

#### Outline

- Modern architectures.
- Delay slots.
- Introduction to instruction scheduling.
- List scheduling.
- Resource constraints.
- Interaction with register allocation.
- Scheduling across basic blocks.
- Trace scheduling.
- Scheduling for loops.
- Loop unrolling.
- Software pipelining.

Advanced Compiler Techniques 07/05/20 http://lamp.epfl.ch/teaching/advancedCompile

#### Simple Machine Model

- Instructions are executed in sequence.
  - Fetch, decode, execute, store results.
  - One instruction at a time.
- For branch instructions, start fetching from a different location if needed.
  - ◆ Check branch condition.
  - Next instruction may come from a new location given by the branch instruction.

Advanced Compiler Techniques 07/05/2004 p://lamp.epfl.ch/teaching/advancedCompiler/

### Simple Execution Model

5 Stage pipe-line:

Cycle: 1 2 3 4 5

| fetch | decode | execute | memory | write back

Fetch: get the next instruction.

Decode: figure out what that instruction is.

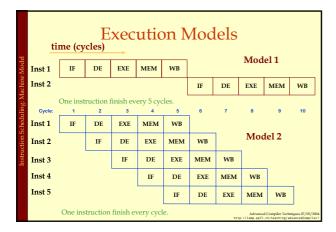
Execute: perform ALU operation.

address calculation in a memory op

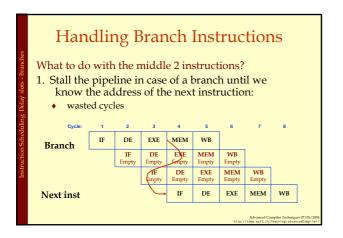
Memory: do the memory access in a mem. op.

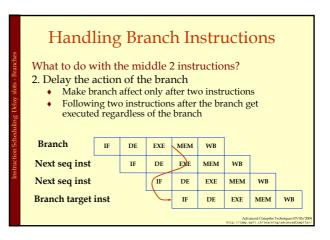
Write Back: write the results back.

Advanced Compiler Techniques 07/05/200 to://lamp.epfl.ch/teaching/advancedCompiler



#### **Handling Branch Instructions** Problem: We do not know the location of the next instruction until later. • after DE in jump instructions • after EXE in conditional branch instructions Branch DE MEM WB EXE ??? IF DE EXE MEM WB ??? DE EXE MEM Inst IF DE EXE MEM WB What to do with the middle 2 instructions?





MIPS has a branch delay slot

The instruction after a conditional branch gets executed even if the code branches to target

Fetching from the branch target takes place only after that

ble r3, foo

Branch delay slot

What instruction to put in the branch delay slot?

Filling the Branch Delay Slot

Simple Solution: Put a no-op.

Wasted instruction, just like a stall.

ble r3, lbl
nop

Branch delay slot

Filling the Branch Delay Slot

Move an instruction from above the branch.

prev\_instr
ble r3, lbl
prev\_instr
Branch delay slot

Moved instruction executes iff branch executes.
Get the instruction from the same basic block as the branch.
Don't move a branch instruction!
Instruction need to be moved over the branch.

• Branch does not depend on the result of the inst.

Filling the Branch Delay Slot

Move an instruction dominated by the branch instruction.

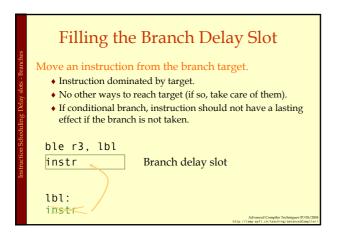
ble r3, lbl

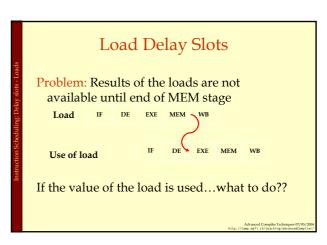
dom\_instr

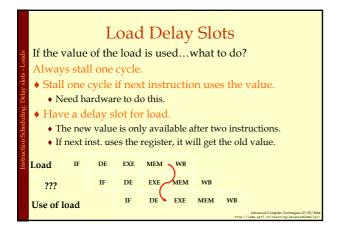
Branch delay slot

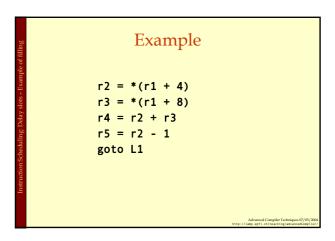
lbl:

dom\_instr









```
Example

r2 = *(r1 + 4)

r3 = *(r1 + 8)

noop

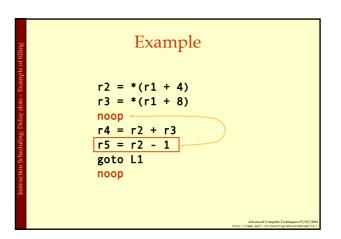
r4 = r2 + r3

r5 = r2 - 1

goto L1

noop

Assume 1 cycle delay on branches
and 1 cycle latency for loads
```



## Example r2 = \*(r1 + 4) r3 = \*(r1 + 8) r5 = r2 - 1 r4 = r2 + r3 goto L1 noop

#### Example

r2 = \*(r1 + 4) r3 = \*(r1 + 8) r5 = r2 - 1 goto L1 r4 = r2 + r3

Final code after delay slot filling

Advanced Compiler Techniques 07/05/2i

## From a Simple Machine Model to a Real Machine Model

- Many pipeline stages.
  - ♦ MIPS R4000 has 8 stages.
- Different instructions take different amount of time to execute.

mult 10 cycles
div 69 cycles
ddiv 133 cycles

 Hardware to stall the pipeline if an instruction uses a result that is not ready.

Advanced Compiler Techniques 07/05/2

#### Real Machine Model cont.

- Most modern processors have multiple execution units (superscalar).
  - If the instruction sequence is correct, multiple operations will take place in the same cycles.
  - Even more important to have the right instruction sequence.

Advanced Compiler Techniques 07/05/2004

#### **Instruction Scheduling**

Goal: Reorder instructions so that pipeline stalls are minimized.

#### Constraints on Instruction Scheduling:

- Data dependencies.
- Control dependencies .
- Resource constraints.

Advanced Compiler Techniques 07/05/20 p://lamp.epfl.ch/teaching/advancedCompile

#### Data Dependencies

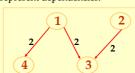
- If two instructions access the same variable, they can be dependent.
- Kinds of dependencies:
  - ♦  $True: write \rightarrow read$ . (Read After Write, RAW)
  - lacktriangle Anti: read  $\rightarrow$  write. (Write After Read, WAR)
  - ♦ Anti (Output): write → write. (Write After Write, WAW)
- What to do if two instructions are dependent?
  - ◆ The order of execution cannot be reversed.
  - Reduce the possibilities for scheduling.

#### Computing Data Dependencies

- For basic blocks, compute dependencies by walking through the instructions.
- Identifying register dependencies is simple.
  - is it the same register?
- For memory accesses.
  - simple: base + offset1 ?= base + offset2
  - ♦ data dependence analysis: a[2i] ?= a[2i+1]
  - interprocedural analysis: global ?= parameter
  - pointer alias analysis: p1 ?= p

#### Representing Dependencies

- Using a dependence DAG, one per basic block.
- Nodes are instructions, edges represent dependencies.



Edge is labeled with latency:

 $v(i \rightarrow j)$  = delay required between initiation times of i and j minus the execution time required by i.

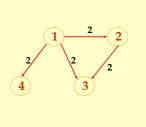
#### Example

1: 
$$r2 = *(r1 + 4)$$

2: 
$$r3 = *(r2 + 4)$$

$$3: r4 = r2 + r3$$

$$4: r5 = r2 - 1$$



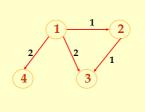
#### Another Example

1: 
$$r2 = *(r1 + 4)$$

$$2: *(r1 + 4) = r3$$

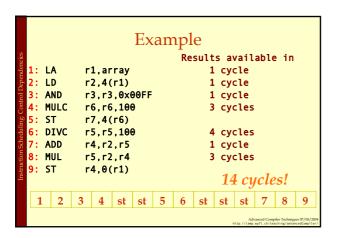
3: r3 = r2 + r3

4: r5 = r2 - 1



#### Control Dependencies and **Resource Constraints**

- ♦ For now, let's only worry about basic blocks.
- For now, let's look at simple pipelines.



#### List Scheduling Algorithm

- ♦ Idea:
- ◆ Do a topological sort of the dependence DAG.
- Consider when an instruction can be scheduled without causing a stall.
- Schedule the instruction if it causes no stall and all its predecessors are already scheduled.
- Optimal list scheduling is NP-complete.
  - Use heuristics when necessary.

Advanced Compiler Techniques 07/05/2004 ttp://lamp.epfl.ch/teaching/advancedCompiler/

#### List Scheduling Algorithm

- Create a dependence DAG of a basic block.
- ◆ Topological Sort.

READY = nodes with no predecessors.

Loop until READY is empty.

Schedule each node in READY when no stalling READY += nodes whose predecessors have all been scheduled.

Advanced Compiler Techniques 07/05/20 http://lamp.epfl.ch/teaching/advancedCompile

#### Heuristics for selection

Heuristics for selecting from the READY list:

- 1. pick the node with the longest path to a leaf in the dependence graph.
- 2. pick a node with the most immediate successors.
- 3. pick a node that can go to a less busy pipeline (in a superscalar implementation).

Advanced Compiler Techniques 07/05/200 to://lamp.epfl.ch/teaching/advancedCompiler

#### Heuristics for selection

Pick the node with the longest path to a leaf in the dependence graph

Algorithm (for node x)

- If x has no successors  $d_x = 0$
- $d_x = MAX(d_y + c_{xy})$  for all successors y of x.

Use reverse breadth-first visiting order

Advanced Compiler Techniques 07/05/2004

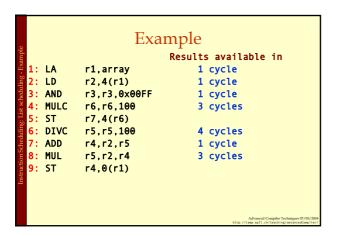
#### Heuristics for selection

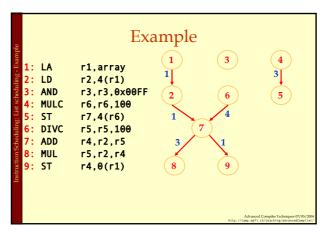
Pick a node with the most immediate successors.

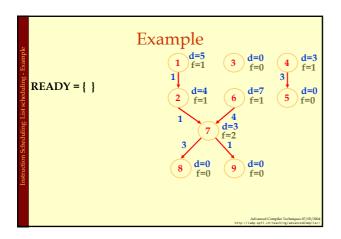
Algorithm (for node x):

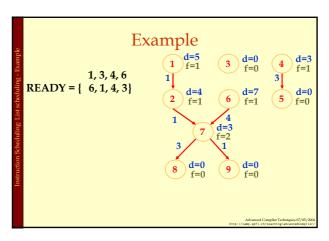
•  $f_x$  = number of successors of x

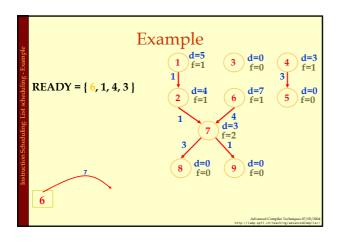
Advanced Compiler Techniques 07/05/200 //lamp.epfl.ch/teaching/advancedCompiler

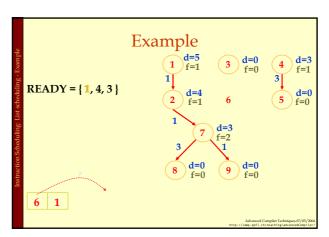


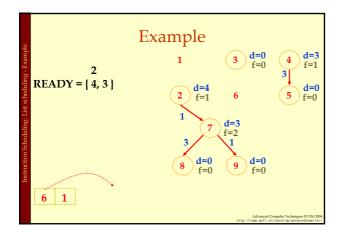


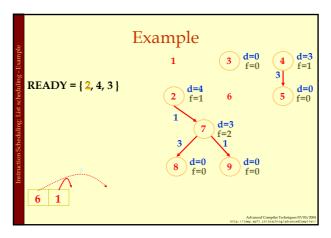


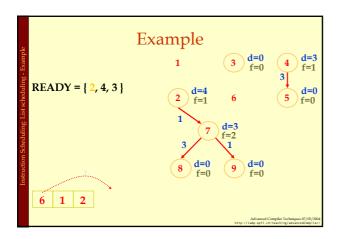


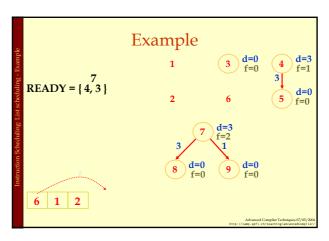


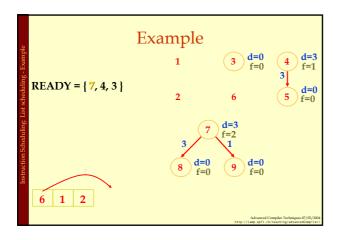


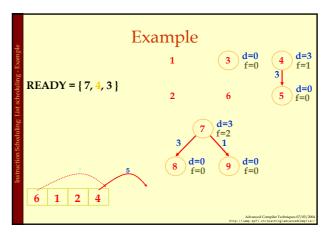


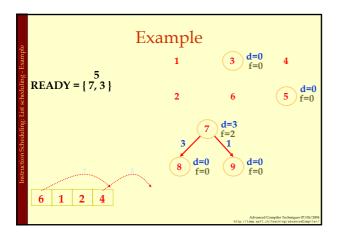


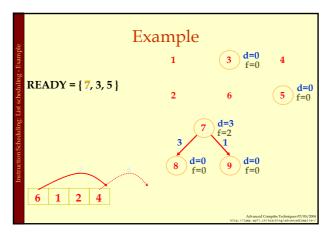


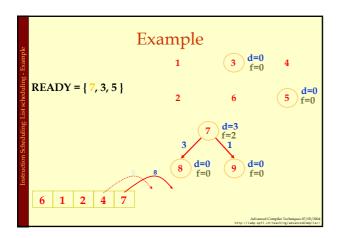


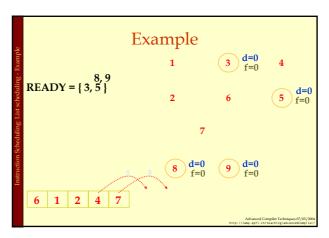


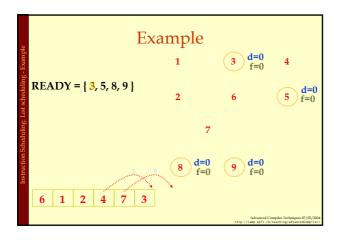


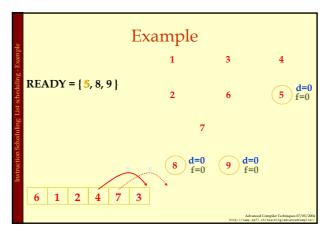


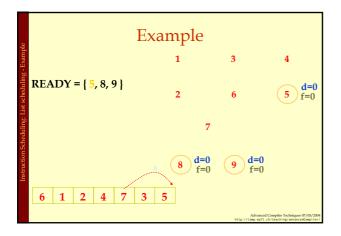


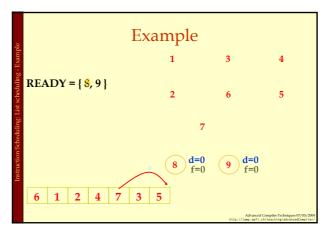


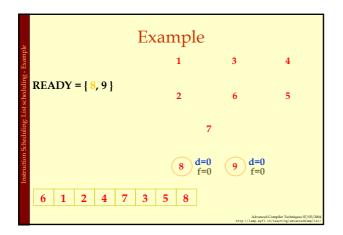


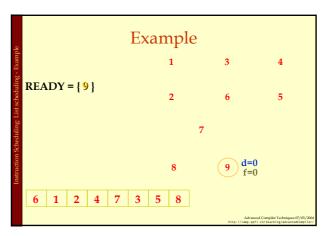


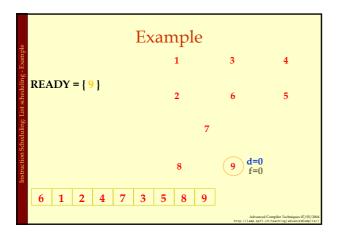


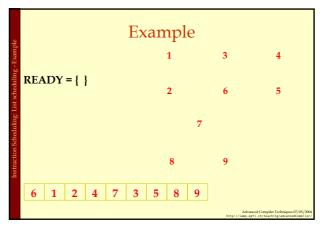


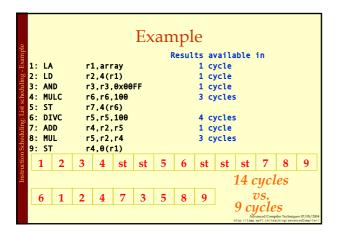












#### **Resource Constraints**

- Modern machines have many resource constraints.
- ◆ Superscalar architectures:
  - can run few parallel operations.
  - but have constraints.

Advanced Compiler Techniques 07/05/200tp://lamp.epfl.ch/teaching/advancedCompiler.

## Resource Constraints of a Superscalar Processor

#### Example:

- ♦1 integer operation, e.g., ALUop dest, src1, src2 # in 1 clock cycle In parallel with
- ♦ 1 memory operation, e.g.,

LD dst, addr # in 2 clock cycles ST src, addr # in 1 clock cycle

> Advanced Compiler Techniques 07/05/20 http://lamp.eofl.ch/teaching/advancedCompile

## List Scheduling Algorithm with Resource Constraints

- Represent the superscalar architecture as multiple pipelines.
  - Each pipeline represents some resource.

Advanced Compiler Techniques 07/05/20 http://lamp.epfl.ch/teaching/advancedCompile

## List Scheduling Algorithm with Resource Constraints

- Represent the superscalar architecture as multiple pipelines
  - Each pipeline represents some resource
- ♦ Example:
  - One single cycle ALU unit.
  - One two-cycle pipelined memory unit.

Volte two-cycle pipelined memory ur								
ALUop								
MEM 1								
MEM 2								

Advanced Compiler Techniques 07/05/2004

## List Scheduling Algorithm with Resource Constraints

- Create a dependence DAG of a basic block.
- ♦ Topological Sort

READY = nodes with no predecessors

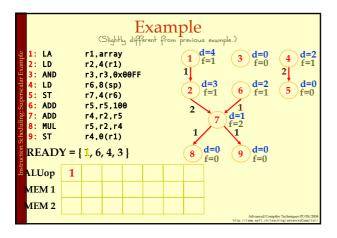
Loop until READY is empty

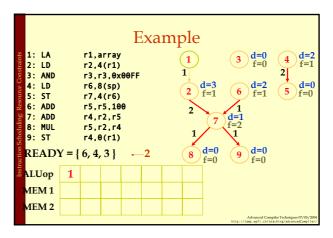
Let  $n \in READY$  be the node with the highest priority

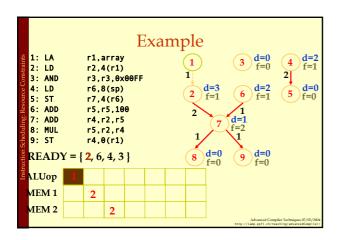
Schedule n in the earliest slot

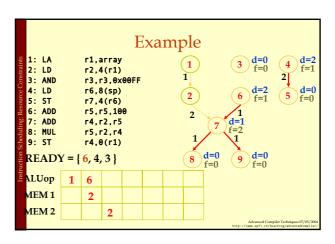
that satisfies precedence + resource constraints Update READY

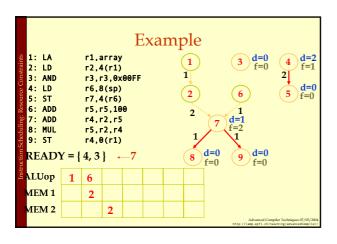
Advanced Compiler Techniques 07/05/

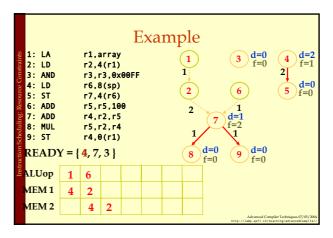


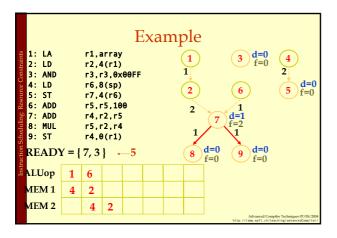


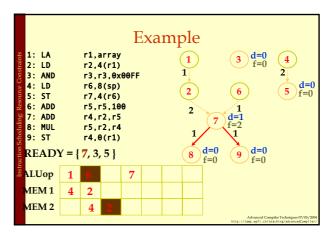


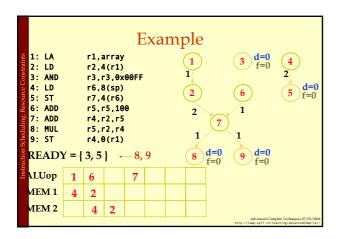


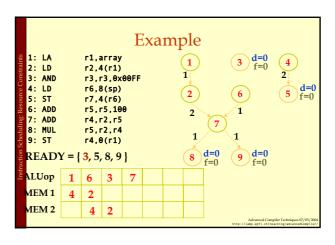


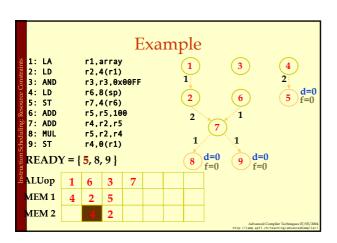


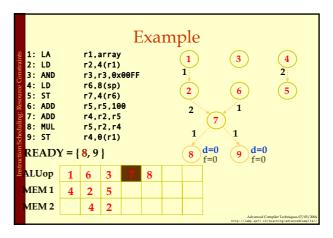


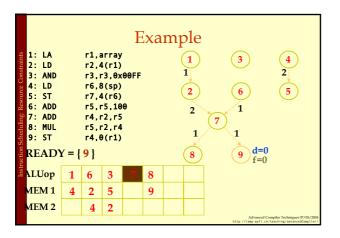


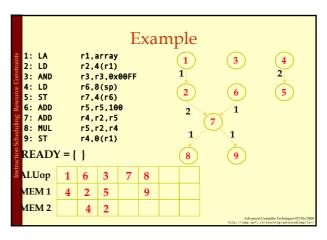


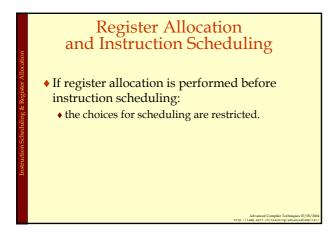


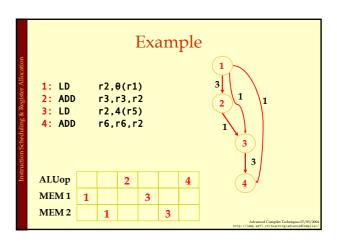


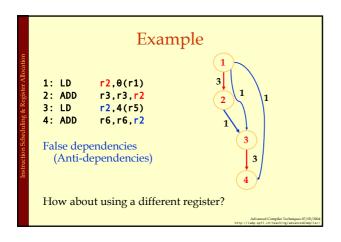


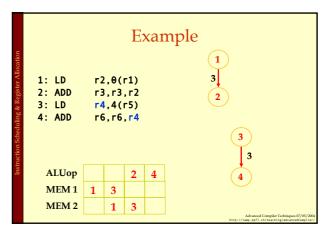












## Register Allocation and Instruction Scheduling

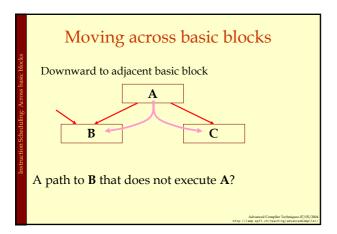
- If register allocation is performed before instruction scheduling:
  - the choices for scheduling are restricted.
- If instruction scheduling is performed before register allocation:
  - register allocation may spill registers.
  - will change the carefully done schedule!!!

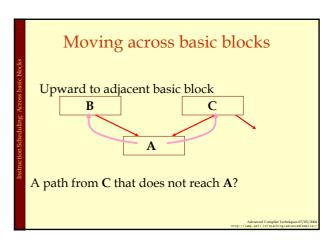
Advanced Compiler Techniques 07/05/2004

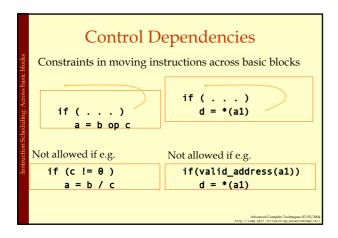
#### Scheduling across basic blocks

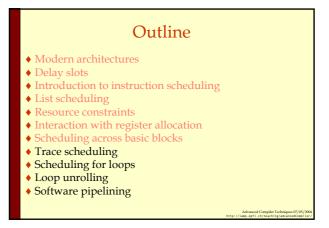
- Number of instructions in a basic block is small
  - Cannot keep a multiple units with long pipelines busy by just scheduling within a basic block.
- Need to handle control dependencies.
  - Scheduling constraints across basic blocks.
  - Scheduling policy.

Advanced Compiler Techniques 07/05/200 http://lamp.epfl.ch/teaching/advancedCompiler





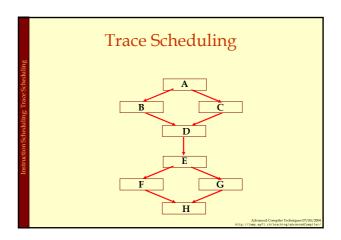


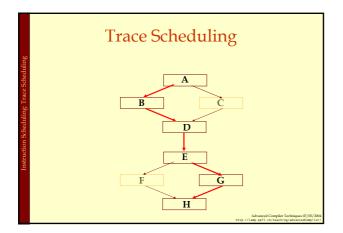


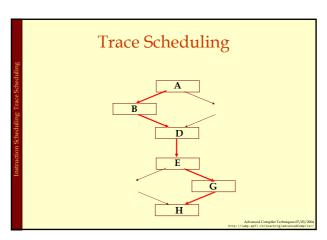
#### Trace Scheduling

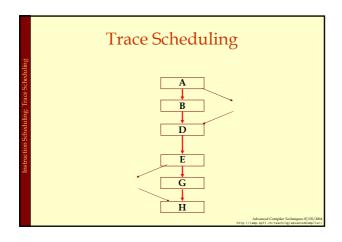
- Find the most common trace of basic blocks.
  - Use profile information.
- Combine the basic blocks in the trace and schedule them as one block.
- ◆ Create compensating (clean-up) code if the execution goes off-trace.

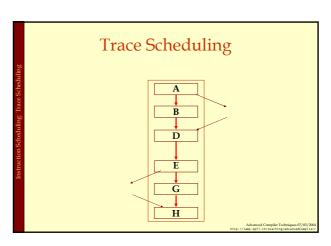
Advanced Compiler Techniques 07/05/ ://lamp.epfl.ch/teaching/advancedCompi

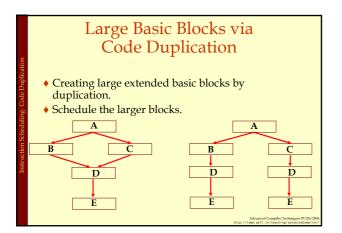












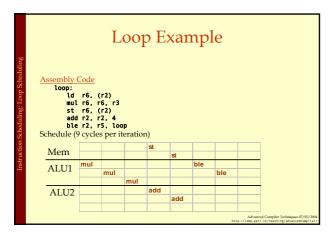
#### Scheduling for Loops

- ♦ Loop bodies are typically small.
- ◆ But a lot of time is spend in loops due to their iterative nature.
- Need better ways to schedule loops.

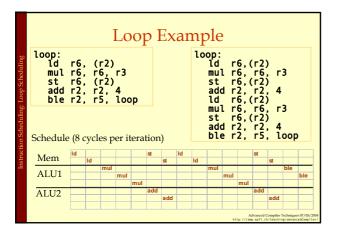
Advanced Compiler Techniques 07/05/2004 http://lamp.epfl.ch/teaching/advancedCompiler/

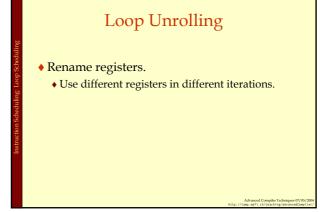
## Machine: • One load/store unit • load 2 cycles • store 2 cycles • Two arithmetic units • add 2 cycles • branch 2 cycles (no delay slot) • multiply 3 cycles • Both units are pipelined (initiate one op each cycle)

### 



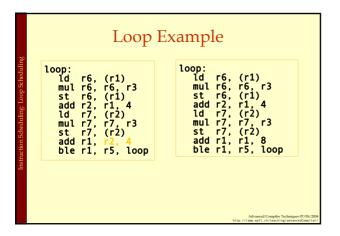
## Oldest compiler trick of the trade: Unroll the loop body a few times Pros: • Creates a much larger basic block for the body. • Eliminates few loop bounds checks. Cons: • Much larger program. • Setup code (# of iterations < unroll factor). • Beginning and end of the schedule can still have unused slots.

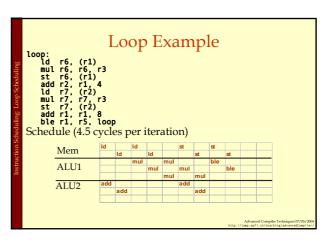




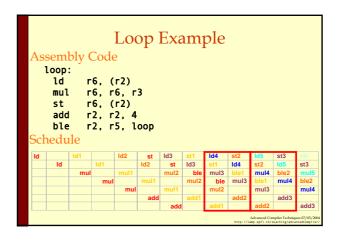
```
    Loop Unrolling
    Rename registers.
    Use different registers in different iterations.

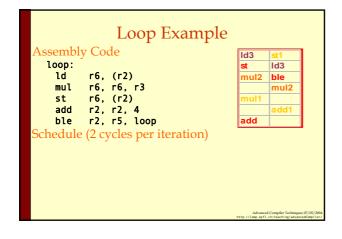
    Eliminate unnecessary dependencies.
    again, use more registers to eliminate true, anti and output dependencies.
    eliminate dependent-chains of calculations when possible.
```

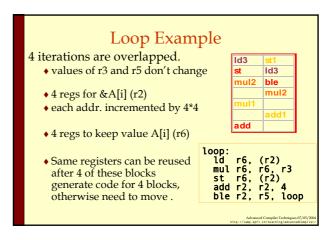




# Software Pipelining Try to overlap multiple iterations so that the slots will be filled. Find the steady-state window so that: all the instructions of the loop body is executed. but from different iterations.







#### Software Pipelining

- Optimal use of resources.
- Need a lot of registers.
  Values in multiple iterations need to be kept.
  Issues in dependencies.
- Issues in dependencies.
   Executing a store instruction in an iteration before branch instruction is executed for a previous iteration (writing when it should not have).
   Loads and stores are issued out-of-order (need to figure-out dependencies before doing this).
   Code generation issues.
   Generate pre-amble and post-amble code.
   Multiple blocks so no register copy is needed.