Thinking Machines Corporation's Connection Machine CM-2

- Distributed-memory SIMD (bit-serial)
- Thinking Machines Corp. founded 1983
  - CM-1, 1986 (1000 MIPS, 4K mem / proc)
  - CM-2, 1987 (2500 MFLOPS, 64K…)
- Programs run on one of 4 front-end processors, which issues instructions to the Parallel Processing Unit (PE array)
  - Control flow and scalar operations run on front-end processors, parallel operations run on PPU
  - A 4x4 crossbar switch (Nexus) connects the 4 front-ends to 4 sections of the PPU
  - Each PPU section is controlled by a sequencer, which receives assembly language (Paris) instructions and broadcasts micro-instructions to each processor

CM-2 Processors

- Each node contains 32 processors (implemented by 2 custom processor chips), 2 floating-point accelerator chips, and memory chips
- Processor chip (contains 16 processors)
  - Contains ALU, flag registers, etc.
  - Contains NEWS interface, router interface, and I/O interface
  - 16 processors are connected in a 4x4 mesh to their N, E, W, and S neighbors
- RAM memory
  - 64Kbits, bit addressable
- FP acceleration (2 chips)
  - First chip is interface, second is FP execution unit
  - 2 chips serve 32 processors

CM-2 Processor Chip

- Instruction
  - All processor chips receive the same instruction from the sequencer
    - Individual processors may be masked using a flag bit for processor activation
  - Produces outputs based on memory / flag lookup table (256 possible functions)
  - Arithmetic is bit-serial
- 16 flag registers
  - 8 bits for general purpose use
  - 8 bits predefined
    - NEWS flag — accessible by neighbors
    - 2 flags for message router data movement and handshaking
    - Memory parity flag
    - Flag for daisy-chaining processors
    - Zero flag (hardcoded)
    - 2 diagnostic flags

CM-2 Interconnect

- Broadcast and reduction network
  - Broadcast
  - Reduction (e.g., bitwise OR, numerically largest, or sum)
  - Scan — collect cumulative results over sequence of processors (e.g., parallel prefix)
  - Spread (scatter)
  - Sort elements
- NEWS grid can be used for nearest-neighbor communication
  - Communication in multiple dimensions: 256x256, 1024x64, 8x8192, 64x32x32, 16x16x16x16, 8x8x4x8x8x4
  - Regular pattern avoids overhead of explicitly specifying destination address

CM-2 Interconnect (cont.)

- The 16-processor chips are linked by a 12-dimensional hypercube
  - Send or get values from arbitrary locations in data-parallel fashion
- Hypercube uses router for point-to-point communication between processor chips
  - Messages move across each of 12 dimensions in sequence
  - If no conflicts, a message will reach its destination within 1 cycle of the sequence
    - All processors can send a message (of any length), all messages are sent and delivered at same time
  - Actual throughput depends on message length and access patterns

CM-2 Nexus

- The nexus is a 4x4 crosspoint switch that connects up to 4 front-end computers to up to 4 sequencers
  - CM can be configured as up to 4 sections, each used separately
  - Any front-end can be connected to any section or combination of sections
  - Example: 64K processors, four 16K sections (1 to one FE, 1 to another FE, 2 to third FE, fourth PE for other tasks)
- Each section connects to one of 8 I/O channels (graphics display frame buffer, or I/O controller)
  - Transfers initiated by front-end computers
  - Data goes into buffers, when buffers are full it goes to a data vault (each has 39 disk drives, total capacity 10GB)

Software

- System software is based on OS used in front-end computers
  - Use familiar OS, languages
  - Front end handles all flow of control, including storing and executing program, and interaction with user and programmer
  - Languages: Paris, *LISP, CM-LISP, C*
- Paris (parallel instruction set)
  - Inter-processor communication, vector summation, matrix multiplication, sorting
  - Front-end processor sends Paris instructions to processor sequencers
    - Functions & subroutines (direct actions of processors, router, I/O, etc., including scan and spread operations), global variables (find out how many processor are available, etc.)
    - Sequencer produces low-level instructs.

DAP Overview

- Distributed-memory SIMD (bit-serial)
- International Computers Limited (ICL)
  - 1976 prototype, deliveries in 1980
  - ICL spun off Actime Memory Technology Ltd in 1986, became Cambridge Parallel Processing Inc in 1992
- Matrix of PEs
  - 32x32 for DAP 500, 64x64 for DAP 600
  - Connection to 4 nearest neighbors (w/ wrap-around), plus column & row buses
  - One-bit PEs with 32Kb–1Mb of memory
- DAP system = host + MCU + PE array
  - Host (Sun or VAX) interacts with user
  - Master control unit (MCU) runs main program, PE array runs parallel code

DAP MCU and HCU

- MCU (Master Control Unit)
  - 32-bit 10 MHz CPU w/ registers, instruction counter, arithmetic unit, etc.
  - Executes scalar instructions, broadcasts others to PE array
- HCU (Host Connection Unit)
  - Gateway between DAP and host
  - Motorola 68020, SCSI port, VME interface, two RS232 serial ports
  - Provides memory boundary protection, has EPROM for code storage, 1MB RAM for data and program storage
  - Data transfers are memory-memory transfers across VME bus
  - Provides medium-speed I/O plus fast data channels (e.g.,to high-resolution color display)

DAP Processing Element

- 3 1-bit registers
  - Q = accumulator, C = carry, A = activity control (can inhibit memory writes in certain instructions)
  - All bits of a register over all PEs is called a “register plane” (32x32 or 64x64 bits)
- Adder
  - Two inputs connect to Q and C registers
  - Third input connects to multiplexor, from PE memory, output of Q or A registers, carry output from neighboring PEs, or data broadcast from MCU
    - A register also get input from this mux
    - Mux output can also be inverted
  - PE outputs (adder and mux) can be stored in memory, under control of A reg
  - D reg for asynchronous I/O, S ref for instructs that both read & write to memory

PE Memory and MCU

- PE Memory
  - Each PE has between 32 Kb and 1 Mb
  - Vector (horizontal) mode: successive bits of a word are mapped onto successive bits of a single row of a store plane
  - Matrix (vertical) mode: successive bits… onto the same bit position in successive store planes
- MCU functionality
  - Instruction fetch, decoding, and address generation
  - Executes scalar instructions and broadcasts instruction streams to PEs
  - Transmits data between PE array memory and MCU registers
  - Transmits data between DAP and host file system or peripherals

Master Control Unit (MCU)

- Code store (memory)
  - 32 bit instructions, between 128 K words and 1 M words
- 32-bit general-purpose registers
  - M0 – M13: general purpose, operated on by arithmetic and logical operations, can be transferred to and from memory array
  - M1 — M7 can be used as “modifiers” for addresses and values
- Machine states
  - Non-privileged, interruptible (user mode)
  - Privileged, interruptible
  - Privileged, non-interruptible
- Datum / limit regs. for address checking

Master Control Unit (MCU) Instructions

- Addresses
  - A 32-bit word, within a row or column, within a store plane
- "DO" instruction
  - No hardware overhead for these loops
  - HW support allows instructions inside the loops to access, in successive iterations, successive bit planes, rows, columns,or words of memory
- Nearest neighbor
  - Specify direction in instruction for shifts
  - For vector adds, specify whether rows or columns are being added, which direction to send carry bit
  - Specify behavior at edge of operation

Gamma II^{Plus}

- Fourth-generation DAP, produced by Cambridge Parallel Processing in 1995
- Gamma II^{Plus} 1000 = 32x32
  Gamma II^{Plus} 4000 = 64x64
- PE memory: 128Kb–1Mb
- PE also contains an 8-bit processor
  - 32 bytes of internal memory
  - D register to transfer data to/from array memory (1-bit data path) and to/from internal memory (8-bit data path)
  - A register, similar to a 1-bit processor
  - Q register, like accumulator, 32 bits wide (any one of which can be selected as an operand), can also be shifted
  - ALU to provide addition, subtraction, and logical operations

Gamma II Plus 4000