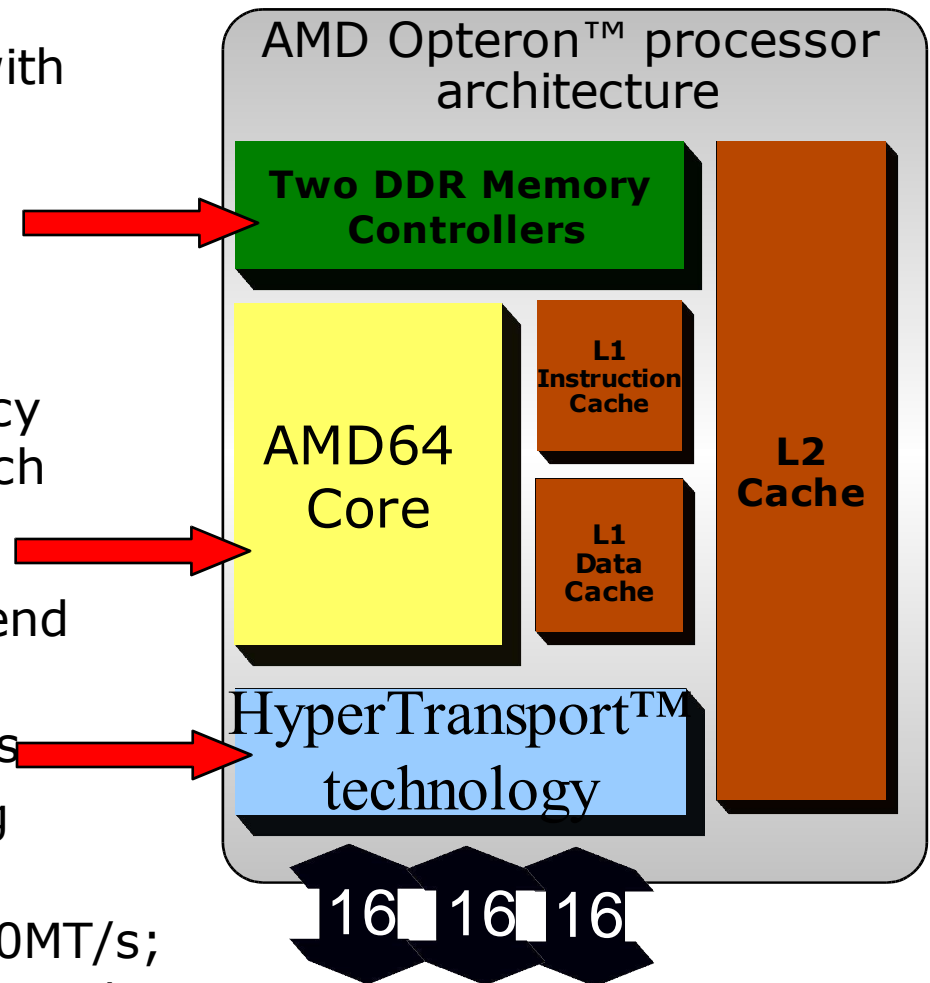
- High-bandwidth integrated memory controller scales with processor frequency and number of processors
- L2 1MB Cache

## ➤ **Compatibility**

- Approximately 10,000 legacy applications at time of launch

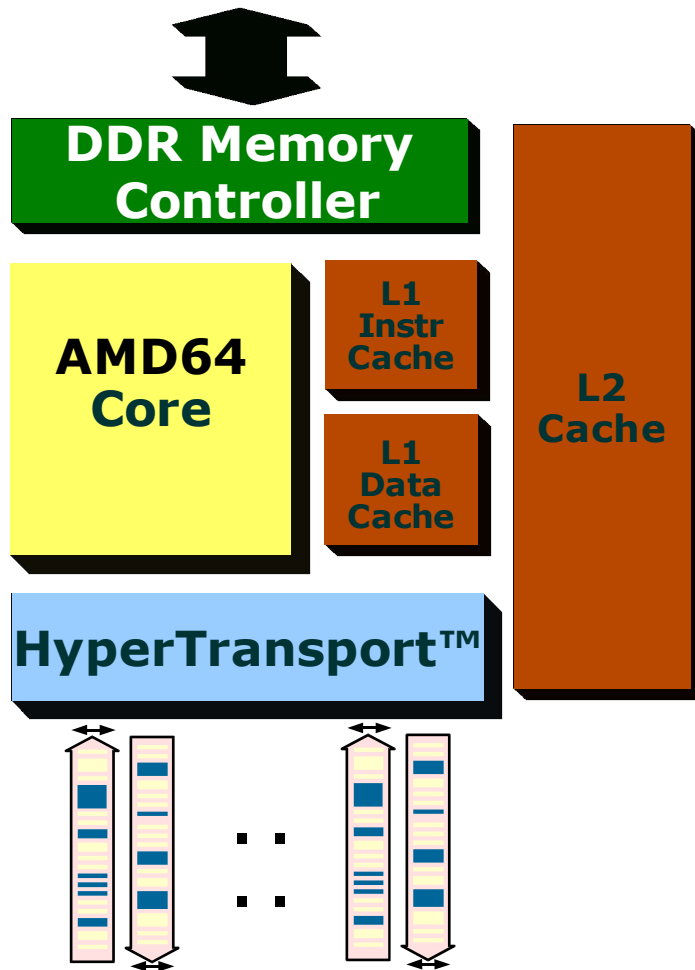
## ➤ **Scalability**

- **Can reduce** costs for high-end systems
- **Can remove** I/O bottlenecks
- Easy multiprocessor scaling
- 16-bit HyperTransport™ technology links are at 1600MT/s; provides 6.4GB/s peak aggregate bandwidth



# AMD Opteron™ Processor

## Integrated Memory Controller



- Designed to run memory controller at processor speeds - *not* FSB speeds
- Designed to dramatically decrease latency
  - AMD Athlon™ processor 1P platforms achieve ~160 ns best-case latency
  - AMD64 architecture is designed to achieve ~80 ns best-case latency
  - Latency generally decreases further as the core frequency increases
- Designed to add intelligence without decreasing performance
- Designed to support multiple DDR memories
  - DDR200, DDR266, and DDR333
  - Registered DIMMs
  - Future processor cores planned to support DDR-II, etc.



## AMD Opteron™ Processor 200 Series:

- 2-way server & workstation processor
- 144-bit DDR interface per CPU: 200, 266, 333 MHz
- Three 16-bit HyperTransport™ technology links per CPU. Typically, two are used to connect to another CPU and I/O

## AMD Athlon™ 64 Processor

- Performance Desktop Processor
- 72-bit DDR interface 200, 266, 333, 400 MHz
- One 16-bit HyperTransport technology link






NOTE: The AMD Athlon 64 and AMD Opteron are processors based on AMD64 technology

## AMD Opteron Processor 800 Series:

- Up to 8-way server processor
- 144-bit DDR interface per CPU: 200, 266, 333 MHz
- Three 16-bit HyperTransport technology links per CPU. Typically all three used to connect to other CPUs & I/O

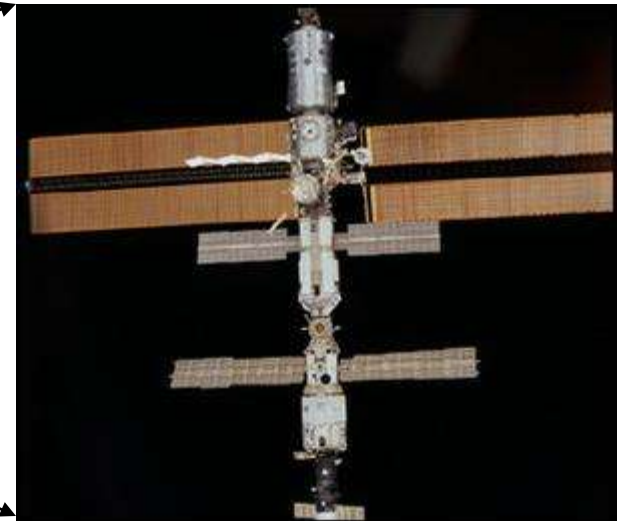
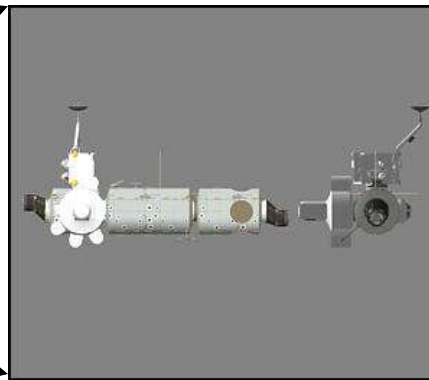
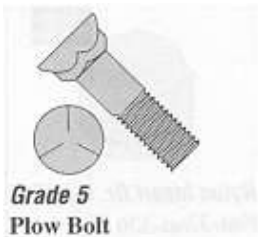
16-bit HyperTransport Links are at 1600MT/s; provides 6.4GB/s Peak Aggregate Bandwidth

- In test after test, AMD64 technology beats the competition
  - 32-bit performance superior to other 32-bit solutions on the market
  - 64-bit performance superior in terms of performance per dollar spent. In most cases more cost-effective to buy multiple AMD64 machines to get the same performance seen from a single competing 64-bit machine.
- See <http://www.amd.com> for the most recent data

	Operating System	Type
	SuSE Linux Enterprise Server (SLES) 8	32 & 64-bit
	SuSE Linux 9.0 Personal & Professional	32-bit & 64-bit
	UnitedLinux Version 1.0 code base by UnitedLinux Consortium	32 & 64-bit
	Conectiva Linux Enterprise Edition	32-bit
	Linux AMD64 kernel patches ( <a href="http://www.x86-64.org">www.x86-64.org</a> )	64-bit
	Mandrake Linux 9.2 (coming soon)	32-bit & 64-bit
	Mandrake Linux Corporate Server 2.1	32-bit & 64-bit
	NetBSD	32 & 64-bit
	Red Hat 9.0	32-bit
	Red Hat Enterprise Linux 3 (coming soon)	32-bit & 64-bit
	Scyld Beowulf Cluster Operating System	32-bit
	Solaris 9 for x86	32-bit
	Turbolinux 8 for AMD64	32 & 64-bit
	Windows® 2000 Server	32-bit
	Windows Server 2003	32-bit & 64-bit

## AMD64 Technology AMD64 Means... Dynamic Scaling

Large-scale simulations and games can be interacted with down to the lowest component level.



From the largest view down to the smallest bolt, designers can maintain accurate physics at all times.

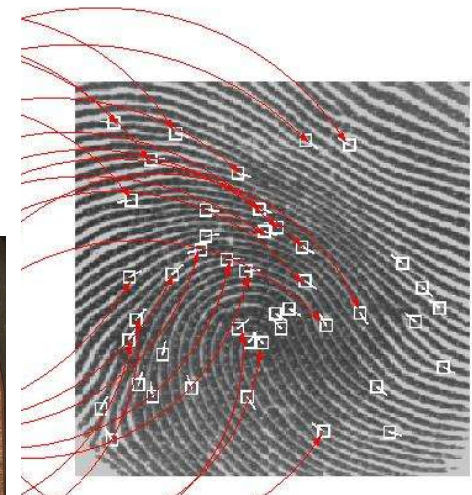
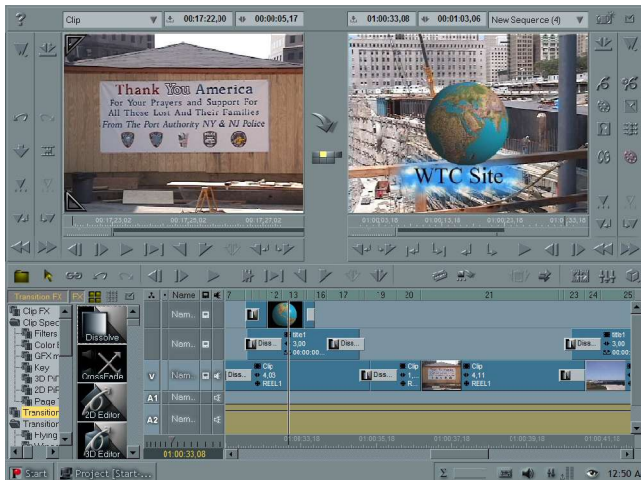
## AMD64 Technology AMD64 Means...

### Large Memory Arrays

### Biometric Identification

### Photo-Realistic Gaming

### Instantaneous Access





## AMD64 Technology AMD64 Means... Additional Registers

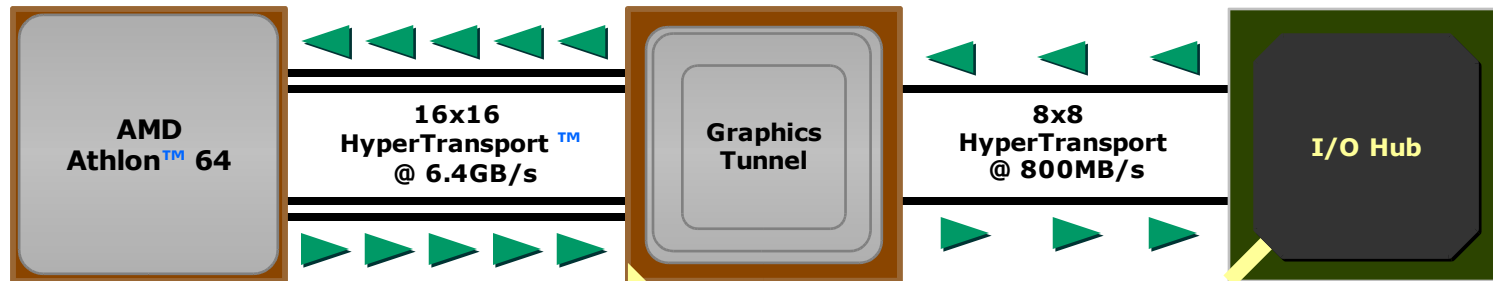
Real-Time Special Effects



Higher Level of Realism



## HyperTransport™ Technology Interface



### HyperTransport™ Technology Interface Attributes

- **Unidirectional**
- **DDR-like performance (800MHz = 1600MT/sec)**
- **4 bytes wide ... 6.4GB/sec bandwidth**

## Overview

---

### **Reliability and Stability**

- ECC Protection
  - L1 data cache
  - L2 tags and data
  - Main memory DRAM (optional)
- Hardware Scrubbing
- Thermal Protection
  - ThermTrip
    - Shuts down processor without motherboard intervention
  - Thermal Diode
    - Works with motherboard circuitry to monitor CPU temperature and work with thermal control hardware (i.e, temp controlled fans, etc.)



## Chipsets

AMD  
Launch  
Chipsets

**AMD-8151™**  
**Graphics Tunnel**  
8x AGP

**AMD-8111™**  
**I/O Hub**  
ATA 133, USB 2.0  
10/100 Ethernet

Discrete  
Graphics

**VIA**  
**K8T400M+VT8235**  
8x AGP, ATA 133  
8x V-Link, USB 2.0  
10/100 Ethernet

**SiS**  
**755 + 963**  
8x AGP, ATA 133  
USB 2.0, 1394A

**NVIDIA**  
**CrushK8**  
8x AGP, ATA 133  
Two 10/100 Ethernet  
USB 2.0

**NVIDIA**  
**CrushK8S**  
8x AGP, S-ATA, RAID  
Gigabit Ethernet  
USB 2.0, 1394A

Integrated  
Graphics

**VIA**  
**K8M400+VT8235**  
8x AGP + integrated gfx  
ATA 133, USB 2.0  
10/100 Ethernet

**SiS**  
**760 + 963**  
8x AGP + Ultra256 gfx  
ATA 133  
USB 2.0, 1394A

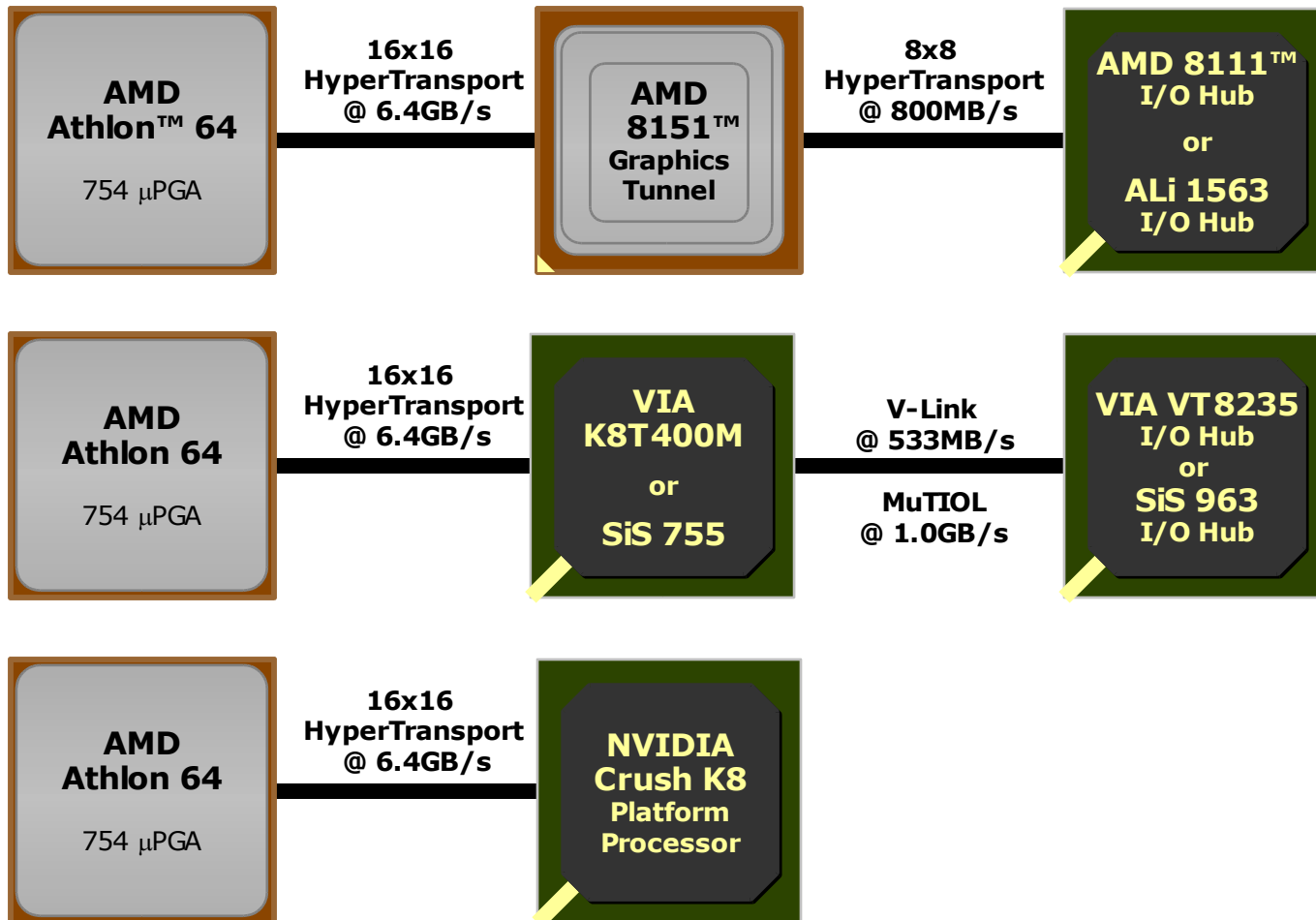
**NVIDIA**  
**CrushK8G**  
8x AGP + GeForce4i gfx  
ATA 133, SATA  
Two 10/100 Ethernet  
USB 2.0, 802.11b

I/O Hub

**Ali**  
**1563**  
ATA 133, USB 2.0  
10/100 Ethernet

**Please contact the respective 3rd party vendors directly for latest schedules and information.**

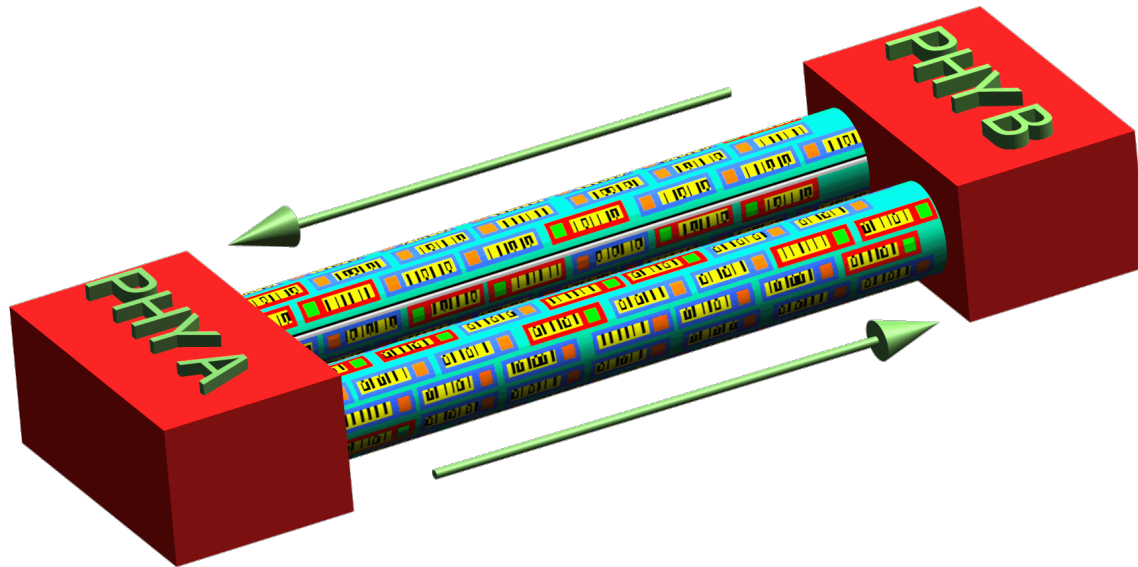
## Chipset Interfaces



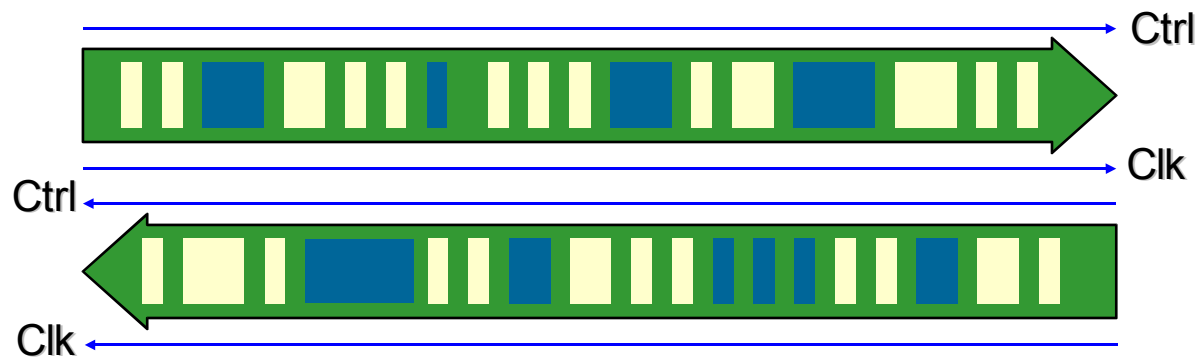
# HyperTransport™ Technology and Server Chipset Highlights

# HyperTransport™ Technology Basics

- HyperTransport™ Technology buses have two unidirectional point-to-point links:
  - The links can be 2-, 4-, 8-, 16-, or 32-bits wide in each direction
  - HyperTransport™ links have a data rate up to 1.6 Gigabits/second per pin-pair (800 MHz clock)
  - Total Aggregate Bandwidth = 12.8 Gbytes/second at 32 bits wide
  - AMD Opteron™ supports three 16-bit HyperTransport™ links
    - Provides 19.2 Gbytes/second on total data bandwidth

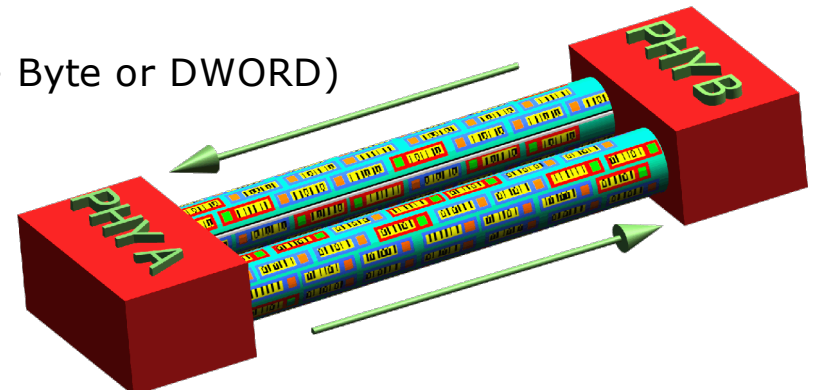
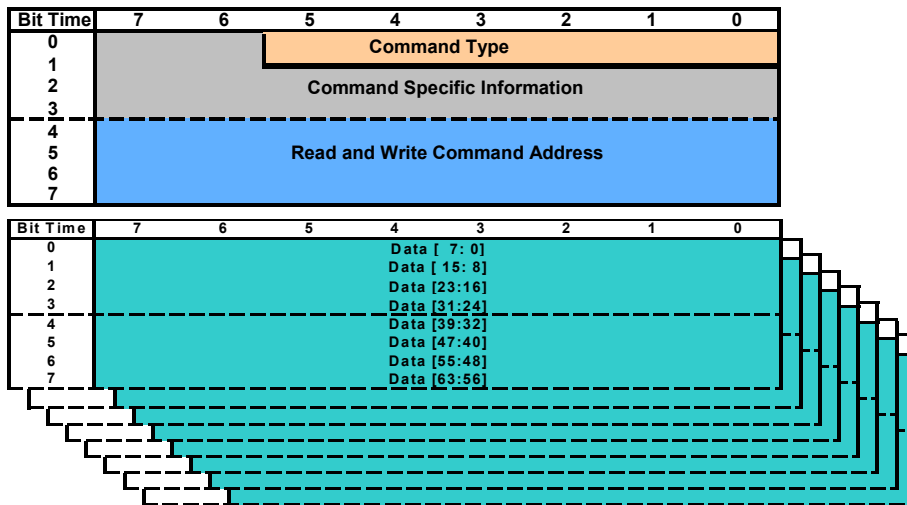


# HyperTransport™ Technology Clock and Control Signals



- **Asynchronous clock forwarding**
  - One clock is forwarded for each eight bits in each direction
  - Clocks are double pumped; a 800 MHz clock is used for 1600 Mbit data rate
- **Control line distinguishes command packets**
  - De-asserted during data packets
- **In-band system management & legacy signal transport**
  - Eliminates sideband wires, interrupts use messages instead of wires
- **Embedded code in back channel messages used for flow control**
  - Code indicates how many buffers are available for each virtual channel

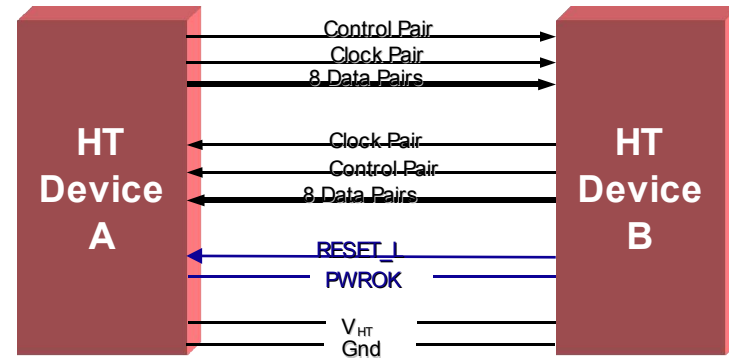
- Commands and interrupts are realized as a 32 bit command word
- Address and Data is preceded by a 64-bit header
  - ✓ 6-bit type field – Write, Read, Read Response, Fence & Flush
  - ✓ 26-bit Command specific field
  - ✓ 32-bit address field (command specific – Byte or DWORD)



- ❑ At 800MHz DDR it takes:
  - 1.25ns to send a request (32-bits)
  - 22.5ns to send a 64B block
- ❑ A PCI-X write of one 64 byte block takes ~290ns + PCI X I/O latency

# HyperTransport™ Technology Pin count

- ❑ Additional control signals
  - Power OK (PWROK)
  - Reset (RESET\_L)
- ❑ Signal to ground ratio is conservatively 4:1
- ❑ Optional link power down signals for mobile systems
  - LDT\_Stop
  - DevReq
- ❑ Power per pin-pair is nil when a HyperTransport™ technology device is stopped (LDT\_Stop)



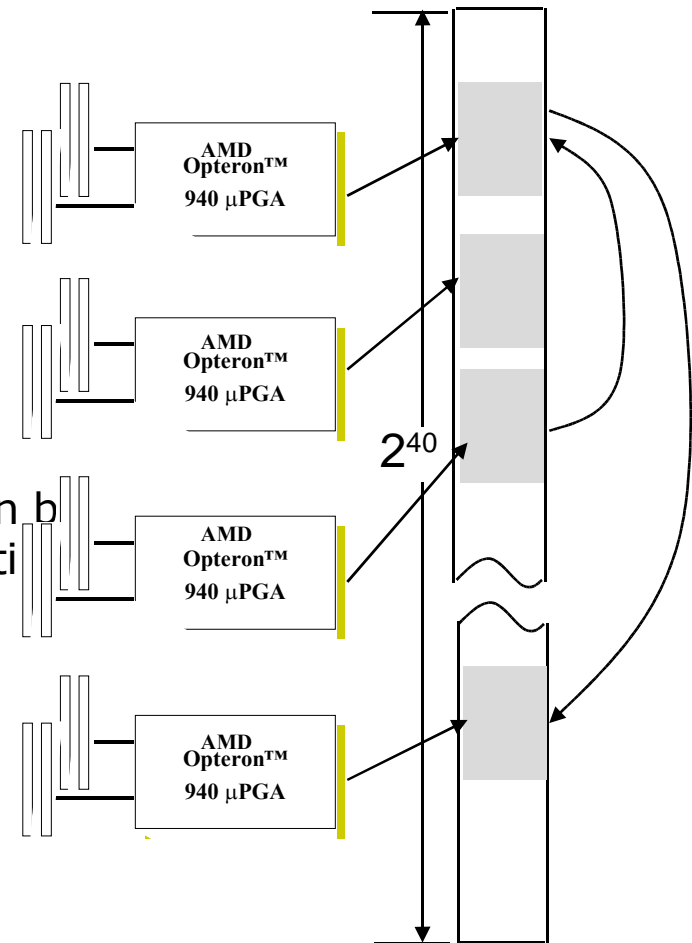
PWROK, RESET\_L required for proper reset & init  
 $V_{HT}$  routed between devices is required for proper common mode range

Bus Width (Both Ways)	2	4	8	16	32
Data Pins (total)	8	16	32	64	128
Clock Pins (total)	4	4	4	8	16
Control Pins (total)	4	4	4	4	4
<b>Subtotal (high speed)</b>	<b>16</b>	<b>24</b>	<b>40</b>	<b>76</b>	<b>148</b>
VHT	2	2	3	6	10
GND	4	6	10	19	37
PWROK	1	1	1	1	1
RESET_L	1	1	1	1	1
<b>Total Pins</b>	<b>24</b>	<b>34</b>	<b>55</b>	<b>103</b>	<b>197</b>

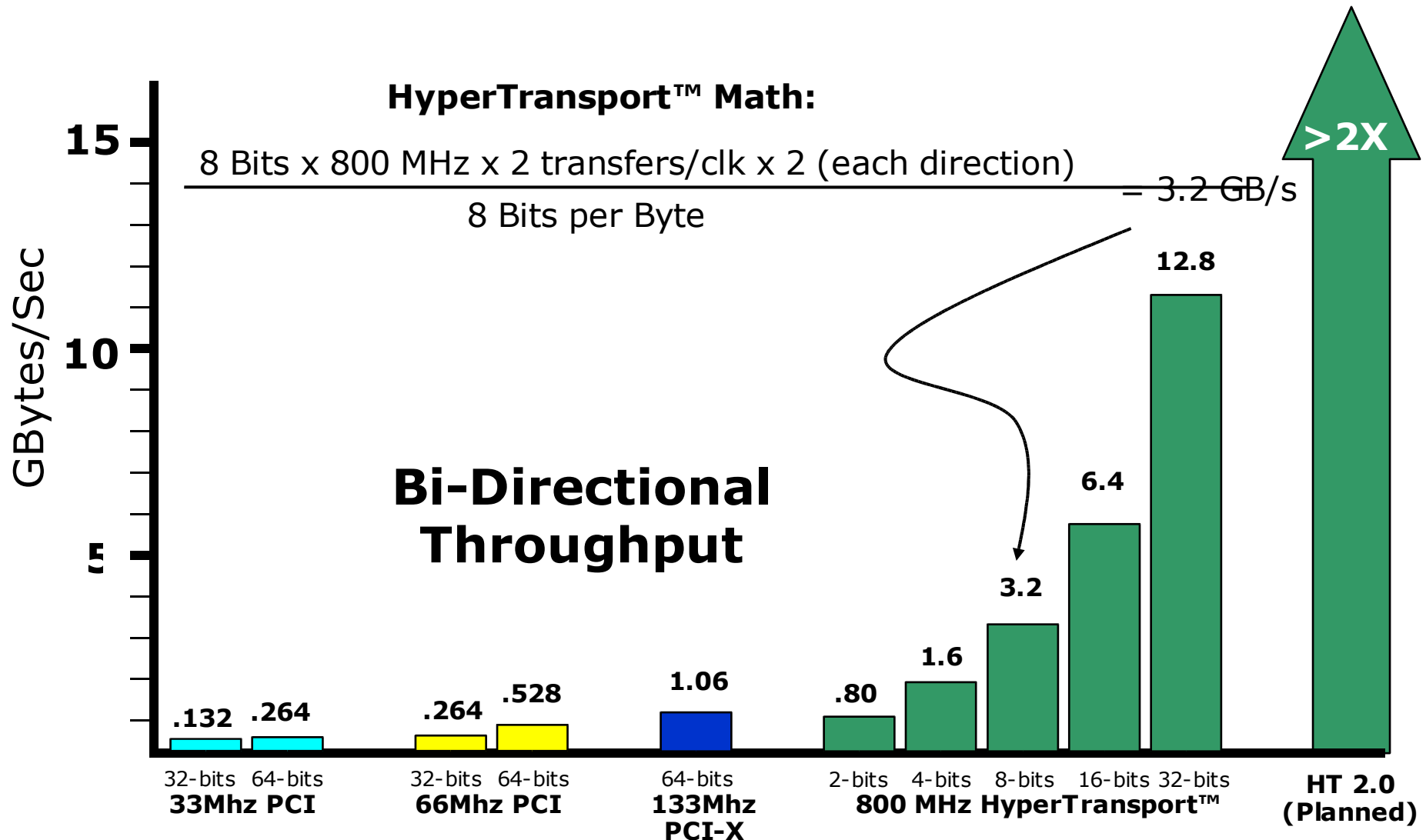
DC Power per Pin-Pair: 4 - 9 mW, 6 mW<sub>typical</sub>  
 Signal to  $V_{LDT}/Gnd$  Ratio: 4:1

# HyperTransport™ Technology Intelligence

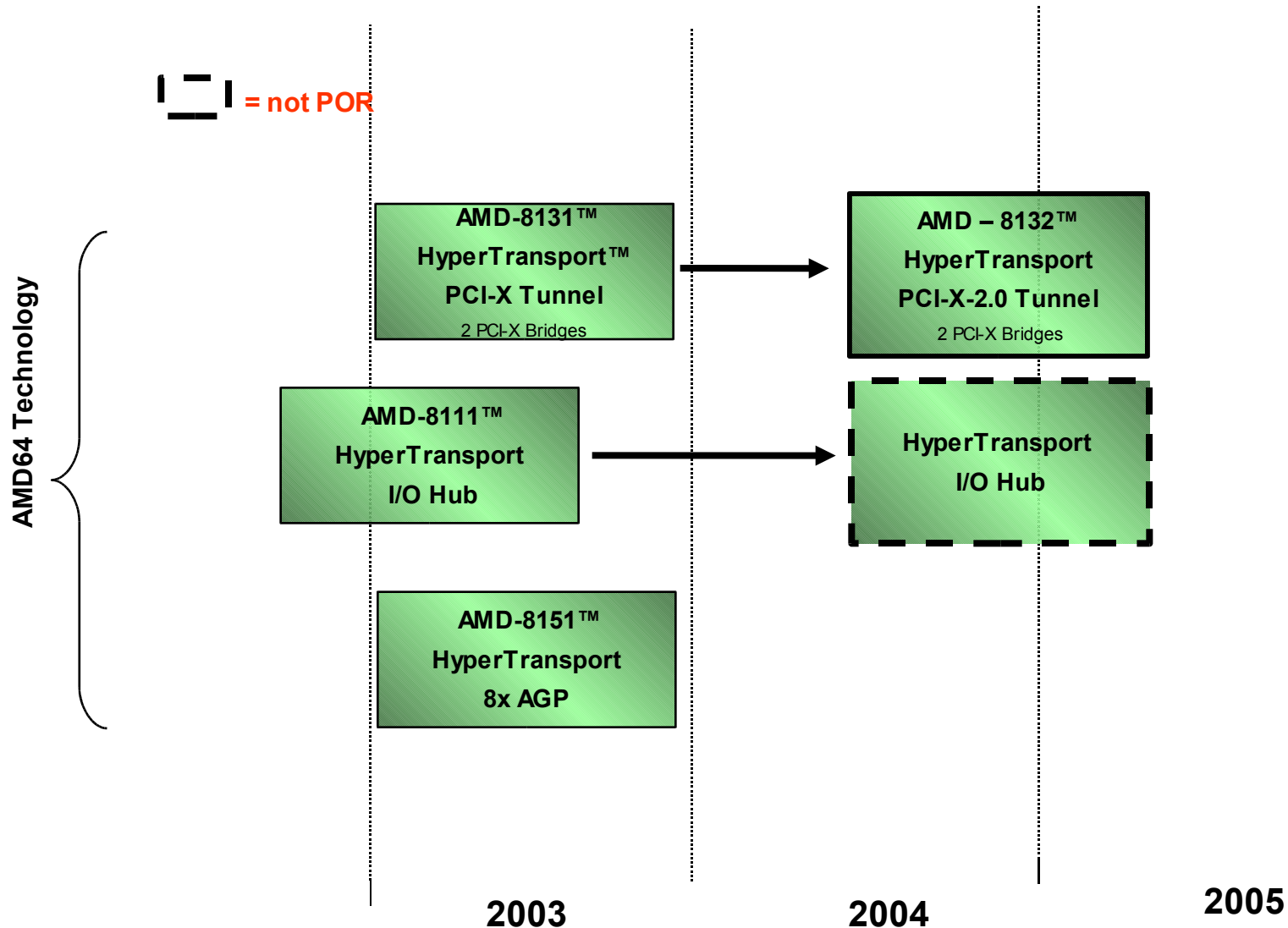
- Data movement over the HyperTransport™ bus does not use any CPU machine cycles.
- External device can write to any address within the processor's physical 40-bit address range without CPU intervention.
- In cases where there are multiple HyperTransport™ technology ports, data can be passed between ports without CPU intervention.
- Because all devices reside within one physical  $2^{40}$  linear space all I/O devices have access to all processors and their associated memory & I/O.



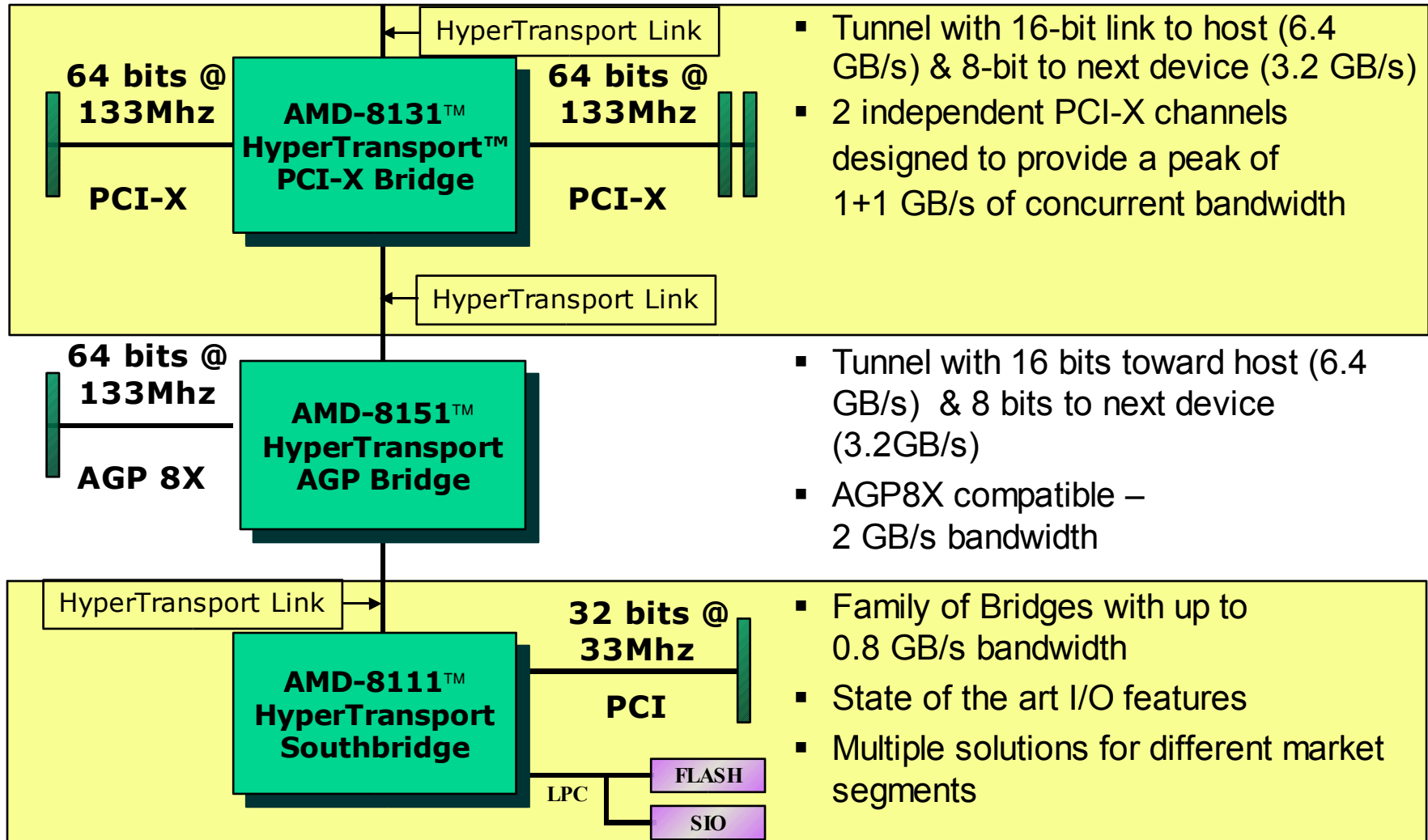




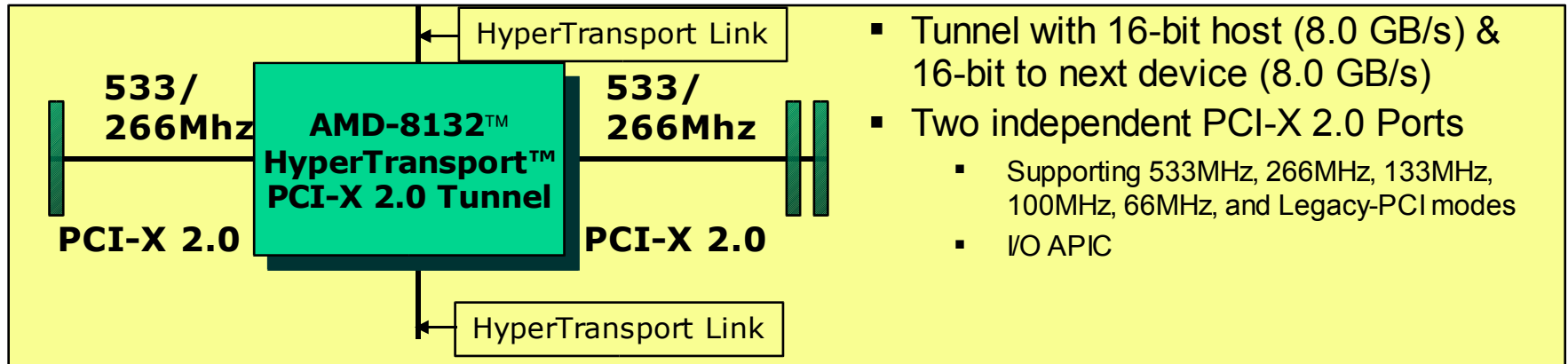
# AMD Opteron™ Chipset Roadmap



## Building Blocks



## *Future Building Blocks*



# HyperTransport™ Technology

Consortium ...[www.hypertransport.org](http://www.hypertransport.org)





AMD, the AMD Arrow logo, AMD Athlon, AMD Opteron, 3DNow! and combinations thereof, AMD-8111, AMD-8131, AMD-8132, and AMD-8151 are trademarks of Advanced Micro Devices, Inc. HyperTransport is a licensed trademark of the HyperTransport Technology Consortium. Microsoft and Windows are registered trademarks of Microsoft Corporation in the U.S. and/or other jurisdictions. Pentium and MMX are registered trademarks of Intel Corporation in the U.S. and/or other jurisdictions. SPEC and SPECfp are registered trademarks of Standard Performance Evaluation Corporation in the U.S. and/or other jurisdictions. Alpha is a trademark of Digital Equipment Corporation. MIPS is a registered trademark of MIPS Technologies, Inc. Other product and company names used in this presentation are for identification purposes only and may be trademarks of their respective companies.