POWER CHALLENGEarray Architecture

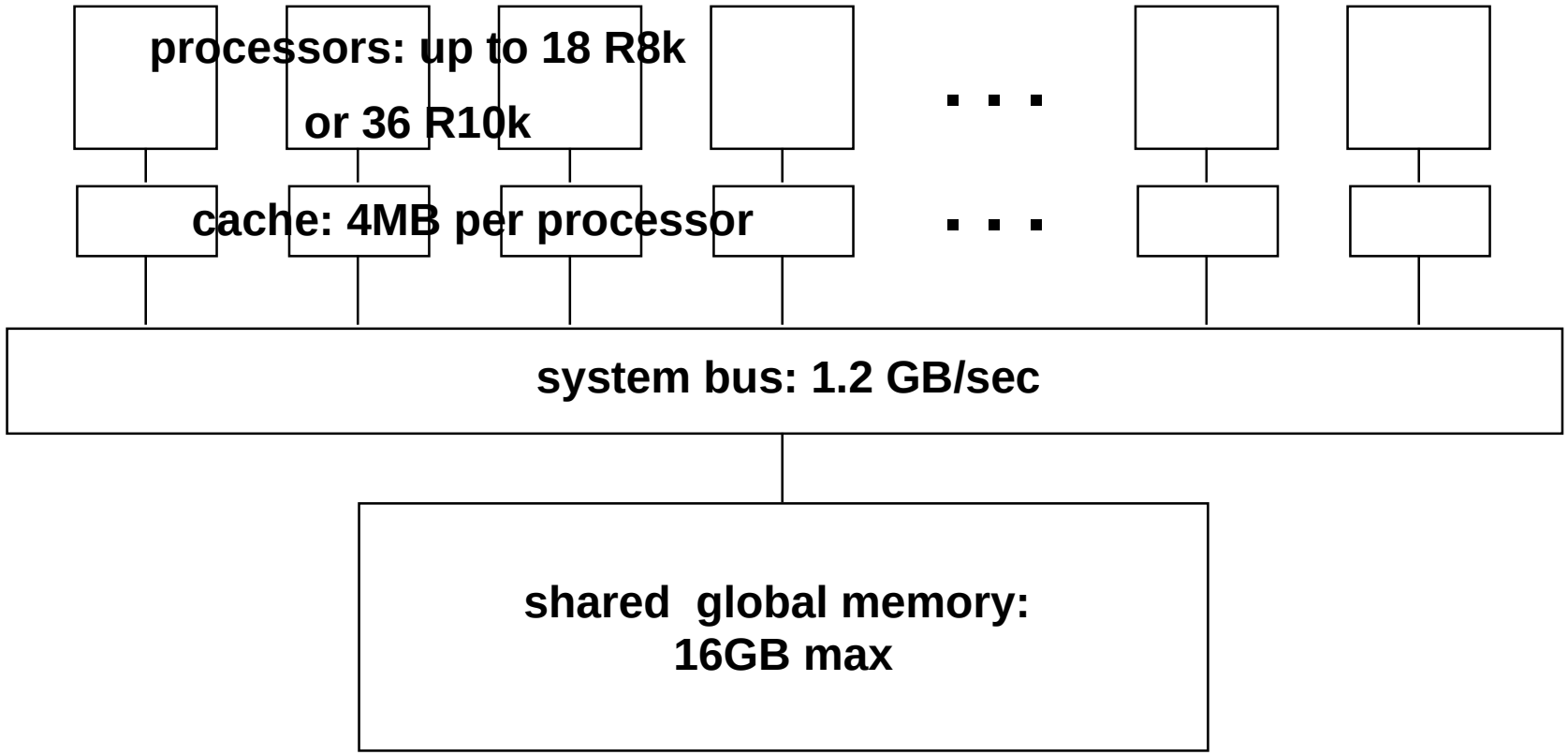
Distributed Memory System

- **SMP nodes (Power Challenge)**
- **High Speed Interconnect**
  - **HIPPI**

HIPPI

- **Point-to-Point connection network**
  - **only one connection to another system at a time**
  - **fast creation/destruction of connections**
- **All nodes equidistant**

SMP Nodes - Power Challenge



R8000 Processors

- 2-chip design
  - separate FP units
- 75 or 90 Mhz Clock (90Mhz at NCSA)
- MIPS IV instruction set
- 4-way superscalar
- 300 or 360 Mflop/s peak
  - 2 mult-adds/cycle
- 16KB instruction and 16KB integer data caches
- 4MB cache
  - primary FP cache; secondary int/instr cache

R8000 Processors

- **64 bit processor**
  - **64 bit registers**
  - **64 bit integer and floating point operations**
  - **64 bit virtual addresses**
  - **40 bit physical addresses (1 TB)**
  - **Binary compatible with earlier MIPS processor**

R8000 Processors

- **Superscalar -- up to 4 inst per cycle**
  - **up to 2 floating point inst per cycle**
  - **up to 2 integer ops inst cycle**
  - **up to 2 load/stores inst cycle**
    - **max of 2 integer and load/stores combined per cycles**
- **Pipelined (5 stages)**
- **384 entry TLB**

R8000 Processors

- **Peak Performance (at 90 Mhz)**
  - **360 Mflop/s: 2 mult-adds per cycle**
  - **360 Mips: 2 integer + 2 floating point instructions per cycle**

Cache (R8k)

- **External (L2) Cache**
  - **primary floating point cache**
  - **secondary instruction and integer data cache**
  - **"Data Streaming Cache" in SGI documentation**
  - **4MB per processor**
  - **4-way set associative**
  - **cache line 128 bytes**
  - **write-back policy**

Cache (R8k)

- **External Cache**
  - **2 way interleaved -- even/odd doubleword (64 bits per doubleword)**
  - **600 MB/sec/bank cache-to-processor**
    - **1 doubleword per bank per cycle**
  - **1.2 GB/sec overall possible**
    - **2 doublewords per cycle**
  - **miss penalties**
    - **roughly 50 cycles to main memory**
    - **roughly 100 cycles to another cpu's cache**

Cache (R8k)

- **External Cache**
  - **coherence maintained in hardware**
  - **snoopy write-invalidate protocol**

R10000 Processors

- 195 Mhz Clock
- MIPS IV instruction set
- 4-way superscalar
- 390 Mflop/s peak
  - 1 mult-add/cycle (chained) OR
  - 2 FP instructions
- 32KB floating point and 32KB integer/instruction data caches
- 2MB secondary cache
  - 2-way set associative

R10000 Processors

- 64 bit processor
  - 64 bit registers
  - 64 bit integer and floating point operations
  - 64 bit virtual addresses
  - 40 bit physical addresses (1 TB)
  - Binary compatible with earlier MIPS processor

R10000 Processors

- **Superscalar -- up to 4 instructions per cycle**
  - **up to 2 floating point inst per cycle**
  - **up to 2 integer inst per cycle**
  - **up to 1 load/stores per cycle**
- **5 independent pipelined functional units**
- **Out-of-order instruction execution**
- **64 entry TLB**

R10000 Processors

- **Peak Performance (at 195 Mhz)**
  - **390 Mflop/s: 2 floating point instructions per cycle**
  - **780 Mips: 2 integer + 2 floating point instructions per cycle**

Cache R(10k)

- **External (L2) Cache**
  - **secondary floating point cache**
  - **secondary instruction and integer data cache**
  - **2MB per processor**
  - **2-way set associative**
  - **cache line 128 bytes**
  - **write-back policy**

Cache (R10k)

- **External Cache**
  - **2 way interleaved -- even/odd doubleword (64 bits per doubleword)**
  - **1.6 GB/sec/bank cache-to-processor**
    - **1 doubleword per bank per cycle**
  - **3.2 GB/sec overall possible**
    - **2 doublewords per cycle**
  - **miss penalties**
    - **roughly 50 cycles to main memory**
    - **roughly 100 cycles to another cpu's cache**

Cache (R10k)

- **External Cache**
  - **coherence maintained in hardware**
  - **snoopy write-invalidate protocol**

Buses

- **System Bus -- 1.2 GB/sec**
  - **separate address and data buses**
  - **128 bit wide data bus**
  - **40 bit wide address bus**
  - **distributed arbitration algorithm**
- **I/O Buses -- 320 MB/sec peak**
  - **supports VME, FDDI, HiPPI**

IRIX 6

- 64-bit UNIX
  - IRIX 6.1 -- full 64 bit file system
- standards compliant
  - System V Release 4 ABI
    - System V Interface Definition 3 (SVID3)
  - also MIPS ABI Group
- upwardly compatible revision of IRIX

IRIX 6

- **Process Scheduling**
  - **non-preemptive (only kernel preempts)**
  - **processor cache affinity heuristic**
  - **gang scheduling**
- **Lightweight Processes**
  - **sproc -- shared virtual address space**
  - **POSIX compliant threads**
    - **Supported as a patch to IRIX 6.2**
    - **Fully supported in IRIX 6.4**

IRIX 6 -- High Performance I/O

- **Asynchronous I/O (Posix)**
  - **multiple concurrent I/O ops per process**
- **Memory Mapped I/O (SVID 3)**
  - **disk files visible in program address space**
  - **no explicit I/O ops required**
- **Direct I/O (SGI defined)**
  - **unbuffered transfers to/from device**
  - **performance approaches raw device interface**