

AMD's Direct Connect Architecture

Pat Conway

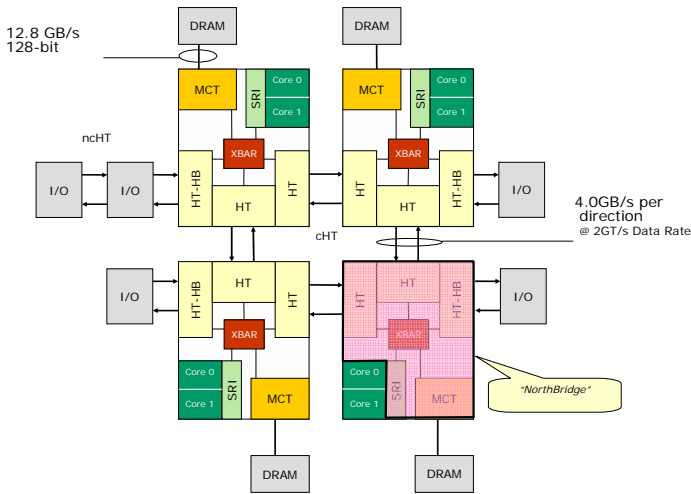
Principal Member of Technical Staff

Glueless Multiprocessing with Operton CMP processors

This poster illustrates the variety of Operton system topologies built to date using HyperTransport and AMD's Direct Connect Architecture for glueless multiprocessing and summarizes the latency/bandwidth characteristics of these systems.

Building and using Operton systems has provided useful, real world lessons which expose the challenges to be addressed when designing future system interconnect, memory hierarchy and I/O to scale up both the number of cores and sockets in future x86 CMP Architectures.

4 Node Operton MP Architecture

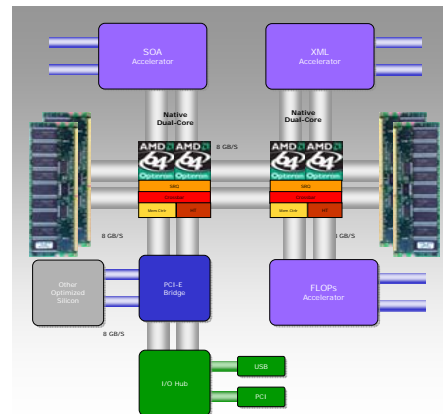
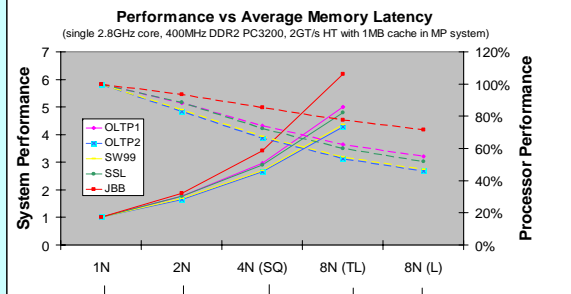


Lessons Learned #1

Memory Latency is the Key to Application Performance!

Torrenza - HyperTransport-based Accelerators

Imagine it, Build it

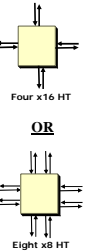


- Open platform for system builders
 - 3rd Party Accelerators
 - Media
 - FLOPs
 - XML
 - SOA
- AMD Operton™ Socket or HTX slot
- HyperTransport interface is an open standard see www.hypertransport.org
- Coherent HyperTransport interface available if the accelerator caches system memory (under license)
- Homogeneous/heterogeneous computing
- Architectural enhancements planned to support fine grain parallel communication and synchronization

"Fully Connected 4 Node and 8 Node Topologies are becoming Feasible"

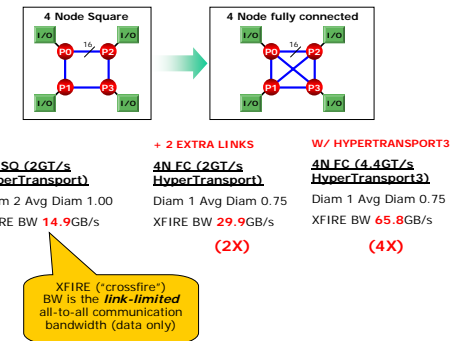
Additional HyperTransport™ Ports

- Enable Fully Connected 4 Node (four x16 HT) and 8 Node (eight x8 HT)
- Reduced network diameter
 - Fewer hops to memory
 - cHT packets visit fewer links
 - HyperTransport3 up to 5.2GT/s
- Benefits
 - Low latency because of lower diameter
 - Evenly balanced utilization of HyperTransport links
 - Low queuing delays



Low latency under load

Fully Connected 4 Node Performance



Fully Connected 8 Node Performance

