High Performance Multiprocessor System

Requirements:

- Large Number of Processors (4)

- Large **WriteBack** Caches for Each Processor. Less Bus Traffic => Higher Performance

- Large Shared Main Memories (128MB - 1GB)

MIPS ARCHITECTURE MEETS THESE REQUIREMENTS

RESULT: RISC PROCESSOR IN MOST MP SYSTEMS SHIPPED TODAY

- Ardent, Fermi, RC, SGI are announced

- Several designs yet to announce
MIPS MULTIPROCESSOR SUPPORT

- ARCHITECTURAL SUPPORT
  - PHYSICAL CACHE
  - ARCHITECTURAL PARTITIONING WELL SUITED FOR MP
    - ON-CHIP CACHE CONTROL AND MMU
    - OFF-CHIP SEAMLESS FP
  - TOTAL 128 K BYTE CACHE PER PROCESSOR

-R3000 HARDWARE SUPPORT FOR CACHE COHERENCE
  Snooper watches Bus transactions
  If Address matches with any cache entry:
    MP_STALL STALLS THE PROCESSOR
    MP_INVALIDATE SIGNAL CAUSES CPU TO WRITE THE CACHE LOCATION AS INVALID.
LOW COST DESIGN

- DUPLICATE TAG RAMS FOR SNOOPING

- WRITE THROUGH CACHES

- SNOOPER COMPARES WRITE ADDRESSES IN DUPLICATE TAG

- IF MATCHES, MP_STALL ASSERTED

- ENTRY INVALIDATED FOR THAT ENTRY
SNOOPER LOGIC

DATA

MP BUS

ADDRESS

INDEX

TAG

SECONDARY CACHE

OE*

LATCH 1

LATCH 2

COMPARATOR

COMPARATOR

DATA

TAG

MP INVALIDATE

MPSTALL
HIGH PERFORMANCE MP DESIGN

- TWO LEVELS OF CACHE

- WRITE THROUGH PRIMARY

- WRITE BACK SECONDARY

- EXTERNAL WRITE BACK LOGIC AND SNOOP CONTROL
MP DESIGN - 2

R3000

CACHE BUS

I-CACHE

D-CACHE

SECONDARY CACHE & CONTROL

SNOOPER

MP BUS

READ BUFFER

WRITE BUFFER

READ BUFFER

WRITE BUFFER

D-CACHE

I-CACHE

SECONDARY CACHE & CONTROL

SNOOPER

R3000
WRITE BACK LOGIC

EACH BLOCK IN SECONDARY CACHE IN ONE OF 4 STATES:

INVALID : NO DATA
VALID : CLEAN, POTENTIALLY SHARED
RESERVED : WRITTEN JUST ONCE, WRITTEN THROUGH'D TO MAIN MEMORY AND ONLY COPY IN ANY CACHE
DIRTY : WRITTEN MORE THAN ONCE AND ONLY COPY IN ANY CACHE

WRITE HIT:
IF STATE == DIRTY, THEN WRITE TO CACHE
IF STATE == RESERVED
   THEN 1. WRITE TO CACHE
   2. CHANGE STATE TO DIRTY
IF STATE == VALID
   THEN 1. WRITE THROUGH TO MAIN MEMORY
   2. CHANGE STATE TO RESERVED
SNOOP CONTROLLERS FOR OTHER RESERVED:
   IF THE ENTRY MATCHES, INVALIDATE THE CACHE

WRITE MISS:
IF SNOOP CONTROLLERS FIND THE BLOCK IN ANOTHER CACHE
   THEN
      IF STATE == DIRTY
         THEN 1. SUPPLY THE DATA
         2. WRITE TO MEMORY
         3. CHANGE STATE TO VALID
      IF STATE == RESERVED
         THEN 1. SUPPLY THE DATA
         2. CHANGE STATE TO DIRTY
      IF STATE == VALID
         THEN SUPPLY THE DATA
   ELSE READ FROM MEMORY

READ MISS:
IF SNOOP CONTROLLERS FIND THE BLOCK IN ANOTHER CACHE
   THEN
      IF STATE == DIRTY
         THEN 1. SUPPLY THE DATA
         2. WRITE TO MEMORY
         3. CHANGE STATE TO VALID
      IF STATE == RESERVED
         THEN 1. SUPPLY THE DATA
         2. CHANGE STATE TO DIRTY
      IF STATE == VALID
         THEN SUPPLY THE DATA
   ELSE READ FROM MEMORY

WRITE MISS:
IF SNOOP CONTROLLER FINDS A COPY IN ANOTHER CACHE
   THEN
      IF STATE == DIRTY
         THEN 1. SUPPLY THE DATA TO THE CACHE
         2. INVALIDATES ITS COPY
         3. AFTER LOADING THE BLOCK, CHANGE STATE TO DIRTY
   ELSE LOAD THE BLOCK FROM MEMORY AND CHANGE STATE TO DIRTY
SNOOP CONTROLLERS FOR OTHER CACHES, IF FIND THE COPY, INVALIDATES THE BLOCK
MP DESIGN - 2 SECOND LEVEL CACHE DESIGN

1. ASSUME FIRST LEVEL CACHE IS WRITE THRU AND SECOND LEVEL CACHE IS WRITE BACK

2. BOTH FIRST AND SECOND LEVEL CACHES ARE PHYSICALLY ADDRESSED (NO REVERSE TRANSLATION NEEDED)

3. MAIN MEMORY READS AND WRITES ALWAYS GO THROUGH THE SECOND LEVEL CACHE (NOT THE WRITE BUFFERS)

4. FIRST AND SECOND LEVEL CACHES MAINTAIN COHERENCY. FIRST LEVEL CACHE IS A SUBSET OF THE SECOND LEVEL CACHE

5. PROCESSOR GETS DATA FROM THE FASTER FIRST LEVEL CACHE IF IT HITS
MP DESIGN - 2 CACHE COHERENCY PROTOCOL

• THE ILLINOIS SCHEME IS USED - WRITE INVALIDATE

• FIRST LEVEL CACHE HAS 2 STATES
  VALID, INVALID

• SECOND LEVEL CACHE HAS 4 STATES
  INVALID
  PRIVATE CLEAN (UNMODIFIED, ONLY COPY)
  SHARED CLEAN (UNMODIFIED, POSSIBLY OTHER COPIES)
  DIRTY (MODIFIED, ONLY COPY, WRITE BACK TO MAIN MEMORY)
ILLINOIS SCHEME STATE DIAGRAM

INVALID

PRIVATE
CLEAN

DIRTY

SHARED
CLEAN

WRITE HIT

WRITE MISS

WRITE HIT

READ MISS

READ MISS

WRITE HIT

BUS WRITE MISS

BUS READ MISS

BUS WRITE MISS

BUS READ MISS

BUS READ MISS

PROCESSOR-BASED TRANSITION

BUS-INDUCED TRANSITION

READ MISS

(FROM MEMORY)

READ MISS

(FROM CACHE)
ESTIMATED PARTS LIST

READ BUFFER : FCT374A

WRITE BUFFER : MIPS R3020 OR FCT521 FIFO'S FROM IDT

PRIMARY I & D CACHE : 64K X 1

SECONDARY D CACHE : 8K X 8

WRITE BACK LOGIC : APPROX 10 PALS

SNOOPER LOGIC : 1 BUFFER
               1 COMPARATOR
               2 PALS

APPROXIMATE BOARD SPACE : 100 Square inches
per processor module
Multiprocessor System Shipping Today

Dual Processor Board
2 Modules + 100 IC's, 15" x 15"

Module

CPU
FPU
64KB I-Cache
64KB D-Cache
(16K x 4's)
Rd/Wr Buffers

Module

CPU
FPU
64KB I-Cache
64KB D-Cache
(16K x 4's)
Rd/Wr Buffers

64KB Secondary Cache
(8K x 8's, to be 32K x 8's)
Snoopy Logic
Write Back Logic

64KB Secondary Cache
(8K x 8's, to be 32K x 8's)
Snoopy Logic
Write Back Logic

Common Bus Interface Logic
MIPS MULTIPROCESSOR SUMMARY

• MINIMUM LOGIC NEEDED TO PROVIDE SUPPORT WITHOUT PENALIZING UNIPROCESSORS.

• EXISTENCE PROOF: MP SYSTEMS SHIPPING TODAY BY MULTIPLE VENDORS.

• ARCHITECTURAL AND HARDWARE SUPPORT:
  - PHYSICAL CACHE, USEFUL PARTITIONING, LARGE CACHES PLUS...
  - TWO SIGNALS ALLOW THE CPU TO CONTROL THE CACHE AT HIGH CLOCK RATES & ALSO INVALIDATE WHEN SNOOPING DETECTS A HIT.

• FLEXIBLE MP SYSTEM CONFIGURATIONS:
  - SUPPORTS MINIMUM SYSTEM WITH DUPLICATE TAGS AND LARGE SYSTEMS WITH SECONDARY CACHES.