x86 in High Performance Computing

- The Six System Challenges

#6: Watt density:

• x86 is the most widely installed instruction set in the world.
• Instruction set not relevant to CPU performance ("to first order").
• What is important:

  #6: Watt density:

  • With clusters exceeding 10,000 processors, watt density is an important issue. As cluster size expands, cooling capacity and costs can be significant.

#5: The I/O infrastructure:

• Design the lowest watts/Gig Cycle solution leveraging state-of-the-art AMD64 architecture and silicon-on-insulator process.

#4: Addressable memory:

#3: Memory bandwidth:

#2: Cost per processing node:

#1: Backward compatible to x86-32:

- There is an 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.

• 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.
Advanced AMD Opteron™ Processor System Architecture

AMD64 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

Typical System

- 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
Typical Multiprocessing System

- 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
Intel Xeon – Light Load
Intel Xeon – Heavy Load
AMD Opteron™ Processor 4P 800 Series

Processor-based Server

- 200-333MHz 144-Bit Reg DDR
- 200-333MHz 144-Bit Reg DDR

- 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

- [1] = 16x16 Coherent HyperTransport™ @ 1600MT/s
- [2] = 16x16 HyperTransport @ 1600MT/s
- [3] = 8x8 HyperTransport @ 400MT/s
- [4] = 8x8 HyperTransport @ 1600MT/s
AMD64 Technology

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

START

BOOT UP
Using 32 bit BIOS

Load 32 bit OS

Run 32 bit Applications

Look at OS

Load 64 bit OS

Run 32 & 64 bit apps
AMD Athlon™ 64 Processor Technical Overview

AMD64 Technology

- 32-bit thread
  - 32-bit Application
  - THUNKING LAYER

- 64-bit thread
  - 64-bit Application
  - 64-bit Operating System
    - 64-bit Device Drivers
AMD64 Technical Overview

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
AMD64 Technology

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 Computing Strategy (2)

• 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
AMD64 Processor Overview

- **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
• 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.
## AMD64 Processors And Target Systems

### 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

### 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

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

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 Support

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

AMD64 Technology

AMD64 Means...

Large Memory Arrays

Biometric Identification

Photo-Realistic Gaming

Instantaneous Access
AMD Athlon™ 64 Processor Technical Overview

**AMD64 Technology**
**AMD64 Means...**
**Additional Registers**

Higher Level of Realism

Real-Time Special Effects
HyperTransport™ Technology Interface Attributes

- Unidirectional
- DDR-like performance (800MHz = 1600MT/sec)
- 4 bytes wide ... 6.4GB/sec bandwidth
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.)
AMD Athlon™ 64 Infrastructure Support

**Chipsets**

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

**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.
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)

<table>
<thead>
<tr>
<th>Bus Width (Both Ways)</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>16</th>
<th>32</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Pins (total)</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>64</td>
<td>128</td>
</tr>
<tr>
<td>Clock Pins (total)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>Control Pins (total)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

| Subtotal (high speed) | 16 | 24 | 40 | 76 | 148 |

<table>
<thead>
<tr>
<th>VTH</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>6</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>GND</td>
<td>4</td>
<td>6</td>
<td>10</td>
<td>19</td>
<td>37</td>
</tr>
<tr>
<td>PWROK</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RESET_L</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

| Total Pins           | 24 | 34 | 55 | 103| 197|

DC Power per Pin-Pair: 4 - 9 mW, 6 mW\text{ Typical}

Signal to $V_{\text{LDT/Gnd}}$ 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.
HyperTransport™ Technology
Scalable Bandwidth

HyperTransport™ Math:
8 Bits x 800 MHz x 2 transfers/clk x 2 (each direction)
= 3.2 GB/s

Bi-Directional Throughput

32-bits  64-bits
33Mhz PCI  .132  .264
66Mhz PCI  .264  .528
133Mhz PCI-X  1.06
2-bits  4-bits  8-bits  16-bits  32-bits
800 MHz HyperTransport™  .80  1.6  3.2  6.4  12.8

HT 2.0 (Planned) >2X
AMD Opteron™ Chipset Roadmap

- **AMD-8111™**
  - HyperTransport™
  - I/O Hub

- **AMD-8131™**
  - HyperTransport™
  - PCI-X Tunnel
  - 2 PCI-X Bridges

- **AMD-8151™**
  - HyperTransport™
  - 8x AGP

- **AMD - 8132™**
  - HyperTransport™
  - PCI-X-2.0 Tunnel
  - 2 PCI-X Bridges

= not POR

<table>
<thead>
<tr>
<th>Year</th>
<th>Product</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td>AMD-8111™</td>
<td>HyperTransport™ I/O Hub</td>
</tr>
<tr>
<td>2004</td>
<td>AMD-8131™</td>
<td>HyperTransport™ PCI-X Tunnel, 2 PCI-X Bridges</td>
</tr>
<tr>
<td>2005</td>
<td>AMD - 8132™</td>
<td>HyperTransport™ PCI-X-2.0 Tunnel, 2 PCI-X Bridges</td>
</tr>
</tbody>
</table>
HyperTransport™ Technology

Building Blocks

- **AMD-8131™ HyperTransport™ PCI-X Bridge**
  - 64 bits @ 133Mhz
  - Tunnel with 16-bit link to host (6.4 GB/s) & 8-bit to next device (3.2 GB/s)
  - 2 independent PCI-X channels designed to provide a peak of 1+1 GB/s of concurrent bandwidth

- **AMD-8151™ HyperTransport™ AGP Bridge**
  - 64 bits @ 133Mhz
  - Tunnel with 16 bits toward host (6.4 GB/s) & 8 bits to next device (3.2 GB/s)
  - AGP8X compatible – 2 GB/s bandwidth

- **AMD-8111™ HyperTransport™ Southbridge**
  - 32 bits @ 33Mhz
  - Family of Bridges with up to 0.8 GB/s bandwidth
  - State of the art I/O features
  - Multiple solutions for different market segments
HyperTransport™ Technology

Future Building Blocks

- HyperTransport Tunnel
  - 16-bit host interface (8.0GB/s)
  - 16-bit next device interface (8.0GB/s)

- Two independent PCI-X 2.0 Ports
  - Supporting 533MHz, 266MHz, 133MHz, 100MHz, 66MHz, and Legacy-PCI modes

- I/O APIC

![Diagram showing HyperTransport Link and PCI-X 2.0 Tunnel connections]

- Tunnel with 16-bit host (8.0 GB/s) & 16-bit to next device (8.0 GB/s)
- Two independent PCI-X 2.0 Ports
  - Supporting 533MHz, 266MHz, 133MHz, 100MHz, 66MHz, and Legacy-PCI modes
  - I/O APIC
HyperTransport™ Technology
Consortium ...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.