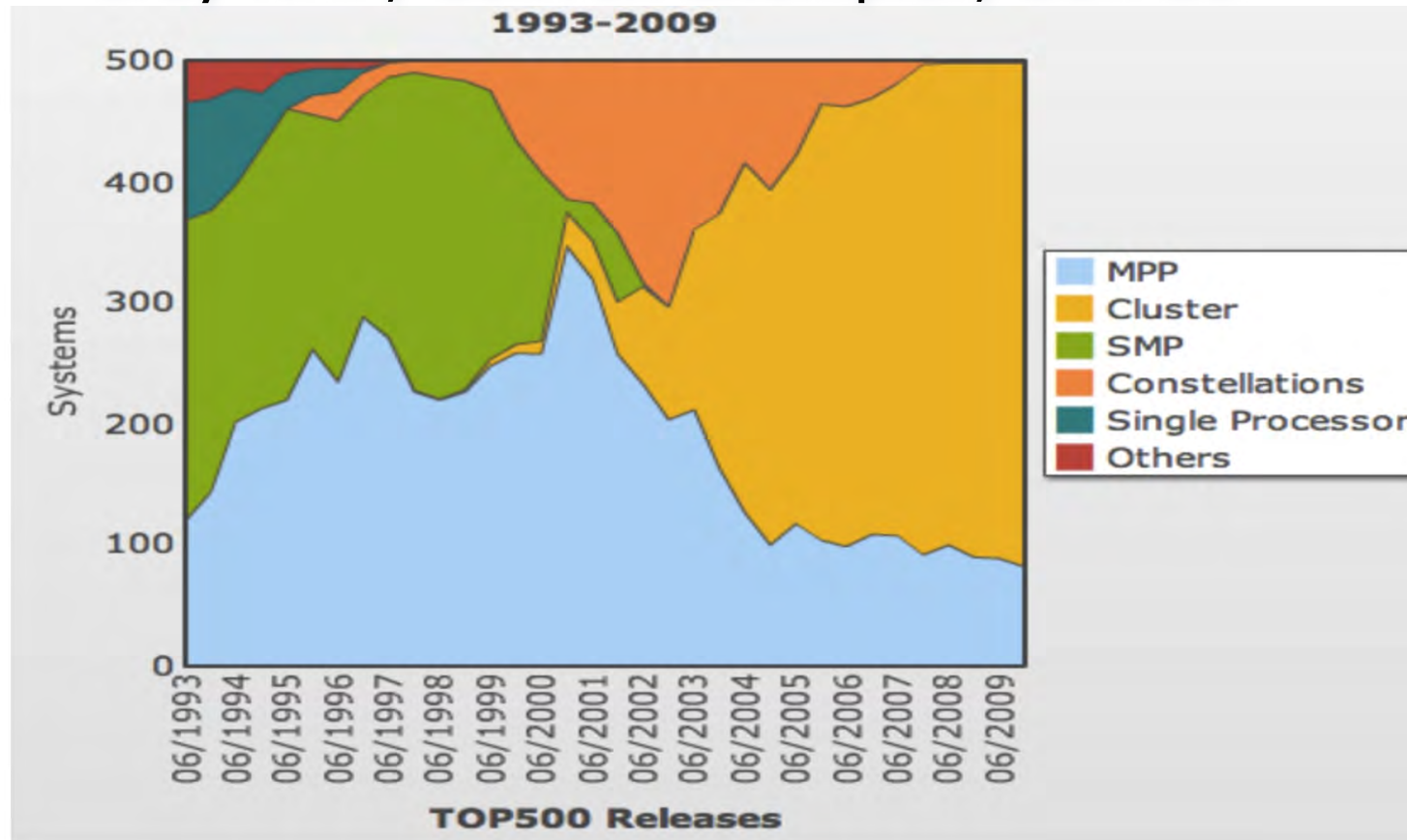


numascale

The Return of the SMP

ISC Hamburg, May31, 2010

- ▶ The expensive SMPs used to rule:
 - Cray XMP, Convex Exemplar, Sun ES



- ▶ NOW, the Clusters are dominating

- ▶ In the early 1990s, expensive SMPs ruled in HPC
 - Cray MPs, Convex Exemplar, Sun ES
- ▶ The MPPs were always in the shadow
 - Intel Paragon, Thinking Machines CM, CRAY T3
 - Similar price, more complex programming
- ▶ Then came the Clusters (Early 2000s)
 - $\approx 50x$ cheaper, complex programming (MPI)
- ▶ Clusters made HPC affordable and widespread

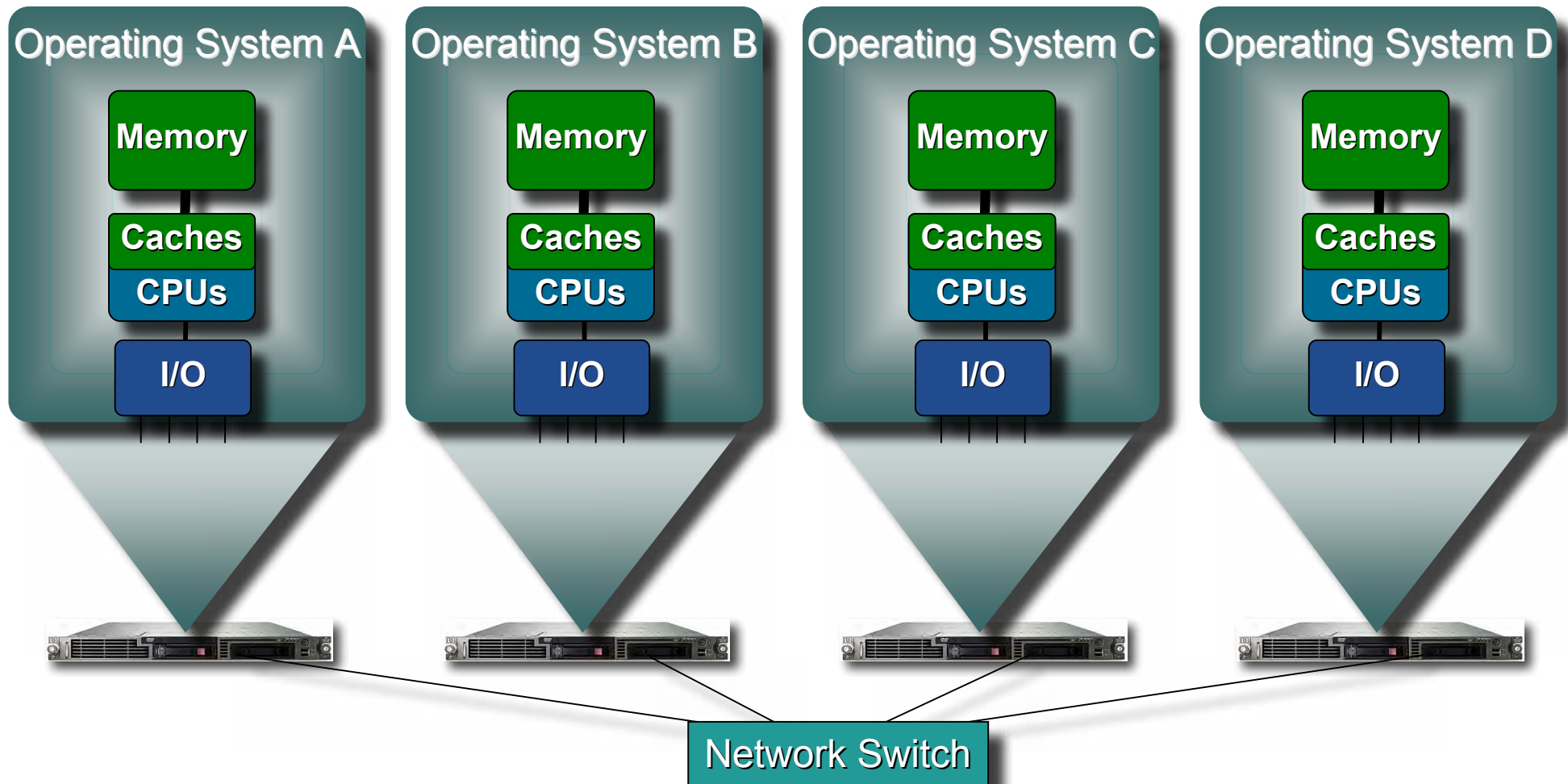
- ▶ Inexpensive multi-core CPUs
- ▶ Compelling programming model
 - Less code
 - Large memories - less effort, no data domain decomposition
- ▶ NumaCluster
 - Same programming model across an entire system
 - Reduced effort for system management
- ▶ Virtualization
 - More efficient utilization of Resources
 - ▶ CPUs, Memory, I/O

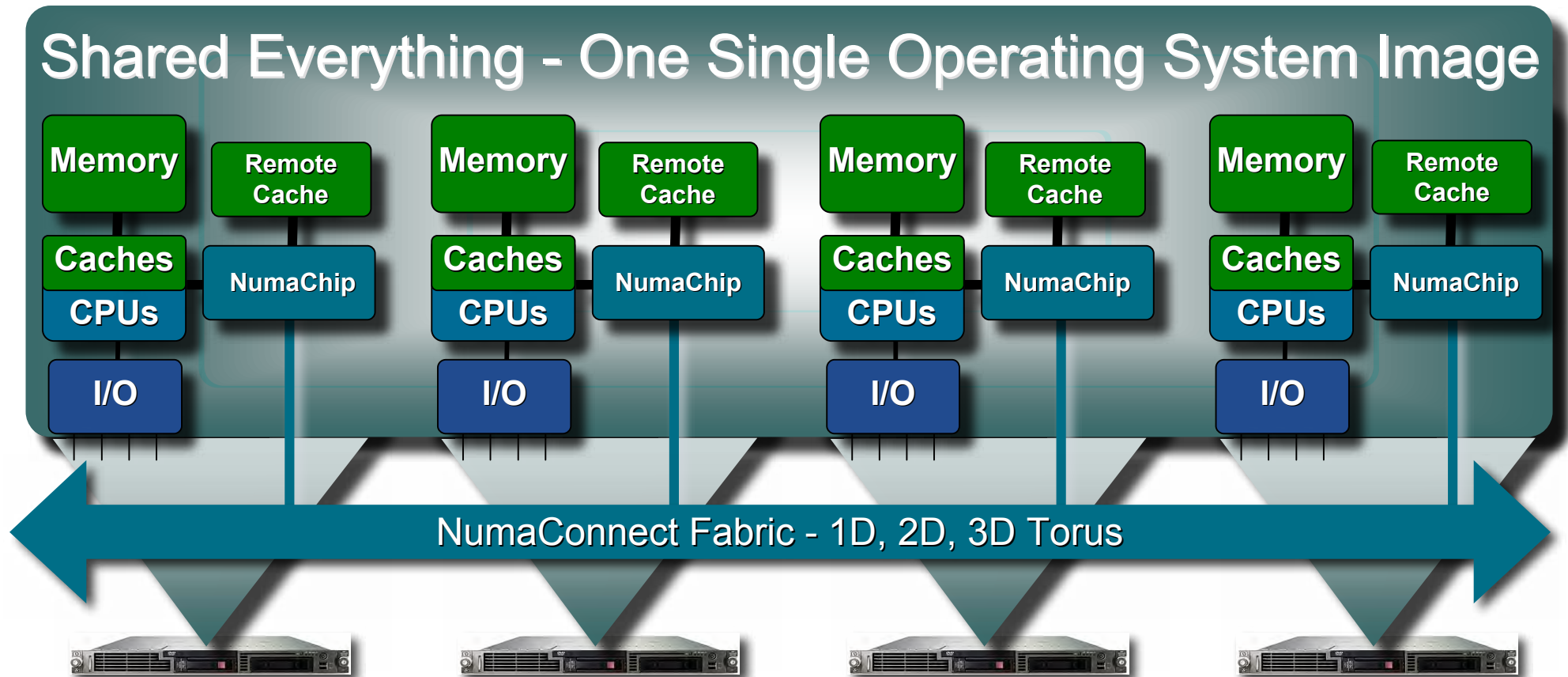
- ▶ System-wide cache coherency in hardware
 - 64Byte cache line granularity
- ▶ Standard Linux (or any other x86 - 64 OS)
 - Virtual Memory
- ▶ Runs any application - shared memory or message passing
- ▶ Virtualization
 - All system resources can be used by all processors
 - Run any number of virtual OS instances

- ▶ Convex Exemplar (Acquired by HP)
 - First implementation of the CC-NUMA architecture from Dolphin in 1994
- ▶ Data General Aviiion (Acquired by EMC)
 - Designed in 1996 with deliveries from 1997 - 2002
 - Used Dolphin's chips with 3 generations of processor/memory buses
- ▶ I/O Attached Products for Clustering



Individual Instances of the Operating System

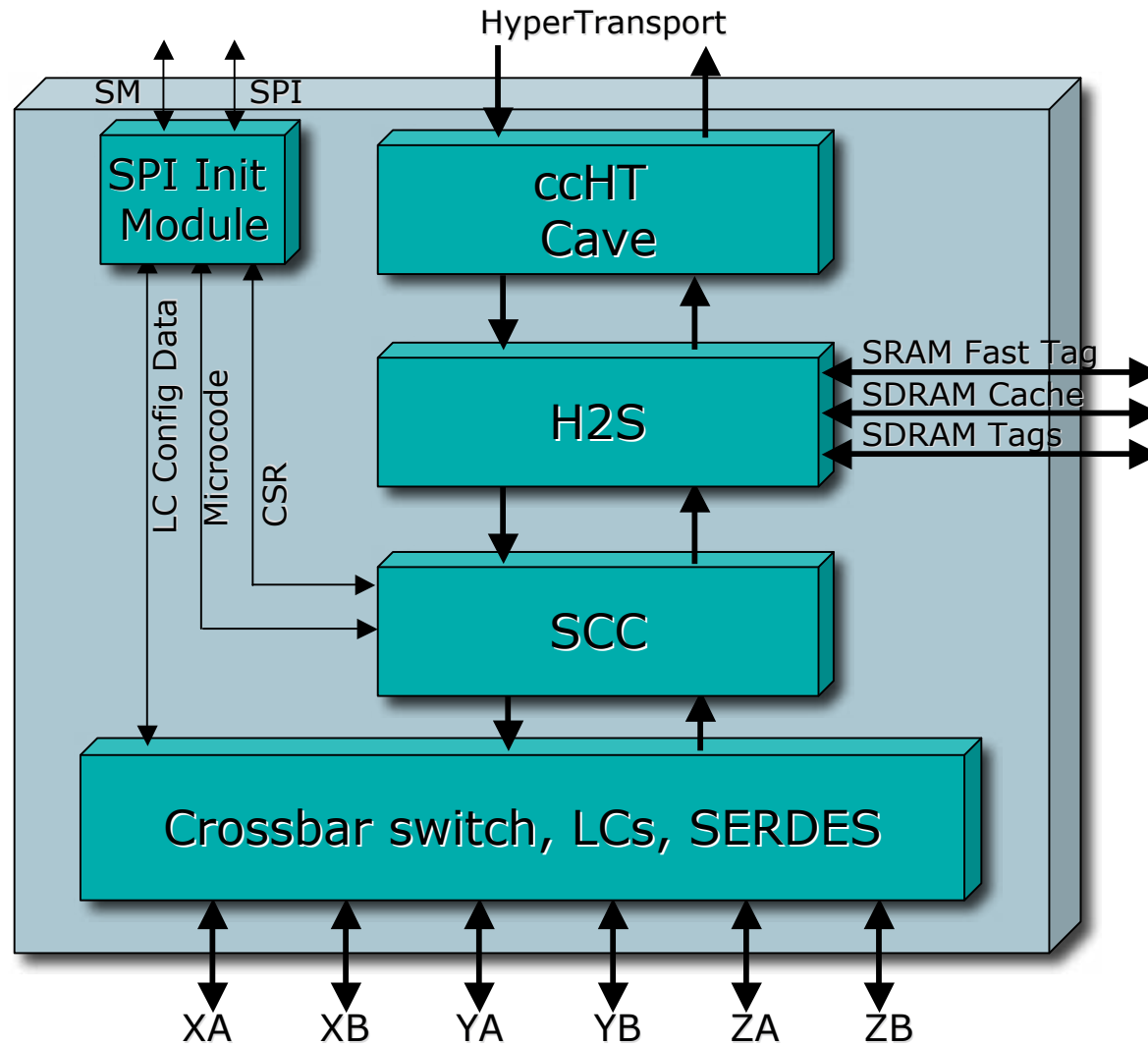




Capabilities like Mainframe - Price like Cluster

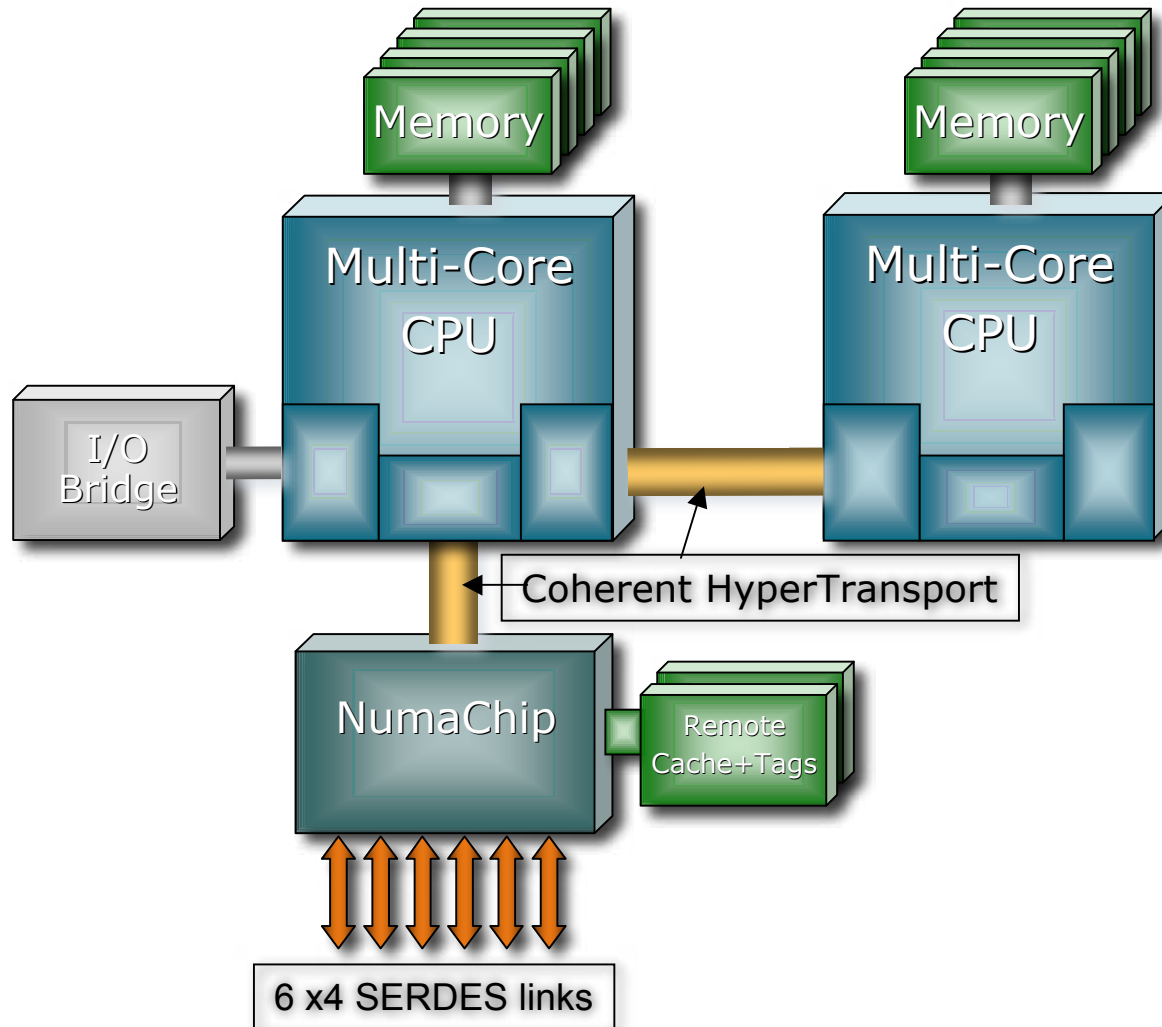
- ▶ 256TBytes physical address space
- ▶ Scalable, directory based cache coherency protocol
- ▶ Scalable On-Chip switch fabric (2-D, 3-D Torus)
- ▶ Configurable Cache for remote data (1 - 16GB/node)

NumaChip Top Block Diagram

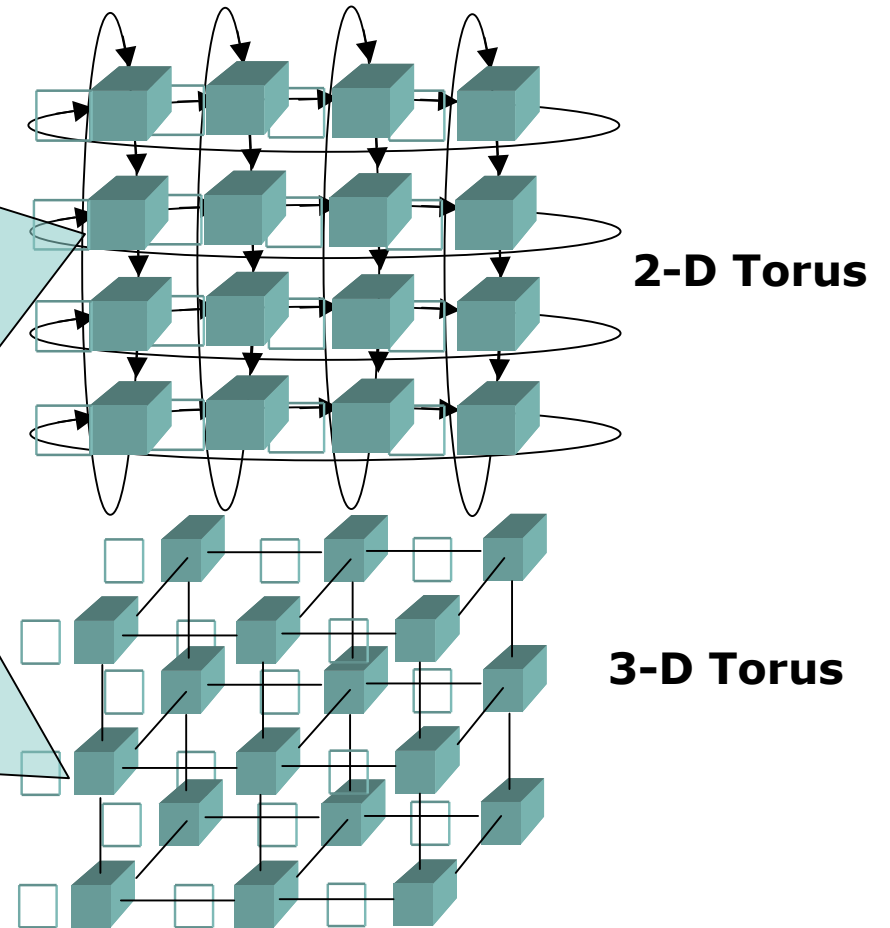
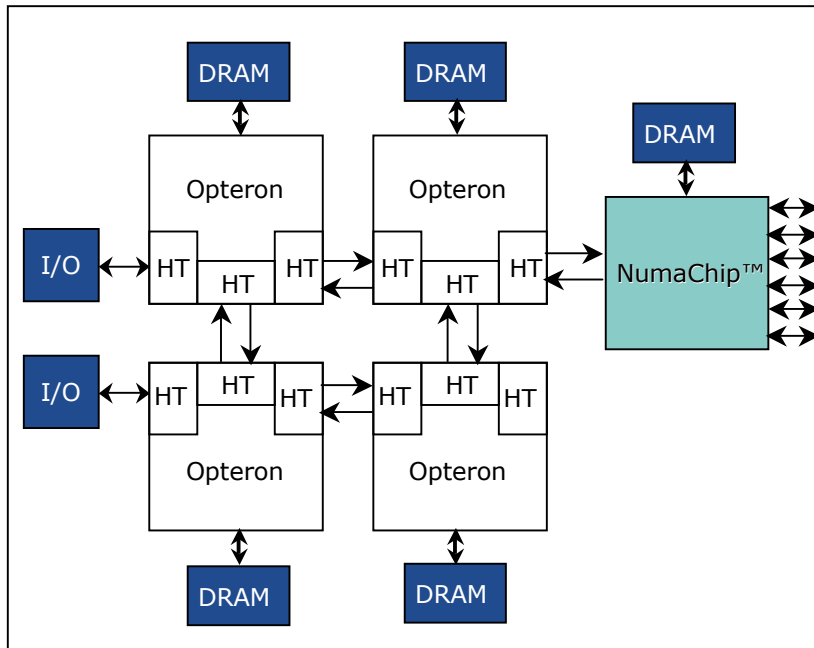


NumaChip™ Server Configuration

numascale



Multi-socket Node



6 links allow flexible system configurations in multi-dimensional topologies

- ▶ Operating System
 - Single System Image
 - Multiple System Image Partitions
 - ▶ Individual nodes
 - ▶ Multiple nodes

- ▶ User Applications
 - Shared, ccNUMA
 - Shared, NUMA (non-coherent)
 - Pure Message Passing (MPI or others)
 - or any combination

- ▶ Linux
- ▶ Windows Server
- ▶ Solaris
- ▶ Unix

- ▶ Use any Programming Model Available for the Node on the whole System
 - OpenMP
 - MPI
 - Threads
 -

- ▶ NO Application Changes Required!

- ▶ Visit our Website

- www.numascale.com

- ▶ Contact us for more information

- Einar Rustad, VP Business Development

- ▶ er@numascale.com

- Jon Snilsberg, VP Sales and Marketing

- ▶ js@numascale.com