Cray T3D Overview

- Hundreds or thousands of DEC Alpha RISC processors arranged in a 3D torus
- Each processor has a local memory, but the memory is globally addressable
  - NUMA = Non Uniform Memory Access = can access all memory, but can access local memory faster than remote memory
  - Can “prefetch” data before it is needed, so it is ready when necessary
- A multiprocessor Cray Y-MP C90 or Model E Cray Y-MP C90 is used as a UNIX server and used with the Cray T3D
  - OS is MAX, Massively Parallel UNIX
  - I/O attached to Y-MP is available to T3D
  - Shared file system between Y-MP & T3D
  - Some applications run on the Y-MP, others on the T3D, some on both

Cray T3D Organization

- Processors can be divided into partitions
  - System administrator can define a set of processors as a pool, specifying batch use, interactive use, or both
  - User can request a specific number of processors from a pool for an application, MAX selects that number of processors and organizes them into a partition
- 3D torus is virtual, and the physical layout includes redundant nodes that can be mapped into the torus if a node fails
- Programming models
  - Message passing based on PVM
  - High-Performance FORTRAN — implicit communication
  - MPP FORTRAN — implicit and explicit communication

Cray T3E Overview

- T3D = 1993,
  T3E = 1995 successor (300 MHz, $1M),
  T3E-900 = 1996 model (450 MHz, $.5M)
- T3E system = 6–2048 processors,
  3.6–1228 GFLOPS,1–4096 GB memory
  - PE = DEC Alpha 21164 processor (300 MHz, 600 MFLOPS), local memory, control chip, router chip
  - GigaRing Channel attached to each node and to I/O devices and other networks
  - T3E-900 = same w/ faster processors, up to 1843 GFLOPS
- Ohio Supercomputer Center (OSC)
  - T3E with 128 PEs (300 MHz), 76.8 GFLOPS, 128 MB memory / PE
  - T94 (T90 w/ 4 450 MHz processors), 8 GFLOPS

Convex Exemplar SPP-1000 Overview

- Shared-memory MIMD multiprocessor
  - 4–128 HP PA 7100 RISC processors, up to 25 GFLOPS
  - 256 MB – 32 GB memory
  - Logical view: processors – crossbar – shared memory – crossbar – peripherals
- System made up of “hypernodes”, each of which contains 8 processors and 4 cache memories (each 64–512MB) connected by a crossbar switch
  - NUMA = hardware support for global shared memory access
    - Also caches in the processors
  - Hypernodes connected via Coherent Toroidal Interconnect, an implementation of IEEE standard 1596-1992, Scalable Coherency Interface — keeps all the caches consistent with each other

Convex Exemplar SPP-1000 Processors

- HP PA7100 RISC processor
  - 555,000 transistors, 0.8µ
- 64 bit wide, external 1MB data cache & 1MB instruction cache
  - Reads take 1 cycle, writes take 2
- Can execute one integer and one floating-point instruction per cycle
- Floating Point Unit can multiply, divide / square root, etc. as well as multiply-add and multiply-subtract
  - Most fp operations take two cycles, divide and square root take 8 for single precision and 15 for double precision
- Supports multiple threads, and hardware semaphore & synchronization operations