HyperTransport™ Technology Tutorial

Prof. José Duato
Technical University of Valencia, Spain
Simula Research Laboratory, Oslo, Norway
HyperTransport Technology Consortium

Hot Chips Symposium
August 23, 2009
With us Today and Happy to Address Your Questions

Brian Holden
VP and Chair, Technical Working Group
HyperTransport Technology Consortium
brian.holden@hypertransport.org
408-472-6310
Topics:
HyperTransport Technology

Mario Cavalli
General Manager
HyperTransport Technology Consortium
mario.cavalli@hypertransport.org
925-968-0220
Topics:
HyperTransport Market Positioning
HyperTransport Consortium
Topics

- Scope and Design Goals
- HyperTransport Defined
  - Host Interface
  - Connecting Device to Host
  - Connecting Multiple Devices to Host
  - Interconnecting Multiple Hosts
- AMD Cache Coherence Support
- Beyond Motherboards
- New in HT3
- Beyond HT3
- Beyond Conventional
- HyperTransport Technology Consortium
Scope and Design Goals

• System Area Network Supporting Cache-Coherent Shared-Memory Multiprocessors and I/O Devices
  • High-Performance Replacement for Processor Front Side Bus (Point-to-Point Links vs. Bus)
Scope and Design Goals (cont.)

- System Area Network Supports Cache-Coherent Shared-Memory Multiprocessors and I/O Devices
  - High-Performance Replacement for Processor Front Side Bus (Point-to-Point Links vs. Bus)

- HyperTransport’s Distinction: **Processor-Native**
  - Integrated in Processor Architectures
Scope and Design Goals (cont.)

- Lowest-Latency, High Bandwidth, Cost-Effective, Reliable Motherboard-Level Interconnect
  - SMP Programming Model

- Unified Interface For Local and Remote Memory

- Self-Configuring Topology and Link Speed

- HT3 Enhancements
  - Increased Bandwidth and Reliability
  - Link Splitting
  - Dynamic Power Management
  - AC Mode
  - Hot Plugging
Typical Server Architecture
HyperTransport™ Defined

Nine Years of Fine Tuning, Perfecting, Polishing
Host Interface

Integrated North Bridge

DRAM Controller

Memory Controller

Crossbar Switch

System Request Queue (SRQ)

HyperTransport Technology BUS

Cache, Load/Store & Bussing Unit

1MB

Bussing Unit

128-bits wide

Load/Store Unit

128-bits wide

FP Unit Execution Unit Fetch Scan Align

L1 Data Cache

L1 Instruction Cache

L2 Cache

AMD64 CPU

AMD64 CPU
Host Interface (cont.)

- Single Interface for All Cores (SRQ/ SRI)
- On-Chip Crossbar and Routing
- Host Bridge for I/ O Device Chain
# Northbridge Architecture

<table>
<thead>
<tr>
<th>Virtual channel</th>
<th>Use</th>
<th>Data</th>
</tr>
</thead>
</table>
| Request         | • Read   
                    • Nonposted write 
                    • Cache block commands | Y    |
| Posted request  | Posted write         | Y    |
| Probe           | Broadcast probes     | N    |
| Response        | • Read response 
                    • Probe response 
                    • Completion     | Y    |

```
<table>
<thead>
<tr>
<th>Victim buffer (8-entry)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write buffer (8-entry)</td>
</tr>
<tr>
<td>Instruction MAB (2-entry)</td>
</tr>
<tr>
<td>Data MAB (8-entry)</td>
</tr>
<tr>
<td>Core 1</td>
</tr>
<tr>
<td>SRI queue (24-entry)</td>
</tr>
<tr>
<td>Address MAP and GART</td>
</tr>
<tr>
<td>HT0 input</td>
</tr>
<tr>
<td>Router 10-entry buffer</td>
</tr>
<tr>
<td>Router 18-entry buffer</td>
</tr>
<tr>
<td>Router 16-entry buffer</td>
</tr>
<tr>
<td>Router 16-entry buffer</td>
</tr>
<tr>
<td>Router 12-entry buffer</td>
</tr>
<tr>
<td>Crossbar</td>
</tr>
<tr>
<td>MCT queue (20-entry)</td>
</tr>
<tr>
<td>To DCT</td>
</tr>
</tbody>
</table>
```
Connecting Device to Host

- **Approach: High Speed Point-to-Point Parallel Link**
  - Point-to-Point Link Minimizes Parasitic Capacitance
  - Clock Forwarding Removes Clock Recovery Overhead
  - Parallel Link Delivers High Bandwidth, Low Latency

- **Control and Data Packets Interleaved on Each Link**
  - CTL Signal Distinguishes Between Control and Data
  - Two Additional System Signals: PWROK and RESET
HT Physical Layer

- Low-Voltage Differential Signalling (LVDS)

- Pre-Emphasis Supports Higher Clock Rates
HT Physical Layer (cont.)

- Clock Rate: From 200 MHz to 3.2 GHz
- Link Width: 2, 4, 8, 16, 32-Bit
HT Physical Layer (cont.)

• Support for Asymmetric and Mixed Link Width
**HT Transaction/ Data Link Layer**

HyperTransport Packet Format

- **Header**: 8 or 12 Bytes
- **DATA**: 4-64 Bytes

**TRANSACTION/DATA LINK LAYER**
- 8-12 bytes overhead

**PHYSICAL LAYER**
- 0 % overhead
HT Basic Read/Write Sequences

HyperTransport Data Write Sequence

- Write Request Control Packet
- Data Packet
  - 8-bytes
  - 4-64 bytes of data

HyperTransport Data Read Sequence

- Read Request Control Packet
- Read Response Control Packet
- Data Packet
  - 8-bytes
  - 4-bytes
  - 4-64 bytes of data
# HT Request Packet Format

![HyperTransport Request Packet Packet Format With Address](image)

<table>
<thead>
<tr>
<th>Bit-Time</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td>SeqID[3:2]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td>SeqID[1:0]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **SeqID[3:2]**: Sequential ID
- **Cmd[5:0]**: Type of command
- **Pass PW**: Pass Through Wait
- **UnitID[4:0]**: Unit ID
- **Addr[15:8]**
- **Addr[23:16]**
- **Addr[31:24]**
- **Addr[39:32]**

---

Copyright HyperTransport Consortium, 2009
### HT Request Packet Format (cont.)

#### HyperTransport Request Packet Format with Address

<table>
<thead>
<tr>
<th>Bit-Time</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>SeqID[3:2]</td>
<td>Cmd[5:0]=type of command</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Pass PW</td>
<td>SeqID[1:0]</td>
<td>UnitID[4:0]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Command-Specific</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Command-Specific</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Addr[15:8]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Addr[23:16]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Addr[31:24]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Addr[39:32]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Typical Request Types

<table>
<thead>
<tr>
<th>Command Type</th>
<th>Packet Type/Size in bytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Write Request</td>
<td>Req/Addr/Data (8)</td>
</tr>
<tr>
<td>- Nonposted/posted</td>
<td></td>
</tr>
<tr>
<td>- data length = byte/doubleword</td>
<td></td>
</tr>
<tr>
<td>- normal/isochronous</td>
<td></td>
</tr>
<tr>
<td>- noncoherent/coherent</td>
<td></td>
</tr>
<tr>
<td>Write Request with Extended Address</td>
<td>Req/Addr/Data (12)</td>
</tr>
<tr>
<td>Read Request</td>
<td>Req/Address (8)</td>
</tr>
<tr>
<td>- Ordering - pass/no pass</td>
<td></td>
</tr>
<tr>
<td>- data length = byte/doubleword</td>
<td></td>
</tr>
<tr>
<td>- normal/isochronous</td>
<td></td>
</tr>
<tr>
<td>- noncoherent/coherent</td>
<td></td>
</tr>
<tr>
<td>Broadcast Message</td>
<td>Req/Address (8)</td>
</tr>
<tr>
<td>Atomic Read-Modify-Write</td>
<td>Req/Addr/Data (8)</td>
</tr>
</tbody>
</table>
### HT Request Packet Format (cont.)

#### Typical Request Types

<table>
<thead>
<tr>
<th>Command Field</th>
<th>Command Type</th>
<th>Packet Type/ (size in bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cmd[5:0]</td>
<td>Write Request</td>
<td>Req/Addr/Data (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Nonposted/posted</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- data length = byte/doubleword</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- normal/isochronous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- noncoherent/coherent</td>
</tr>
<tr>
<td></td>
<td>Write Request with Extended Address</td>
<td>Req/Addr/Data (12)</td>
</tr>
<tr>
<td></td>
<td>Read Request</td>
<td>Req/Address (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Ordering - pass/no pass</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- data length = byte/doubleword</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- normal/isochronous</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- noncoherent/coherent</td>
</tr>
<tr>
<td></td>
<td>Broadcast Message</td>
<td>Req/Address (8)</td>
</tr>
<tr>
<td></td>
<td>Atomic Read-Modify-Write</td>
<td>Req/Addr/Data (8)</td>
</tr>
</tbody>
</table>
HT Read Response Packet Format

HyperTransport Read Response Packet Format

<table>
<thead>
<tr>
<th>Bit-Time</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Isoc</td>
<td>Rsv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cmd[5:0]: 110000</td>
</tr>
<tr>
<td>1</td>
<td>Pass PW</td>
<td>Bridge</td>
<td>Rsv</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>UnitID[4:0]</td>
</tr>
<tr>
<td>2</td>
<td>Count[1:0]</td>
<td>Error0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SrcTag[4:0]</td>
</tr>
<tr>
<td>3</td>
<td>Rsv/RqUID</td>
<td>Error1</td>
<td>Rsv/RspVCSet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Count[3:2]</td>
</tr>
</tbody>
</table>

Important Read Response Packet Fields

Cmd[5:0]=110000

- Command Type
- Error0
- Error1 (Error Status)

<table>
<thead>
<tr>
<th>Error Codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>E0</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>
## Common HT Command Types

<table>
<thead>
<tr>
<th>Command Field</th>
<th>Command Type</th>
<th>Packet Type/ (size in bytes)</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>Cmd[5:0]</code></td>
<td>Write Request</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- Nonposted/posted</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- data length = byte/doubleword</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- normal/isochronous</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- noncoherent/coherent</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Write Request with Extended Address</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Read Request</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- Ordering - pass/no pass</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- data length = byte/doubleword</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- normal/isochronous</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>- noncoherent/coherent</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Read Response</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Target Done</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Broadcast Message</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Atomic Read-Modify-Write</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Address Extension</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Flush posted writes</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Fence for posted requests</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Extended Flow Control for VCSets 0-7</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td>Link Synchronization and Error Packet</td>
<td><img src="image.png" alt="Image" /></td>
</tr>
</tbody>
</table>
Connecting Multiple Devices to Host
HyperTransport I/O Device Configurations

- "Cave" Single-Link
  - Host → Endpoint
  - Primary

- "Tunnel" Dual-link
  - Host → Daisy Chain
  - Primary

- "Bridge" with Tunnel
  - Host → Multiple Daisy Chains with bridge to other I/O protocols
  - Primary
  - Secondary

- "Bridge" without Tunnel
  - Host → Daisy Chain
  - Primary
  - Secondary

Upstream ▲
Downstream ▼
Routing to Target Device

HyperTransport Request Packet Format With Address

<table>
<thead>
<tr>
<th>Bit-Time</th>
<th>7</th>
<th>6</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>SeqID[3:2]</td>
<td>Cmd[5:0]=type of command</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Pass PW</td>
<td>SeqID[1:0]</td>
<td>UnitID[4:0]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Command-Specific</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Command-Specific</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Addr[15:8]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Addr[23:16]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Addr[31:24]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Addr[39:32]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Pipelining Multiple Requests
Communication Between Two I/O Devices
Communication Between Two I/O Devices (cont.)
Transmission Error Handling
Priority Request Interleaving™

1. Data transfer 1 under way
2. PRI While transfer 1 carries on…
3. Data transfer 2 initiates while data transfer 1 still under way

Lowest Achievable Latency
Interconnecting Multiple Hosts

- Coherent vs. Non-Coherent HyperTransport
- cHT-Enabled Links Configurable at Boot Time

HT = HyperTransport™ technology
Cache Coherence Support
Proprietary Technology
AMD cHT Basics

- On-Chip Support for Up to 8 CPUs
- Broadcast-Based 3-Hop Invalidation Cache Coherence Protocol
AMD cHT Read Request Example

Step 1

Memory 0

P0

P1

Read Cache Line

P2

P3

Memory 2

Memory 1

Memory 3
AMD cHT Read Request Example (cont.)

Step 2

Analysis of the diagram:
- **Step 2**: This step involves a read cache line operation.
- **Memory 0** is connected to **P0**.
- **P0** is connected to **P1** via a red arrow labeled "Read Cache Line".
- **P1** is connected to **P2**.
- **P2** is connected to **P3**.
- **P3** is connected to **Memory 3**.
- **Memory 1** is connected to **P1**.
- **Memory 2** is connected to **P0**.

The diagram illustrates the flow of the read request from **Memory 0** via **P0** to **P1**, then to **P2** and finally to **P3**.
AMD cHT Read Request Example (cont.)

Step 3

Read Cache Line

Snoop Request P0

Snoop Request P2

P0

Memory 0

P1

Memory 1

P2

P3

Memory 2

Memory 3
AMD cHT Read Request Example (cont.)

Step 4

Memory 0

P0

Snoop Response P0

P1

Memory 1

P2

Memory 2

Snoop Request P3

P3

Memory 3
AMD cHT Read Request Example (cont.)

Step 5

Memory 0

Read Response M0

P0

P1

Memory 1

Snoop Response P0

P2

Snoop Response P2

P3

Memory 2

Memory 3
AMD cHT Read Request Example (cont.)

Step 6

Memory 0

P0

P1

Read Response M0

P2

P3

Snoop Response 1

Memory 1

Memory 2

Memory 3
AMD cHT Read Request Example (cont.)

Step 7

Memory 0

P0

Memory 1

P1

Read Response M0

P2

Memory 2

P3

Memory 3
AMD cHT Read Request Example (cont.)

Step 8

Diagram showing a network of nodes labeled P0, P1, P2, and P3, connected by lines, with memory nodes labeled Memory 0, Memory 1, Memory 2, and Memory 3.
AMD cHT Read Request Example (cont.)

Step 9

Memory 0

P0

Memory 1

P1

P2

Source Done to M0

Memory 2

P3

Memory 3
AMD cHT Read Request Example (cont.)

Step 10

Source Done to M0

Memory 0
P0

Memory 1
P1

Memory 2
P2

Memory 3
P3
AMD cHT HT-Assist (Probe Filter)

- Old cHT Broadcast Protocol Broadcasts Probes to Invalidate Copies even if Memory Line is Clean
AMD cHT HT-Assist (Probe Filter) (cont.)

- Sparse Directory Cache Next to Memory Controller
- Rule: If a Line is Cached, it has an Entry in PF
  - Replacement Policy Makes Room for New Lines
- Enhanced behavior:
  - No Probing for Uncached Lines
  - Directed Probe to Request Copy of Cached Line
- Benefits:
  - Significantly Less Bandwidth Use
  - Shorter Access Latency, Mainly for Uncached Lines

For More Details:

"Blade Computing with The AMD Magny-Cours Processor"
Presented by AMD - Pat Conway, Hot Chips 2009
AMD cHT HT-Assist (Probe Filter) (cont.)

**PF – clean data**

- **Home Node**
- **Req Node**

**PF Lookup**

**DRAM Resp**

**PF – dirty data**

- **Home Node**
- **Req Node**

**Directed Probe**

**Cache Resp**
Beyond Motherboards
HTX Connector

- Low Latency CPU-to-High-Perf. Subsystem Direct Connect
- Removes Performance Bottlenecks in Compute-Intensive Data Processing and Acceleration Functions
- Complements PCI-Class Interconnects
- Link Splitting Capability
### HTX Specification Evolution

<table>
<thead>
<tr>
<th>Feature</th>
<th>HTX</th>
<th>HTX3</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Clock Rate</td>
<td>800 MHz</td>
<td>2.6 GHz</td>
<td>12” Trace length</td>
</tr>
<tr>
<td>Max Bandwidth x Lane</td>
<td>1.6 GT/s</td>
<td>5.2 GT/s</td>
<td>Bi-directional</td>
</tr>
<tr>
<td>Max Bandwidth Aggregate</td>
<td>6.4 GB/s</td>
<td>20.8 GB/s</td>
<td>Bi-directional 16-Bit HT link</td>
</tr>
<tr>
<td>HT3 Link Splitting Support</td>
<td>NO</td>
<td>YES</td>
<td>HT link can be 1x 16-Bit or 2x 8-Bit for multi-CPU support</td>
</tr>
<tr>
<td>HT3 Extended Power Management</td>
<td>NO</td>
<td>YES</td>
<td>LDTREQ# Signal Added to participate in x86 power states</td>
</tr>
<tr>
<td>Extended FPGA Guidelines</td>
<td>NO</td>
<td>YES</td>
<td>Incorporated field-proven recommendations</td>
</tr>
<tr>
<td>Full Backward Compatibility</td>
<td>--</td>
<td>YES</td>
<td>Level shifters and signal allocation</td>
</tr>
</tbody>
</table>
New in HyperTransport™ 3
HT3 - Link Splitting

All Links or Individual Links

1x 4-Bit → 2x 2-Bit
1x 8-Bit → 2x 4-Bit
1x 16-Bit → 2x 8-Bit
1x 32-Bit → 2x 16-Bit
HT3 - Link Splitting (cont.)

Extended SMP Topologies Enabled
HT3 – Dynamic Power Management

- Dynamic Link Width, Clock Rate and Voltage Scaling

- Partial Link Shutdown via Link Splitting
HT3 – AC Mode (Optional – Enabled if Needed)

- 8b/10b Encoding
  - Lower Bandwidth, Higher Latency than DC Mode
- DC/AC Autoconfiguration
- TX Equalization

In-System

DC Mode

HT3 Spec at <= 12 ln

AC Mode

HT3 Spec at <= 3 Ft

Backplane

Chassis-to-Chassis

Transmit with Pre- and Post-Cursor De-Emphasis
HT3 - AC Mode + Link Splitting

Maximize Multi-Processor Expansion Capability

Chassis 1

To other CPU subsystems

CPU

CPU

CPU

CPU

HTX Connector

I/O

Chassis n

To other CPU subsystems

CPU

CPU

CPU

CPU

HTX Connector

I/O

16-Bit DC HyperTransport Links

8-Bit DC HyperTransport Links

16-Bit AC HyperTransport Links
HT3 – Hot Plugging

- Add/ Remove Devices from HT Fabric Without Disrupting Other Operations
  - Defined Link Termination Methods
  - Transaction Termination Behaviors
  - Sync Flood Isolation
  - Link Training Times
- Parameter Configuration Mechanism

High-Availability Applications
Server and Storage Markets
HT Specifications Evolution

- **HT 1.0**: 2001
- **HT 1.1**: 2002
- **HT 2.0**: 2004
- **HTX**: 2004
- **HT 3.0**: 2005
- **HT 3.1**: 2008
- **HNC 1.0**: 2009

**Chip-to-Chip**
- 17.7M HT-Based Systems Shipped (Note 1)

**Chip-to-Chip and Beyond**
- 62.7M HT-Based Systems Shipped (Note 2)

**Notes**
- Note 1: by end of 2003 – Source InStat
- Note 2: by end of 2008 – Source InStat
- Note 3: High Node Count Specification 1.0 - Accessible/Useable by Promoter and Contributor Members Only
## HT Specifications Evolution (cont.)

### Special Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>HT 1.x</th>
<th>HT 2.0</th>
<th>HT 3.0</th>
<th>HT 3.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Max Clock Speed</td>
<td>800 MHz</td>
<td>1.4 GHz</td>
<td>2.6 GHz</td>
<td>3.2 GHz</td>
</tr>
<tr>
<td>Max Aggregate Bandwidth (32-bit links)</td>
<td>12.8 GB/s</td>
<td>22.4 GB/s</td>
<td>41.6 GB/s</td>
<td>51.2 GB/s</td>
</tr>
<tr>
<td>AC Operation – Capacitive Coupling (optional) with AC/DC Autosensing, Autoconfiguration</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Link Splitting (un-ganging) Each HT Link Split into 2x Half-Width Links</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Hot-Plugging</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Dynamic Link Clock/Width Adjustment</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DirectPackets™ Data Streaming - +16 Virtual Channels (22 total), Peer-to-Peer Support, Native Packet Handling</td>
<td>HT 1.1 only</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>PCI Express Mapping</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Common Features

- 2 x Unidirectional, Low-Voltage-Differential-Signaling Links per HT Bus
- Scalable Link Width (2-bit to 32-bit)
- Asymmetric Link Support (different link widths)
- Asynchronous Link Operation
- Clock Forwarding (no SerDes latency penalty)
- DC Operation (direct signal coupling)
- PCI, PCI-X Mapping
- Priority Request Interleaving™ (lowest I/O latency)
- Virtual Channels Support (6)

Error Retry
Beyond HT3
High Node Count (HNC) Specification Complements HT3

HyperTransport Link Specifications with HNC Extended Addressing Features as Options

100% Backward Compatible
HNC Model

• Physically Distributed, Logically Shared Memory System (PGAS)
• Cache Coherence Not Mandatory, but Multiple Coherent Domains Can be Part of it
• Nest: Each HNC-Clustered Component
HNC Model (cont.)

- DestNest Identifier Location in Request and Response Packets was Carefully Designed to Minimize Routing Time
- Optimized for Most Common Cases
- Minimal Extra Latency and Bandwidth Use
HNC Extensions

- Interconnecting Large Numbers of Hosts Requires:
  - Addressing Scheme: Ability to Address Large Number of Devices. Affects Packet Format
  - Network Topology: Support for High Connectivity Topologies that Enable More Concurrent Transmissions
  - Routing Mechanisms

Addressing Scheme and Packet Format

Implementation-Dependent

Network Topology and Routing Mechanisms
HNC Specification Removes

Latency Overhead of Message-Passing Protocols

Through HyperTransport Protocol Encapsulation

Using System-Wide HNC Addressing Scheme
**HNC Specification Enables**

Scalable Global Resource-Sharing & Partitioning

- **Flash / SSD**
- **HD-DRAM**
- **DRAM**

Localized Compute, Memory and Peripheral Resources Serve the Whole Cluster
HNC Specification Delivers

Resource-Sharing & Modularity Resulting in
Less Over-Provisioning
Best Resource Utilization
Maximum Energy Efficiency
Beyond Conventional
HT Everywhere
HyperTransport Empowers More Processor Architectures and Market Segments than its AMD Roots May Suggest
In Video Gaming

"If the HyperTransport link in the Xbox was a pipe carrying water, and every bit of information equalled one gallon, the pipe would fill up the Pacific Ocean every second!"

NVIDIA Website
In Communications Processors

BCM1125H 64-bit MIPS Processor
BCM 1250 64-bit MIPS Processor
BCM 1280 64-bit MIPS Processor
BCM1480 4-Core 64-bit MIPS Processor

BCM 1480 x 4
16 Cores
Coherent HT SoC Cluster

RM9150 64-bit MIPS Processor
RM9200/9100 64-bit MIPS Dual/Single Processor
RM11200 64-bit MIPS Multiprocessor
RM9220/9224 64-bit MIPS Multiprocessor
In PowerPC Computing
In PowerPC Embedded Systems

IBM

HyperTransport

South bridge

North bridge

Elastic interface 0 (EIO)

PCI-X_0

PCI-X_1

HyperTransport

PCI

USB

LPC

Super I/O

Boot ROM

DDR SDRAM

HyperTransport

I2C

PowerPC 970FX_0

PowerPC 970FX_1

Service processor
In Knowledge-Based Processors
In Network/Media Processors

Chesapeake
In Reconfigurable Accelerators

RPU 110
RPU 100
Opteron Socket Plug-In

RCHTX
Accelerator Board
HTX Slot Connector

XD 2000F
XD 1000
Opteron Socket Plug-In
In Routing Processors

Classification and action engine capable of consistent association between CPU cores and their data streams, resulting in efficient multi-core system load balancing of network traffic processing and lower power consumption.
In State-of-The-Art Supercomputers

XT5

Cray SeaStar2+™
3D Torus Interconnect
9.6 GB/s

Cray

HyperTransport
6.4 GB/s

XR1
Reconfigurable Blades

FPGA
FPGA

HyperTransport
6.4 GB/s

HyperTransport
Consortium

Copyright HyperTransport Consortium, 2009
In x86 Emulation Processors

- Petascale Performance Target
- 4-Core MIPS64-Based with 200+ More Instructions for x86 Translation and Acceleration
- 16 GFLOPS at 1GHz
- 10W of Power
Also in Computing Platforms in which...

One Would Not Expect to Find HyperTransport!
Vibrant Open Standard Consortium

- Founded 2001
- Controls, Promotes and Licenses HyperTransport Technology to Global Industry
- 60 Members-Strong, Including Technology Leaders
Vast HT Products Ecosystem Fosters Technology Strength

From IP to Software
From Technology Licensor to Members Business Partner

- HT Enablement
- Product Validation Services

Online Design, Simulation and Test Support

- HT Enablement
- Product Validation Services
All About HT on Consortium Web Portal

www.hypertransport.org

Members General & Extended HT Technical Support Database

Public Web Portal

HyperTransport Specifications

Design Support

White Papers

Members Material

HT Link

HT High Node Count

HTX and DUT Interface Connectors
New HyperTransport Book

700-Page Must Have HT Tutorial

Covers All HT Link and HTX Specification Releases

Available Online from MindShare in Paper and eBook Formats

www.mindshare.com
Additional Information

• HyperTransport™ I/O Link Specification 3.10a
• HyperTransport HTX™ Slot Connector Specifications
• HyperTransport™ High Node Count Specification 1.0
• White papers
• Dally & Towles, “Principles and Practices of Interconnection Networks”
• Duato, Yalamanchili & Ni, “Interconnection Networks: An Engineering Approach”
  • All three books published by Morgan Kaufmann
Thank You!