Connection Machines – Thinking Machines Corporation

SIMD Connection Machines

- CM-1

  12D hypercube with 16 single bit PEs at each node.

  Designed to exploit parallel processing for Artificial Intelligence.

- CM-2

  Added 1 *14 Mflop* floating point accelerator for every 32 PEs.

  In 1989 CM-2 won IEEE Gorgon Bell Award for 5.6 GFlops Sustained Performance on a seismic modelling problem[1].

MIMD Connection Machine

- CM-5

  A massively parallel MIMD computer which draws on the success of its SIMD ancestry.

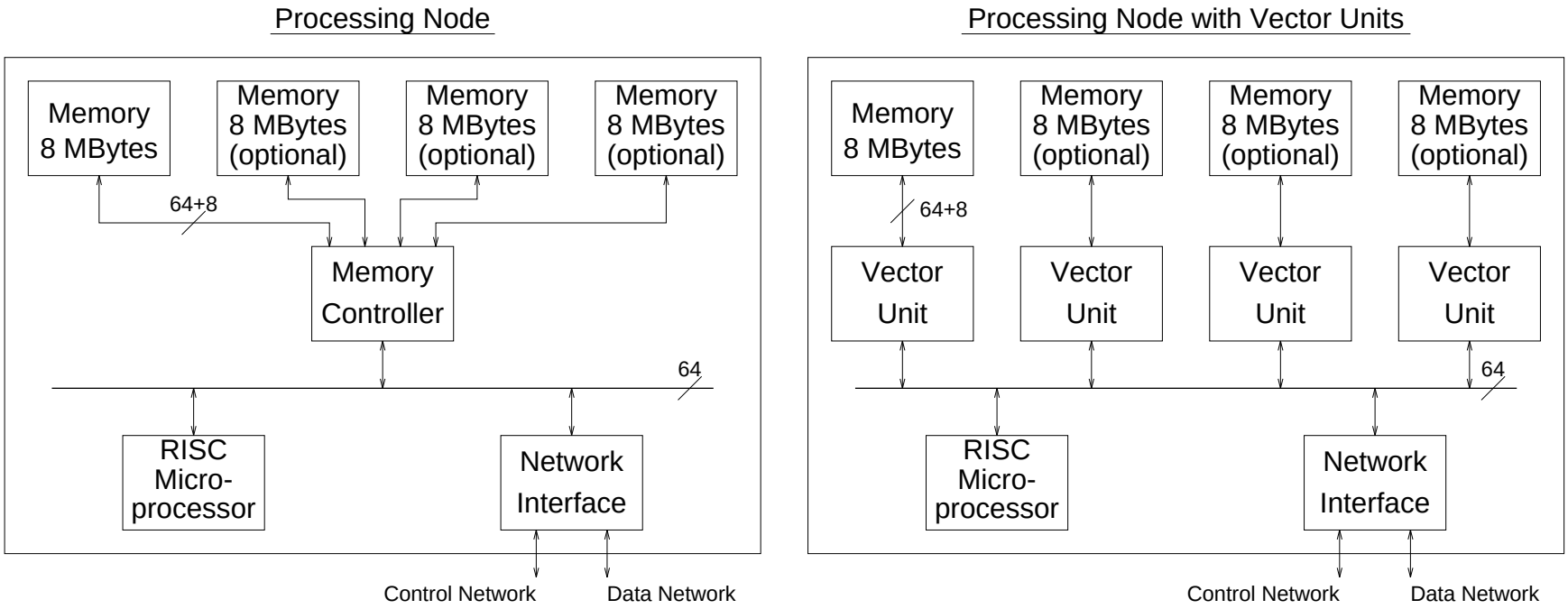
[1] 8-Pipe Cray Y-MP has a peak rating of 3 GFlops!

CM-5 A Usable TerraFlop Computer

- Power
  - 16384 Nodes
  - Each node 128 MFlops.
  - Total: 2 TerraFlops
- Memory Bandwidth
  - Up to 512 MBytes/sec per node
- Communications
  - 20 MBytes/sec per node for nearest neighbour communications (over simulated 2D Grid).
  - Bisection bandwidth is 5 MBytes/sec per node.
- Programming Paradigm
  - *Data Parallel on an MIMD machine!*

Connection Machine CM-5

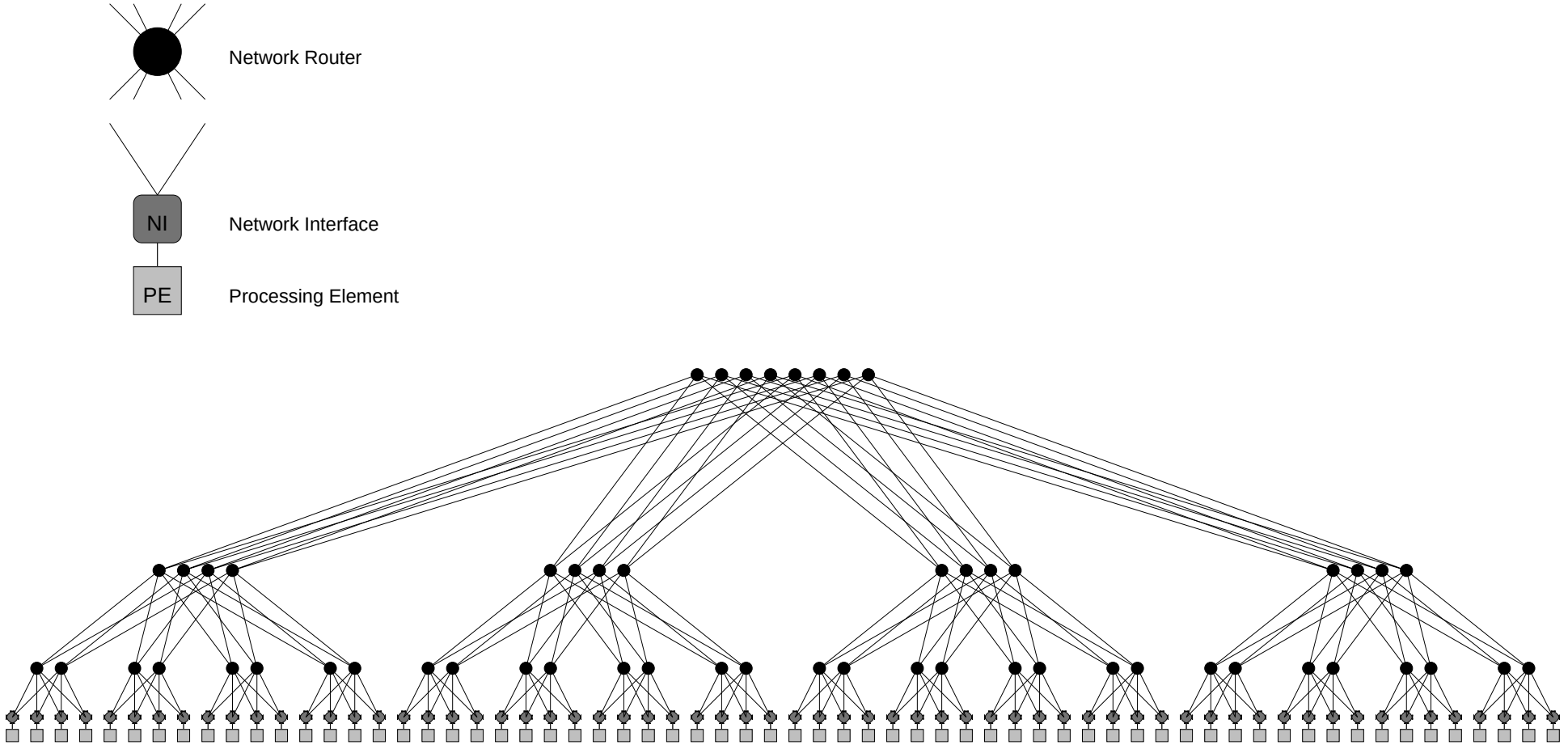
PE Architecture



- RISC microprocessor is a SPARC.

Connection Machine CM-5

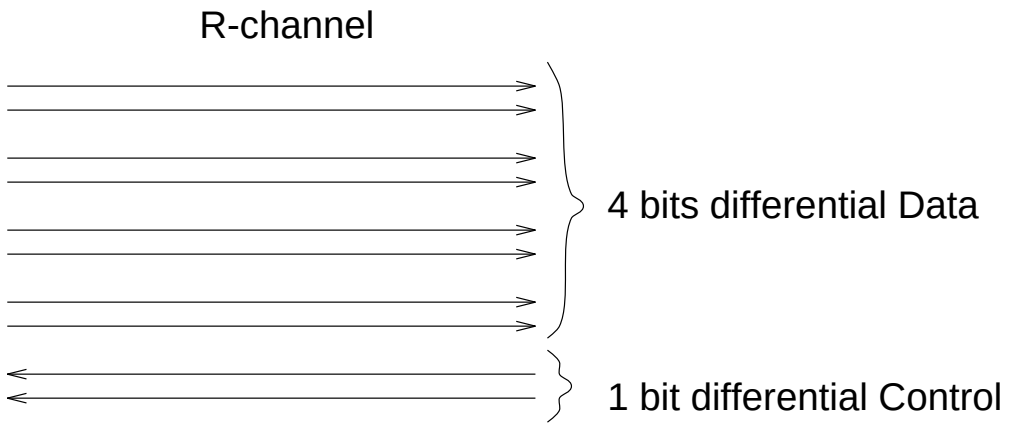
CM-5 Architecture (Data Network)



16004

Inter-node Links

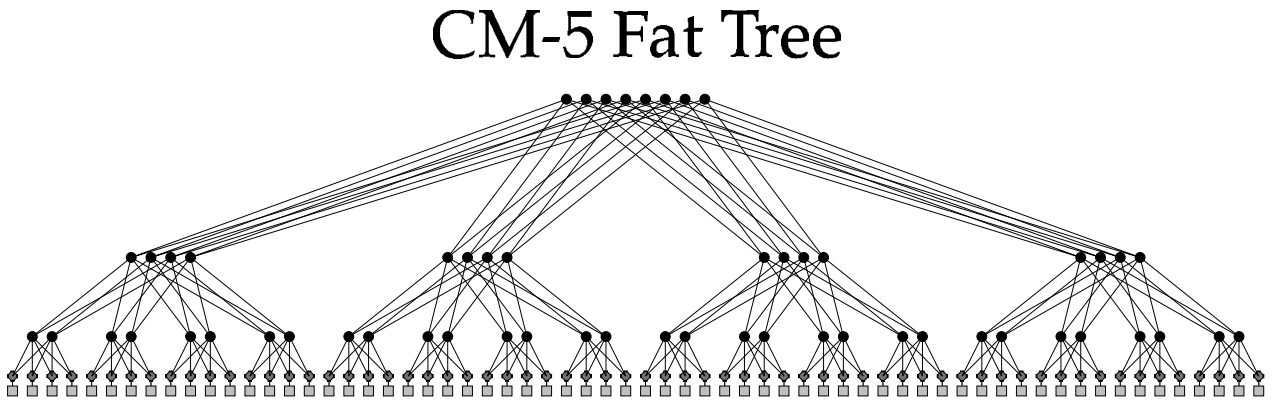
- Each bi-directional link consists of two R-channels, one in each direction.
- An R-channel consists of 4 differential pairs for data and 1 differential pair for control.



- Communication is synchronous - *all PEs and routing nodes share the same clock!*

Connection Machine CM-5

The CM-5 employs Virtual Cut-Through Routing over a quaternary fat tree network.



- Quaternary fat tree gives the same diameter as hypercube.
- For the first 3 levels of the tree each node has only 2 parents[2]. This explains the limited bisection bandwidth but also helps to reduce the number of routers required in a full system from 114,688 to 22,528.

[2]nodes in subsequent levels have the expected 4 parents

Scalability

- Bisection Bandwidth

  Increases in proportion to the number of processors.

  - scalable from 64-16384 nodes.
  - for small machines, 16-64 nodes, scaling is lost due to topology of lower tree.
  - large machines are expensive - increased proportion of routing nodes.

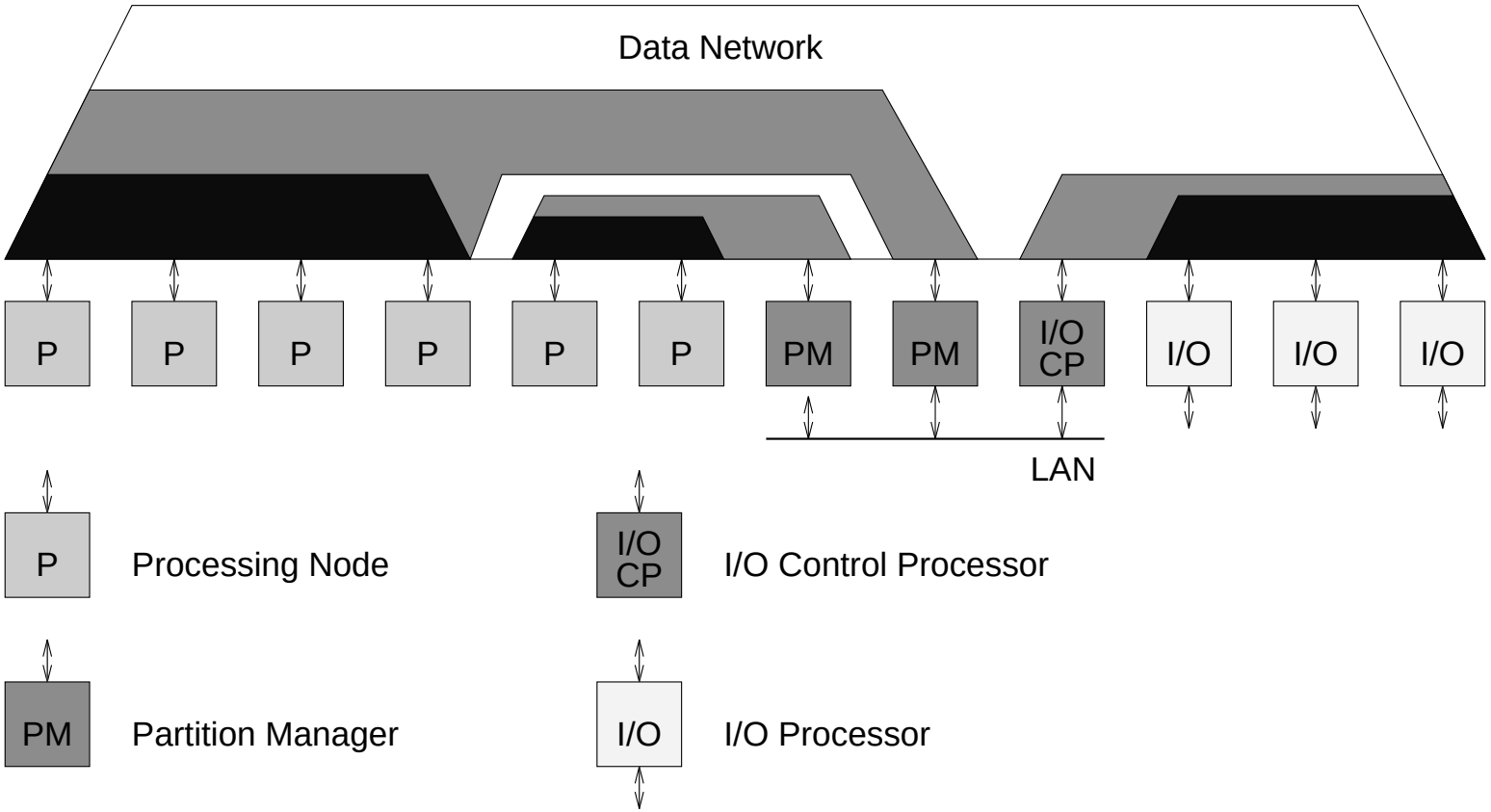
- Actual Machines

  A full 16384 node, 2 TerraFlop machine would have cost $100,000,000 this is possibly why none were built.

	CM-5	
No of PEs	1056	
Theoretical Peak	135.1	GigaFlops
Measured (Linpack-MP)	59.7	GigaFlops
Year	1993	

Connection Machine CM-5

Partitions & I/O



Data Parallel Programming

- Although each *Processing Node* is capable of independent *Process Parallel* type operation, the software for CM-5 supports a *Data Parallel* SIMD type programming style.
- Supports CM-1/2 Data Parallel languages; CM Fortran, C* & *Lisp.
- The processing nodes can be divided into partitions.
- Each partition is controlled by a *Partition Manager*
  - a standard SPARCstation with a *Network Interface* and software to drive it.
- Each Partition Manager may execute a number of *Data Parallel* processes[3] on its own partition, any of which may communicate with a process from another partition.

[3]The different processes are timesliced in a UNIX type fashion.