



Krste Asanovic  
May 2, 2001  
6.823, L21-1

## Microprocessor Evolution: 4004 to Pentium Pro

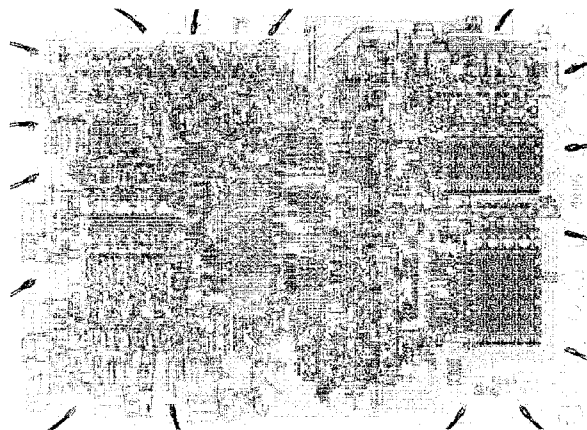
Krste Asanovic  
Laboratory for Computer Science  
M.I.T.

<http://www.csg.lcs.mit.edu/6.823>



Krste Asanovic  
May 2, 2001  
6.823, L21-2

## First Microprocessor Intel 4004, 1971



- 4-bit accumulator architecture
- 8 $\mu$ m pMOS
- 2,300 transistors
- 3 x 4 mm<sup>2</sup>
- 750kHz clock
- 8-16 cycles/inst.



## Microprocessors in the Seventies

### Initial target was embedded control

- First micro, 4-bit 4004 from Intel, designed for a desktop printing calculator

### Constrained by what could fit on single chip

- Single accumulator architectures

### 8-bit micros used in hobbyist personal computers

- Micral, Altair, TRS-80, Apple-II

### Little impact on conventional computer market until VISICALC spreadsheet for Apple-II (6502, 1MHz)

- First “killer” business application for personal computers



## DRAM in the Seventies

### Dramatic progress in MOSFET memory technology

1970, Intel introduces first DRAM (1Kbit 1103)

1979, Fujitsu introduces 64Kbit DRAM

=> By mid-Seventies, obvious that PCs would soon have > 64KBytes physical memory



## Microprocessor Evolution

### Rapid progress in size and speed through 70s

- Fueled by advances in MOSFET technology and expanding markets

#### Intel i432

- Most ambitious seventies' micro; started in 1975 - released 1981
- 32-bit capability-based object-oriented architecture
- Instructions variable number of *bits* long
- Severe performance, complexity, and usability problems

#### Intel 8086 (1978, 8MHz, 29,000 transistors)

- "Stopgap" 16-bit processor, architected in 10 weeks
- Extended accumulator architecture, assembly-compatible with 8080
- 20-bit addressing through segmented addressing scheme

#### Motorola 68000 (1979, 8MHz, 68,000 transistors)

- First production microprocessor with microcode
- 32-bit general purpose register architecture (24 address pins)
- 8 address registers, 8 data registers



## Intel 8086

<u>Class</u>	<u>Register</u>	<u>Purpose</u>
Data:	AX,BX	"general" purpose
	CX	string and loop ops only
	DX	mult/div and I/O only
Address:	SP	stack pointer
	BP	base pointer (can also use BX)
	SI,DI	index registers
Segment:	CS	code segment
	SS	stack segment
	DS	data segment
	ES	extra segment
Control:	IP	instruction pointer (lower 16 bit of PC)
	FLAGS	C, Z, N, B, P, V and 3 control bits

- Typical format  $R \leftarrow R \text{ op } M[X]$ , many addressing modes
- **Not a GPR organization!**



## IBM PC, 1981

Krste Asanovic  
May 2, 2001  
6.823, L21-7

### Hardware

- Team from IBM building PC prototypes in 1979
- Motorola 68000 chosen initially, but 68000 was late
- IBM builds “stopgap” prototypes using 8088 boards from Display Writer word processor
- 8088 is 8-bit bus version of 8086 => allows cheaper system
- Estimated sales of 25,000
- 100,000,000s sold

### Software

- Microsoft negotiates to provide OS for IBM. Later buys and modifies QDOS from Seattle Computer Products.

### Open System

- Standard processor, Intel 8088
- Standard interfaces
- Standard OS, MS-DOS
- IBM permits cloning and third-party software



## The Eighties: Microprocessor Revolution

Krste Asanovic  
May 2, 2001  
6.823, L21-8

### Personal computer market emerges

- Huge business and consumer market for spreadsheets, word processing and games
- Based on inexpensive 8-bit and 16-bit micros: Zilog Z80, Mostek 6502, Intel 8088/86, ...

### Minicomputers replaced by workstations

- Distributed network computing and high-performance graphics for scientific and engineering applications (Sun, Apollo, HP,...)
- Based on powerful 32-bit microprocessors with virtual memory, caches, pipelined execution, hardware floating-point

### Massively Parallel Processors (MPPs) appear

- Use many cheap micros to approach supercomputer performance (Sequent, Intel, Parsytec)



## The Nineties

**Distinction between workstation and PC disappears**

**Parallel microprocessor-based SMPs take over low-end server and supercomputer market**

**MPPs have limited success in supercomputing market**

**High-end mainframes and vector supercomputers survive “killer micro” onslaught**

**64-bit addressing becomes essential at high-end**

- In 2001, 4GB DRAM costs <\$5,000

**CISC ISA (x86) thrives!**



## Reduced ISA Diversity in Nineties

**Few major companies in general-purpose market**

- Intel x86 (CISC)
- IBM 390 (CISC)
- Sun SPARC, SGI MIPS, HP PA-RISC (all RISCs)
- IBM/Apple/Motorola introduce PowerPC (another RISC)
- Digital introduces Alpha (another RISC)

**Software costs make ISA change prohibitively expensive**

- 64-bit addressing extensions added to RISC instruction sets
- Short vector multimedia extensions added to all ISAs, but without compiler support

**=> Focus on *microarchitecture* (superscalar, out-of-order)**

**CISC x86 thrives!**

- RISCs (SPARC, MIPS, Alpha, PowerPC) fail to make significant inroads into desktop market, but important in server and technical computing markets

***“RISC advantage” shrinks with superscalar out-of-order execution***



## Intel Pentium Pro, (1995)

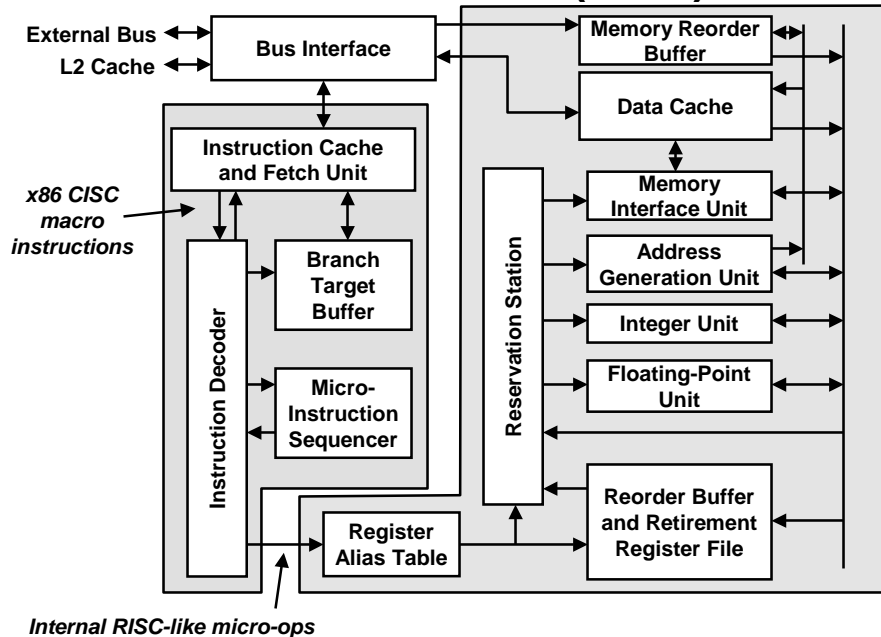
Krste Asanovic  
May 2, 2001  
6.823, L21-11

- During decode, translate complex x86 instructions into RISC-like micro-operations (uops)
  - e.g., “ $R \leftarrow R \text{ op Mem}$ ” translates into
    - load T, Mem # Load from Mem into temp reg
    - $R \leftarrow R \text{ op T}$  # Operate using value in temp
- Execute uops using speculative out-of-order superscalar engine with register renaming
- Pentium Pro family architecture (P6 family) used on Pentium-II and Pentium-III processors



## Intel Pentium Pro (1995)

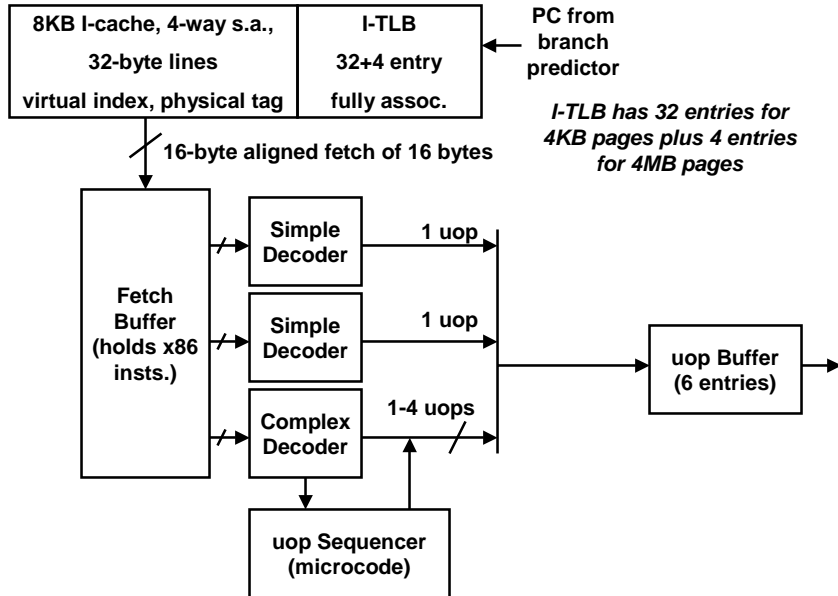
Krste Asanovic  
May 2, 2001  
6.823, L21-12





## P6 Instruction Fetch & Decode

Krste Asanovic  
May 2, 2001  
6.823, L21-13



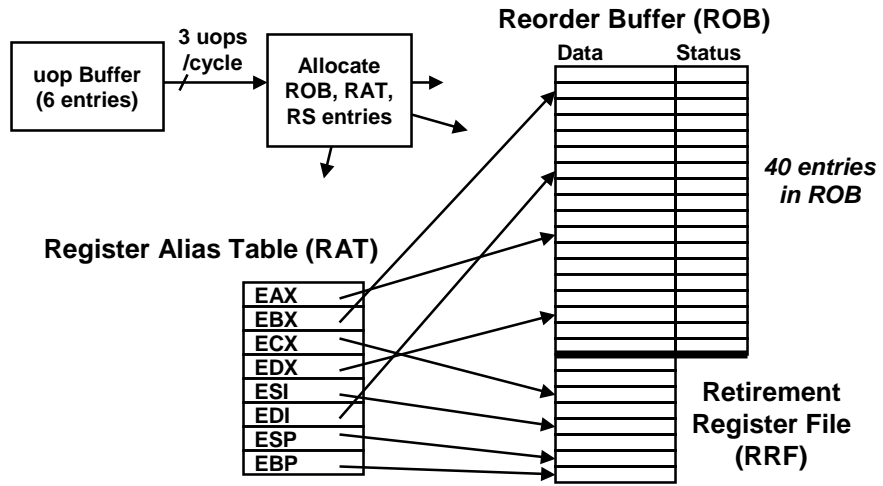
## P6 uops

Krste Asanovic  
May 2, 2001  
6.823, L21-14

- **Each uop has fixed format of around 118 bits**
  - opcode, two sources, and destination
  - sources and destination fields are 32-bits wide to hold immediate or operand
- **Simple decoders can only handle simple x86 instructions that map to one uop**
- **Complex decoder can handle x86 translations of up to 4 uops**
- **Complicated x86 instructions handled by microcode engine that generates uop sequence**
- **Intel data shows average of 1.2-1.7 uops per x86 instruction on SPEC95 benchmarks, 1.4-2.0 on MS Office applications**



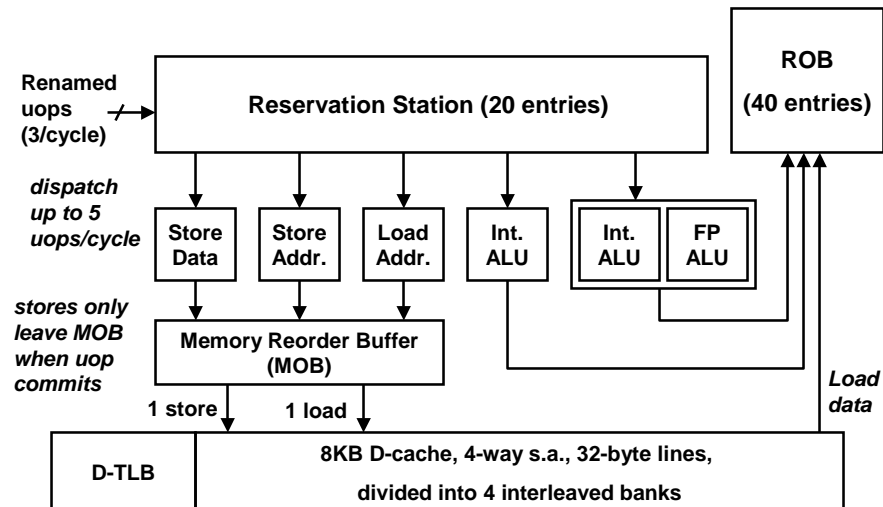
## P6 Reorder Buffer and Renaming



Values move from ROB to architectural register file (RRF) when committed



## P6 Reservation Stations and Execution Units



D-TLB has 64 entries for 4KB pages fully assoc., plus 8 entries for 4MB pages, 4-way s.a.



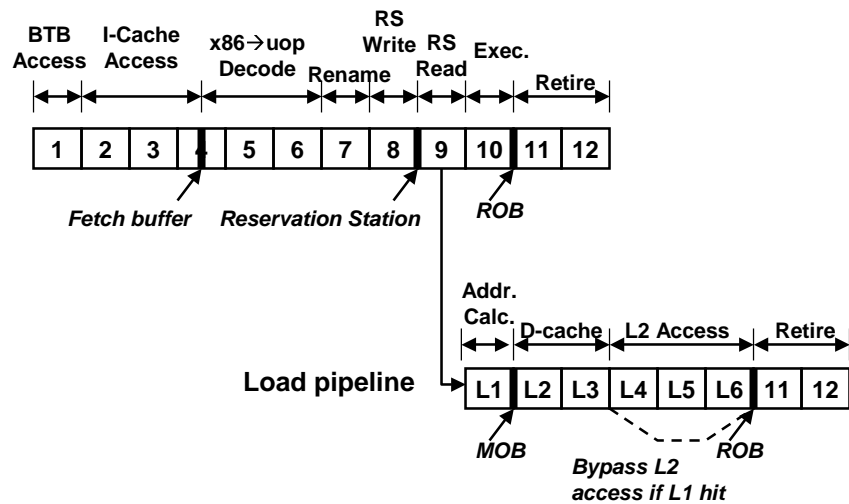


## P6 Retirement

- After uop writes back to ROB with no outstanding exceptions or mispredicts, becomes eligible for retirement
- Data written to RRF from ROB
- ROB entry freed, RAT updated
- uops retired in order, up to 3 per cycle
- Have to check and report exceptions at valid x86 instruction fault points
  - complex instructions (e.g., string move) may generate thousands of uops

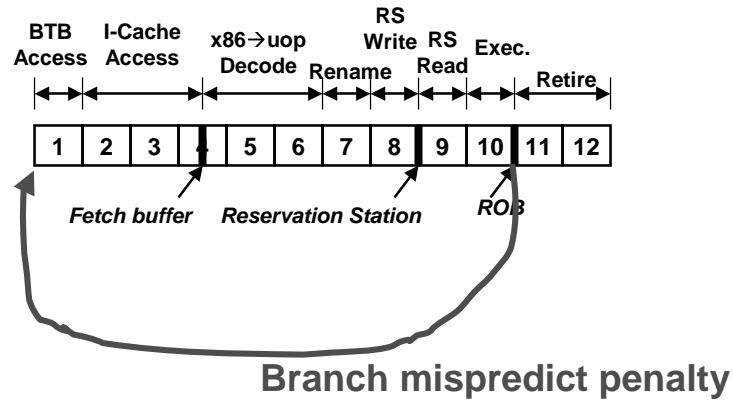


## P6 Pipeline





## P6 Branch Penalties



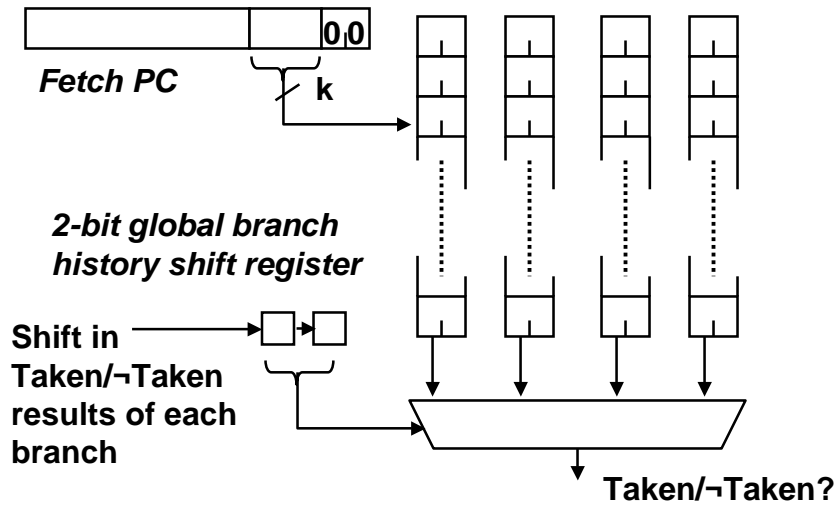
## P6 Branch Target Buffer (BTB)

- 512 entries, 4-way set-associative
- Holds branch target, plus two-level BHT for taken/not-taken
- Unconditional jumps not held in BTB
- One cycle bubble on correctly predicted taken branches (no penalty if correctly predicted not-taken)



## Two-Level Branch Predictor

*Pentium Pro uses the result from the last two branches to select one of the four sets of BHT bits (~90-95% correct)*

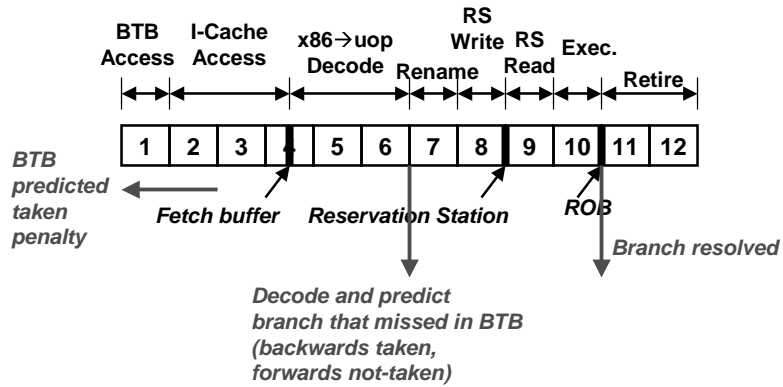


## P6 Static Branch Prediction

- If a branch misses in BTB, then static prediction performed
- Backwards branch predicted taken, forwards branch predicted not-taken



## P6 Branch Penalties



## P6 System

