

Chapter 3: SuperScalar Concepts

Prof. Ben Lee

School of Electrical Engineering and Computer Science
Oregon State University

Chapter Outline

- Elements of a SuperScalar Processor
- SuperScalar Execution Model
- Complexity of SuperScalar Processors
- Case Study - Pentium 4
- Microprocessor Comparison
- The Core Microarchitecture



Elements of a SuperScalar Processor

Chapter 3: SuperScalar Concepts

3



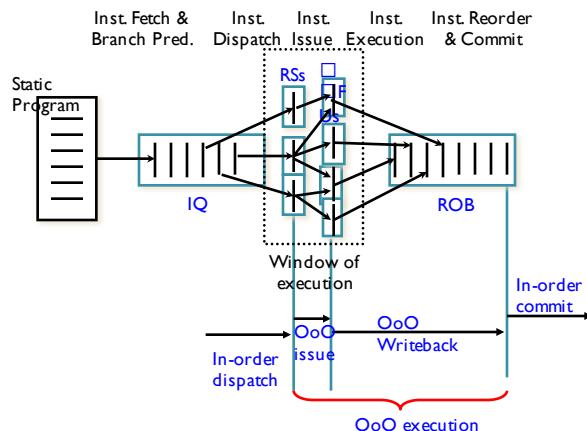
Elements of a SuperScalar Processor

- All the elements listed below are integrated in a cohesive, almost seamless, manner:
 - Fetching multiple instructions and by predicting the outcomes of, and fetching beyond, conditional branches.
 - Determining RAW dependences among register values.
 - Issuing multiple instructions OoO and resolving WAR and WAW hazards.
 - Multiple FUs and memory hierarchy capable of simultaneous servicing memory references.
 - Committing the process state in correct order (precise interrupt).

Chapter 3: SuperScalar Concepts

4

Basic SuperScalar Design (1)



Chapter 3: SuperScalar Concepts

5

Basic SuperScalar Design (2)

- **Instruction Window** – Consists of Instruction Queue (IQ), and Reservation Stations (RS), which allows multiple instructions to be executed out-of-order, i.e., OoO issue and OoO writeback, on multiple FUs.
- **Register Renaming** - Using additional operand buffers in RS and/or ROB eliminates WAR and WAW hazards.
- **Reorder Buffer (ROB)** - Allows the process state to be committed in correct order by dispatching in-order and committing in-order.

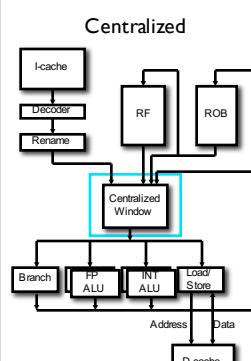
Chapter 3: SuperScalar Concepts

6

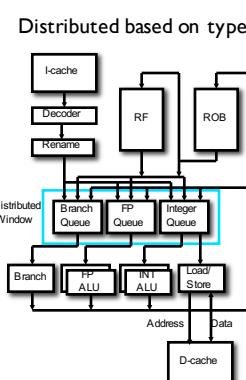
Instruction Window

- Instruction Window (IW) resides between the instruction decoder and the FUs. The decoder places instructions into the window, and the instruction-issue logic examines instructions in the window to appropriately select instructions to issue to the FUs.
- There are three ways to implement the instruction window:
 - Centralized
 - Distributed based on type
 - Distributed based on FUs

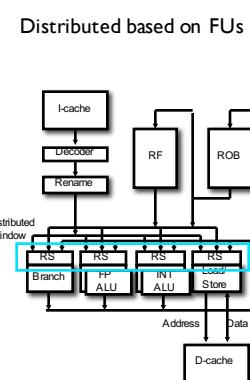
Different IW Implementations



Examples: UltraSparc,
Pentium Pro (II & III), Core



Examples: Alpha, MIPS, Pentium
4, Power5



Examples: PowerPC

Register Renaming

- The processor allocates a new register for every new value produced, i.e., for every instruction that writes a register.

$$\begin{array}{ll} \textcircled{R3} = R3 \text{ op } R5 & R3_b = R3_a \text{ op } R5_a \\ \textcircled{R4} = \textcircled{R3} + I & R4_b = R3_b + I \\ \textcircled{R3} = R5 + I & R3_c = R5_a + I \\ \textcircled{R7} \neq \textcircled{R3} \text{ op } R4 & R7_b = R3_c \text{ op } R4_b \end{array}$$

- Need more **physical registers** than **logical registers**.
- Renaming avoids WAW and WAR hazards; turns instruction stream into a "single assignment" form.

Register Renaming in Tomasulos' Approach (I)

- In the Tomasulo Algorithm, we saw that renaming was implemented by
 - Operand buffers in the RSs that serve as extra registers.
 - This eliminates the need to get the operand from the RF (which caused WAR and WAW in the first place).
 - A pending instruction designates the RS that will provide its input.
 - As instructions are issued, the register specifiers for pending operands are renamed to the names of the RS.
 - When successive writes to a register appear, only the last one is actually used to update the register.

Register Renaming in Tomasulos' Approach (2)

Example: WAR hazard

Instruction	Issued at	Execute During	Write Result	Comments
L.D F6, 34(R2)	0	1-2	3	Load unit is busy until cycle 3
L.D F2, 45(R1)	1	4-5	6	Load unit is busy until cycle 6
MUL.D F0, F2, F4	2	7-16	17	wait for F2; MULTI busy until cycle 17
SUB.D F8, F6, F2	3	7-8	9	wait for F2; ADD1 busy until 9
DIV.D F10, F0, F6	4	18-57	58	wait for F0 of MULTF; MULTI busy until 58
ADD.D F6, F8, F2	5	10-11	12	wait for F8 of SUBF; ADD1 busy until 12

- Both DIV.D and ADD.D were issued, even though there is a WAR hazard on F6.
 - If the L.D has completed, then V_k (source operand F6) will store the result.
 - If the L.D had not completed, then Q_k (RS that will produce F6) would point to the Load1 reservation station.

Chapter 3: SuperScalar Concepts

11

Register Renaming in Tomasulos' Approach (3)

Example: WAW hazard

Loop:	LF F0, 0(R1)	LF F0, 0(R1)	
	MULTF F4, F0, F2	MULTF F4, F0, F2	
	SF 0(R1), F4	SF 0(R1), F4	
	SUBI R1, R1, #8	WAW SUBI R1, R1, #8	• Will be in different RSs and the RS that will produce results will be different
	BNEZ R1, Loop	BNEZ R1, Loop	
	LF F0, 0(R1)	LF F0, 0(R1)	
	MULTF F4, F0, F2	MULTF F4, F0, F2	• This is done using Q_i in the RF that indicates which RS will produce the result.
	SF 0(R1), F4	SF 0(R1), F4	
	SUBI R1, R1, #8	SUBI R1, R1, #8	
	BNEZ R1, Loop	BNEZ R1, Loop	
	...		

Chapter 3: SuperScalar Concepts

12

Register Renaming Using Physical Registers (1)

- Since not all processors use RSs, another way to do it is by have physical registers > logical registers.
- Example:

```
L.D    F6, 34(R2)
L.D    F2, 45(R3)
MUL.D F0, F2, F4
SUB.D F7, F6, F2
DIV.D F1, F0, F6
ADD.D F6, F7, F2
```

Assume 8 logical, 16 physical registers.

Chapter 3: SuperScalar Concepts

13

Register Renaming Using Physical Registers (2)

Initial Mapping		Final Mapping
Logical	Physical	
0	12	⇒ 2
1	13	⇒ 4
2	15	⇒ 1
3	14	
4	9	
5	7	
6	6	⇒ 0
7	8	⇒ 5 ⇒ 3
Free Pool: 0, 1, 2, 3, 4, 5, 10, 11		

Chapter 3: SuperScalar Concepts

14

Register Renaming Using Physical Registers (3)

- Code with physical register assignments

```
L.D    P0, 34(R2)
L.D    P1, 45(R3)
MUL.D P2, P1, P9
SUB.D P3, P0, P1
DIV.D P4, P2, P0
ADD.D P5, P3, P1
```

- Afterwards, reclaim the physical registers that are no longer used.

- Ignoring precise interrupts, a register can be returned after the last read is done.
- Considering precise interrupts, a register can be returned after it is logically over-written.

Register Renaming Using ROB (1)

- The ROB is managed as a FIFO queue:
 - When an instruction is decoded, it is allocated an entry at the tail of the ROB so that result value of this instruction can be written into the allocated entry after it completes.
 - When the value reaches the head of the ROB, it is written into the RF (if no exception occurred). If the instruction is not complete when its entry reaches the bottom of the ROB, the ROB does not advance until the instruction completes.
- ROB keeps the original lexicographical order of the instructions and allowed out-of-order execution while preserving in-order writing of results to registers and memory => **precise interrupts!**

Register Renaming Using ROB (2)

- Renaming with ROB:
 - Logical registers == physical registers
 - ROB commits state in order
 - ROB holds “renamed” logical registers
 - Registers read only at dispatch
 - Otherwise, values come from FUs or ROB
- Reorder buffer while FP divide in progress:

		<u>Entry</u>	<u>Content</u>
LD	F6,34(R2)	0	F6
LD	F2,45(R3)	1	F2
MULD	F0,F2, F4	2	F0
SUB.D	F7,F6, F2	3	F7
DIV.D	F1,F0, F6	4	F1
ADD.D	F6,F7, F2	5	F6

Chapter 3: SuperScalar Concepts

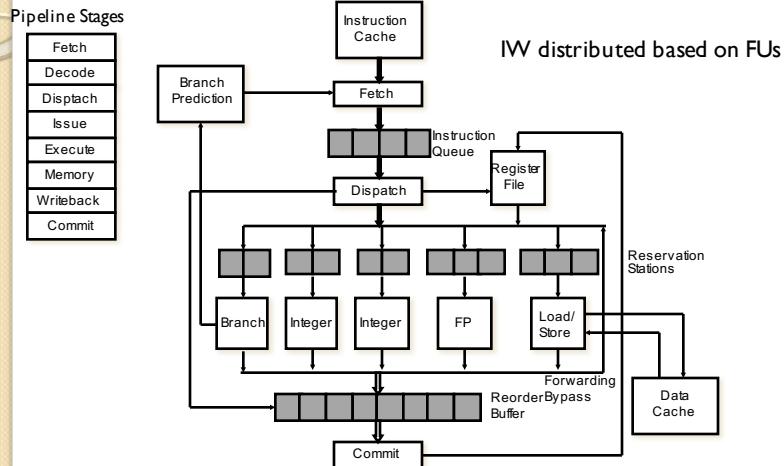
17

• SuperScalar Execution Model

Chapter 3: SuperScalar Concepts

18

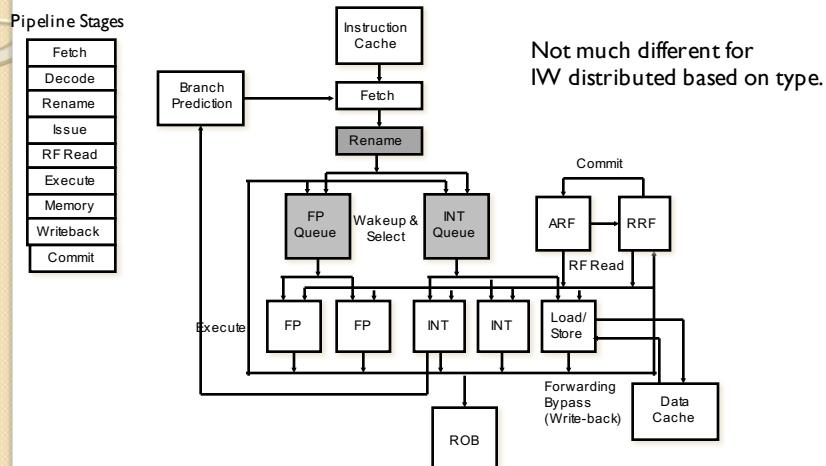
SuperScalar Execution Model



Chapter 3: SuperScalar Concepts

19

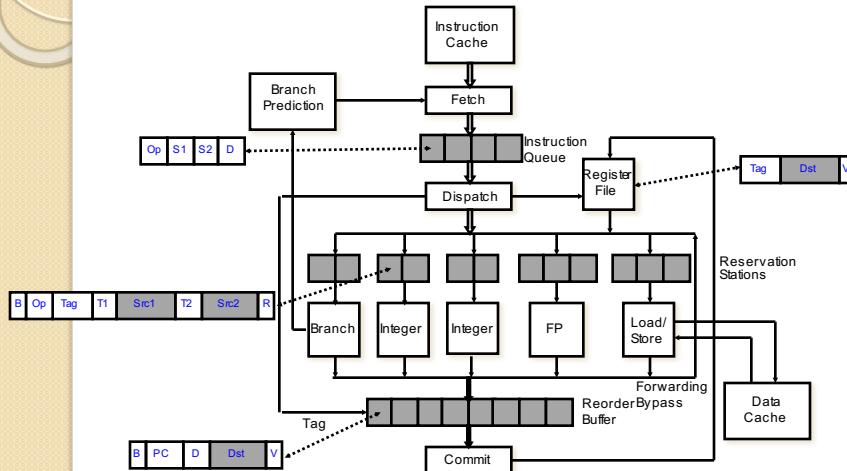
SuperScalar Execution Model



Chapter 3: SuperScalar Concepts

20

Information in IQ, RS, & ROB



Chapter 3: SuperScalar Concepts

21

Information in Various Components

Instruction Queue:

Op Opcode
SI/2 Source I/2 register number.
D Destination register number.

Reservation Station:

B Indicates the RS has been allocated.
Op Opcode
Tag The pointer to the ROB entry where the result of this instruction will be written.
Src1/2 Source operand.
TI/2 The pointer to the ROB entry which will receive the result needed by the source registers Src1/2.
R Indicates the instruction is ready to be issued.

FU:

B Indicates FU is busy.
Op Opcode
Tag The pointer to the ROB entry where the result of this instruction will be written to.
Src1/2 Source operand.

Reorder Buffer:

B Indicates the ROB entry is busy.
PC Address of the instruction
D Destination register number.
Dst Result value
V When true, signifies that Dst field holds a valid result.

Register File:

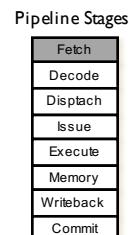
Tag The pointer to the ROB entry that will contain the result.
Dst Result value.
V When true signifies that Dst field holds a valid result.

Chapter 3: SuperScalar Concepts

22

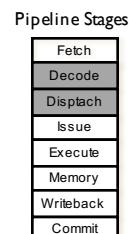
Fetch Stage

- Instructions are fetched from the I-cache and enqueued into the IQ in program order.
- For an n -issue superscalar operation, the Fetch stage needs to fetch at least n -instructions.



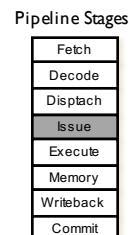
Decode/Dispatch Stage

- Instructions are decoded and then dispatched when RSs are available:
 - Up to n instructions can be decoded and dispatched per cycle.
- For each instruction dispatched
 - Allocate a RS and set it to busy.
 - Allocate the next sequential entry in the ROB and set it to busy. Place the Tag value pointing to this entry into the Tag field of the allocated RS.
 - For destination register D, update its Tag with the Tag of the just allocated ROB entry and set Dst to invalid.
 - For each of the source registers S1 and S2, check its Dst in RF.
 - If valid, read Dst from register and update Src1/Src2.
 - Else, check if Dst of ROB entry pointed to by its Tag is valid:
 - If valid, update Src1/Src2 with Dst;
 - Else, update T1/T2 with Tag from the register



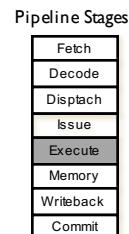
Issue Stage

- Issue all the instructions when both source operands and FUs are available (i.e., R-bit is set). This is done by
 1. Copy Op, Tag, Src1, and Src2 from RS to FU. Set FU to busy.
 2. Free RS.



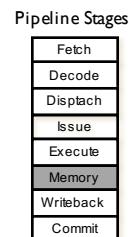
Execute Stage

- Execute instructions.
 - Latency will depend on the instruction



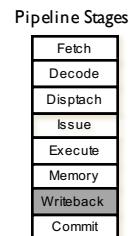
Memory Stage

- Only load and store instructions perform this stage.



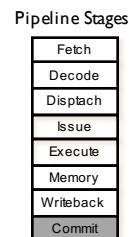
Write-back Stage

- When an instruction completes, post its result in the ROB and forward.
 - Write the result into the Dst field of the entry in the ROB pointed to by the Tag field, and set it to valid.
 - Forward the result back to RSs where the Tag field of the completed instruction matches source register tags T1 and T2 in RSs. For each match, copy the result into Src1/Src2 fields.



Commit Stage

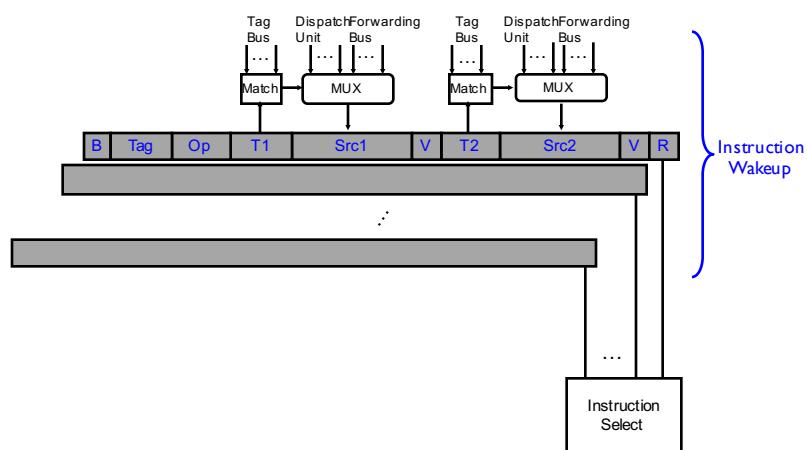
- If Dst of the entry at the front of the ROB is valid (i.e., instruction has completed), write it to the destination register
- D. Remove this ROB entry by clearing B.
- Up to n instructions can be committed per cycle.



Chapter 3: SuperScalar Concepts

29

Reservation Stations



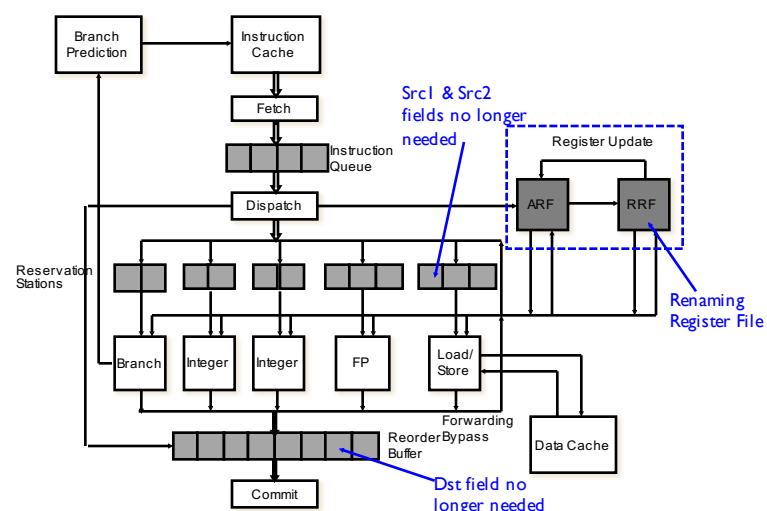
Chapter 3: SuperScalar Concepts

30

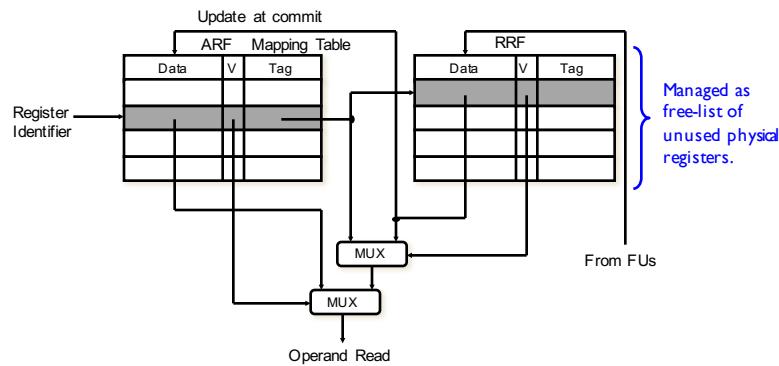
Register Renaming

- Register renaming in our example thus far has been based on ROB:
 - Dst field in ROB acts as a renaming register.
 - Tag field in RF points to the renaming register in ROB.
 - A renaming register is allocated for every instruction in flight.
- However, not all instructions require a renaming register:
 - e.g., Branch instructions
 - May be wasteful.
- Some processor use a separate register file for renaming:
 - Architected Register File (ARF) - logical registers
 - Renaming Register File (RRF) - physical registers
 - ROB no longer has Dst field for data
 - Data will be held in one of the renaming register in RRF
 - Data from FUs will be forwarded and written to renaming registers in RRF.

Renaming Register File (I)



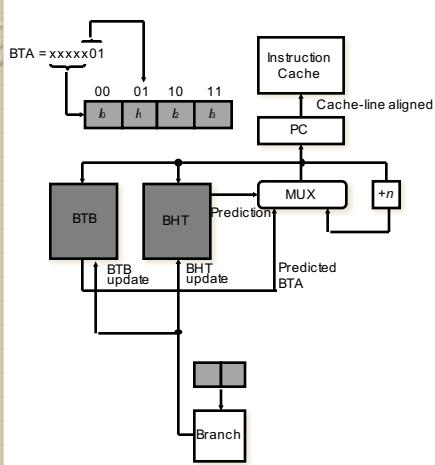
Renaming Register File (2)



Chapter 3: SuperScalar Concepts

33

Branch Prediction



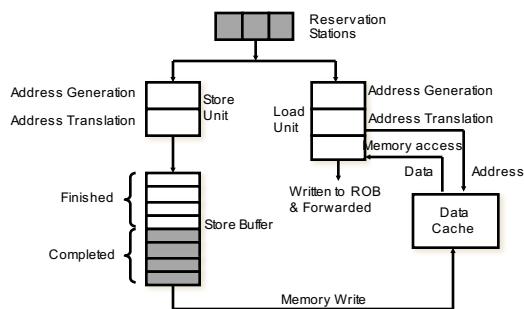
- BTB and BHT accessed during Fetch stage.
- Instruction fetched is cache aligned.
- Predicted BTA will be the first branch in fetched group of instructions.
- When branch is resolved
 - BHT is updated (if needed).
 - BTB entry is either left alone (taken) or deleted (not taken).
- Can speculate up to the number of RSs available.
 - After branch is resolved, speculative branch instructions are either made non-speculative or invalidated.
 - If branch is invalidated, ROB entries occupied by the branch and its succeeding insts are all deallocated.

Chapter 3: SuperScalar Concepts

34

Memory Dataflow

- Assumes loads/stores are issued in-order from RSs.

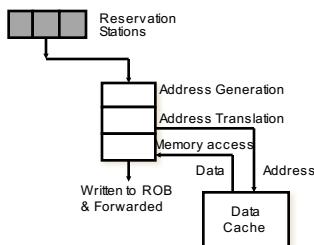


Chapter 3: SuperScalar Concepts

35

Loads

- When rs of lw rt, disp(rs) is available, inst. is issued.
- Address = disp + [rs]
- Virtual to Physical address translation (TLB).
- Data Cache accessed and data written to ROB and forwarded.
- When load instruction is at the head of ROB, RF is updated (i.e., committed).



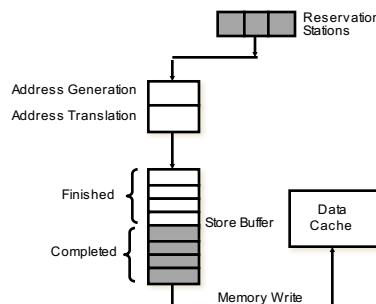
Chapter 3: SuperScalar Concepts

36

Stores

Stores are processed differently:

- When both rs and rt of sw rt,disp(rs) is available, inst. is issued.
- Address = disp + [rs]
- Virtual to Physical address translation (TLB).
- Address and data is put into Store Buffer and is considered ***finished***.
- When store instruction is at the head of ROB, it is considered ***completed***.
- When ***memory bus is available***, completed store exists the store buffer and memory is updated (i.e., committed).

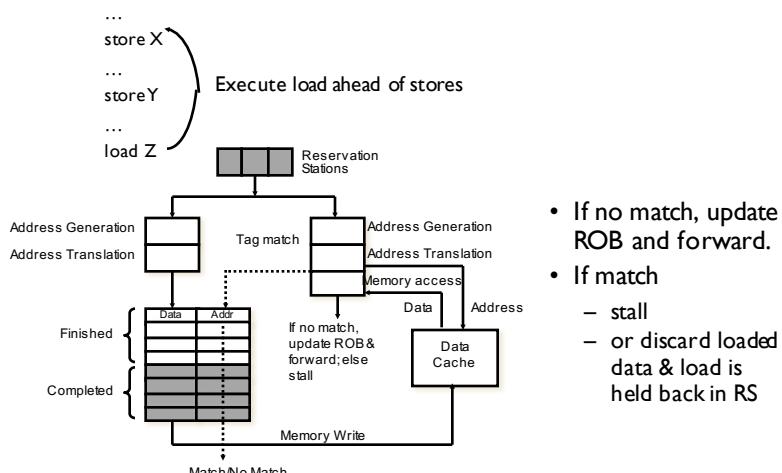


Chapter 3: SuperScalar Concepts

37

Load Bypassing Stores

- Why have store buffers but no load buffers? Consider the following:

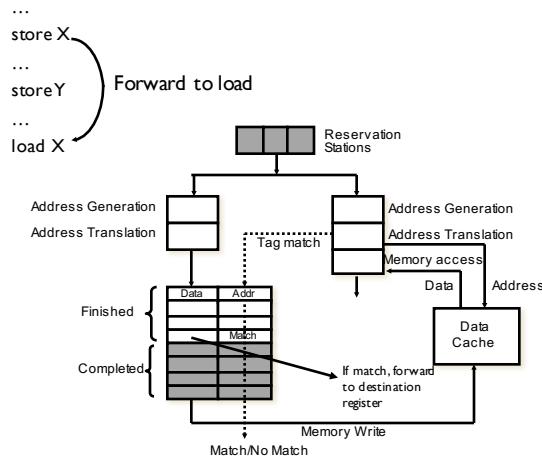


Chapter 3: SuperScalar Concepts

38

Load Forwarding

- If load depends on a previous store...

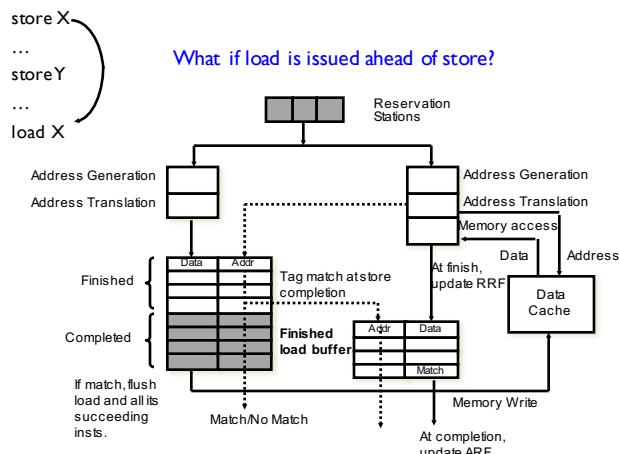


Chapter 3: SuperScalar Concepts

39

OoO Loads and Stores

- We need to be careful since loads and stores can be issued OoO from the RSs.



Chapter 3: SuperScalar Concepts

40

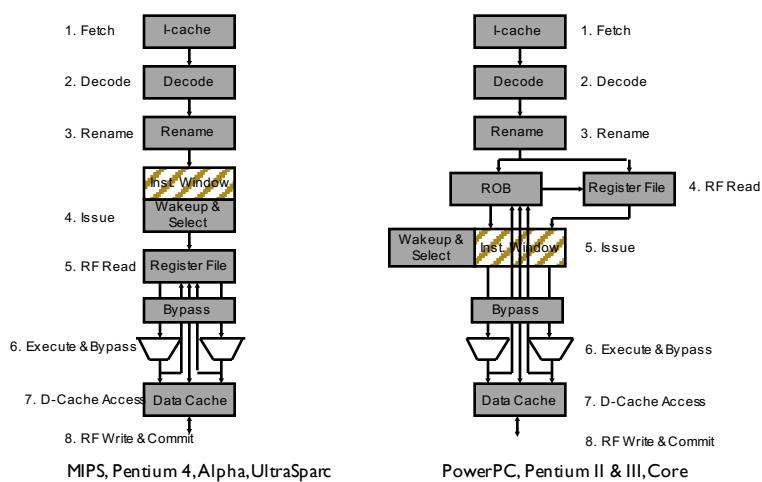
Complexity of SuperScalar Processors

Chapter 3: SuperScalar Concepts

41

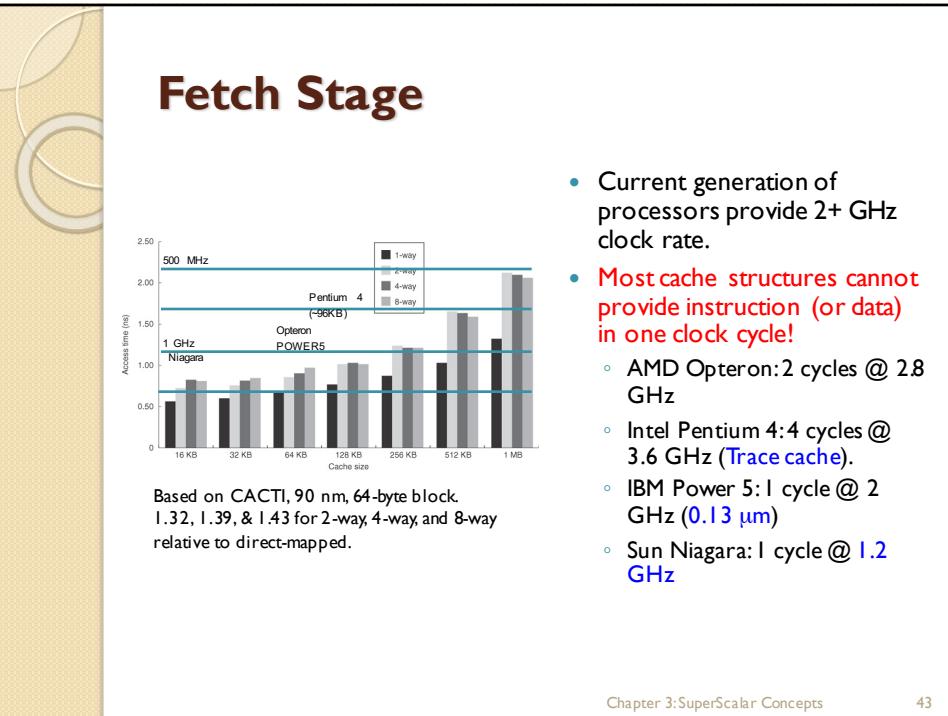
Complexity of SS Processors

Primary Delay Structures



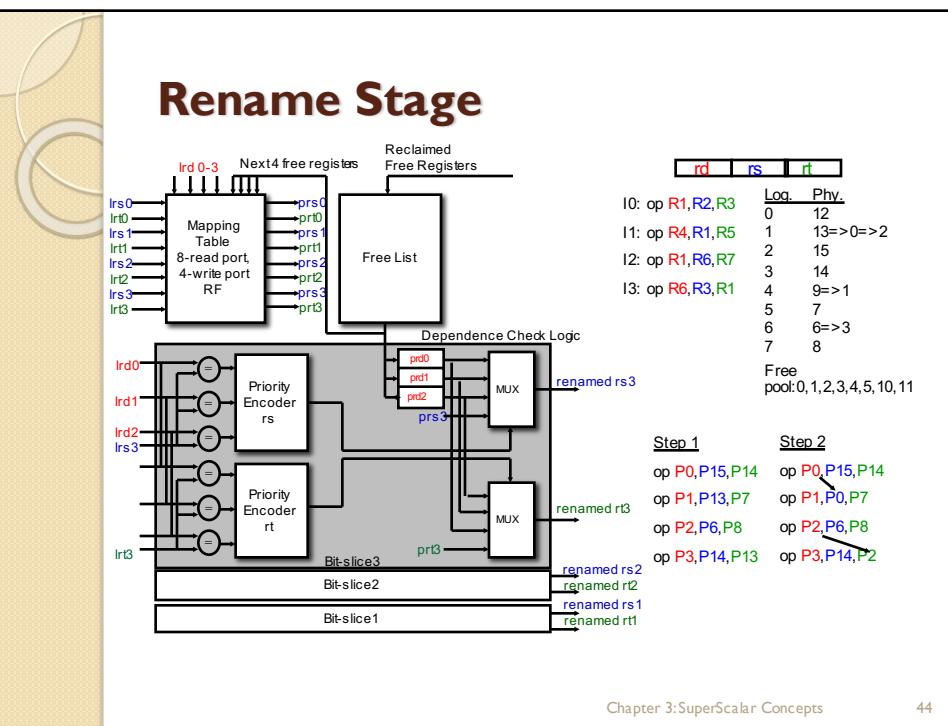
Chapter 3: SuperScalar Concepts

42



Chapter 3: SuperScalar Concepts

43



Chapter 3: SuperScalar Concepts

44

Delay Analysis of Rename Stage

The critical path is the time it takes to read the Mapping Table (register file). Delay through the output MUX is minimal.

$$T_{\text{rename}} \propto pR^{1/2}$$

where

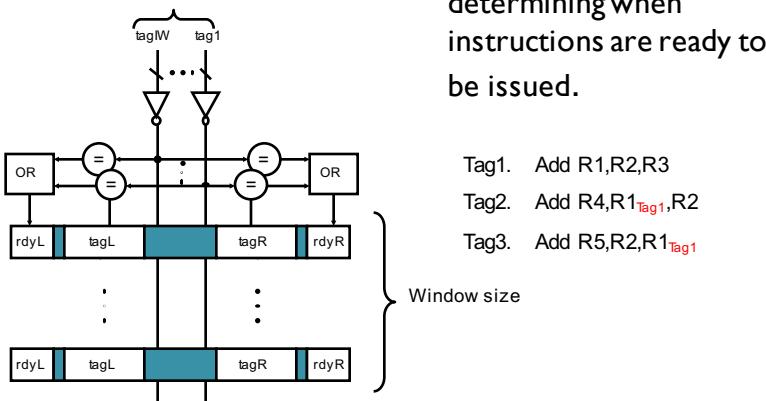
p - number of ports

R - number of entries

Usually, R is small (≤ 32), thus rename delay is not critical.

Wakeup Logic

From outputs of FUs
of FUs = Issue Width



- Responsible for determining when instructions are ready to be issued.

Tag1. Add R1,R2,R3

Tag2. Add R4,R1^{Tag1},R2

Tag3. Add R5,R2,R1^{Tag1}

Window size

Delay Analysis of Wakeup Logic

Delay consists of buffers to drive the tag bits, the comparators to match tags, and OR the individual match signals:

$$T_{\text{wakeup}} = T_{\text{tagdrive}} + T_{\text{tagmatch}} + T_{\text{matchOR}}$$

T_{tagdrive} depends on the length of tag lines and the number of comparators on the tag lines (both depend on WINSIZE).

$$T_{\text{wakeup}} = c_0 + (c_1 + c_2 \times IW) \times \text{WINSIZE} + (c_3 + c_4 \times IW + c_5 \times IW^2) \times \text{WINSIZE}^2$$

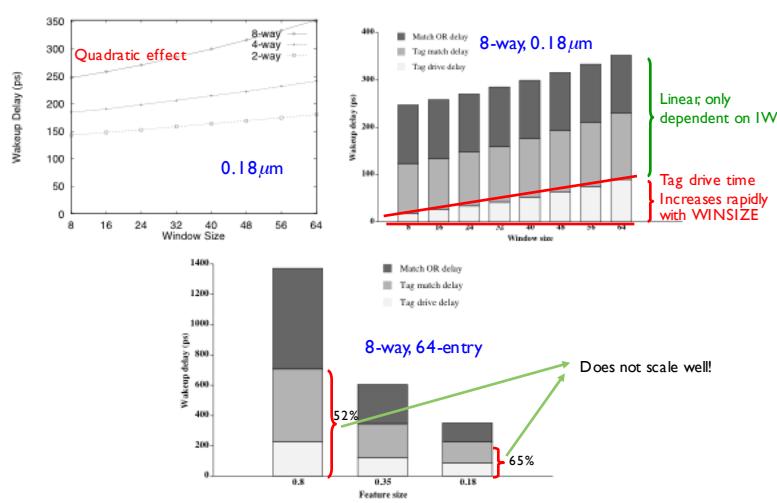
where

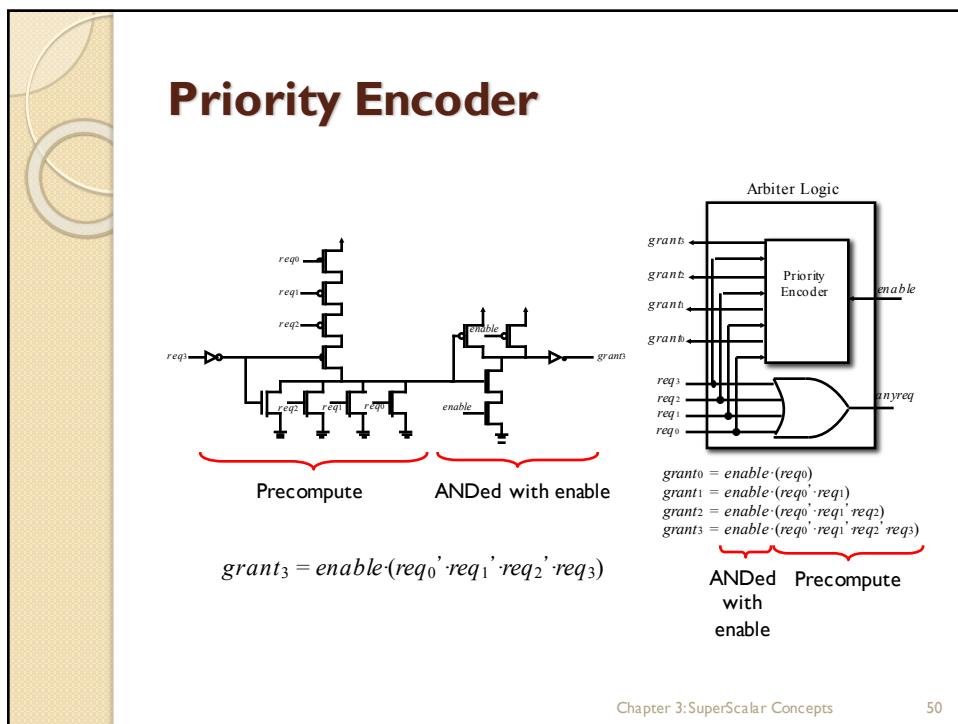
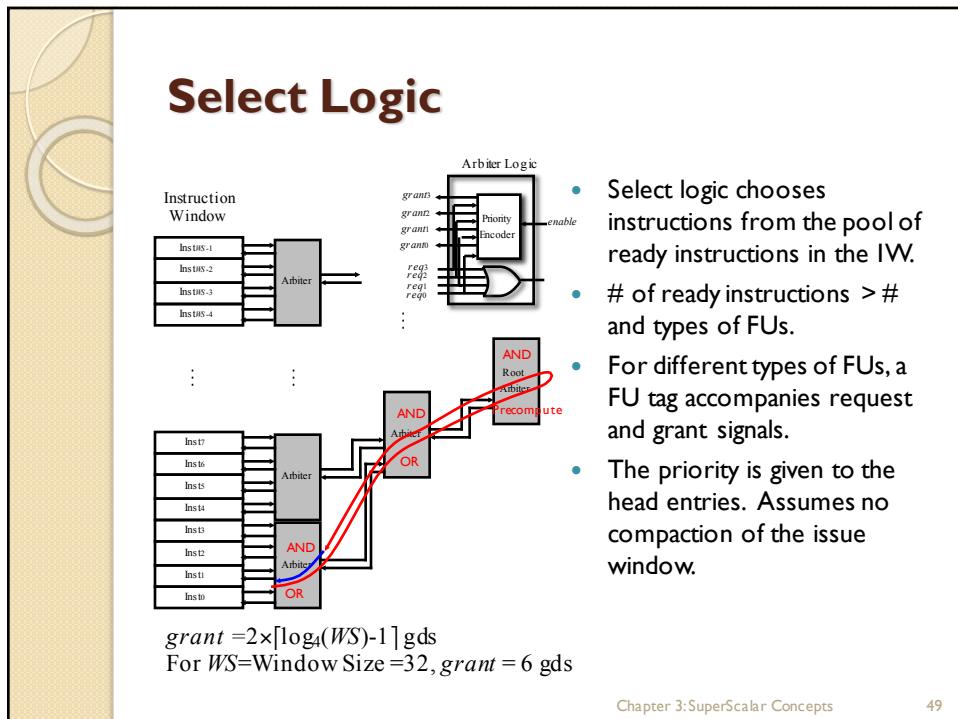
IW - issue width

WINSIZE - window size

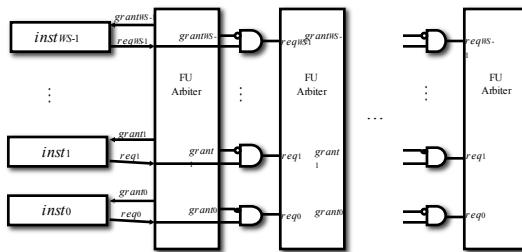
c_{0-5} - fixed constants for given technology & ISA

Delay Analysis of Wakeup Logic





Delay Analysis of Selection Logic



- Can't have too many FUs.
- PowerPC has separate RS for each FU.
- MIPS and Alpha have separate integer and FP IQs.
 - MIPS: 3 IQs for INT, FP, and Load/Store
 - Alpha: 2 IQs for INT and FP
 - Pentium 4: 2 IQs for INT/FP and Load/Store

Chapter 3: SuperScalar Concepts

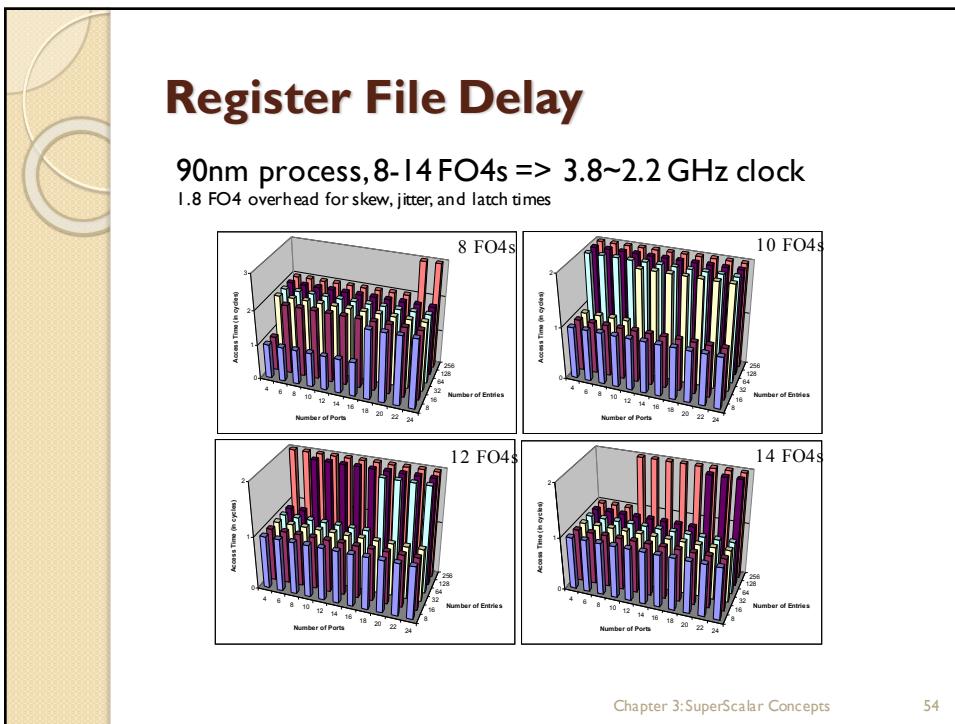
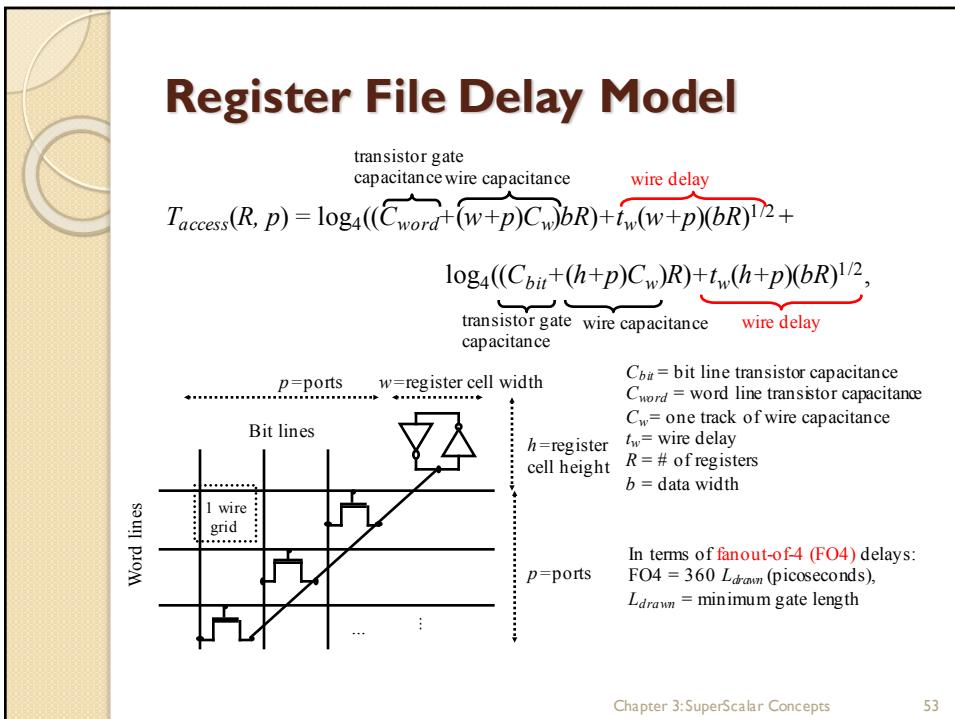
51

Register File Access

- Register file access time depends on number of *ports* and *entries*.
 - n -way issue requires at most $2n$ read ports and n write ports
- RF design
 - Separate physical and logical RFs
 - Combined physical and logical RF
- To support large instruction window and number of in-flight instructions, a physical RF with many entries and ports is required.
 - Pentium 4 has 128 reaming registers.
 - POWER 5 is a 8-way issue.

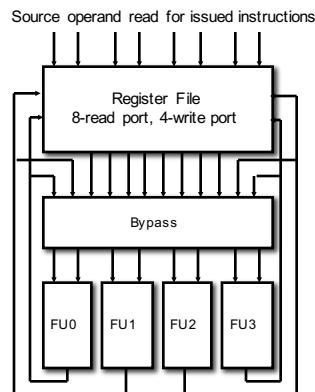
Chapter 3: SuperScalar Concepts

52



Complexity of Bypass Logic

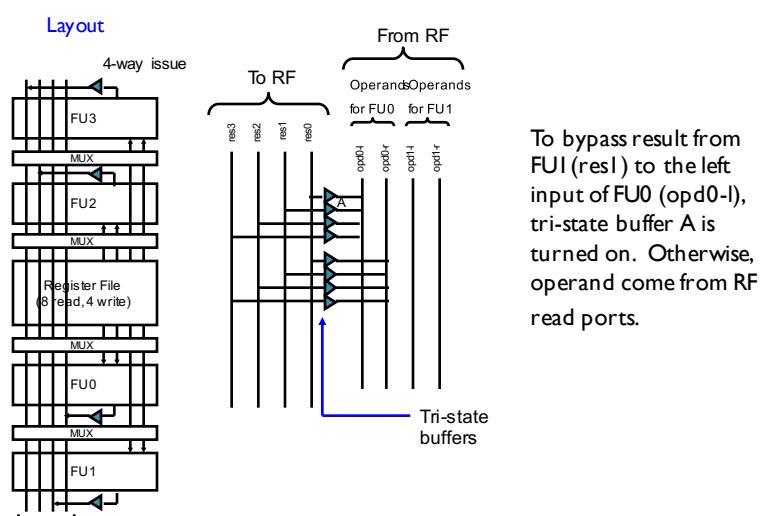
Forwards result values from completed instruction to dependent instruction, bypassing the RF.



Chapter 3: SuperScalar Concepts

55

Complexity of Bypass Logic



Chapter 3: SuperScalar Concepts

56

Delay Analysis of Bypass Logic

Delay is largely determined by the time it takes for the driver at the output of each FU to drive the result value on the result wire, which in turn depends on the length of the result wires.

$$T_{\text{bypass}} = 0.5 \times R_{\text{metal}} \times C_{\text{metal}} \times L^2$$

where

L is the length of result wires

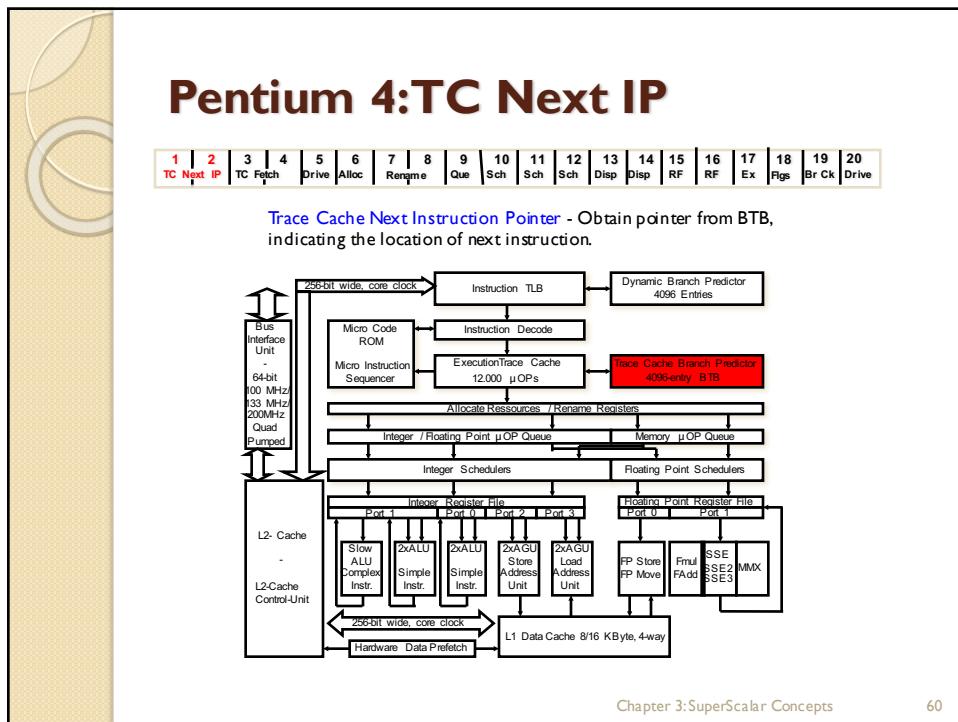
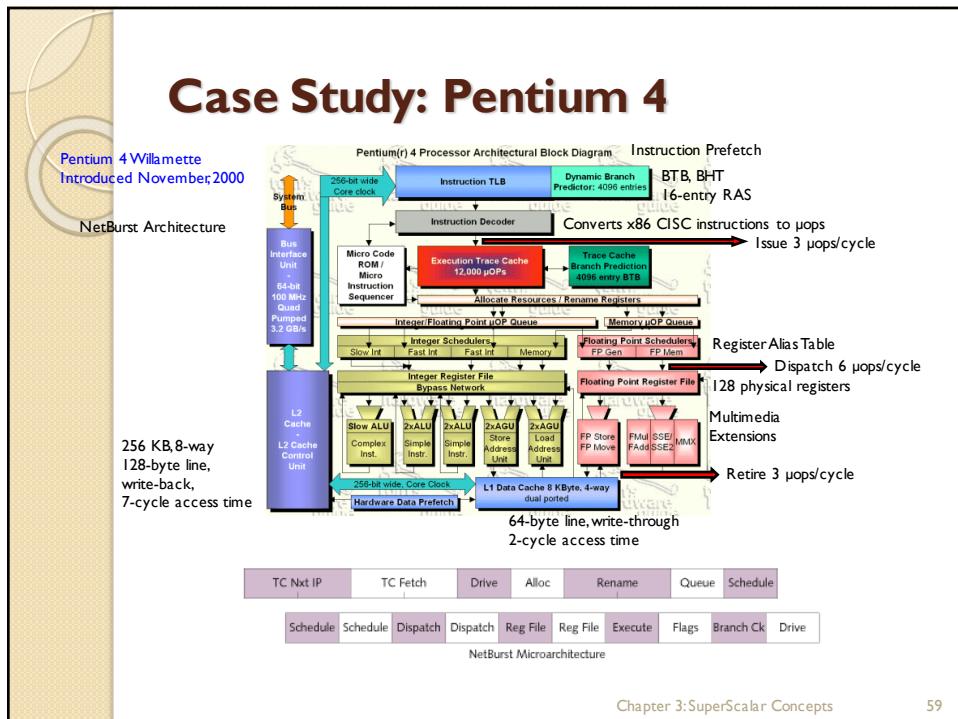
R_{metal} is the resistance per unit length

C_{metal} is the capacitance per unit length

Effect:

- Increasing IW increases # of FUs, which increases L (distance from RF to FUs).
- Increasing IW also increase the fan-in of the operand MUXes connected to a given result wire, which increases the capacitance of the result wires.

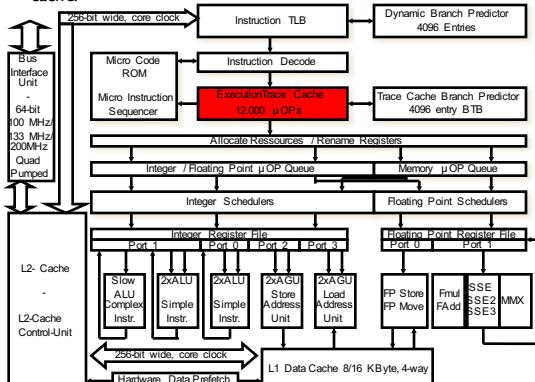
Case Study: Pentium 4



Pentium 4: TC Fetch

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
TC Next IP	TC Fetch	Drive	Alloc	Rename	Que	Sch	Sch	Sch	Disp	Disp	RF	RF	Ex	Fgs	Br Ck	Drive			

Trace Cache Fetch - Read the decoded instructions (μ OPs) out of the execution trace cache.



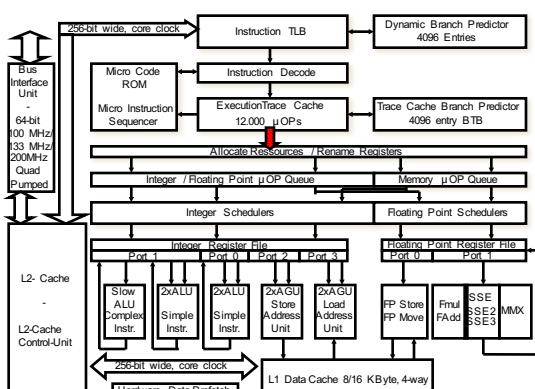
Chapter 3: SuperScalar Concepts

61

Pentium 4: Drive

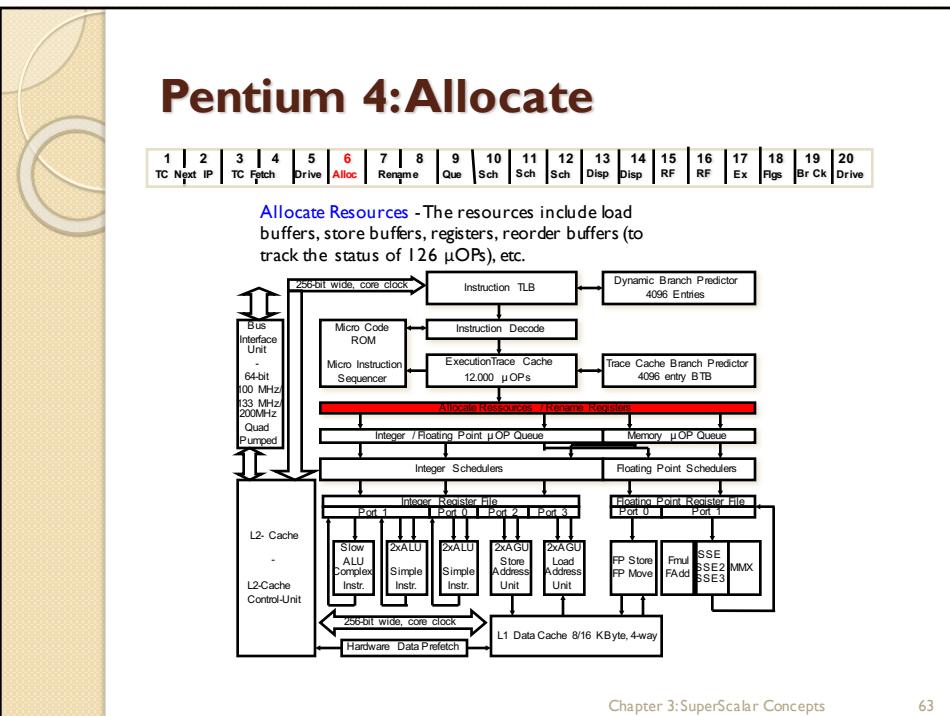
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
TC Next IP	TC Fetch	Drive	Alloc	Rename	Que	Sch	Sch	Sch	Disp	Disp	RF	RF	Ex	Fgs	Br Ck	Drive			

Wire Delay - Drive the μ OPs to the allocator.



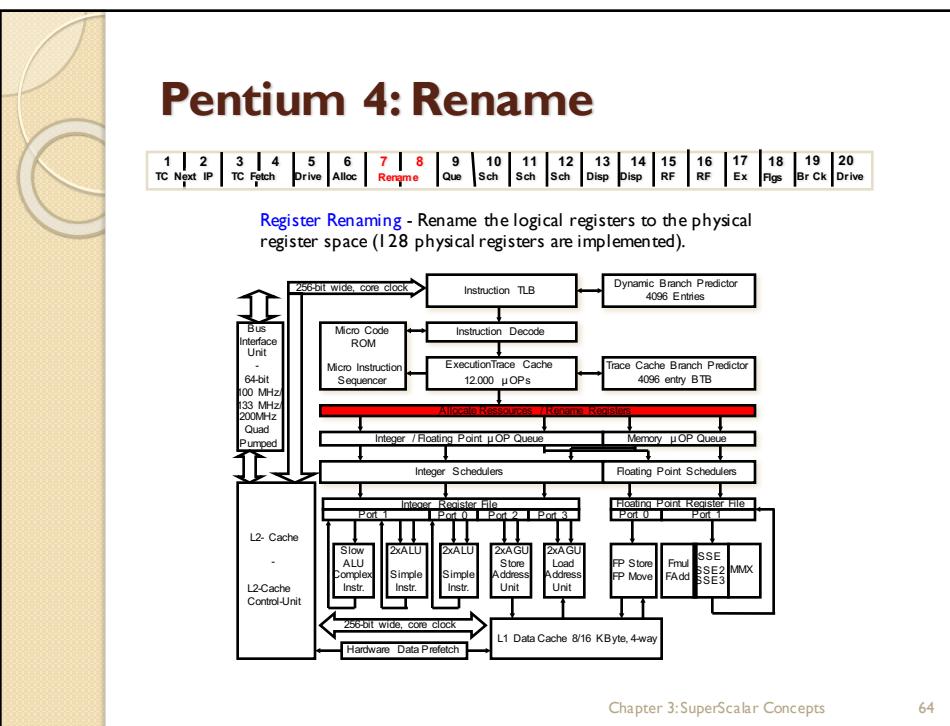
Chapter 3: SuperScalar Concepts

62



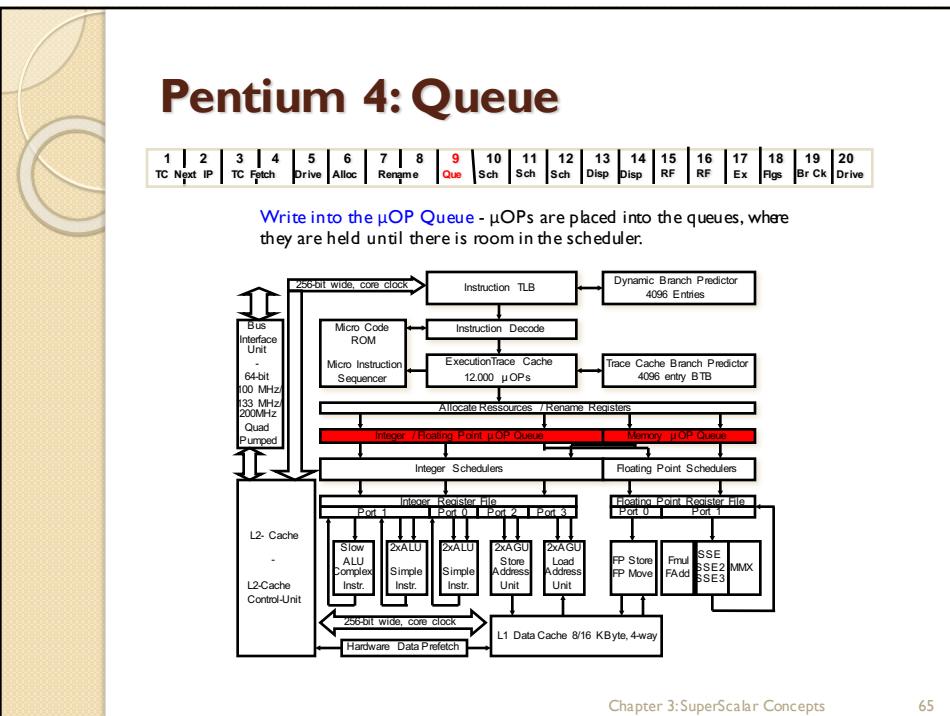
Chapter 3: SuperScalar Concepts

63



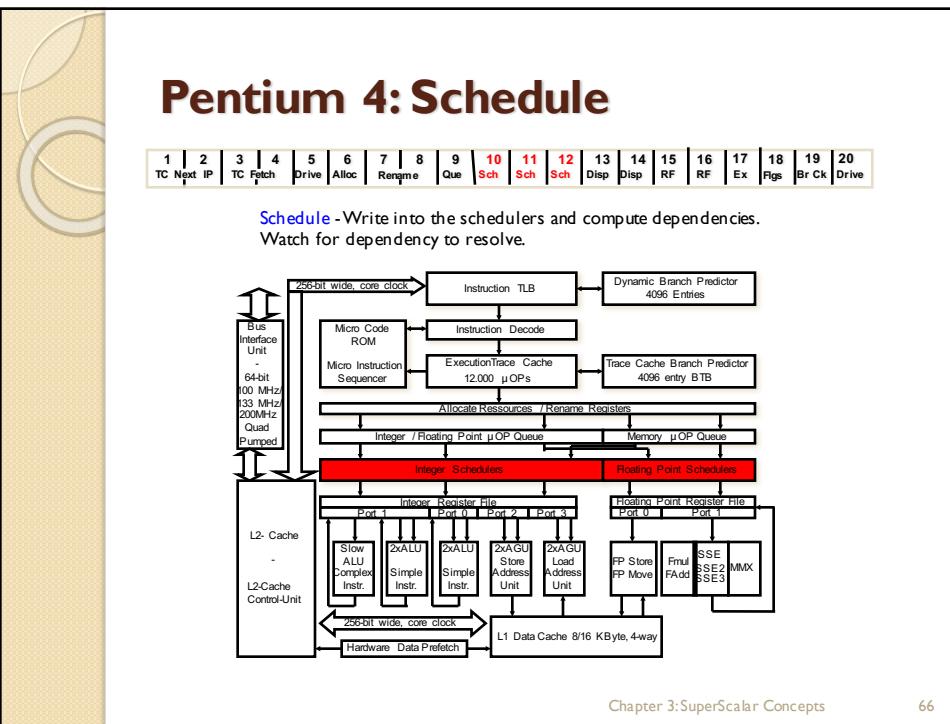
Chapter 3: SuperScalar Concepts

64



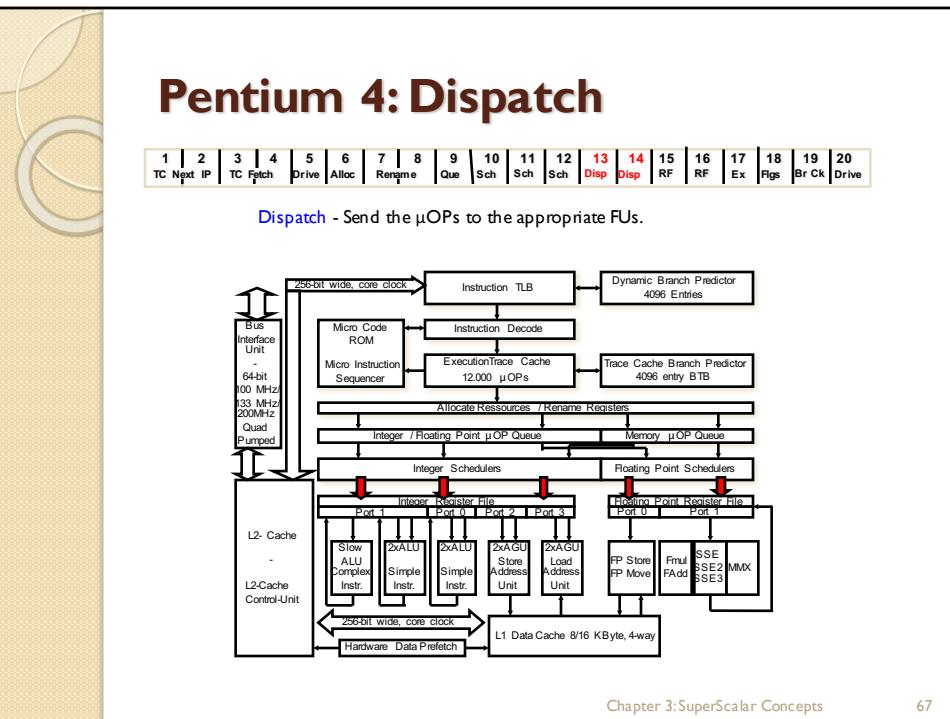
Chapter 3: SuperScalar Concepts

65



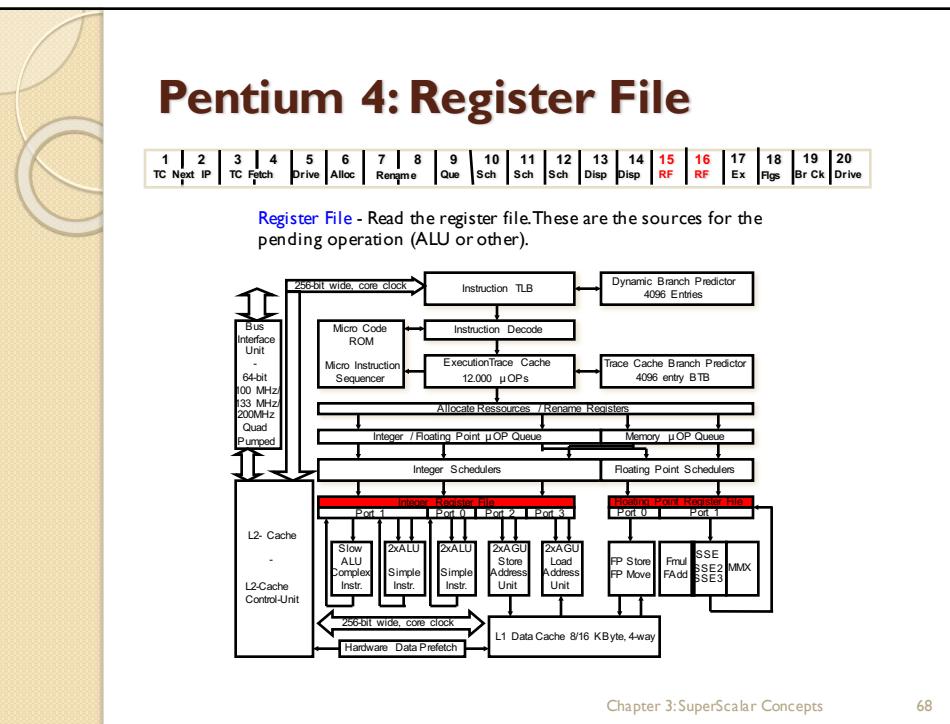
Chapter 3: SuperScalar Concepts

66



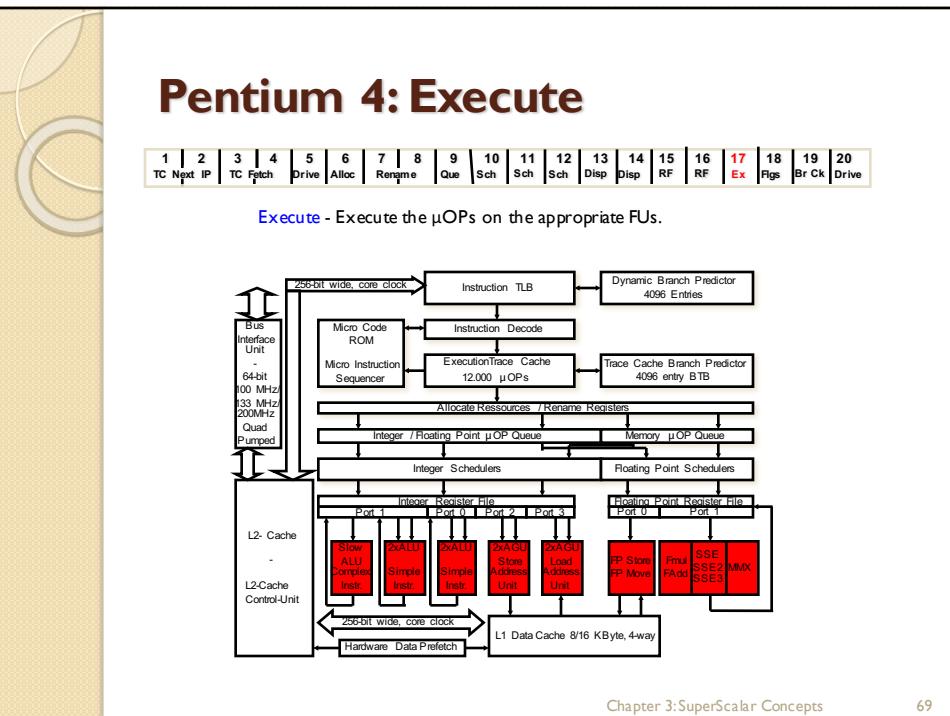
Chapter 3: SuperScalar Concepts

67



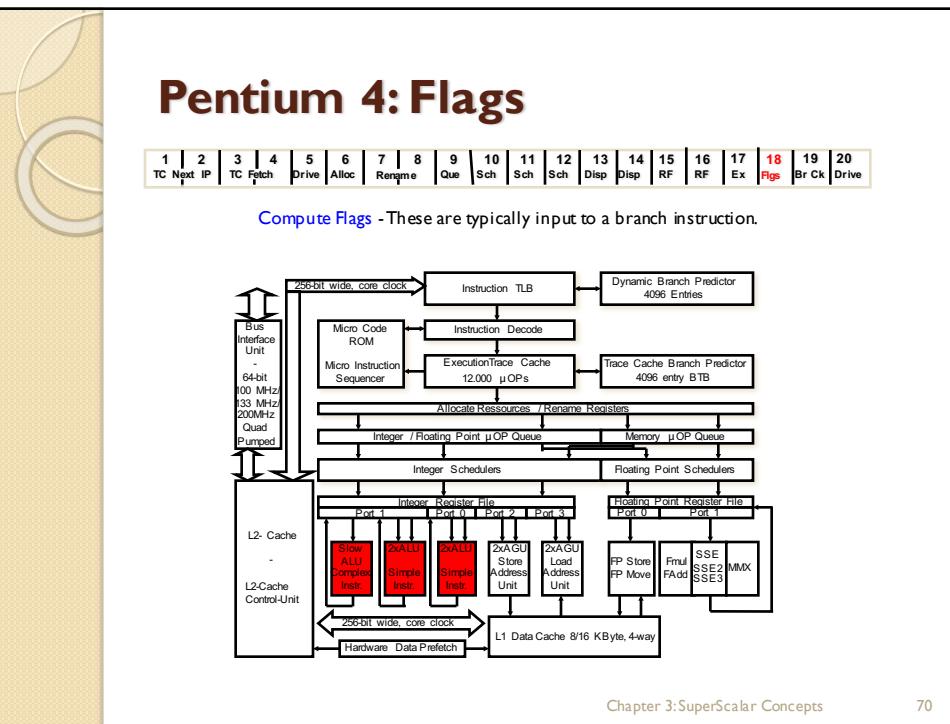
Chapter 3: SuperScalar Concepts

68



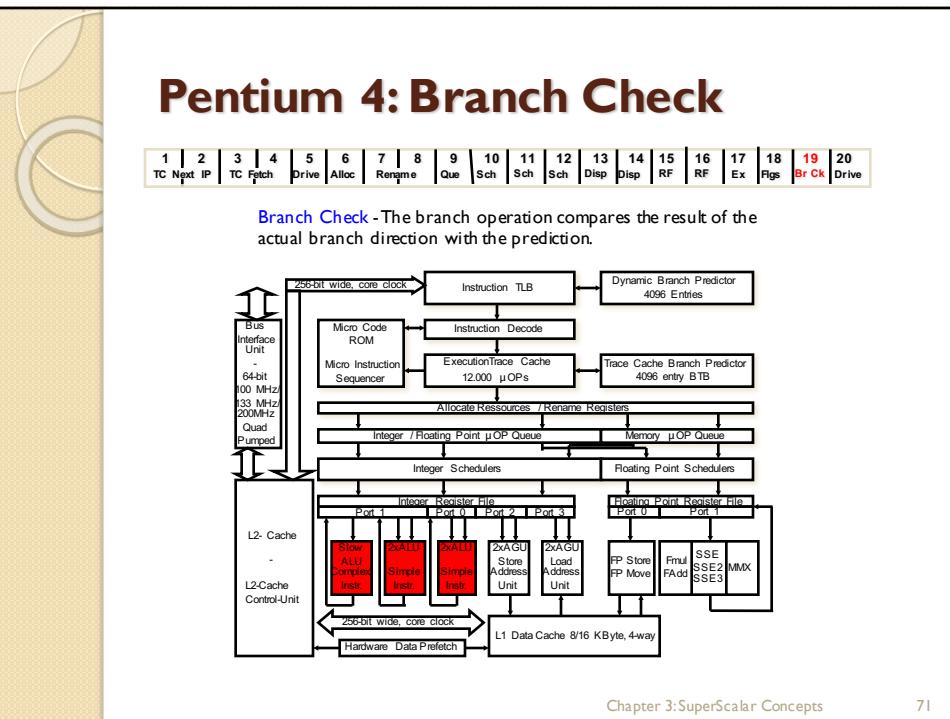
Chapter 3: SuperScalar Concepts

69



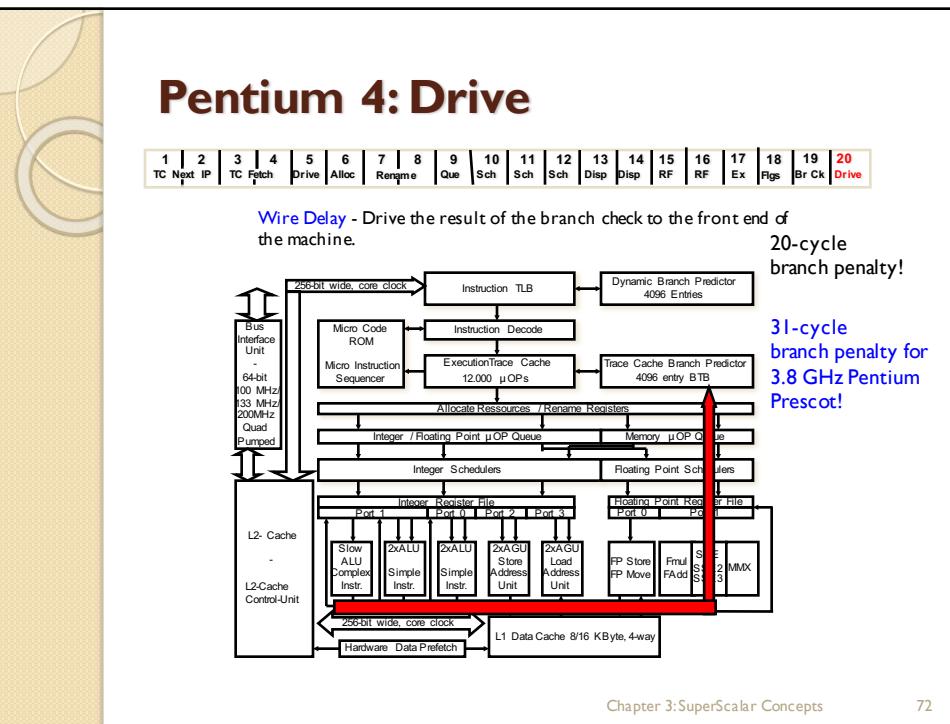
Chapter 3: SuperScalar Concepts

70



Chapter 3: SuperScalar Concepts

71



Chapter 3: SuperScalar Concepts

72

• The Core™ Microarchitecture

Chapter 3: SuperScalar Concepts

73

Evolution of Intel Architectures

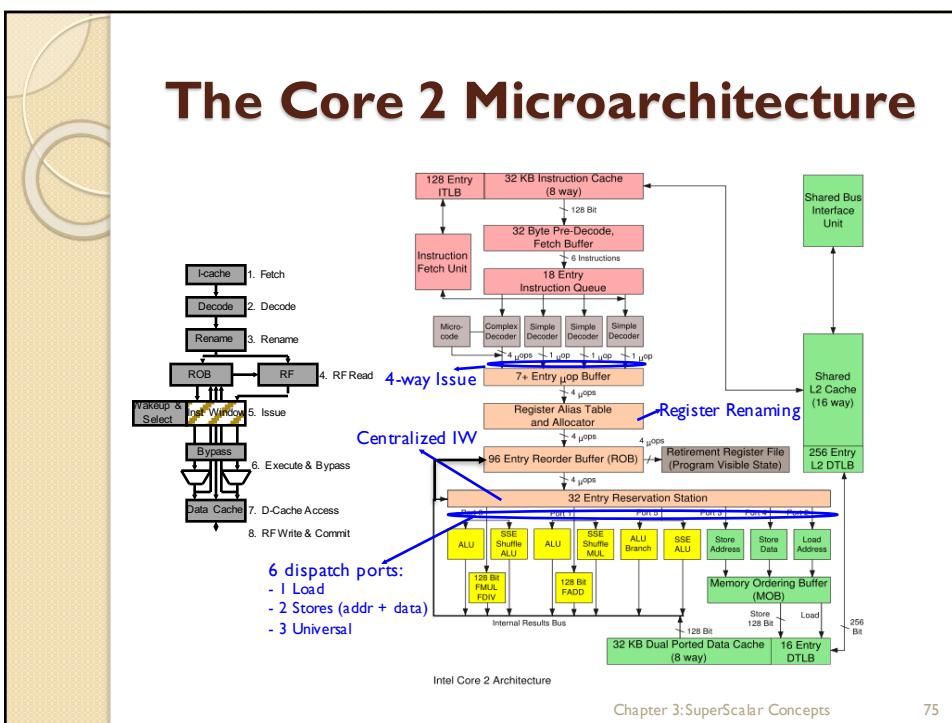
	Pre-multicore era 2000 (180 - 65 nm)	2006 (65 - 45 nm)	2008 (45 nm)	2011 (32 nm)	2013 (22 nm)
Microarchitecture Brand Name	NetBurst™	Intel Core™ Microarchitecture			
Microarchitecture Codename	Willamette, Prescott, etc.	Yonah, Merom, etc.	Nehalem	Sandy Bridge	Haswell
Processor Brand Name	Pentium 4	Core Core Duo Core 2 Duo	Core i7, i5, i3 (Quad-core) Xeon E7, E5, E3 (8-core)		
	20- & 31-stage 6 dispatch ports 2-level cache Hyperthreading	14-stage 6 dispatch ports 2-level cache SSSE3	14-stage 6 dispatch ports 3-level cache Hyperthreading	+ 8 dispatch ports Energy efficiency Intel AVX Integrated memory controller QPI (replaces FSB)	

Supplemental Streaming SIMD Extension 3 (SSSE3)
Advanced Vector Extensions (AVX)
QuickPath Interconnect (QPI)

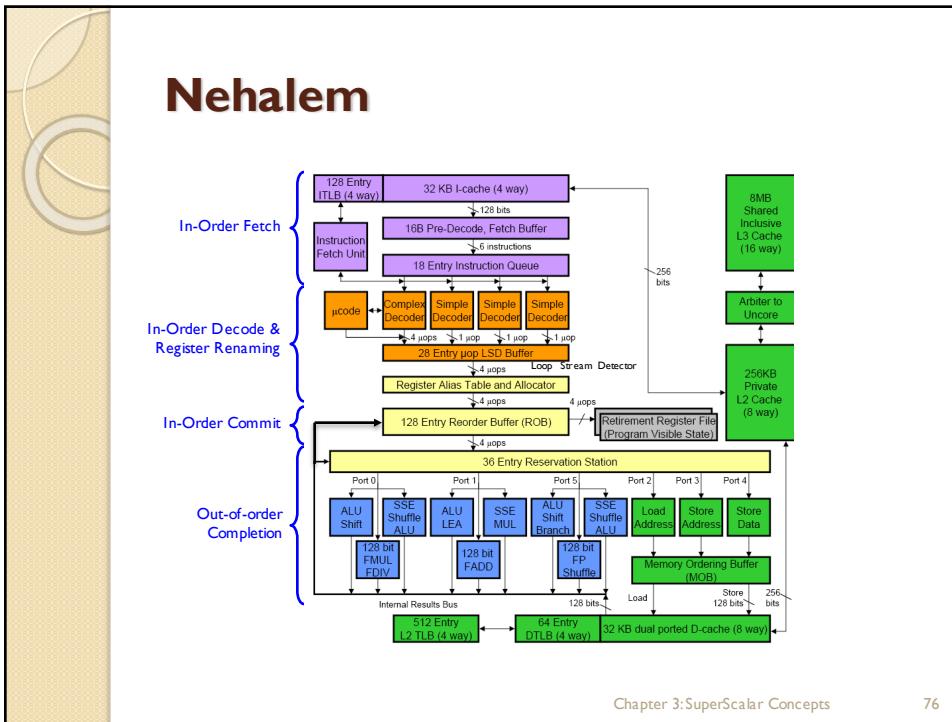
Chapter 3: SuperScalar Concepts

74

The Core 2 Microarchitecture



Nehalem



Haswell

- Evolutionary improvements to Nehalem:
 - Power efficiency
 - AVX2
 - Integrated graphics unit
 - Larger L2 cache
 - Level 3 cache
 - 8 Dispatch ports (2 more than previous models)
 - More buffers

Buffer Size Comparison

	Nehalem	Sandy Bridge	Haswell
ROB Entries	128	168	192
Load Buffer Entries	48	64	72
Store Buffer Entries	32	36	42
RS Entries	36	54	60
Integer RRF	N/A	160	168
FP RRF	N/A	144	168
μ op Decode Queue	28/thread	28/thread	56

Questions?

