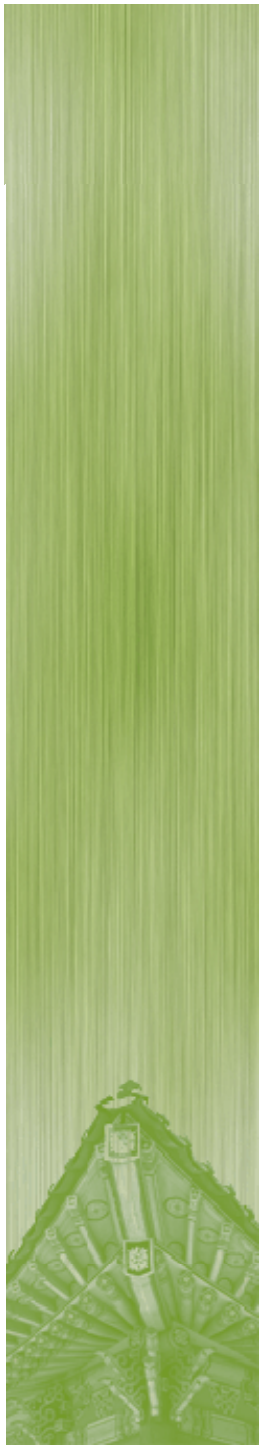


Advanced Processor Architecture

Jin-Soo Kim (jinsookim@skku.edu)
Computer Systems Laboratory
Sungkyunkwan University
http://csl.skku.edu

SKK UNIVERSITY

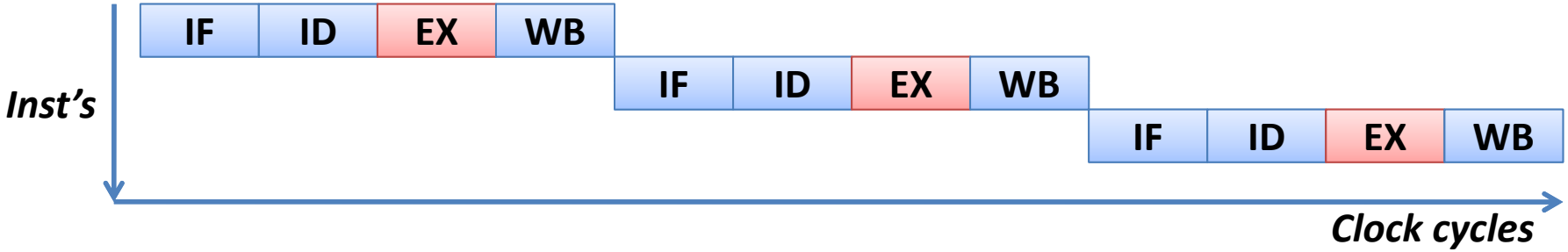
Modern Microprocessors

- More than just GHz

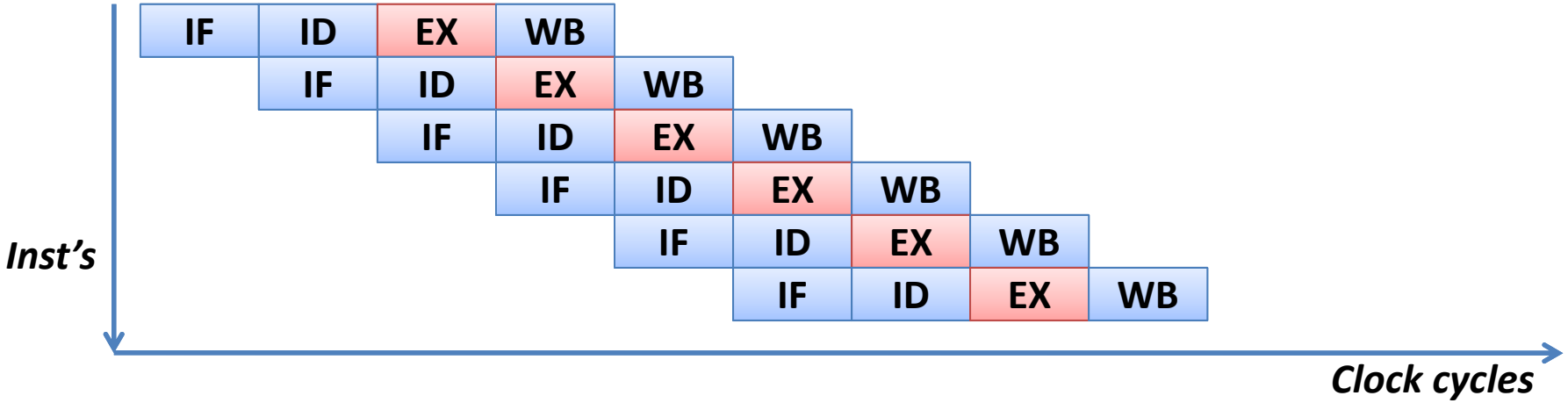
CPU	Clock Speed	SPECint2000	SPECfp2000
Athlon 64 FX-55	2.6GHz	1854	1782
Pentium 4 Extreme Edition	3.46GHz	1772	1724
Pentium 4 Prescott	3.8GHz	1671	1842
Opteron 150	2.4GHz	1655	1644
Itanium 2 9MB	1.6GHz	1590	2712
Pentium M 755	2.0GHz	1541	1088
POWER5	1.9GHz	1452	2702
SPARC64 V	1.89GHz	1345	1803
Athlon 64 3200+	2.2GHz	1080	1250
Alpha 21264C	1.25GHz	928	1019

Pipelining

- Sequential execution

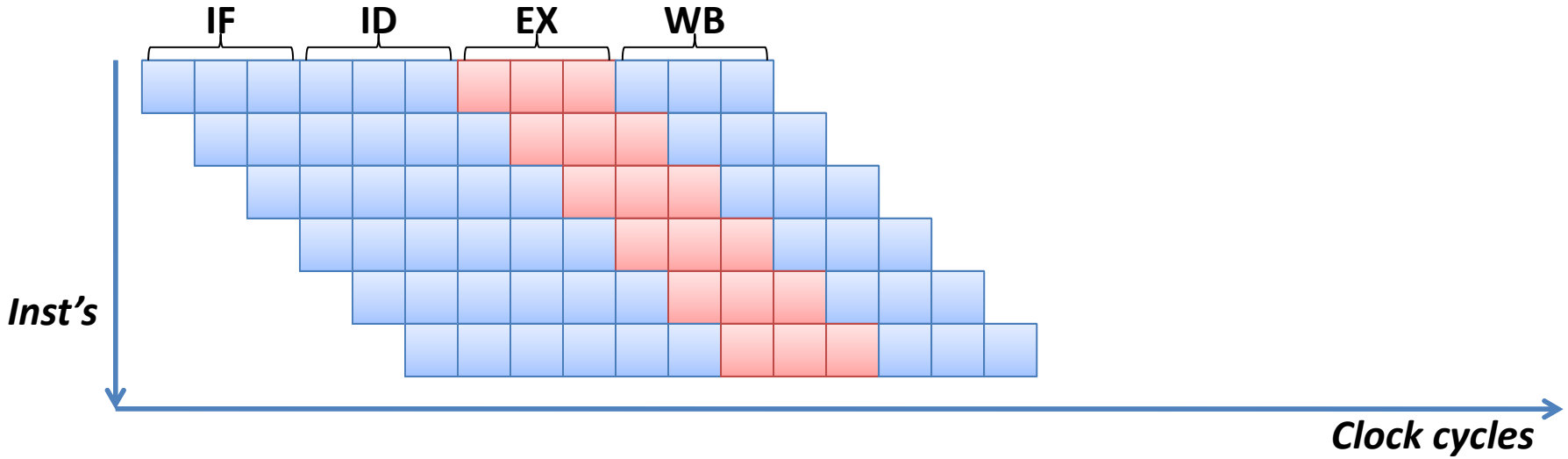


- Pipelining (RISC)



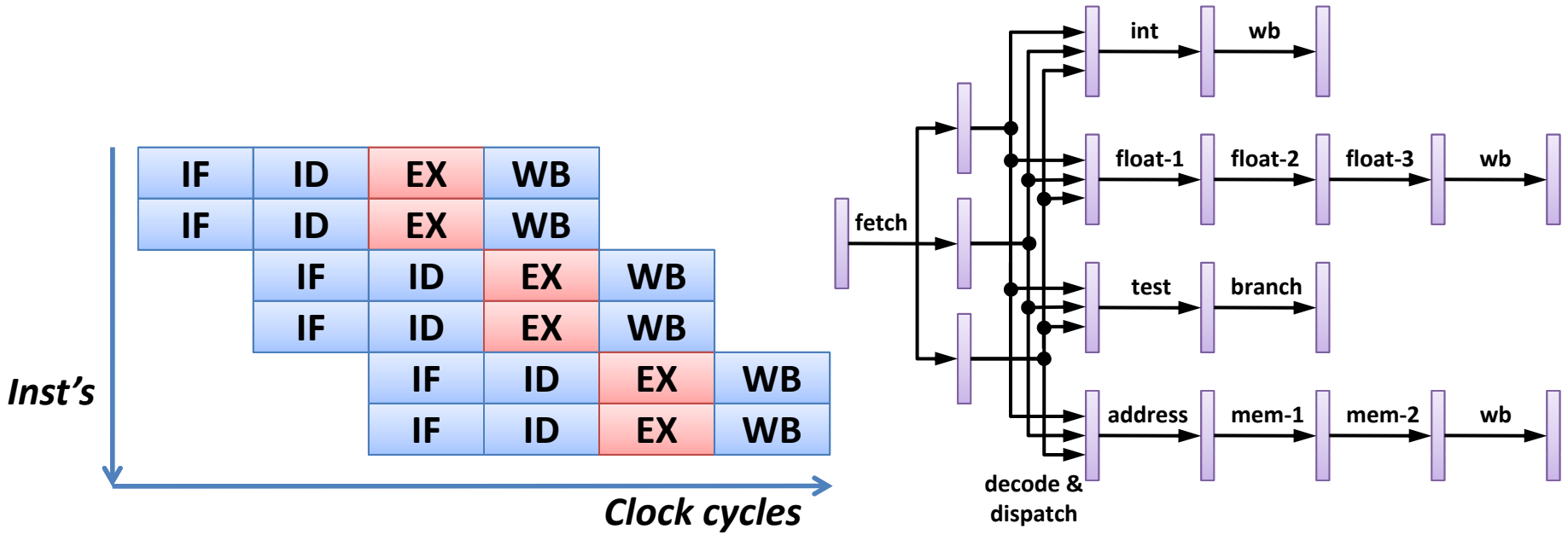
Superpipelining

- **Superpipelining**
  - Subdivide each pipeline stage
  - Higher clock speed
  - 10–15 in Athlon, 12+ in Pentium Pro/II/III, 14 in UltraSparc-III, 16–25 in PowerPC G5, 20+ in Pentium 4



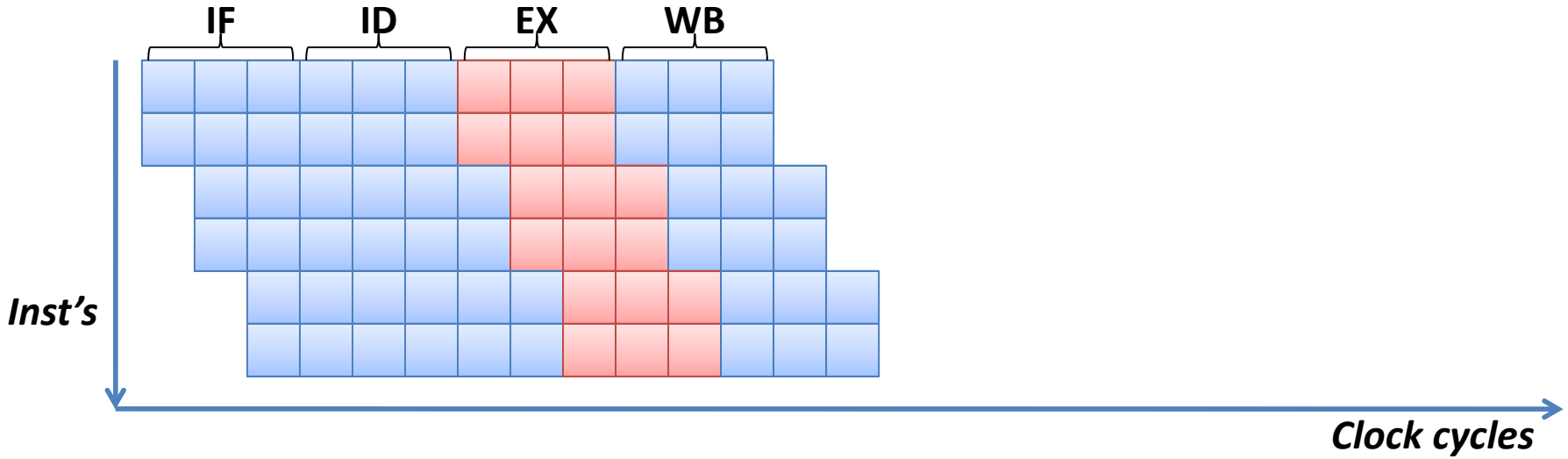
Superscalar

- **Superscalar**
  - The execution stage has a bunch of different functional units.
  - Execute multiple instructions in parallel
  - Pentium: 2-way superscalar



Superpipelined Superscalar (1)

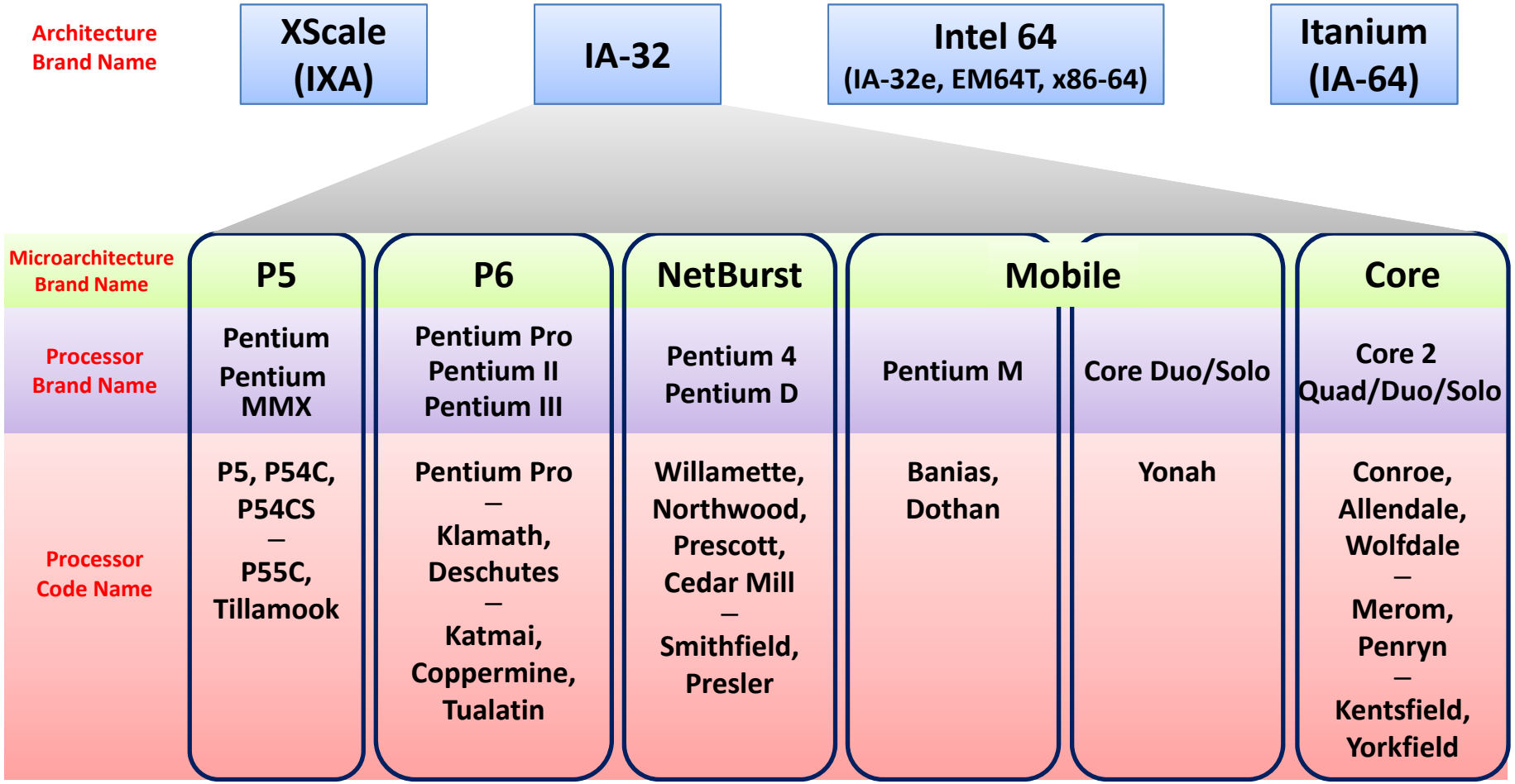
- **Superpipelining + Superscalar**
  - 2-way: MIPS R5000
  - 3-way: PowerPC G3/G4, Pentium Pro/II/III/M/4, Athlon
  - 4-way: UltraSparc, MIPS R10000, PowerPC G4e
    Alpha 21164 & 21264, Core 2 Duo
  - 5-issue: PowerPC G5



Superpipelined Superscalar (2)

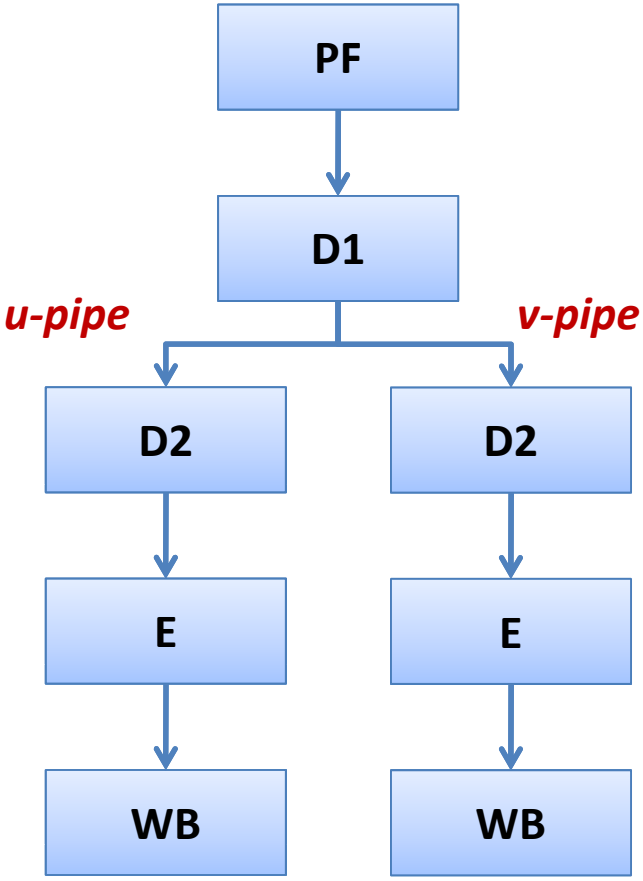
- **Tackling instruction dependencies**
  - Branch prediction + speculative execution
    - Mispredict penalty: 10 – 15 cycles in Pentium Pro/II/III
  - Instruction scheduling
    - In-order execution + compiler optimization
      - » Rearrange the instructions at compile time.
      - » Compiler can see further down the program than the hardware.
      - » SuperSparc, HyperSparc, UltraSparc, Alpha 21064 & 21164
    - Out-of-order execution
      - » Reorder instruction execution sequence in hardware at runtime
      - » Register renaming reduces the dependency further.
      - » MIPS R10000, Alpha 21264, POWER/PowerPC, Pentium Pro, Pentium 4, Core 2 Duo

Intel Architectures



Pentium Integer Pipeline

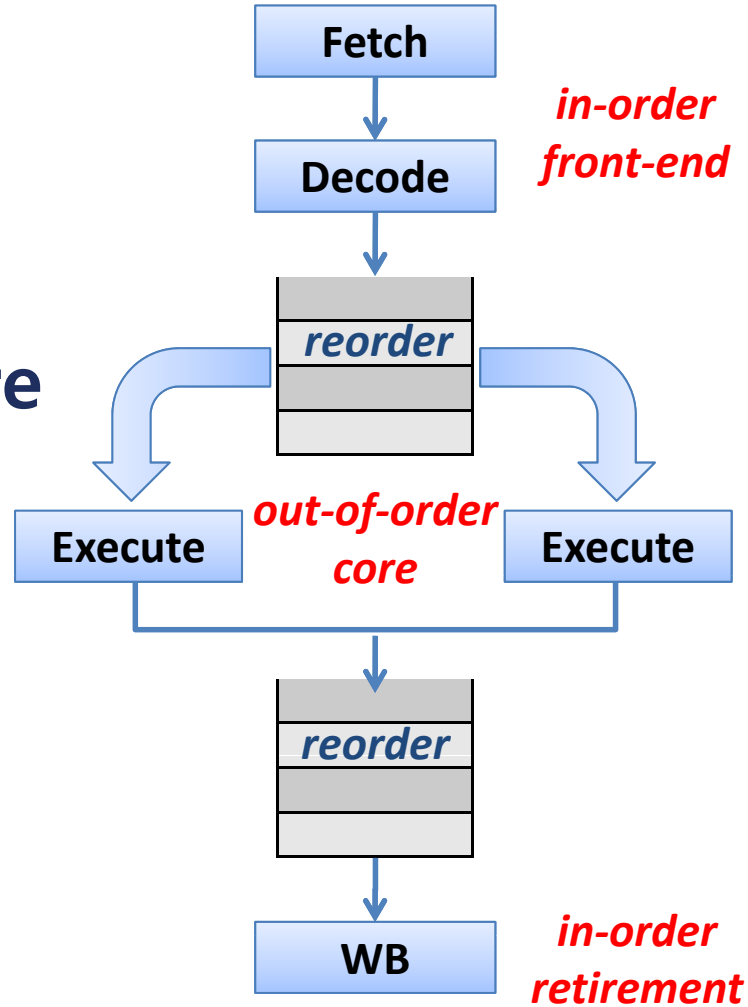
- 2-way superscalar with 5 stages



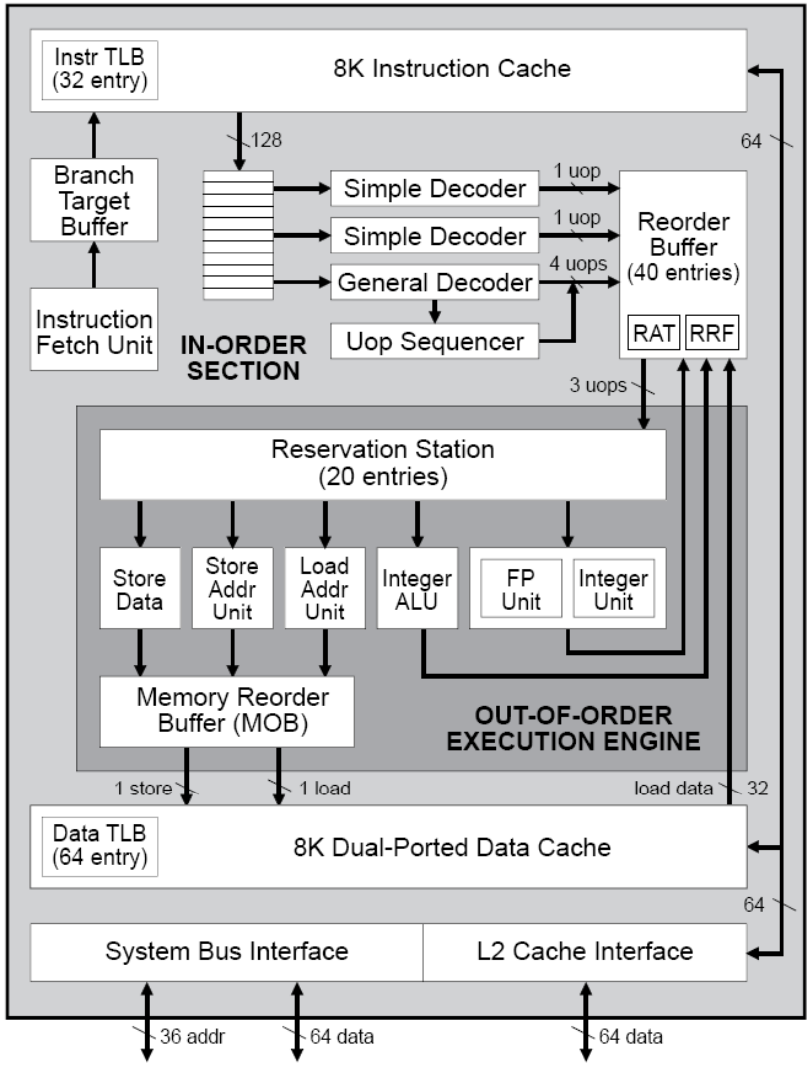
- Fetch and align instruction
- Decode instruction & generate control word
- Check for instruction pairability
- Perform branch prediction
- Decode control word
- Compute memory address
- Access data cache or calculate ALU result
- Write result

Intel Pentium Pro

- **In-order front-end**
  - Multiple branch prediction
  - Micro-operations
  - Register renaming
- **Out-of-order execution core**
  - 3-way superscalar
  - Multiple execution units
  - Dataflow analysis
  - Speculative execution
- **In-order retirement**
  - Precise faulting semantics



P6 Microarchitecture



Intel Pentium 4

- **Netburst microarchitecture**
  - Hyper pipelined for clock rates > 1 GHz
  - Need deeper pipelines
    - Diminishing performance returns
    - Requires deeper buffering to cover the longer pipelines
  - Clock skew, jitter, and latch/wire delay matter
    - Reduced percentage of the clock cycle usable by actual logic
    - Heavily depends on circuit design techniques, process technology, power and thermal constraints, etc.
  - Increases complexity
    - Harder to balance
    - More challenges to architect around
    - Greater validation effort

NetBurst Pipeline

- **Hyper pipelined technology**
  - 20 (Willamette) ~ 31 (Prescott) stages
  - Minimum misprediction penalty: 20 cycles
  - Max clock rate: 3.80 GHz (Prescott)

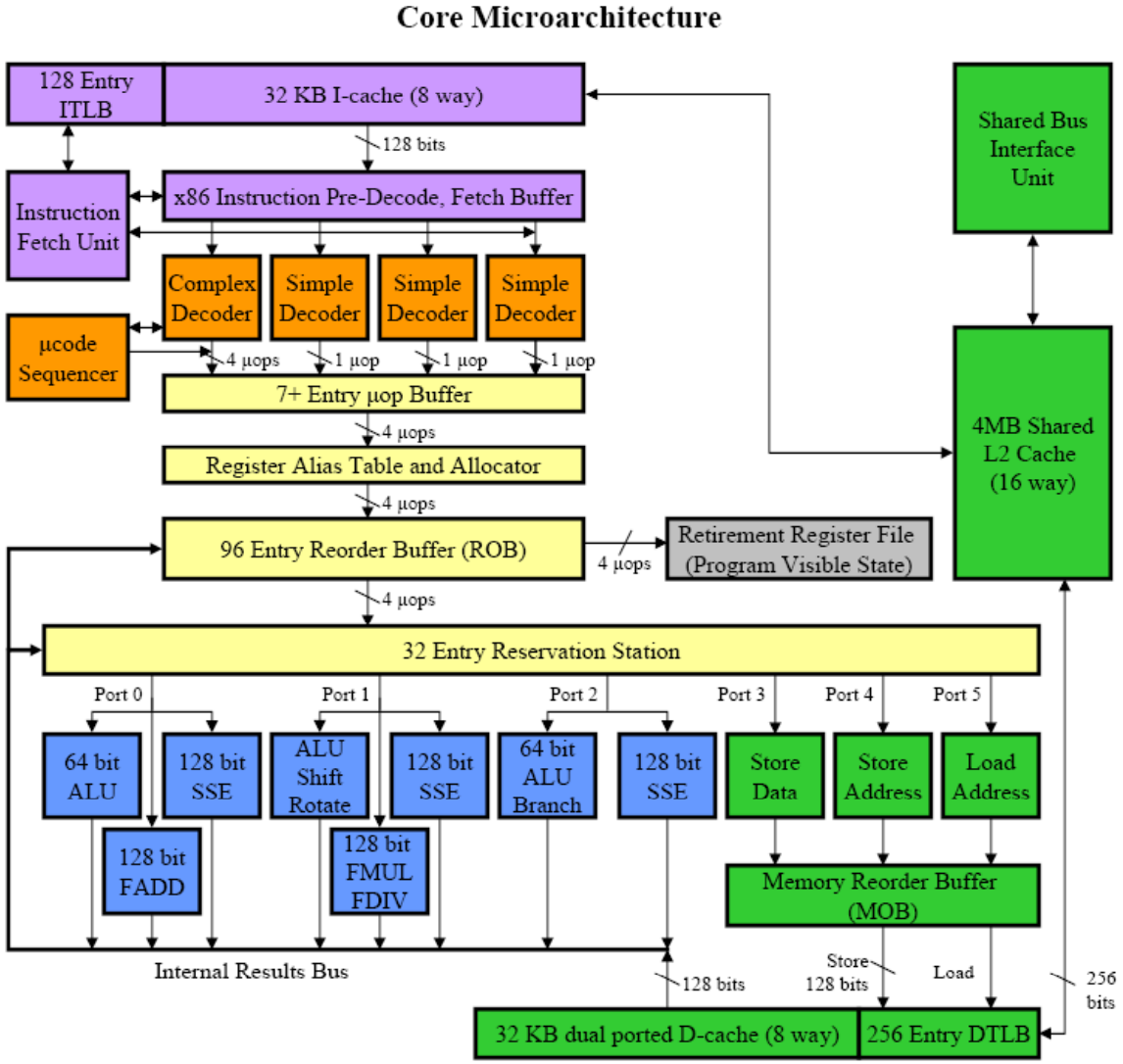
P6 misprediction pipeline (introduced at 733MHz @ 0.18µ)

1	2	3	4	5	6	7	8	9	10
Fetch	Fetch	Decode	Decode	Decode	Rename	ROB Rd	Rdy/Sch	Dispatch	Exec

NetBurst misprediction pipeline (introduced at ≥ 1.4GHz @ 0.18µ)

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

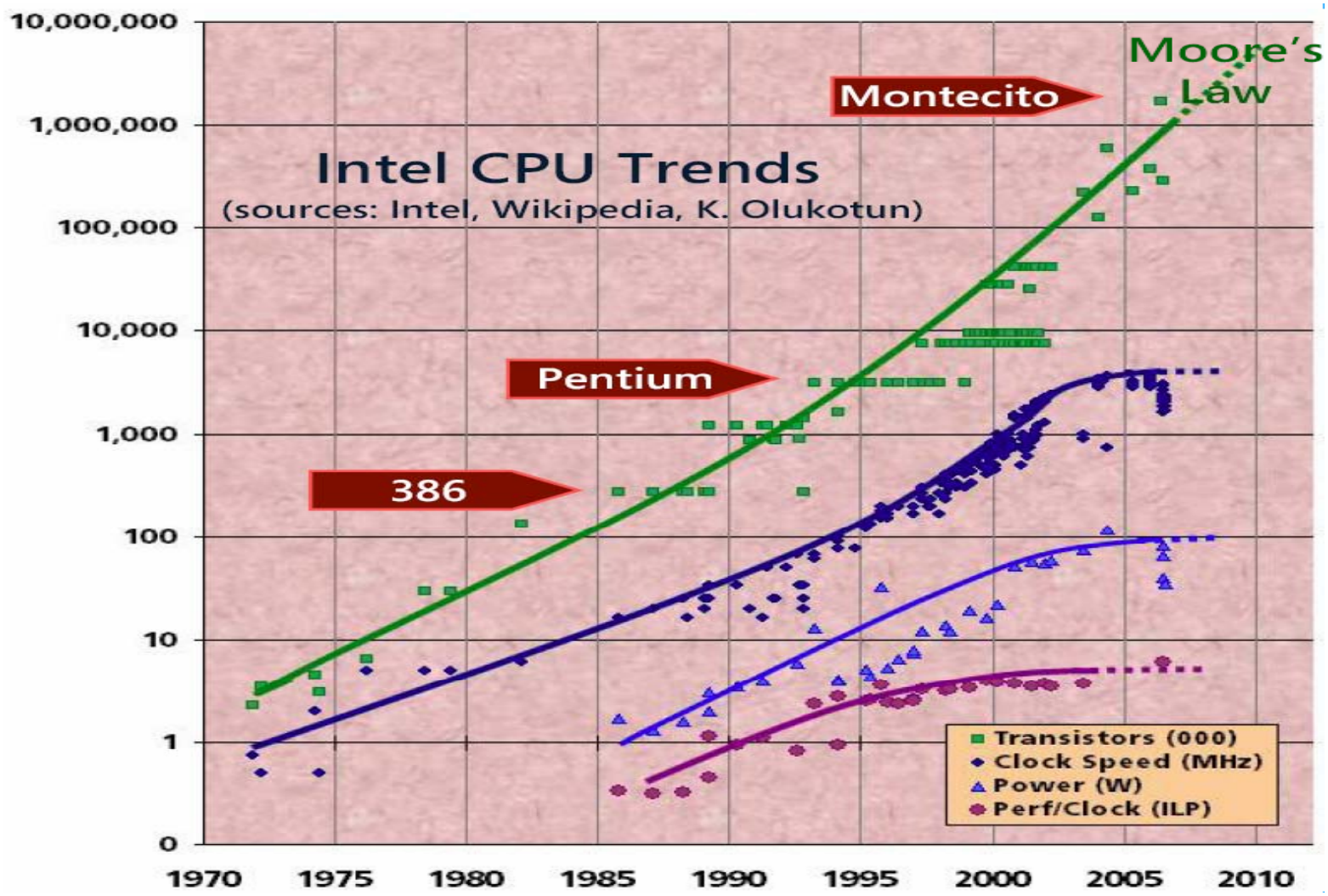
Core Microarchitecture



Multi-core (1)

- **Multi-core**
  - Put two or more processor cores onto a single chip
  - Previously called CMP (Chip Multiprocessor)
  - Examples
    - AMD Opteron: dual-core (Apr. 2005)
    - AMD dual-core Athlon 64 X2: dual-core (May 2005)
    - Intel Core Duo, Core 2 Duo, Xeon: dual-core
    - Sun UltraSparc T1: eight-core, 32 threads (Nov. 2005)
    - Intel Core 2 Extreme Quad-core processor (Sep. 2006)
    - Intel Xeon X7460 Six-core processor (Sep. 2008)
    - Microsoft's Xbox 360: triple-core PowerPC microprocessor

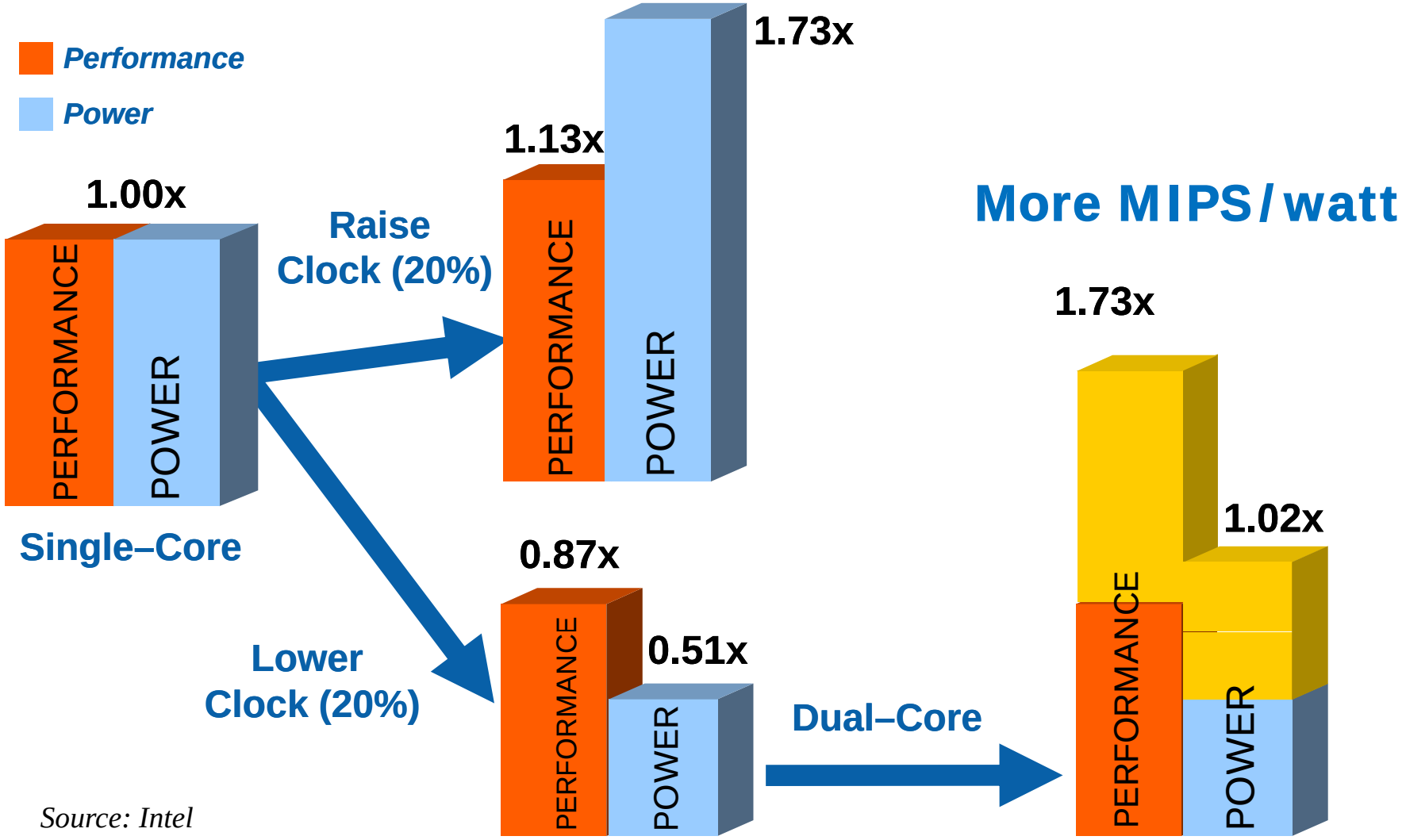
Multi-core (2)


Intel CPU Trends
(sources: Intel, Wikipedia, K. Olukotun)
Moore's Law
Montecito
Pentium
386
10,000,000
1,000,000
100,000
10,000
1,000
100
10
1
0
1970
1975
1980
1985
1990
1995
2000
2005
2010
Transistors (000)
Clock Speed (MHz)
Power (W)
Perf/Clock (ILP)

Multi-core (3)

- **Memory wall**
  - CPU 55%/year, Memory 10%/year (1986 ~ 2000)
  - Caches show diminishing returns.
- **ILP wall**
  - Control dependency
  - Data dependency
- **Power wall**
  - Dynamic power $\propto$ Frequency^3
  - Static power $\propto$ Frequency
  - Total power $\propto$ The number of cores

Multi-core (4)



Source: Intel