COMPUTER ARCHITECURE

Introduction To The Course

Adapted from Patterson material

Course Structure

These notes, and all other notes for the course can also be found at this same location.

Also from this page you can go to the course as taught last year. It will be similar, but wait for the updated slides to appear on this year's version of the course.

So what's in it for me? (1 / 2)

- **In-depth understanding of the inner-workings of modern computers, their evolution, and trade-offs present at the hardware/software boundary.**
  - discussion of fast/slow operations and why they are easy/hard to implement in hardware
  - ``out of order execution'', branch prediction and other techniques to increase the average execution rate of instructions
- **Experience with the *design process* in the context of a large complex (hardware) design.**
  - Functional Spec → Control & Datapath → Physical implementation
  - Modern CAD tools (later on)

So what's in it for me? (2 / 2)

- **Teach Computer Architecture from a software developer's point of view**
  - Starts with what the programmer writes, how it is *translated* to machine language, how the machine *interprets* that program
  - Performance impact of s/w and h/w designs; what makes programs run slowly, what h/w features can speed up the execution of programs.
- **Learn perennial ideas in computer science and engineering**
  - Principle of abstraction, used to study and build systems as layers
  - 5 classic components of any computer
  - Stored program concept: instructions and data stored in memory
  - Raw data (binary bit patterns) can mean anything (integers, floating point numbers, chars, etc): a *program* determines what it is
  - Principle of Locality, exploited via a *memory hierarchy*
  - Compilation *vs*. interpretation to move down the layers of the computer system

What is "Computer Architecture"

Computer Architecture =

Instruction Set Architecture (ISA) +

Machine Organization (MO)

ISA \ Definition of What the Machine Does, *Logical View*

MO \ How Machine Implements ISA, *Physical Implementation*

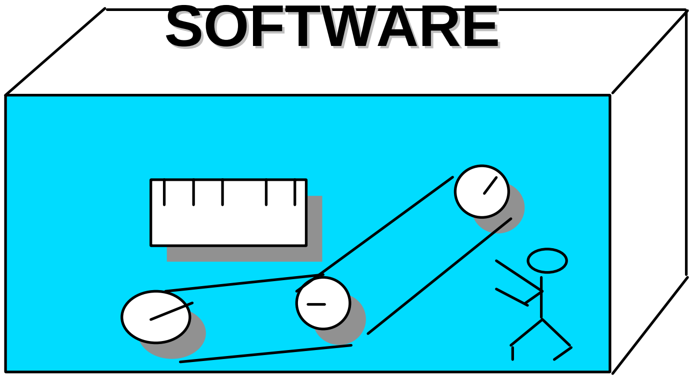
In this course we will study both

Instruction Set Architecture (subset of Computer Arch.)

"... the attributes of a [computing] system as seen by the programmer, *i.e.,* the conceptual structure and functional behaviour, as distinct from the organization of the data flows and controls the logic design, and the physical implementation. "

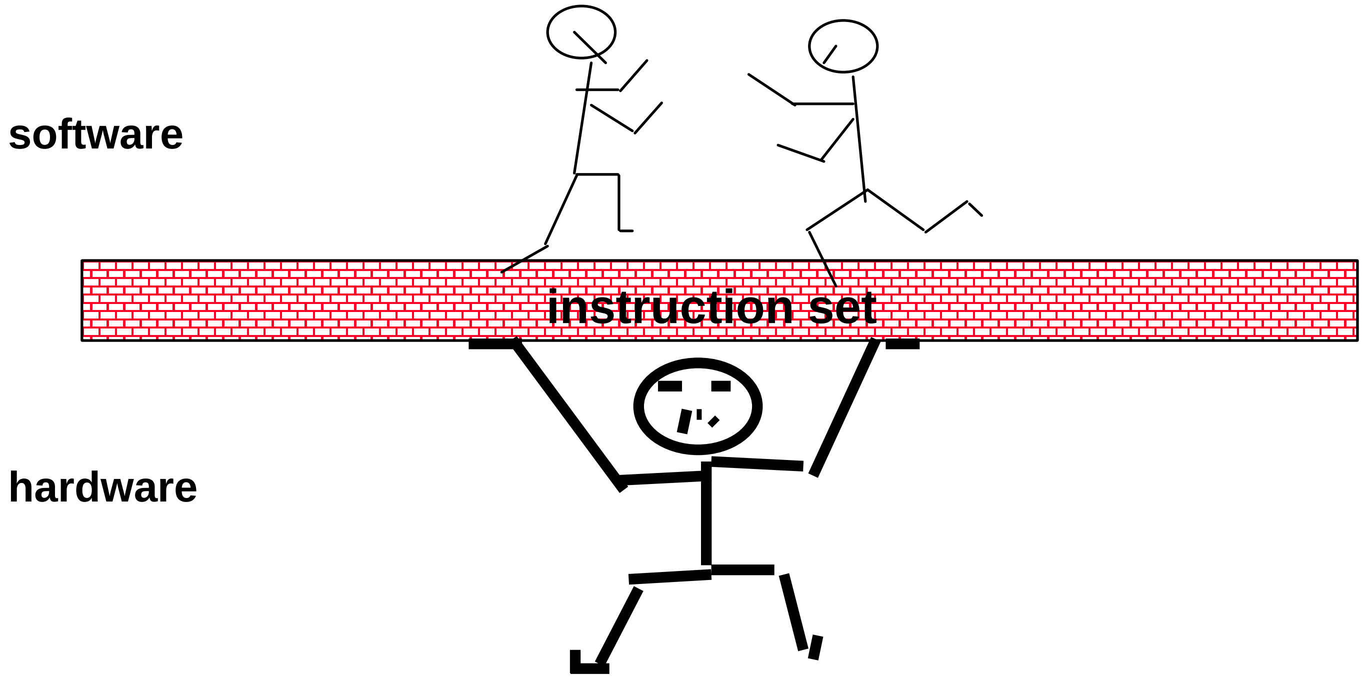
Amdahl, Blaaw, and Brooks, 1964

An ISA encompasses:



- **-- Organization of Programmable Storage (registers, memory)**
- **-- Data Types & Data Structures: Encodings & Representations**
- **-- Instruction Set**
- **-- Instruction Formats**
- **-- Modes of Addressing and Accessing Data Items and Instructions**
- **-- Exceptional Conditions and their Handling Modes**

The Instruction Set:
a (the?) *Critical* Interface



Example ISAs (Instruction Set Architectures)

• Digital Alpha	(v1, v3)	1992-97
• HP PA-RISC	(v1.1, v2.0)	1986-01
• Sun Sparc	(v8, v9, v10, v11)	1987-01
• SGI MIPS	(MIPS I, II, III, IV, V)	1986-01
• Intel	(8086,80286, 80486,Pentium, Pentium)	1978-01
• Intel + HP	EPIC	1998-01

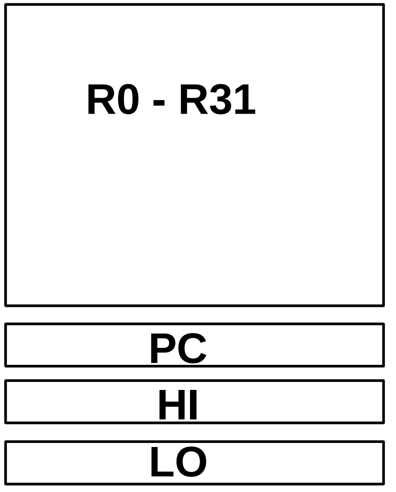
MIPS R3000
Instruction Set Architecture (Summary)

- **Instruction Categories**

  Load/Store
  Computational
  Jump and Branch

  Floating Point - coprocessor
  Memory Management
  Special

Registers



3 Instruction Formats: all 32 bits wide

OP	Rs	Rt	rd	sa	funct

OP	Rs	Rt	Immediate

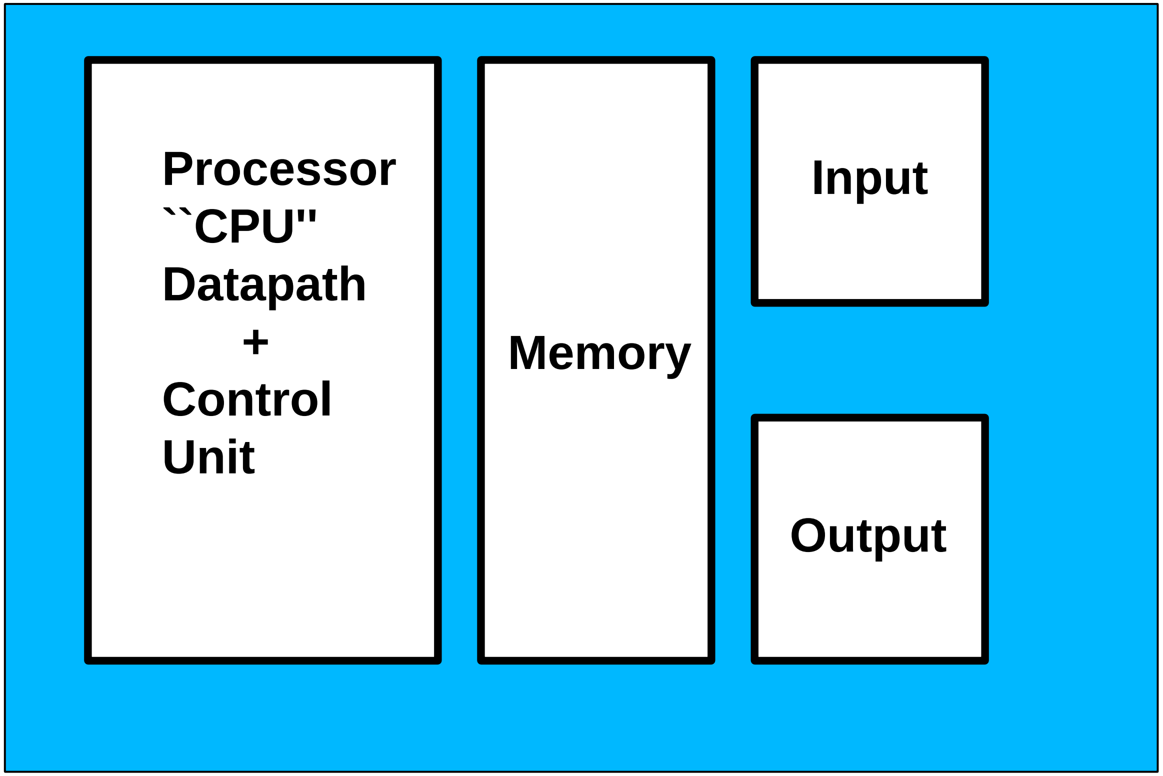
OP	jump target

Impact of changing an ISA

- **Early 1990's Apple switched instruction set architecture of the Macintosh**
  - From Motorola 68000-based machines
  - To PowerPC architecture
  - Upside? Downside?
- **Intel 80x86 Family: many implementations of same architecture**
  - Upside: program written in 1978 for 8086 can be run on latest Pentium chip
  - Downside?

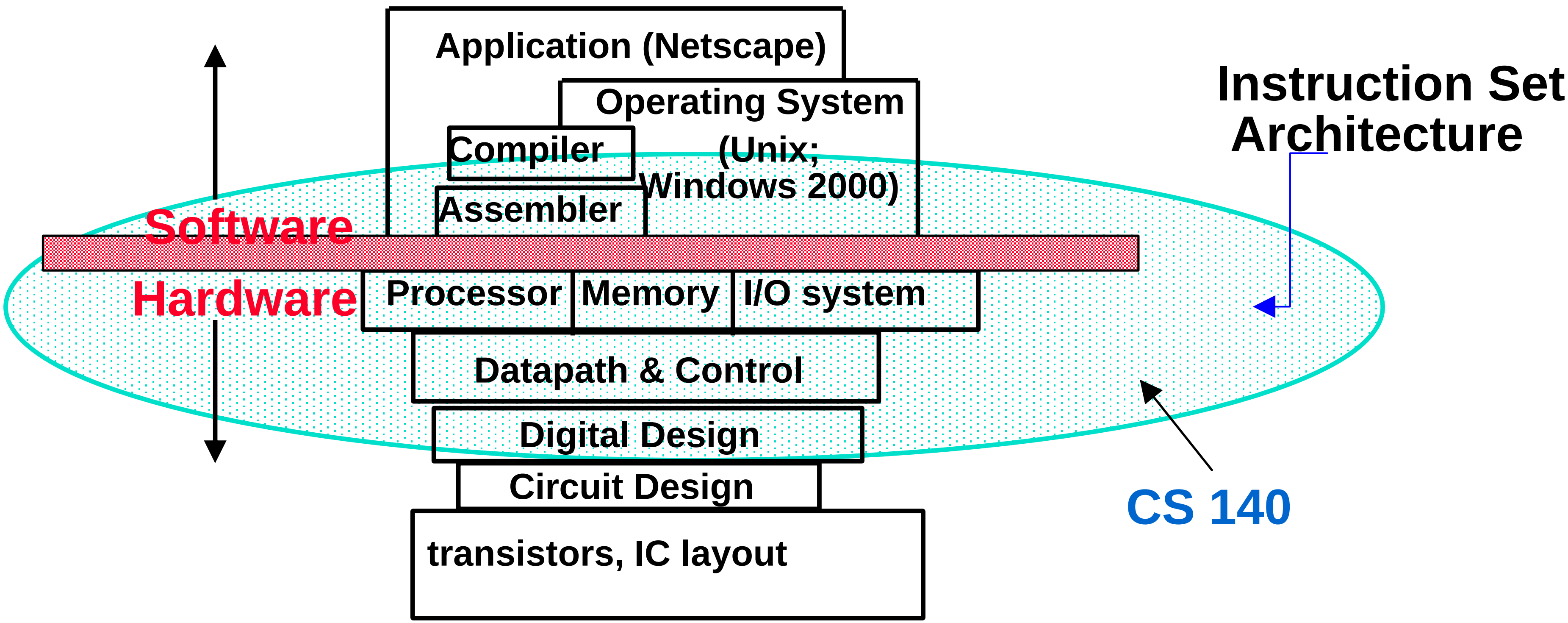
The Big Picture

Since 1946 all computers have had 5 components



Interconnection Structures (buses)

What is ``Computer Architecture''?



- **Co-ordination of many levels of abstraction**
  - hide unnecessary implementation details
  - helps us cope with enormous complexity of real systems
- **Under a rapidly changing set of forces**
- **Design, Measurement, *and* Evaluation**

Forces Acting on Computer Architecture

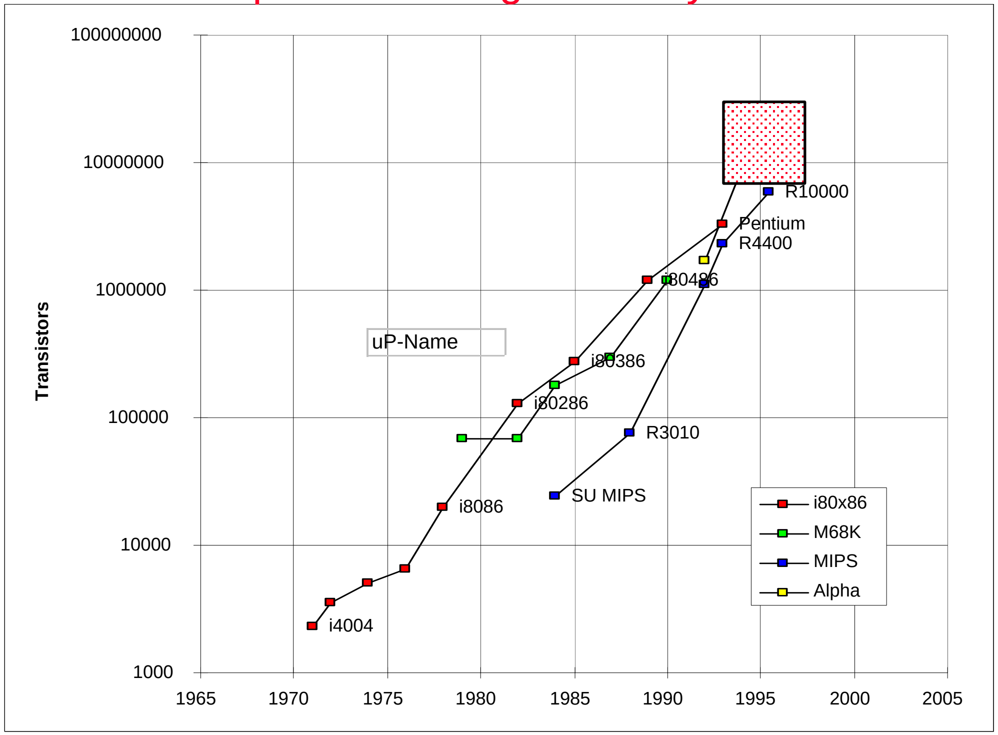
- **R-a-p-i-d Improvement in Implementation Technology:**
  - IC: integrated circuit; invented 1959
  - SSI $\rightarrow$ MSI $\rightarrow$ LSI $\rightarrow$ VLSI: dramatic growth in number transistors/chip $\Rightarrow$ ability to create more (and bigger) Functional Units per processor;
  - bigger memory $\Rightarrow$ more sophisticated applications, larger databases
  - Ubiquitous computing
- **Tomorrow's Science Fiction:**
  - Computers embedded everywhere;
  - Autopilot on your car.
  - Unbreakable encryption on your DVDs
- **New Languages: Java, C++ ...**

Technology Trends

DRAM chip capacity

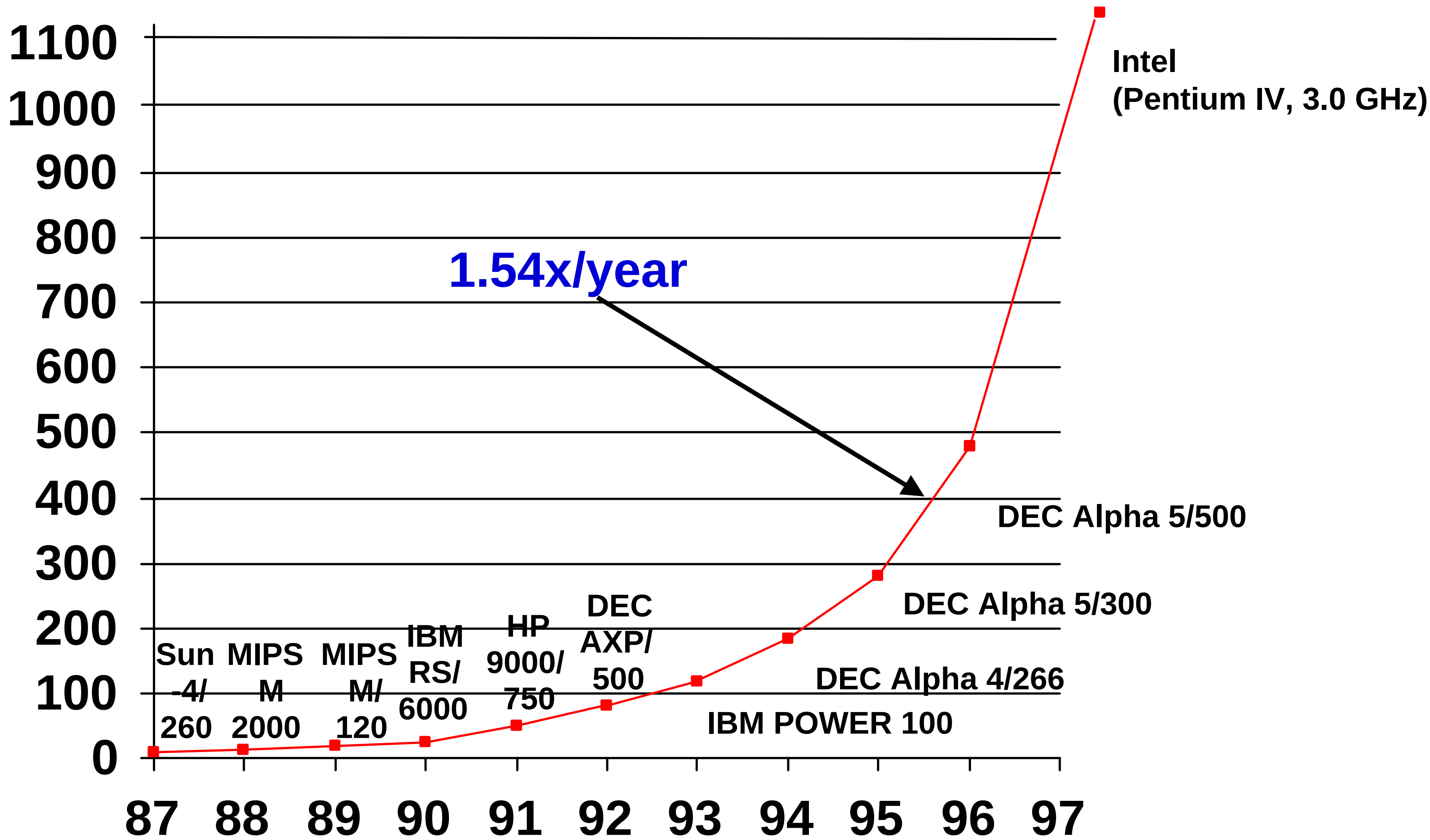
DRAM	
Year	**Size**
1980	64 Kb
1983	256 Kb
1986	1 Mb
1989	4 Mb
1992	16 Mb
1996	64 Mb
1999	256 Mb
2002	1 Gb

Microprocessor Logic Density



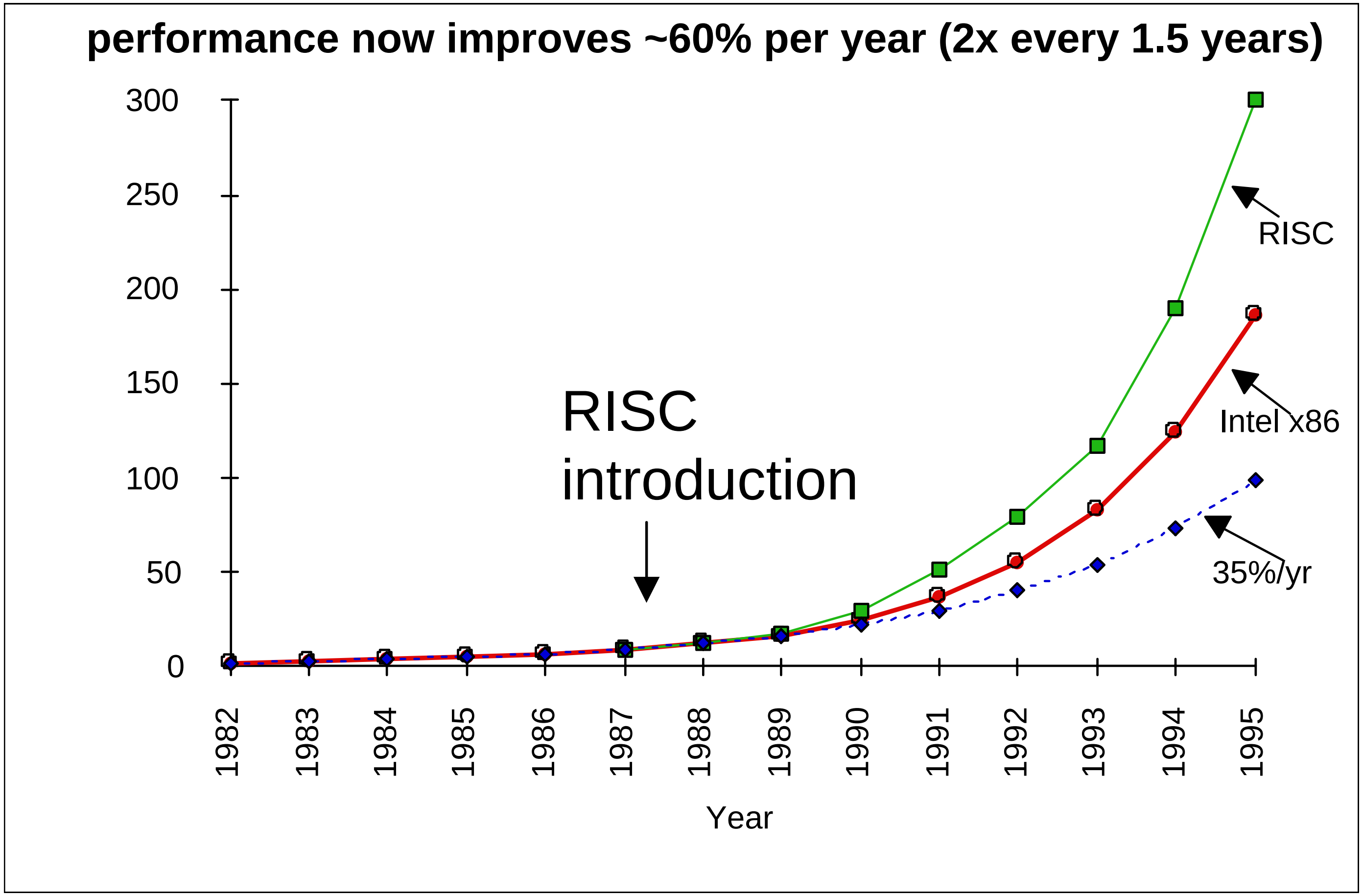
- **In ~1985 the single-chip processor (32-bit) and the single-board computer emerged**
  - Workstations, personal computers, multiprocessors have been riding this wave since
- **In the 2003+ timeframe, these look like mainframes compared to single-chip computer (maybe 2 chips)**

Trends: Processor Performance



- Performance with respect to performance of VAX-11/780

Processor Performance (SPEC)



Did RISC win the technology battle and lose the market war?

OLD PICTURE – BUT THE STORY IS THE SAME

Processor Performance - Capacities

Table 2-2. Key Features of Previous Generations of IA-32 Processors

Intel Processor	Date Intro-duced	Max. Clock Frequency at Intro-duction	Transis-tors per Die	Register Sizes[1]	Ext. Data Bus Size[2]	Max. Extern. Addr. Space	Caches
8086	1978	8 MHz	29 K	16 GP	16	1 MB	None
Intel 286	1982	12.5 MHz	134 K	16 GP	16	16 MB	Note 3
Intel386 DX Processor	1985	20 MHz	275 K	32 GP	32	4 GB	Note 3
Intel486 DX Processor	1989	25 MHz	1.2 M	32 GP 80 FPU	32	4 GB	L1: 8KB
Pentium Processor	1993	60 MHz	3.1 M	32 GP 80 FPU	64	4 GB	L1:16KB
Pentium Pro Processor	1995	200 MHz	5.5 M	32 GP 80 FPU	64	64 GB	L1: 16KB L2: 256KB or 512KB
Pentium II Processor	1997	266 MHz	7 M	32 GP 80 FPU 64 MMX	64	64 GB	L1: 32KB L2: 256KB or 512KB
Pentium III Processor	1999	500 MHz	8.2 M	32 GP 80 FPU 64 MMX 128 XMM	64	64 GB	L1: 32KB L2: 512KB

NOTES:

1. The register size and external data bus size are given in bits. Note also that each 32-bit general-purpose (GP) registers can be addressed as an 8- or a 16-bit data registers in all of the processors
2. Internal data paths that are 2 to 4 times wider than the external data bus for each processor.

Processor Performance - Capacities

Table 2-1. Key Features of Most Recent IA-32 Processors

Intel Processor	Date Intro-duced	Micro-Architecture	Clock Frequency at Intro-duction	Transis-tors Per Die	Register Sizes[1]	System Bus Band-width	Max. Extern. Addr. Space	On-Die Caches[2]
Pentium III and Pentium III Xeon Processors[3]	1999	P6	700 MHz	28 M	GP: 32 FPU: 80 MMX: 64 XMM: 128	Up to 1.06 GB/s	64 GB	32-KB L1; 256-KB L2
Pentium 4 Processor	2000	Intel NetBurst Micro-architecture	1.50 GHz	42 M	GP: 32 FPU: 80 MMX: 64 XMM: 128	3.2 GB/s	64 GB	12K µop Execution Trace Cache; 8KB L1; 256-KB L2
Intel Xeon Processor	2001	Intel NetBurst Micro-architecture	1.70 GHz	42 M	GP: 32 FPU: 80 MMX: 64 XMM: 128	3.2 GB/s	64 GB	12K µop Trace Cache; 8-KB L1; 256-KB L2
Intel Xeon Processor[4]	2002	Intel NetBurst Micro-architecture; Hyper-Threading Technology	2.20 GHz	55 M	GP: 32 FPU: 80 MMX: 64 XMM: 128	3.2 GB/s	64 GB	12K µop Trace Cache; 8-KB L1; 512-KB L2
Intel® Xeon™ Processor MP[4]	2002	Intel NetBurst Micro-architecture; Hyper-Threading Technology	1.60 GHz	108 M	GP: 32 FPU: 80 MMX: 64 XMM: 128	3.2 GB/s	64 GB	12K µop Trace Cache; 8-KB L1; 256-KB L2; 1-MB L3

NOTES

1. The register size and external data bus size are given in bits.
2. First level cache is denoted using the abbreviation L1, 2nd level cache is denoted as L2
3. Intel Pentium III and Pentium III Xeon processors, with advanced transfer cache and built on 0.18 micron process technology, were introduced in October 1999.
4. Hyper-Threading technology is implemented with two logical processors.

Technology --> Dramatic Changes

- **Processor**
  - logic capacity: 2 × in performance every 1.5 years;
  - clock rate: about 30% per year
  - overall performance: 1000 × in last decade
- **Main Memory**
  - DRAM capacity: 2 × / 2 years; 1000 × size in last decade
  - memory speed: about 10% per year
  - cost / bit: improves about 25% per year
- **Disk**
  - capacity: > 2 × in capacity every 1.5 years
  - cost / bit: improves about 60% per year
  - 120 × capacity in last decade
- **Network Bandwidth**
  - Bandwidth: increasing more than 100% per year!

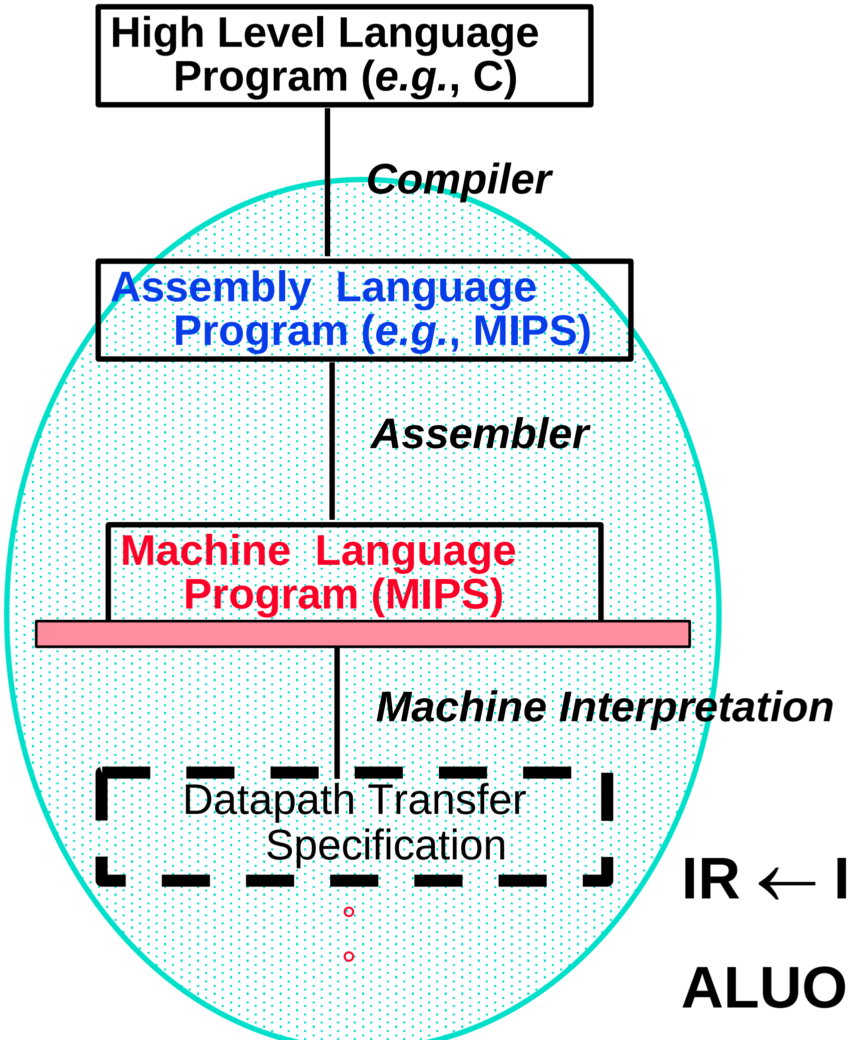
Your PC in 2006

- **State-of-the-art PC(on your desk) three years from now:**
  - Processor clock speed:
    8000 MegaHertz (8.0 GigaHertz)
  - Memory capacity:
    2048 MegaBytes (2.0 GigaBytes)
  - Disk capacity:
    800 GigaBytes (0.8 TeraBytes)
  - Will need new units!
    Mega $\Rightarrow$ Giga $\Rightarrow$ Tera

Applications and Languages

- CAD, CAM, CAE, . . .
- Lotus, DOS, . . .
- Multimedia, . . .
- The Web, . . .
- JAVA, . . .
- The Net $\Rightarrow$ Ubiquitous Computing
- ???

Levels of Abstraction



temp = v[k];
v[k] = v[k+1];
v[k+1] = temp;

C code

Assembly code

```
lw  $15,   0($2)
lw  $16,   4($2)
sw  $16,   0($2)
sw  $15,   4($2)
```

Machine code

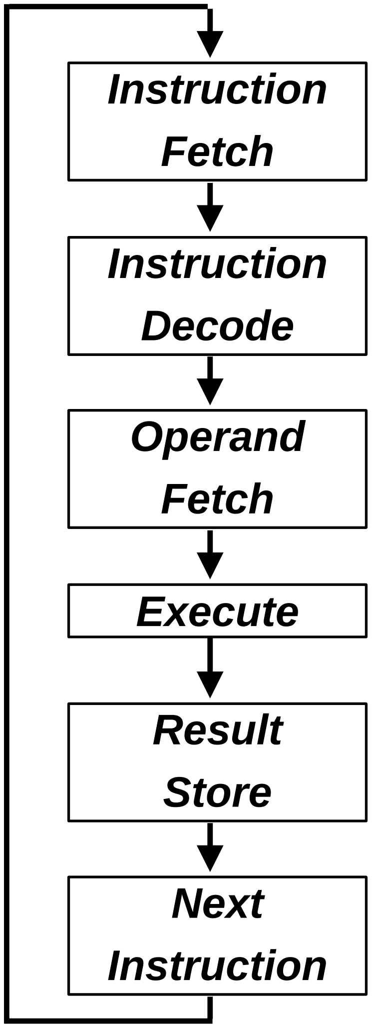
0000 1001 1100 0110 1010 1111 0101 1000
1010 1111 0101 1000 0000 1001 1100 0110
1100 0110 1010 1111 0101 1000 0000 1001
0101 1000 0000 1001 1100 0110 1010 1111

IR $\leftarrow$ Imem[PC]; PC $\leftarrow$ PC + 4

ALUOP[0:3] $\Leftarrow$ InstReg[9:11] & MASK

reasons to use HLL language?

Execution Cycle



Obtain instruction from program storage

Determine required actions and instruction size

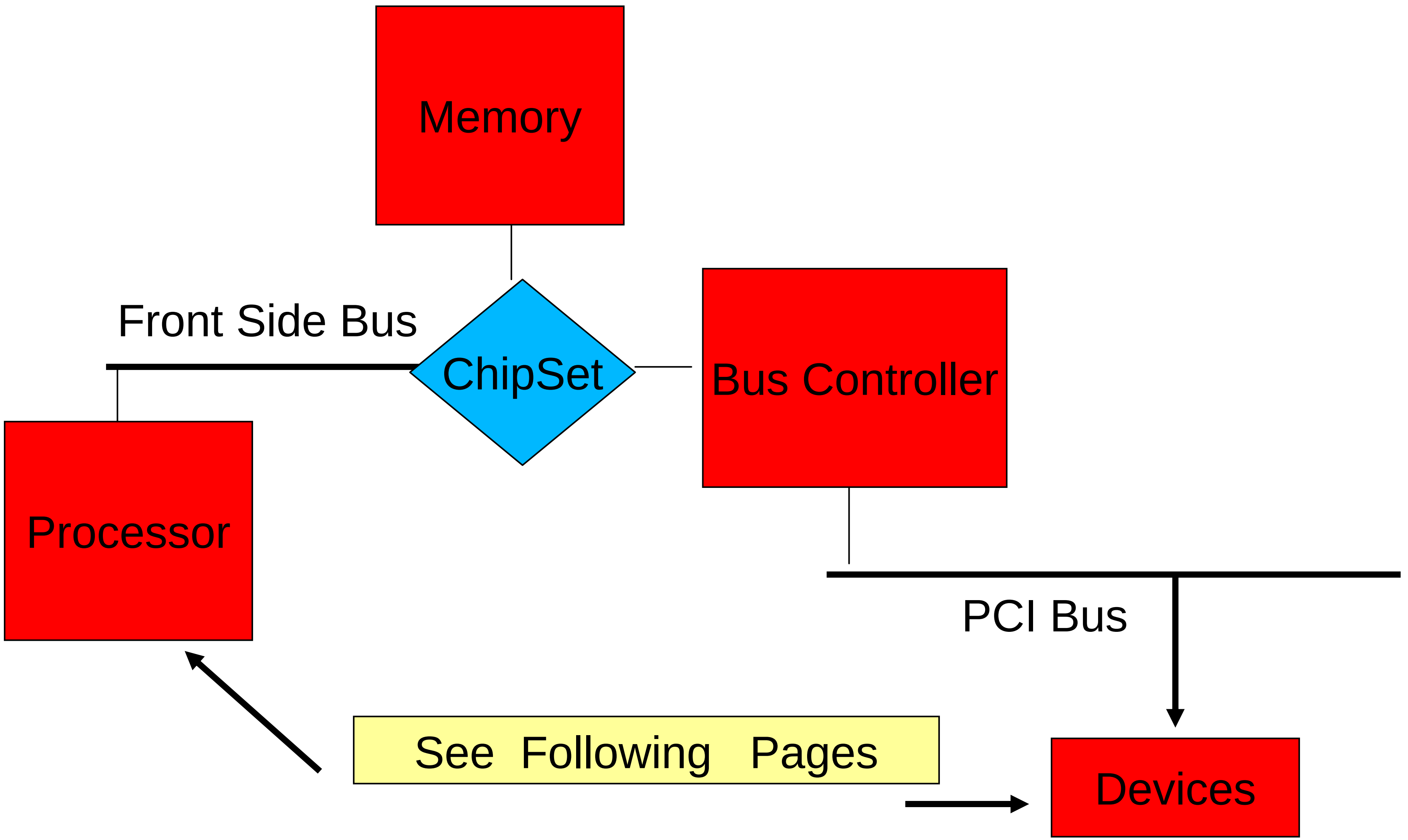
Locate and obtain operand data

Compute result value or status

Deposit results in storage for later use

Determine successor instruction

Overview: Computer System Components



Overview: Processor

Front Side Bus

These pieces implement the instruction cycles

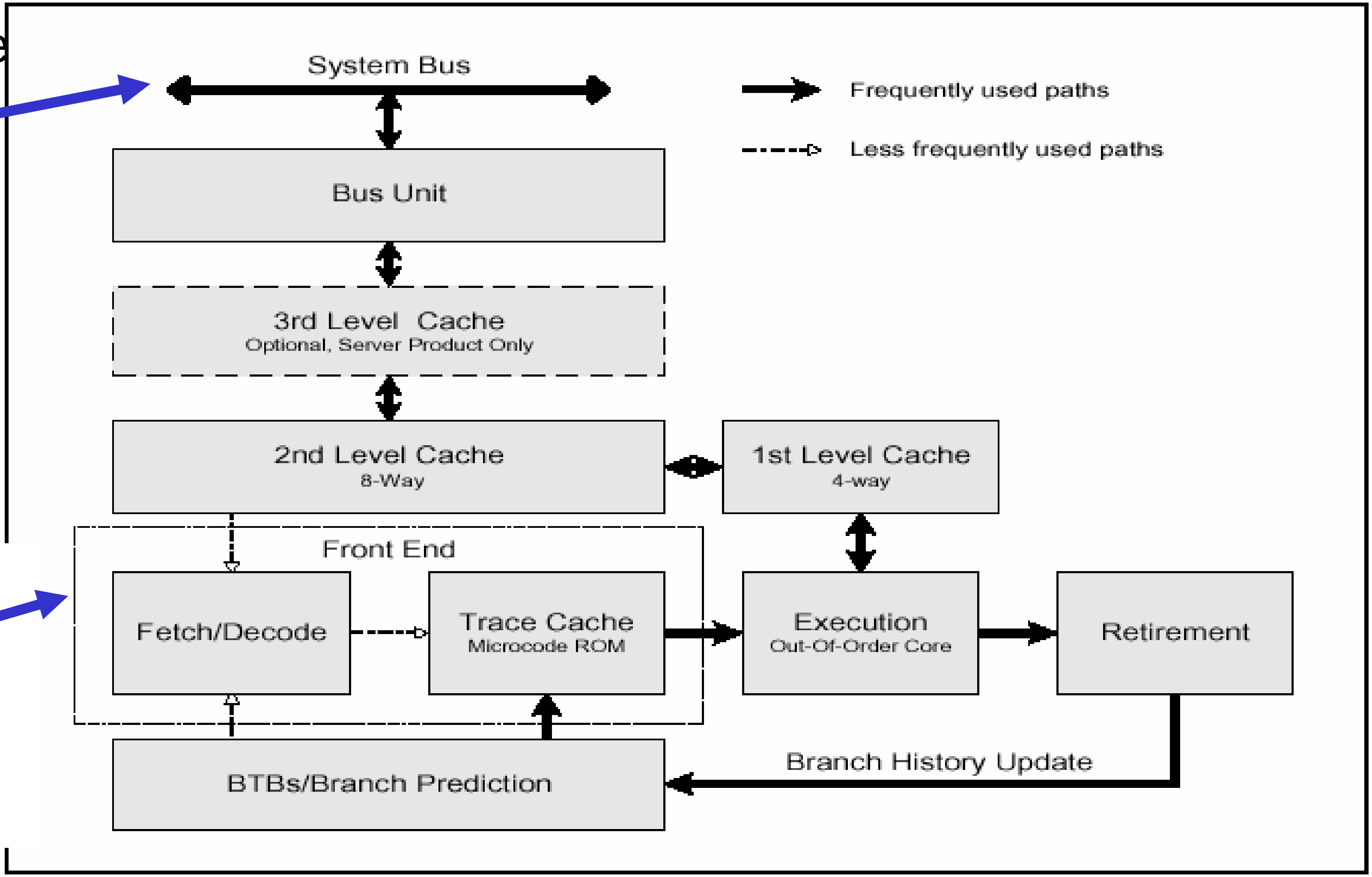
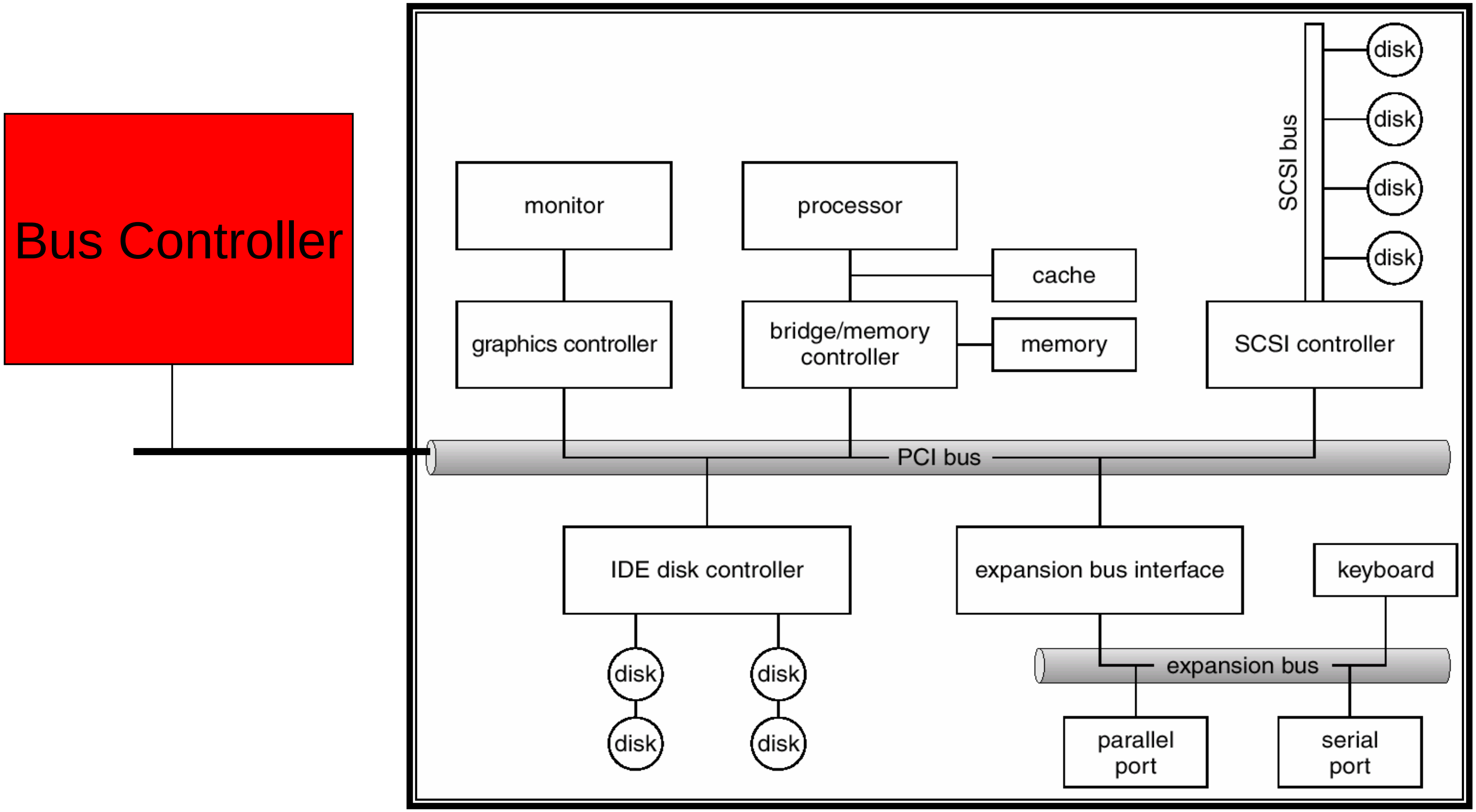


Figure 2-2. The Intel NetBurst Micro-Architecture

Overview: PCI Bus and Devices



- **Each of these busses and adapters has it's own specifications and structure.**

Summary

- **All computers consist of five components**
  - Processor:
  - (1) datapath and (2) control
  - (3) Memory
  - (4) Input devices and (5) Output devices
- **Not all "memories" are created equally**
  - Cache: fast (expensive) memory are placed closer to the processor
  - Main memory: less expensive memory--we can have more
- **Interfaces are where the problems are - between functional units and between the computer and the outside world**
- **Need to design against constraints of performance, power, area and cost**