

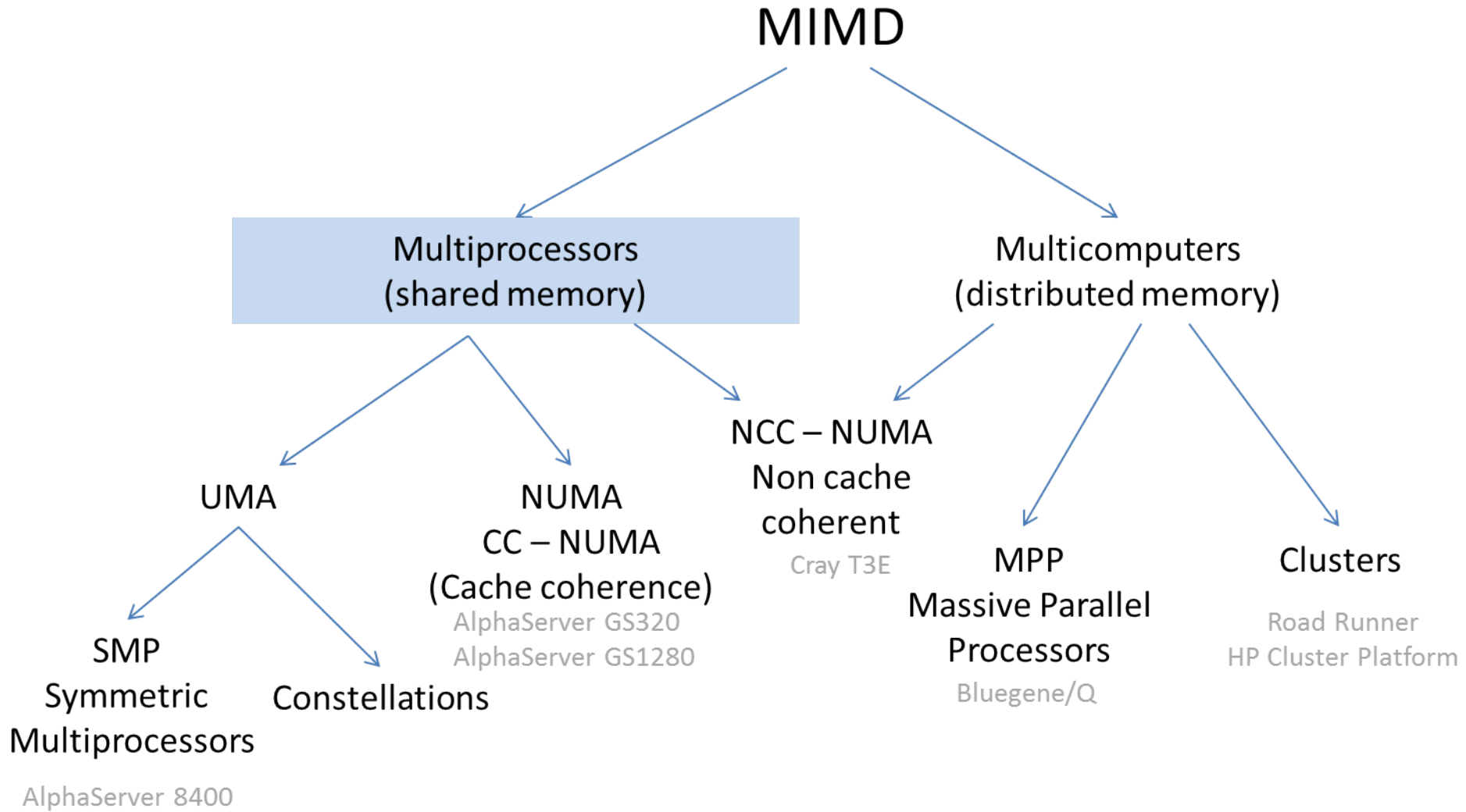


UNIVERSIDAD  
DE BURGOS

# Multiprocessors on the market

José M. Cámara  
([checam@ubu.es](mailto:checam@ubu.es))

v. 1.0





UNIVERSIDAD  
DE BURGOS

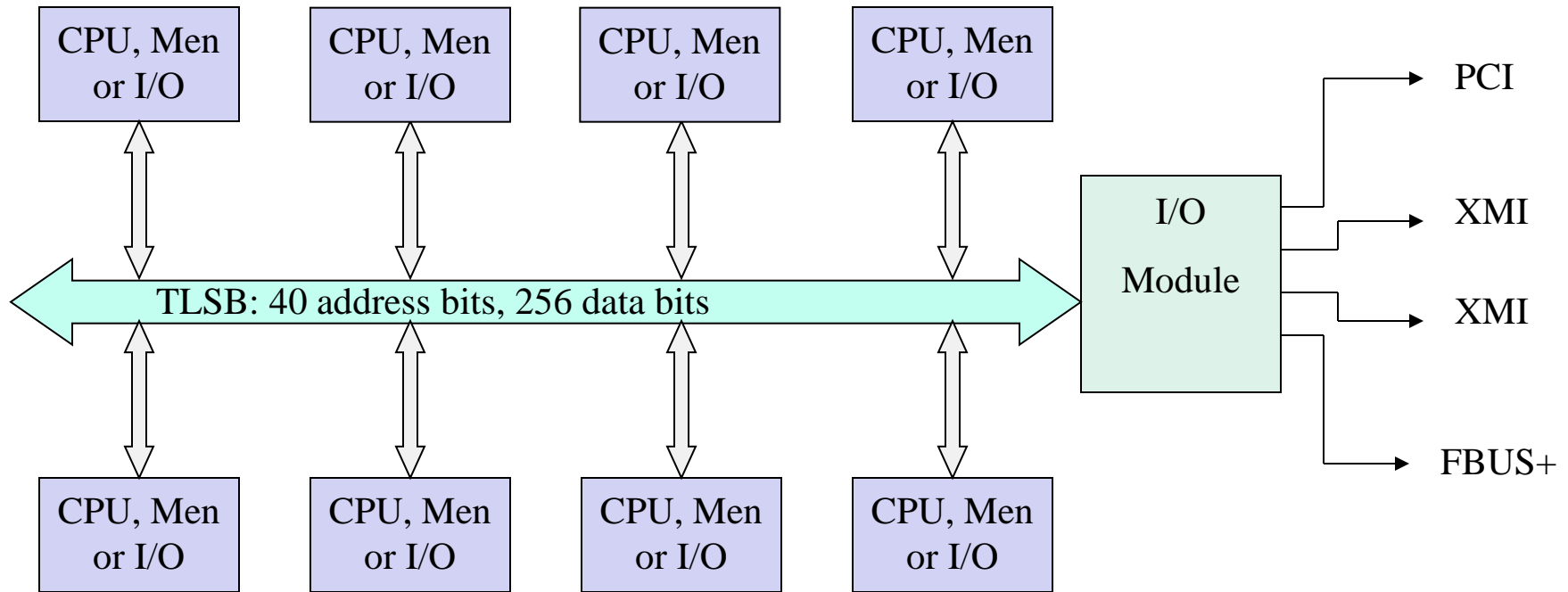
# SMP Architecture

AlphaServer 8400

# General statements

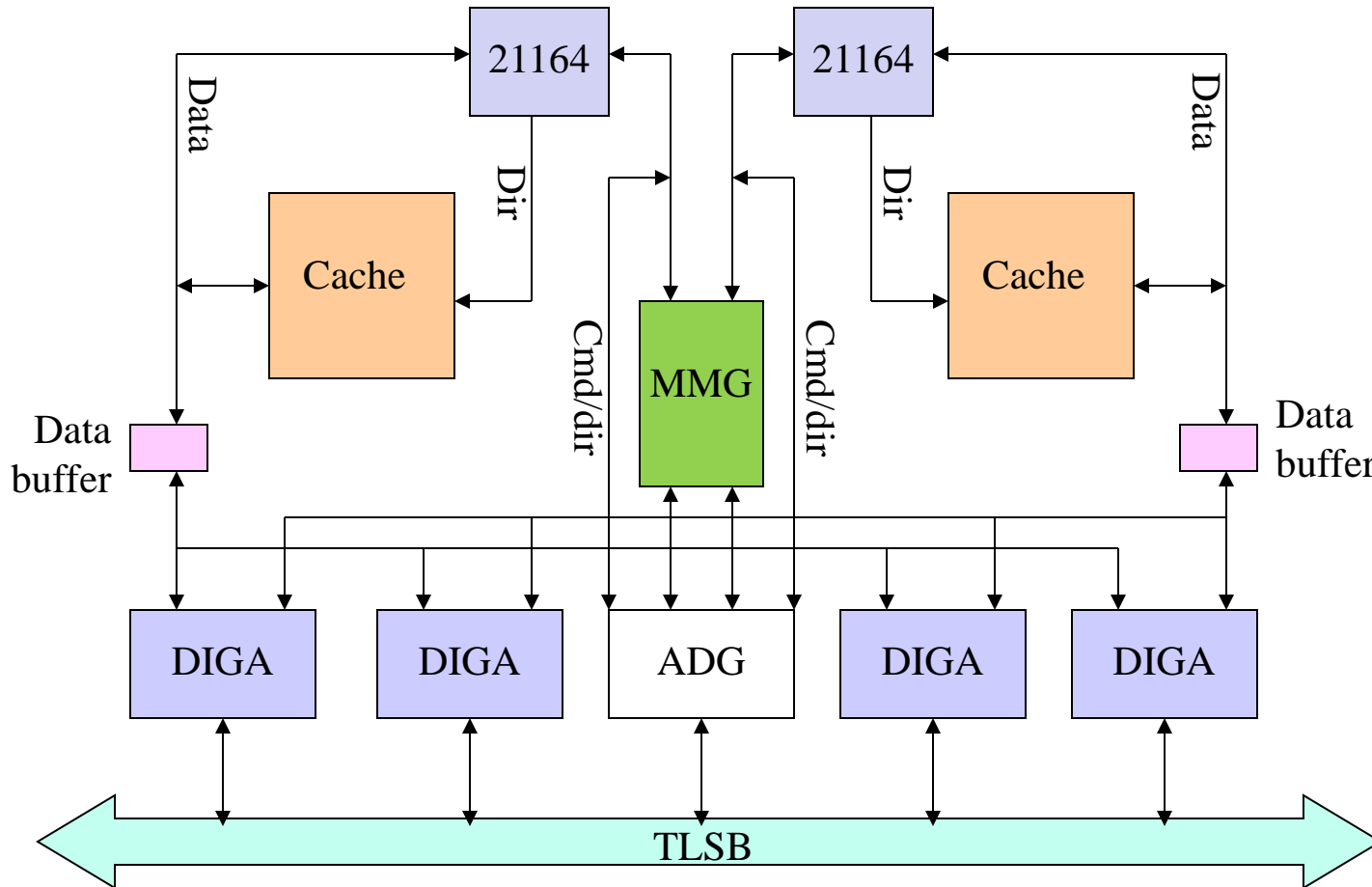
- Uniform memory access (UMA), symmetric multiprocessor (SMP) system.
- Up to 12 CPU Alpha 21164.
- Up to 14 GB main memory.
- 3200MB/s system bus.
- Operating systems: OpenVMS & Digital UNIX.

# System layout



- At least one CPU, one memory and one I/O modules are mandatory.

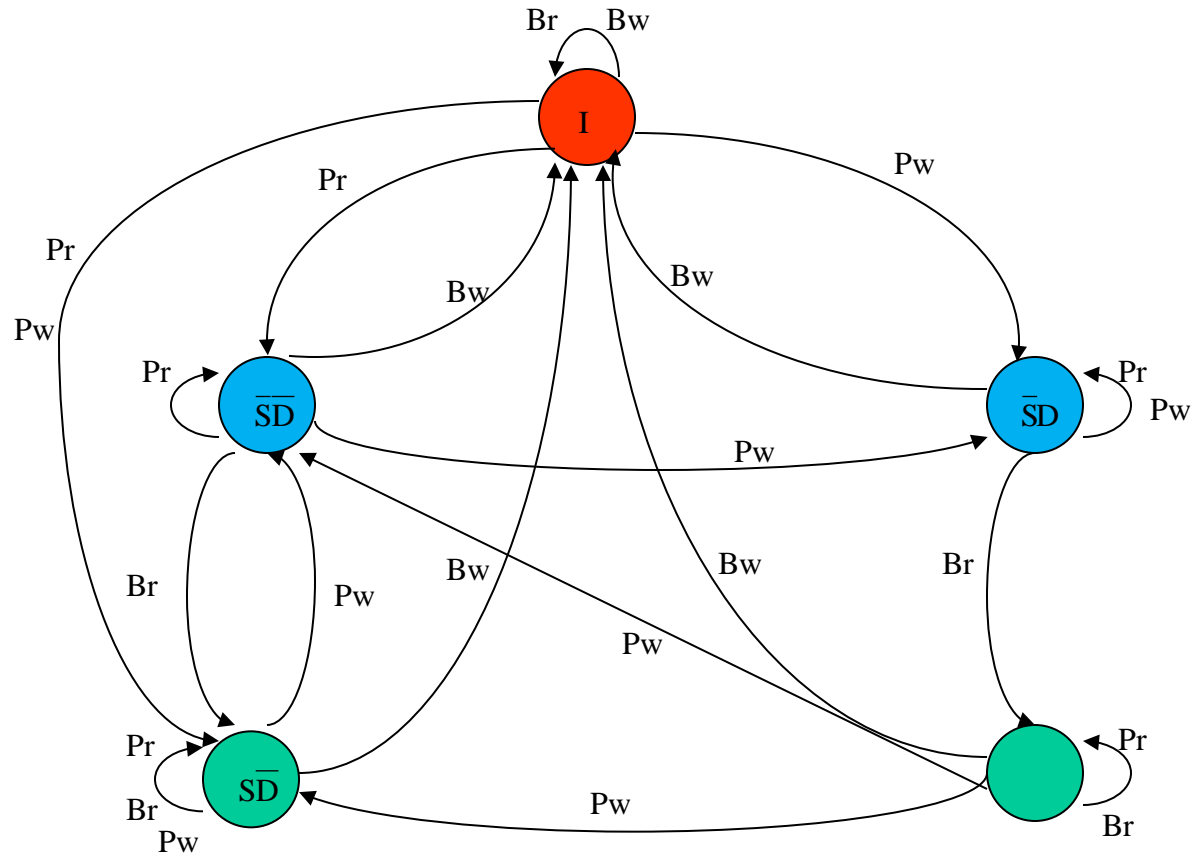
# CPU module



# Alpha 21164

- Vendor: Digital.
- Year: 1996.
- Clock freq.: up to 500 MHz.
- Technology: CMOS 0'35 microns.
- Cache L1: 8 + 8 kB.
- Cache L2: 96 kB.
- Cache L3 external (optional).
- Transistors: 9'3 millions.
- Power: 25W.

# Coherence protocol



- I: invalid
- S: shared
- D: dirty
- Pw in a shared line updates main memory and disables D in the local copy.



# CPU module

- One or two CPUs inside.
- Each CPU works independently and has its own L3 cache.
- CPUs from 142 to 357 MHz are supported.
- L3 is 4MB with 64 bytes lines.
- Has the required duplicated tag space.
- Multiplexors – demultiplexors (DIGA) to adapt CPU's data width (128) to the bus (256). Data buffer is used for temporary storage.
- Address multiplexor (MMG) hold the duplicated tag space as well.
- The address interface receives also commands and manages the duplicated tag space. It is responsible for cache coherence.

# System bus

- Capacity for up to 9 modules.
- One of them must be I/O, another one memory and also a CPU module is required. In case more than 3 CPU modules are installed, at least 3 memory ones are mandatory.
- 40 bits address bus and 256 bits data bus.
- Synchronous bus whose clock is a multiple of the CPU's (33-100 MHz).
- Peak transfer speed: 3200 MB/s.
- Simple parity in addresses and commands and ECC in data.

# Memory module

- From 128 MB to 2 GB.
- Maximum storage capacity = 14 GB.
- ECC.
- Write invalidate protocol.
- Writes on shared lines demand bus control thus updating main memory alongside.

# I/O module

- Provides connection to standard buses:
  - XMI (2)
  - PCI
  - Futurebus+



UNIVERSIDAD  
DE BURGOS

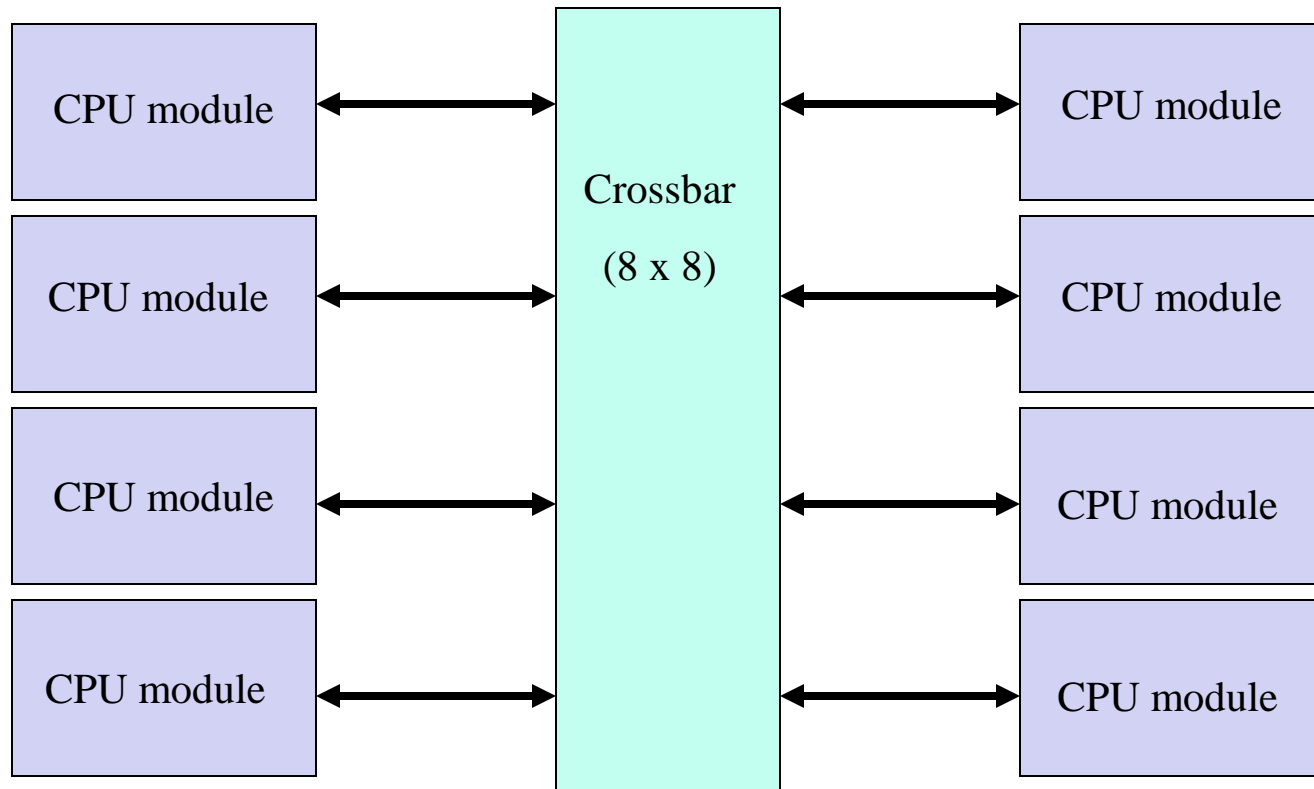
# Non uniform memory access architectures: CC-NUMA

# AlphaServer GS320

# General statements

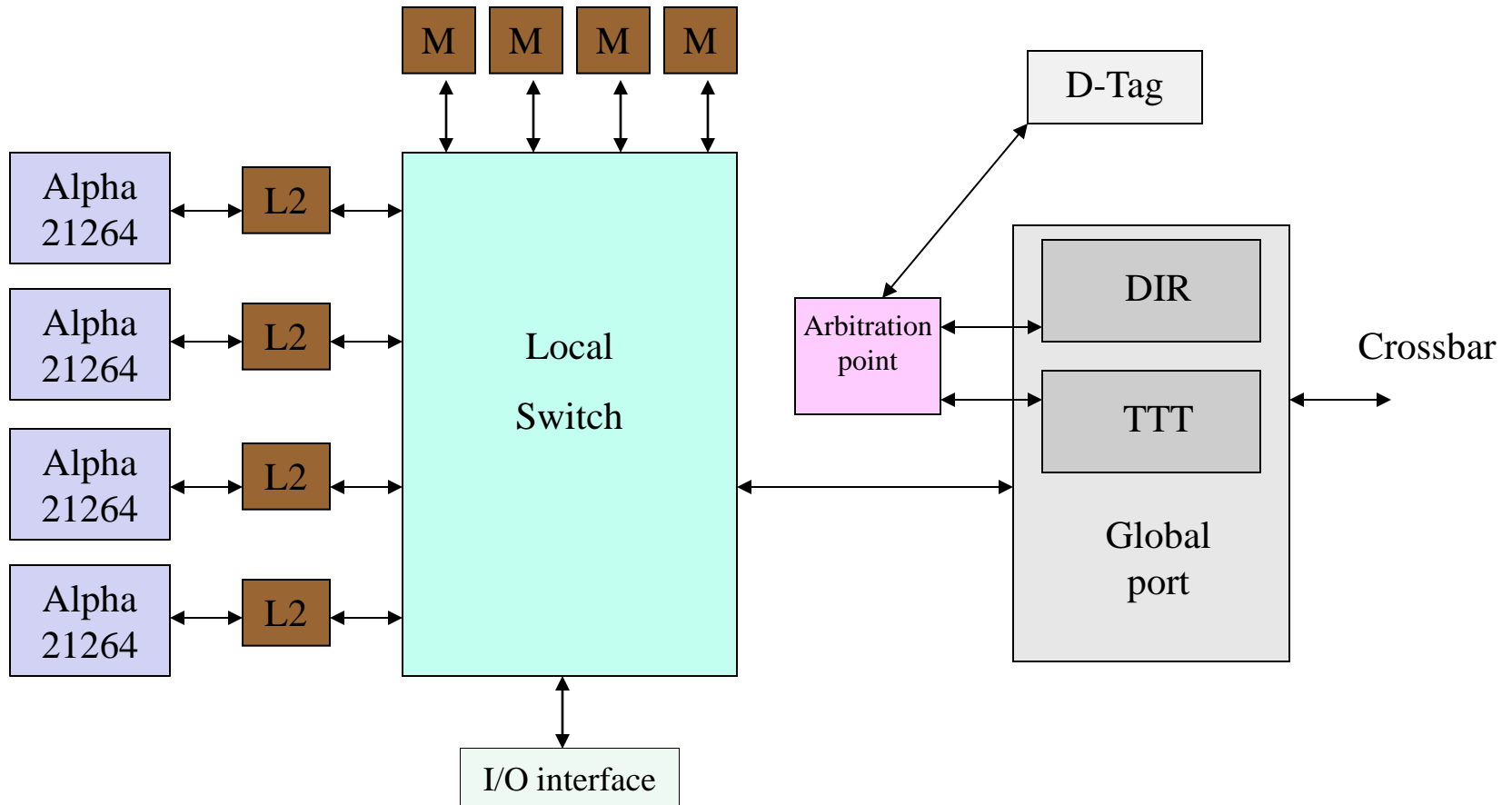
- CC-NUMA architecture.
- 4 processor modules (Alpha 21264).
- Up to 8 modules (32 processors max.)
- Crossbar interconnect 8x8 at 1'6 GB/s.
- Full mapping directories.
- Operating systems: Tru64 UNIX, Open VMS, LINUX

# Architecture





# CPU module



# Alpha 21264

- 731 MHz in this system.
- Cache L1: 64 + 64 kB 2 way set associative.
- Cache L2: external 4 MB direct mapping.
- Out-of-order execution permitted.
- Simplified pipeline.

# Memory

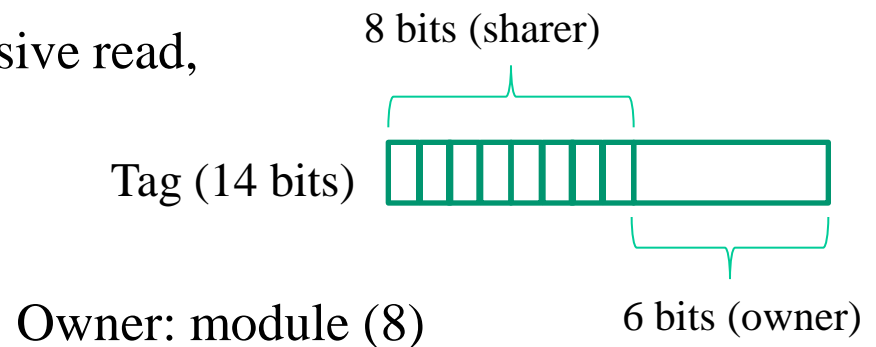
- 4 memory modules: 1 to 8 GB each.
- Maximum capacity per module: 32 GB.
- Each module is 8 way interleaved.
- Overall bandwidth: 6'4 GB/s
- 2 level cache coherence:
  - Intra-module: duplicated tag behaves as a full mapping directory.
  - Inter-module: directory is located in the global port.

# Switch

- Local switch is very much like a crossbar but with an asymmetric structure.
- Not all connections are possible since, for instance, connections between memory modules are not implemented.
- Not all connections are equally fast. Global port's bandwidth is double.

# Coherence

- Involves several elements: DIR, TTT, D-Tag & arbitration point.
- 64 bytes cache lines with 14 bits tag: 6 to locate the owner + 8 to locate copies at module level, not CPU yet.
- Duplicated tag space makes location of copies at CPU level possible.
- The TTT is an associative table meant to hold a list of up to 48 transactions not yet reported to main memory (write-backs).
- 4 possible access requests: read, exclusive read, exclusive, data less exclusive.



Owner: CPU (32) + i/o (8) + memory (1)

# Mismatches

- A memory access request gets to the owner after a write-back.
  - Cache line is kept in a victim buffer until all pending requests in the D-Tag have been served.
  - Afterwards, it is kept in the TTT until main memory acknowledges its reception.
- A memory access request gets to the owner before it receives the data.
  - Owner processor compares this request with its list of cache misses. If there is a coincidence, the request is simply delayed.

# AlphaServer GS1280

# General statements

- 64 CPU Alpha 21364 at 1'5 GHz
- 2-D torus interconnect.
- Adaptive routing with at least 3 virtual channels.
- Each node is formed by a CPU, main memory and I/O.
- Full mapping directories for cache coherence.



# CPU

- On-chip L2 cache 1'75 MB 7 way set associative.
- Integrated memory controller.
- Integrated 2D torus router.
- 0'18 microns technology.
- 152 million transistors.

# Memory

- On-chip L1 cache 64 +64 kB 2 way set associative.
- On-chip L2 cache 1'75 MB 7 way set associative with ECC.
- 8 GB main memory per node with ECC.

# Interconnect

- 2-D torus direct interconnect.
- Bidirectional point to point links 6'2 GB/s (2 x 3'1)
- 3 virtual channels: 2 DOR + 1 adaptive. Packets try to move along the adaptive channel but shift to one DOR if not possible.
- Deadlock free network: inter-dimension because of the DOR, intra-dimension because of the presence of virtual channels.



UNIVERSIDAD  
DE BURGOS

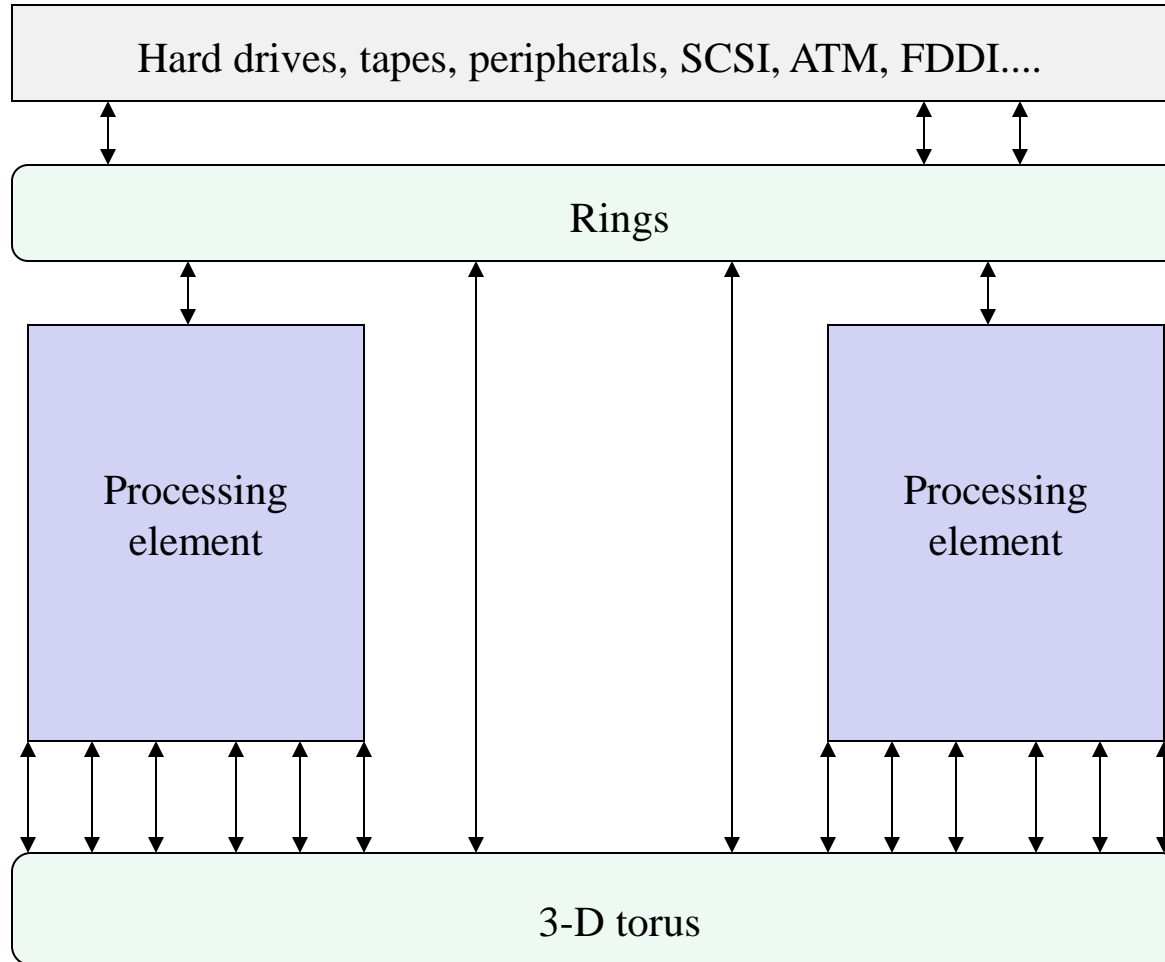
NCC – NUMA Architectures

CRAY T3E

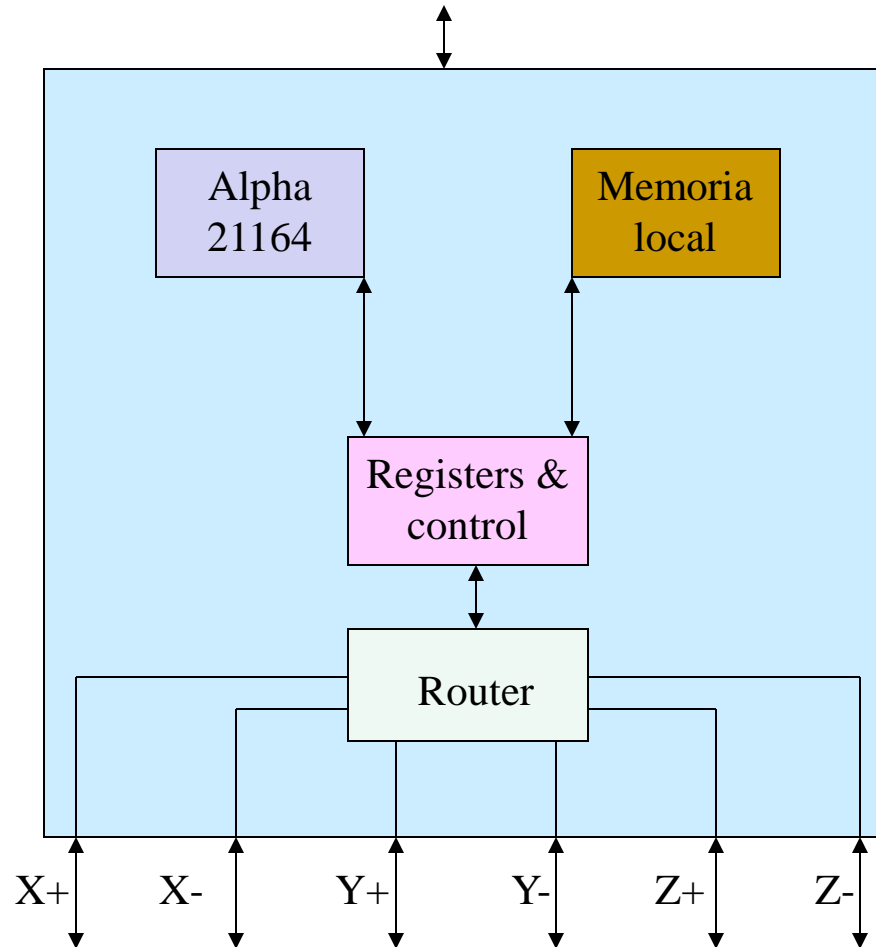
# General statements

- NCC-NUMA architecture.
- Up to 2048 Alpha 21164 nodes.
- 3-D torus interconnect.
- Additional rings for I/O.
- UNICOS operating system.

# Architecture



# Processing element



# Memory

- Embedded L1 & L2 cache.
- Up to 2 GB main memory with ECC.
- No hardware mechanism to guarantee cache coherence.
- Within each processing element a 3 state protocol is used.
- Remote data are accessed individually instead of as cache lines.



# Interconnect

- 3-D torus with bidirectional links 600 MB/s.
- Wormhole flow control.
- Adaptive routing.
- Up to 128 independent rings.
- Rings are made up from 1 GB/s bidirectional links connecting 16 nodes which include processing nodes, peripherals and ports.

# Deadlock

- Deadlock free network.
- Between different dimensions deadlock is avoided by means of a combination of virtual channels and DOR.
- Within each dimension a packet elimination mechanism is used.

# References

- **AlphaServer 8200/8400. System Technical Manual.** Digital Equipment Corporation. 1995. Available at:  
<http://h18002.www1.hp.com/alphaserver/download/t8030tma.pdf>
- Architecture and Design of AlphaServer GS320. Kourosh Gharachorlooy, Madhu Sharma, Simon Steely, and Stephen Van Doren. *Ninth International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS-IX), November 2000.*
- AlphaServer ES47, ES80, and GS1280 Systems. Technical Summary. Available at:  
[http://www.compaq.com/alphaserver/g1280/g1280\\_tech.html](http://www.compaq.com/alphaserver/g1280/g1280_tech.html)