COMP4611 Tutorial 6 Instruction Level Parallelism

1

Instruction Level Parallelism

- Definition
 - An implementation technique whereby multiple instructions are overlapped in execution
- Two separate approaches to exploiting II P
 - Hardware support to help discover and exploit the parallelism dynamically
 - Software technology to find parallelism statically

2

Instruction Level Parallelism

- Few possibilities in a basic block
 - A straight-line code sequence
 - no branches in except to the entry
 - no branches out except at the exit
 - Blocks are small (6-7 instructions)
 - Instructions are likely to depend upon one another
- Goal: Exploit ILP across multiple basic blocks
 - Example: loop-level parallelism

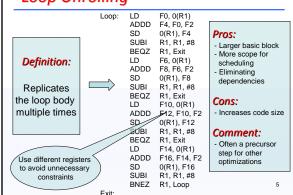
```
for (i = 1000; i > 0; i=i-1)
x[i] = x[i] + s;
```

Basic Schedulingfor (i = 1000; i > 0; i=i-1)

> SUBI R1, R1, #8 BNEZ R1, Loop

Pipelined execution: Scheduled pipelined execution: F0, 0(R1) SUBI stall R1, R1, #8 ADDD F4, F0, F2 ADDD F4, F0, F2 stall stall stall BNEZ R1, Loop SD 0(R1), F4 SD 8(R1), F4 SUBI R1, R1, #8 Data dependency stall

Loop Unrolling



Possible hazards - Data hazard

• RAW (read after write)

BNEZ R1, Loop

stall

i preceding i

Happen when

 j tries to read a source before i writes it, so j incorrectly gets the old value.

Longer operation latency more frequent stalls

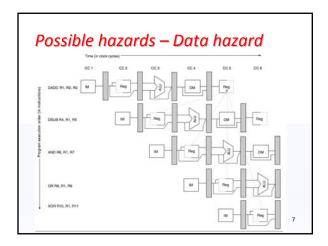
1.write in more than one pipe stage

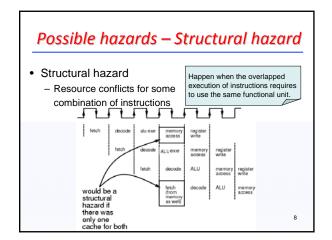
- WAW (write after write)
 - j tries to write an operand before it is written by i

2.one instruction proceed when previous one is stalled

- WAR (write after read)
 - j tries to write a destination before it is read by i, and i incorrectly gets the new value.

Happen when instructions are reordered





Possible hazards - Control hazard

- · Control hazard
 - A control hazard is when we need to find the destination of a branch, and can't fetch any new instructions until we know that destination.

A branch is either taken or not taken

- Taken: PC<=PC+4+immediate
- Not Taken: PC<=PC+4

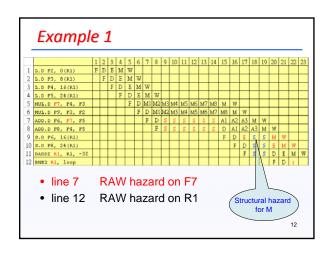
9

Example 1

- Consider a machine with multi-cycle functional units that is an extension of the 5 stage pipeline machine (IF ID EX MEM WB).
- Integer operations have 1 EX stage, FP Add has 3 EX stages and FP multiply has 8 EX stages.
- Each FP unit is pipelined and the pipeline supports full forwarding.
- All other stages in the pipeline complete in one cycle.
 Branches are resolved in the ID stage.
- For WAW hazards, assume that they are resolved through stalls and not through aborting the first instruction.
- List all of the hazards that cause stalls in the following code segment and explain why they occur.

10

Example 1 Loop : 1) L.D F2, O(R1) 2) L.D F3, 8(R1) RAW hazard 3) L.D F4, 16(R1) 4) L.D F5, 24(R1) 5) MUL.D F7, F4, F3 6) MUL.D F9, F3, 2 7) ADD.D F6, F7, F5 8) ADD.D F8, F4, F5 9) S.D F6, 16(R1) 10) S.D F8, 24(R1) RAW hazard 11) DADDI R1, R1, on R1 12) BNEZ R1, Toop 11



Exercise • The following loop is a dot product: L.D F0, 0(R1) ;load X[i] F4, 0(R2) MUL.D F0, F0, F4 ;multiply X[i]*Y[i] ADD.D F2, F0, F2 ;add sum=sum+X[i]*Y[i] ADDUI R1, R1, #-8 ;decrement X index ADDUI R2, R2, #-8 ;decrement Y index ;loop if not done Assume - the pipeline latencies from Table 1, - a 1-cycle delay branch. - a single-issue pipeline. - the running sum in F2 is initially 0. - despite the fact that the loop is not parallel, it can be scheduled with no delays.

Table 1 Instruction Instruction Latency producing result using result in clock cycles FP ALU op FP ALU op 3 FP ALU op Store double 2 Load double FP ALU op 1 Load double Store double 0 Integer op Integer op 0

Exercise (cont.)

- Unroll the loop a sufficient number of times to schedule it without any delays. Show the delay after eliminating any unnecessary branch maintenance instructions.
- Hint: an scheduling of the code is needed

15

```
Exercise: Data Dependency
           L.D
                   F0, 0(R1)
           L.D
                   F4, 0(R2)
           stall
                                    1 clock cycle
           MUL.D
                   F0, F0, F4
           stall
           stall
           stall
                  F2, F0, F2
           ADD.D
           ADDUI
                   R1, R1, #-8
                                   3 clock cycles
                   R2, R2, #-8
           ADDUI
           BNEZ
                   R1, foo
           stall
                                            16
```

```
Unrolling Twice
 L.D
           F0, 0(R1)
                              L.D
                                        F6, -8(R1)
 L.D
           F4, 0(R2)
                              L.D
                                        F8, -8(R2)
 stall
                              stall
          FO. FO. F4
 MUL.D
                              MUL.D
                                        F6, F6, F8
 Stall
                              Stall
 Stall
                              Stall
 stall
                              Stall
 ADD.D
          F2, F0, F2
                              ADD.D
                                        F2, F6, F2
 ADDUI
          R1, R1, #-8
                              ADDUI
                                        R1, R1, #-16
ADDUI
          R2, R2, #-8
                                       R2, R2, #-16
                              ADDUI
 stall
                                   Usina
                               different registers
                                                     17
```

```
New method
    - Calculate the partial sum for even and odd elements separately
   - Combine the result in the end. (F2, F10)
              T. . D
                         F0, 0(R1)
              L.D
                         F6, -8(R1)
F4, 0(R2)
              L.D
                         F8, -8(R2)
              L.D
              MUL.D
                         F0, F0, F4
              MUL.D
                         F6, F6, F8
              ADDUI
                         R1, R1, #-16
                         R2, R2, #-16
              ADDUI
               ADD.D
                         F2, F0, F2
                                            Loop body takes 11 cycles
                         R1, foo
F10, F6, F10
              BNEZ
              ADD.D
                         F2, F2, F10
                                                              18
```

Dynamic Scheduling

- Advantages
 - Enables handling some cases when dependencies are unknown at compile time
 - Allows code that was compiled with one pipeline in mind to run efficiently on a different pipeline
 - Allows Out-of-order execution, Out-of-order completion

DIVD F0, F2, F4 ADDD F10, F0, F8 SUBD F12, F8, F14

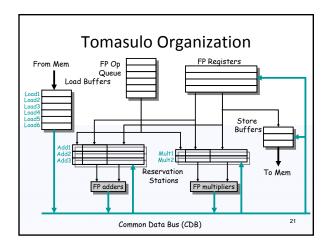
(stall) (have to wait)

19

Tomasulo Algorithm

- Designed to overcome long memory access and floating point delays.
- RAW hazards are avoided by executing an instruction only when its operands are available.
- WAR and WAW hazards arised from name dependencies, are eliminated by register renaming.
- Registers in instructions are replaced by values or pointers to reservation stations.
- The Common Data Bus (CDB) is used to bypass the registers and pass the results from the reservation stations directly to the functional units.

20



Reservation Station Components

- Busy: Indicates reservation station or FU is busy
- Op: Operation to perform in the unit (e.g., + or -)
- Vj, Vk: Value of Source operands
- Qj, Qk: Reservation stations producing source registers (value to be written)
 - Note: Qj,Qk=0 => ready
- A: effective address

22

Three Stages of Tomasulo Algorithm

- 1. Issue —get instruction from FP Op Queue
 - If reservation station free (no structural hazard), control issues the instruction & sends operands (renames registers).
- 2. Execute operate on operands (EX)
 - When both operands ready then execute; if not ready, watch CDB for result
- 3. Write result —finish execution (WB)
 - Write on CDB to all awaiting units; mark reservation station available

23

