

Computer Architecture

Author: Jose 胡冠洲 @ ShanghaiTech

Computer Architecture

[Full-ver. Cheatsheet](#)

[Green Cards](#)

[MIPS Green Card](#)

[RISC-V Green Card](#)

[Links](#)

Full-ver. Cheatsheet

See below (page 2-7).

Green Cards

MIPS Green Card

See below (page 8-9).

RISC-V Green Card

See below (page 10-11).

Links

- [Berkeley CS61C "Great Ideas in Computer Architecture" Course Website](#)
- [RISC-V Web Simulator: Venus](#)
- [MIPS Full IDE + Simulator: MARS](#)

b Great ideas in CA

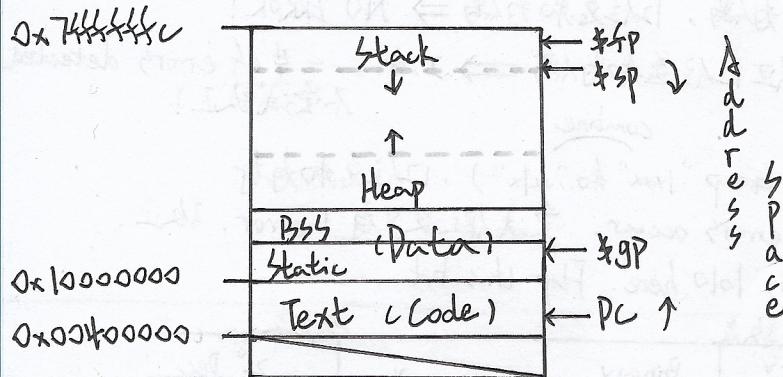
- #1: Abstractions: Layers of Representation & Interpretation
- #2: Moore's Law: 2x Transistors / chip each year.
- #3: Principle of Locality: Memory Hierarchy
- #4: Parallelism: Amdahl's Law
- #5: Performance Measurement & Improvement
- #6: Dependability via. Redundancy.

MSB XXX ... XXX LSB

Unsigned: $0 \sim 2^{32} - 1$, Signed: $-2^{31} \sim 2^{31} - 1$

Two's complement (二进制): 各位取反，再 +1.

One's complement (二进制): 直接各位取反；再 +0.



C: No deep copy for `struct's.

Static variables v.s. On-stack charal

*pt++: * priority > ++

void * 不能自增 / +- !

{sizeof (int * Type) = 8 = sizeof all ptrs.}

{sizeof (int arr [10]) = 10 * sizeof (int).}

{char * s = "..." ⇒ In static section}

{char s [] = "..." ⇒ On function stack}

The remainder of 1b: bit-wise & 1f

(ISA) Instruction Set Architecture

(RISC) Reduced Instruction Set Computing

MIPS: 32 Registers. Each 32-bits wide

{add: cause overflow detection,
Carry into MSB ≠ Carry out of MSB}

addr: No overflow detection

Logical shift:添0, $\times 2^n \Leftrightarrow \text{sh } n$

Arithmetic shift:添sign-bit, $\div 2 \Leftrightarrow \text{srar}$

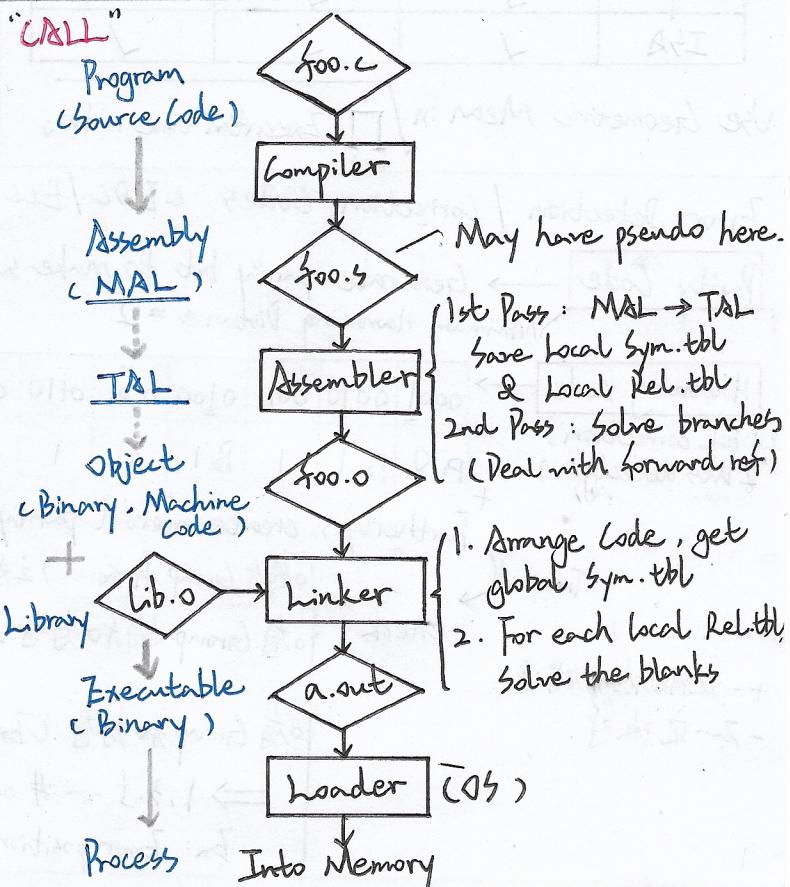
\$at: For Assembler only, × user

Preserved for Function calls Convention:

\$sp, \$fp, \$gp, \$so ~ \$st

Absolute Address Format:

{CPCTH, [31:28], Imm2, 00} → In all
28 bits
represented



PC: Normally +4

On branch: PC = (PC+4) + Imm4

Imm = Label addr. - Addr - 4

On Jump: absolute value as above

Flynn's Taxonomy

		Data	
		Single	Multiple
Instruc-	tions	SISD	SIMD
		MIPS	Intel SSEs
Multiple		MISD	MIMD
		—	Multicore

Cache Replacing: LRU v.s. MRU

Block size [Compulsory ↓
↑ Conflict ↑
Capacity →] Hit time slightly ↓
Miss rate ↓ then ↑
Miss penalty ↑

Associativity [Compulsory →
↑ Conflict ↓
Capacity →] Hit time ↑
Miss rate ↓
Miss penalty →

Cache Capacity [Compulsory ↑
↑ Conflict ↓?
Capacity ↓] Hit time ↑
Miss rate ↓
Miss penalty ↑

Performance Affect

	# of Insts.	C.P.I	CLK rate
Algorithm	✓	✓	
Language	✓	✓	
Compiler	✓	✓	
ISA	✓	✓	✓

Use Geometric Mean: $n \sqrt{\prod_{i=1}^n \text{Execution time ratio}_i}$

Error Detection / Correction Codes (EDC/EC) : Hamming Distance = # of different bits

Parity Code → Generate parity bit to make sum even → Transmit → Check
Minimum Hamming Distance = 2

Hamming Code	→	0001 0010 0011 0100 0101 0110 0111 1000 1001 1010 1011
1 bit correction, 2 bits detection		P ₁ 0 P ₂ 1 1 P ₃ 1 0 1 0 P ₄ 0 1 1 1 0

Furtherly, create overall parity bit P₅ for above 11 bits.
Transmit → Check: 每个 Group ("xx|x") 之和为偶，12位总和为偶 ⇒ NO ERROR!

+ 一定能检出错误。
- 不一定改正。
Check: 有 Group 和为奇，但 12 位总和为偶 ⇒ 2, 4, 6 ... # of errors detected
combine 不尝试改正！

有 Group 和为奇 (Ex: Group "1xx" to "xx|x"), 12 位总和为奇
⇒ 1, 3, 5 ... # of errors occur. 实际假设只有 1 error, 修正。
Ex: Error position is 1010 here. Flip this bit.

Hex 1b	Oct 8	Bin 2	Dec 10
0	00	0000	0
1	01	0001	1
2	02	0010	2
3	03	0011	3
4	04	0100	4
5	05	0101	5
6	06	0110	6
7	07	0111	7
8	10	1000	8
9	11	1001	9
A	12	1010	10
B	13	1011	11
C	14	1100	12
D	15	1101	13
E	16	1110	14
F	17	1111	15

小数点		
i	frac, 2^{-i}	Binary
0	1.0	1.0
1	0.5	0.1
2	0.25	0.01
3	0.125	0.001
4	0.0625	0.0001
5	0.03125	0.00001
6	0.015625	0.000001
7	0.0078125	0.0000001

次方表	
i	2^i Dec
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
20	1024 ²
30	1024 ³
40	1024 ⁴

Boolean 表

$X \cdot \bar{X} = 0$	$X + \bar{X} = 1$
$X \cdot 0 = 0$	$X + 0 = X$
$X \cdot 1 = X$	$X + 1 = 1$
$X \cdot X = X$	$X + X = X$
$X \cdot \bar{X} = \bar{X} \cdot X$	$X + \bar{X} = \bar{X} + X$
$(X \cdot \bar{Y}) \cdot Z = X \cdot (\bar{Y} \cdot Z)$	$(X + Y) + Z = X + (Y + Z)$
$X(X + \bar{Y}) = X \cdot \bar{Y} + X \cdot Z$	$X + \bar{Y} \cdot Z = (X + Y) \cdot (X + Z)$
$X + X \cdot \bar{Y} = X$	$X \cdot (X + \bar{Y}) = X$
$\checkmark X + \bar{X}Y = X + Y$	$\checkmark X \cdot (\bar{X} + Y) = XY$
$\checkmark X \cdot \bar{Y} = \bar{X} + Y$	$\checkmark X + \bar{Y} = \bar{X} \cdot Y$

次方表	
i	2^i Dec
0	1
1	2
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024
20	1024 ²
30	1024 ³
40	1024 ⁴

Starting (Assume x forwarding)

Case.1, write Reg used right after

(Assume W,D done in one cycle)

列对齐 W
D by 重复这条的 D

Case.2, after branch

(Assume Equal get in stage D)

列对齐 D & F by 重复这一条的 F

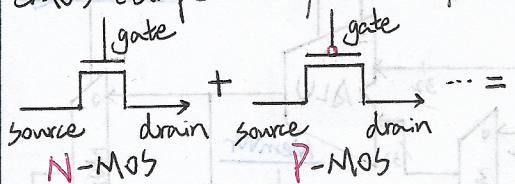
P₂ covers all "xx|1x" positions
s.t. they make even parity
★ $2^P \geq p+d+1$!

Synchronous: Coordinated by a center clock signal. (CLK)

Digital: Represent values in binary.

- Removes noises!

CMOS (Complementary MOS) - pairs!



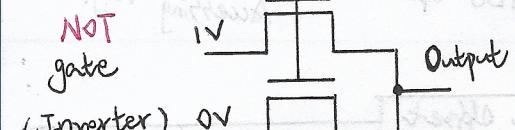
On: $V_{gate} > V_{thre}$ On: $V_{gate} < V_{thre}$

Use to pass 0 Use to pass 1

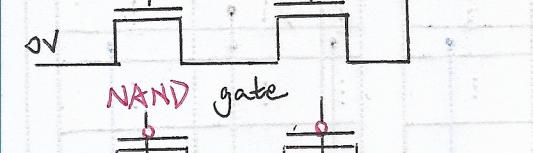
Never leave a wire undriven!

Never create path from $V_{dd} \rightarrow \text{GND}$!

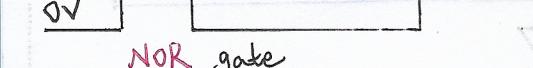
NOT gate (Inverter)



X Y Output



NAND gate



NOR gate

Logical function with N inputs: 2^N

SOP (Sum of Products): $\sum m_i$ Output 1

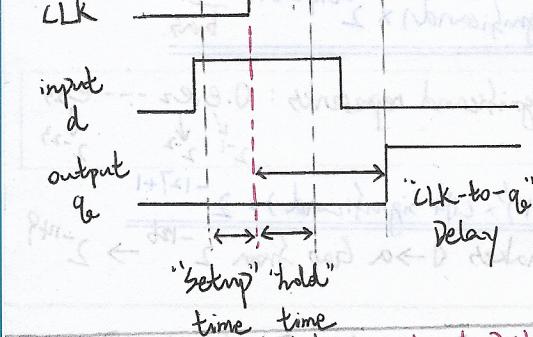
POS: $\prod M_i$ Output 1 (注意前面的反号)

CL (Combinatorial logic):

Function of instant inputs

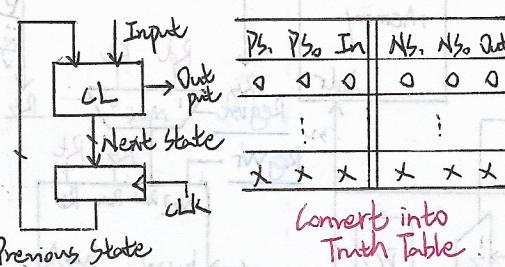
SL (Sequential Logic): registers

Remembers historical information

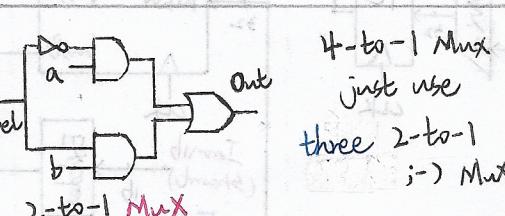


Clock Freq is decided by: Critical Path between 2 registers.

FSM (Finite State Machine)



Convert into Truth Table



Adder logical ALU (Arithmetic logic Unit)

$$\text{Sum}_i = \text{XOR}(a_i, b_i, c_i)$$

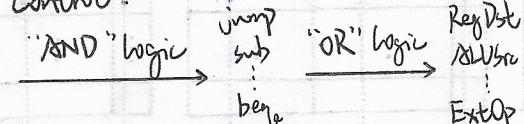
$$\begin{aligned} \text{Carry into } i+1 &= \text{arbit} + \text{arbit} + \text{c}_i \\ &= \text{MAJ}(a_i, b_i, c_i) \end{aligned}$$

Conditional Inverter: XOR

Datapath:

- Can have stages idle
- 'Load' uses all 5 stages
- i use the fewest stages

Control:



Pipelining requisite: Different resources.

- + Overall throughput
- Doesn't help latency of single task

(Actually slowing down since extra registers)

Speed up \rightarrow # of stages

Actually \leq :

- Time to "fill" and "drain"
- Limited by the slowest stage

Tc: time between completion

Tc, pipelined $\geq \frac{T_c, \text{single-cycle}}{\# \text{ of stages}}$

= Only when stages balanced.

1. Structural Hazard

Busy resources shared among stages

- + Can always be solved by adjustment or adding hardware.

- Split Inst. / Data memory.
- Reg File access halves to read + write in one CLK.

2. Data Hazard

Data-flow backwards in time

+ May be solved by forwarding

- MVST stall instruction dependent and right after 'load'

• "nop" / "bubble" / load delay slot

• Reorder code, put an unrelated Inst. right after 'load'

3. Control Hazard

whether to jump / branch?

- May stall every Inst. after them

+ Add branch comparator early

+ Predict, "flush" pipeline when wrong

+ branch delay slot:

• Put a previous unrelated Inst.

Hyperthreading: Duplicate registers but use same CL circuit

Temporal Locality: in time

Spatial Locality: in space

Memory access with cache (#)

Process Address (32-bit)

Tag	Set Index	Offset
-----	-----------	--------

Size of Index = $\log_2(\# \text{ of sets})$

Size of Offset = $\log_2(\# \text{ of bytes per block})$

Valid bit: if 0, always miss

Total capacity = Associativity $\times \# \text{ of sets} \times \text{Bytes per block}$

Average Memory Access Time

AMAT = hit time + miss rate \times miss penalty

Write Through vs. Write Back

+ Buffer + Write Allocate

- Traffic busy - Dirty Bit

+ Simple - Complex

+ Redundancy \Rightarrow Reliable

Direct Mapping vs. Associativity

- Ping-Pong effect - More comparators + # of comparators \Rightarrow Miss rate \downarrow

"3L" Misses

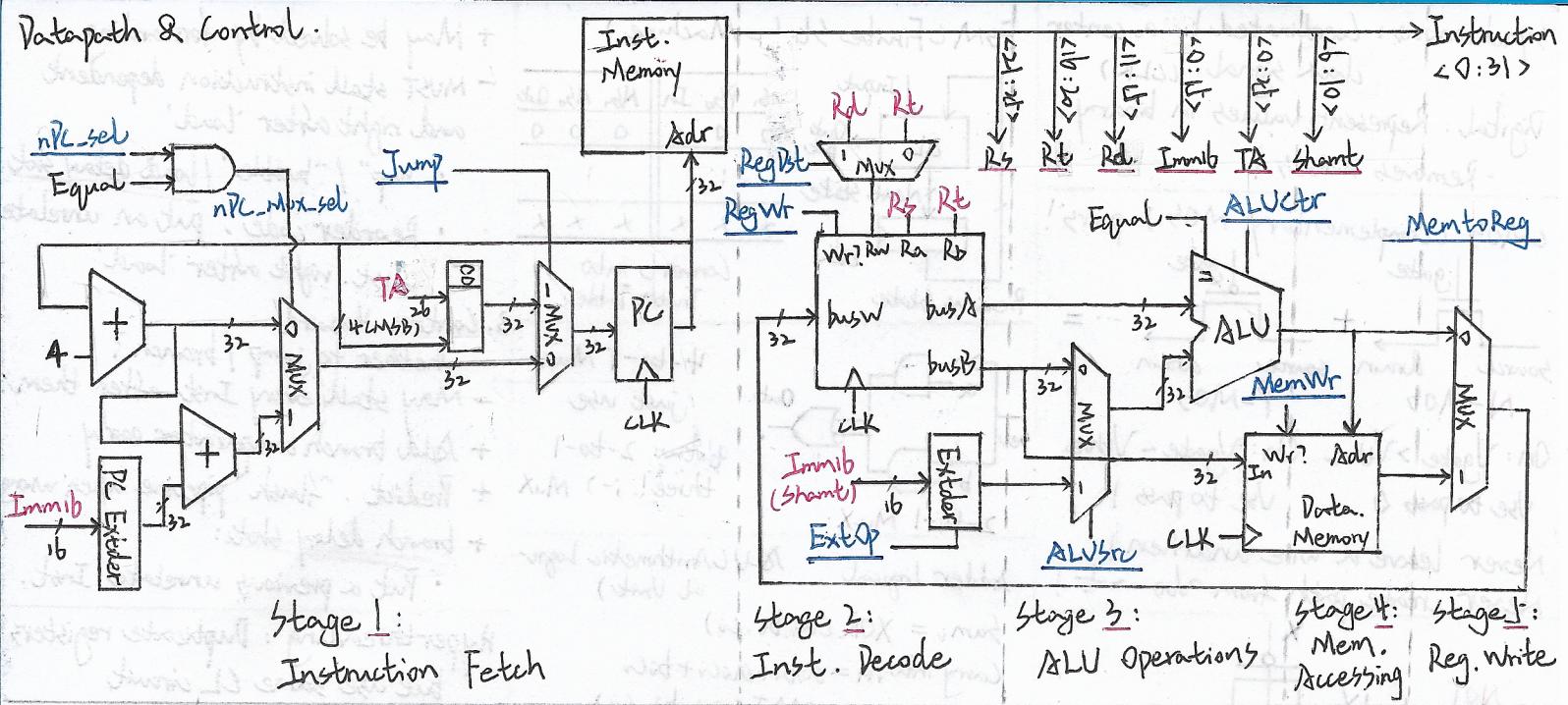
Compulsory: First access must fail.

Capacity: Cache full.

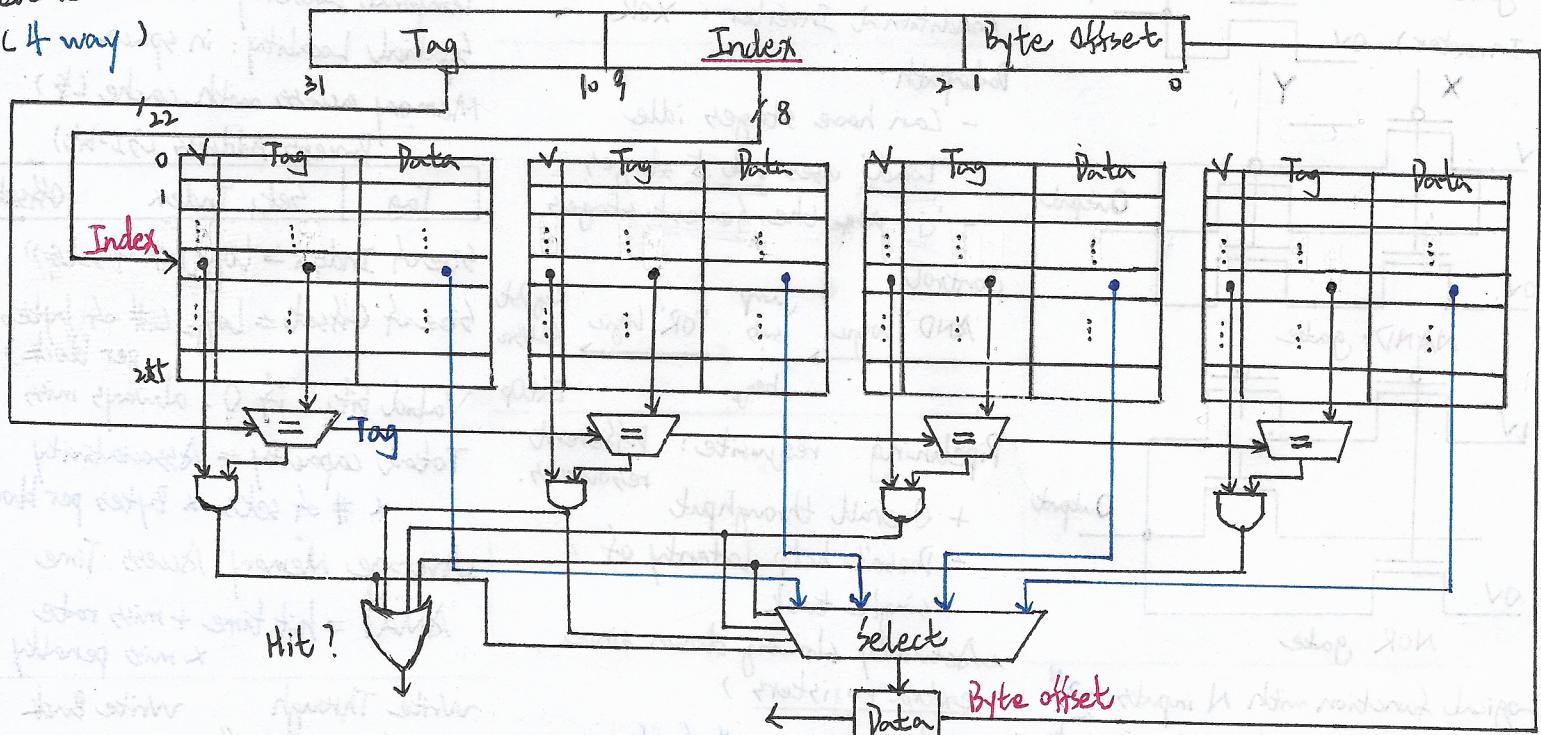
Conflict: Map to same cache block.

Figure Datapath / Cache Architecture

; -)



Cache
(4 way)



Multilevel Cache

$$\text{Local miss rate } R_{L2,\text{local}} = \frac{\# \text{ of } L2 \text{ misses}}{\# \text{ of } L1 \text{ misses}}$$

$$\text{Global miss rate } R_{\text{global}} = \frac{\# \text{ of L2 misses}}{\# \text{ of L1 misses}}$$

$$= R_{\text{global}} + R_{\text{local}}$$

$$AMAT = \text{Time for } \$L1 \text{ hit} + \\ R_{\$1,\text{local}} * (\text{Time for } \$L2 \text{ hit} \\ + R_{\$2,\text{local}} * \$L2 \text{ Miss penalty})$$

Latency: Time for one task

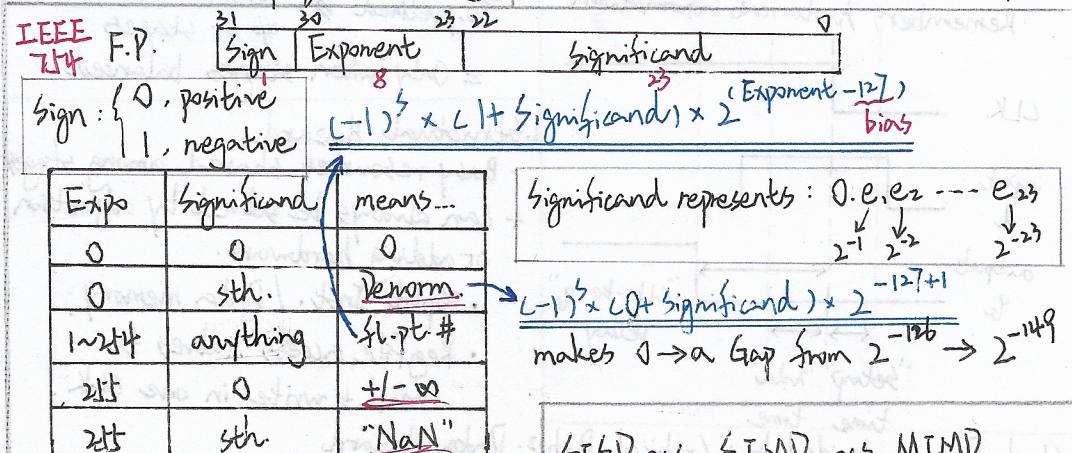
Bandwidth: Tasks per unit time
(throughput)

$$\text{Speedup}_x = \text{Performance}_x = \frac{1}{\text{Execution Time}_x}$$

$$CPU\ Time = \frac{Inst.\ #}{Program} \times \frac{CLK\ Cycles}{Instruction} (CPI) * \text{Time in a cycle}$$

Workload: Set of programs running

Benchmark: Workload chosen for compare



Threads in one process shares memory space!

- Software Multithread \Rightarrow Time-sharing
- Hardware - (Hyperthreading) \Rightarrow Context-Switch
(Looks like 2 processors) Time saved -
- + Multithreading \rightarrow Better Utilization
- + Multicore \rightarrow Duplicate Processors

OpenMP

- + Shared memory, fast
- + Compiler directives, easy to use
- + Serial code no need to be rewritten

-fopenmp

- MUST be shared memory structure
- Compilers MUST support OpenMP

except loop indicator

pragma omp parallel, default is shared
private(*list*)
OR: declared inside parallel region; private
omp_set_num_threads(*n*);, set parallel threads
omp_get_num_threads();, get # of threads
omp_get_thread_num();, get thread ID
pragma omp for \leftarrow Do NOT break this for!
critical; reduction (*t* / *d* ... : *var*)

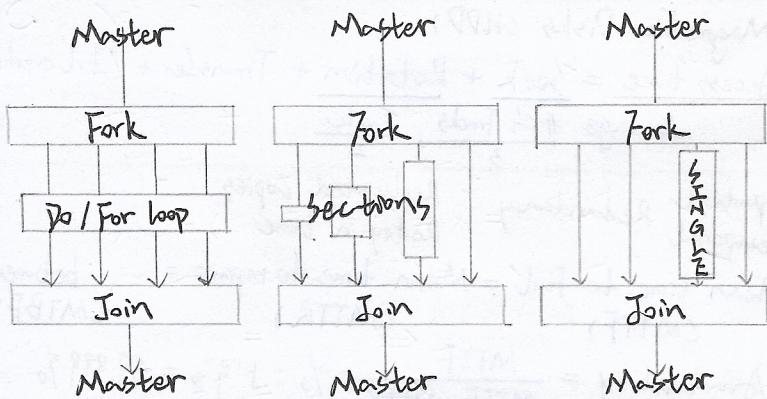
Atomic Hardware-supported Instructions

Test & Set

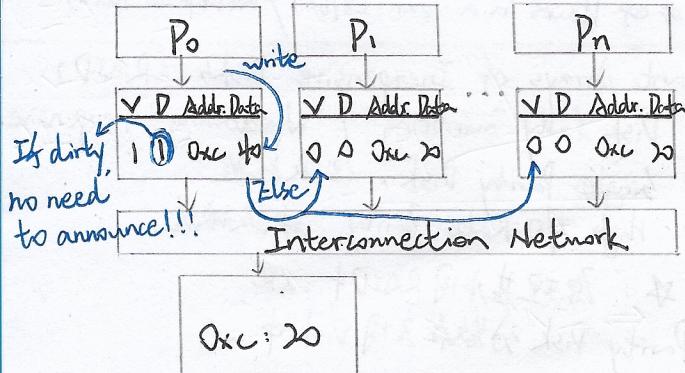
1. Test to see *x* is set?
2. If isn't, Set *x*.
3. Else, return that failed.

Example

ll \$t1, 0(\$ps1)
sc \$t0, 0(\$ps1)
changed? \$0 : 1



Cache Coherency: 4th "C" (Coherence (Communication) Miss)

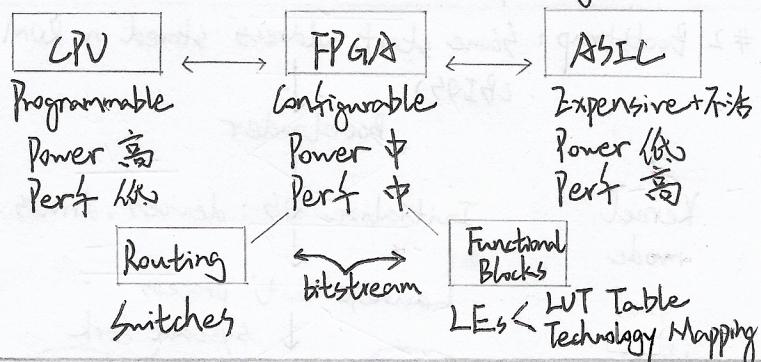


False sharing: Two processors writing to disjoint data which accidentally stay in one block.

Embedded systems \rightarrow General Purpose

- Specially-functioned
- Tightly constrained { Timing performance Power consumption
- Reactive & Real-time
- HW-SW coexistence

$$\text{Dynamic} = L \cdot C \cdot V_{dd} \cdot f_{ck} \xrightarrow{\text{Short-circuit}} \text{Leakage}$$



Warehouse Scale Computing

- + Scale of economy
- High # of Failures

(WSL)

Blades (Clu Server)

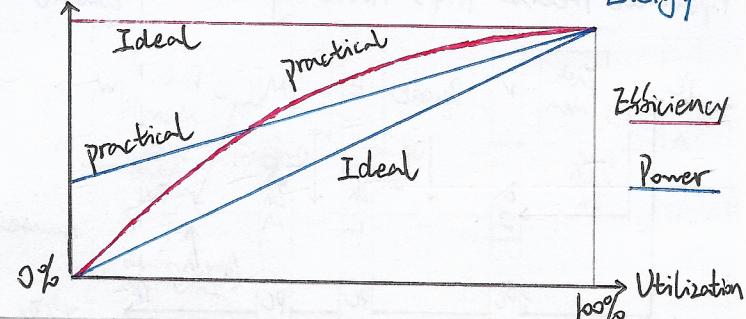
Rack

Array (Clusters)

Power Usage Effectiveness

$$PUE = \frac{\text{Total Building Power}}{\text{IT Equipment Power}} \geq 1.0$$

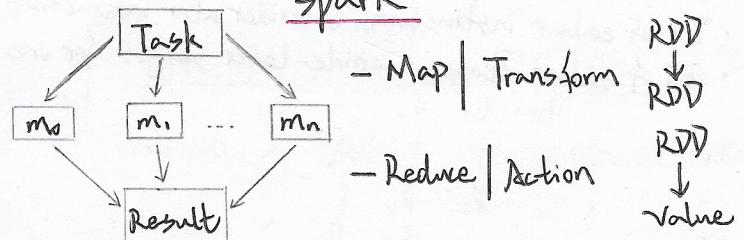
$$\text{Energy} = \text{Power} \times \text{Time}; \text{Efficiency} = \frac{\text{Computation}}{\text{Energy}}$$



Request Level Parallelism (RLP) - Independent!

Data Level Parallelism (DLP) - Map & Reduce

Spark



map, reduce, reduce by key, flatMap, -----

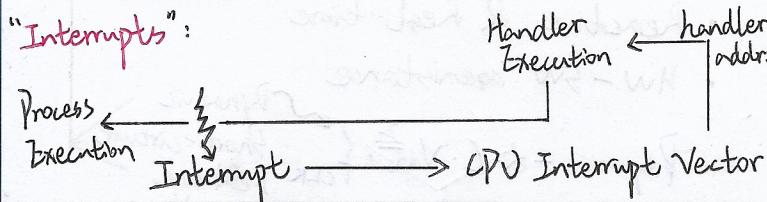
Operating Systems (OS)

- #1: I/O { Special I/O instructions
Memory-mapped I/O

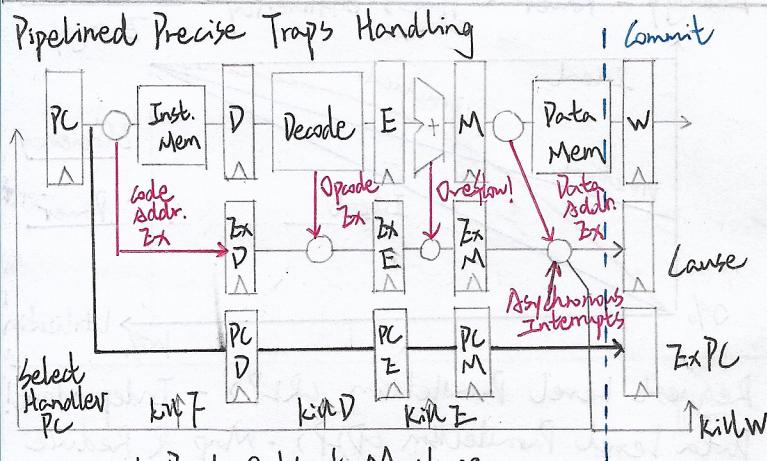
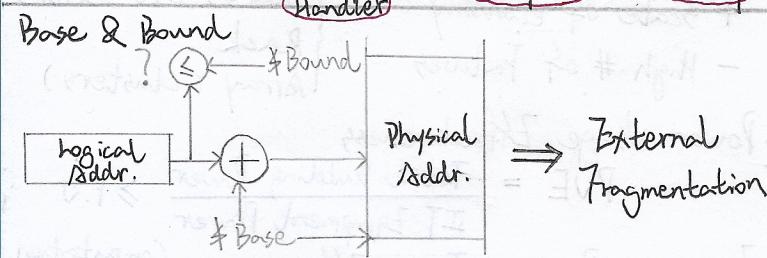
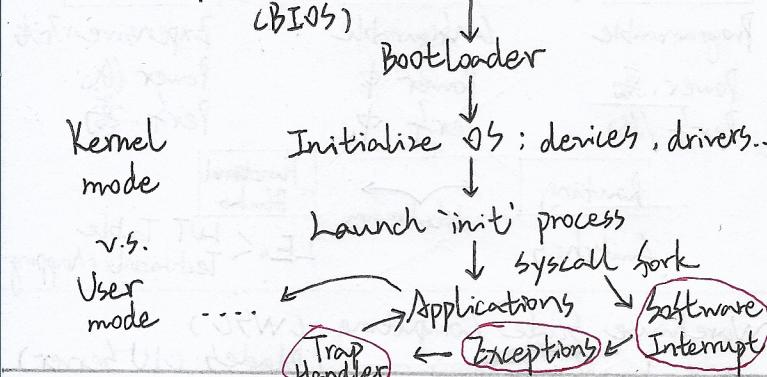
Certain portion of address space corresponds to registers in I/O devices

Speed of CPV \ggg I/O

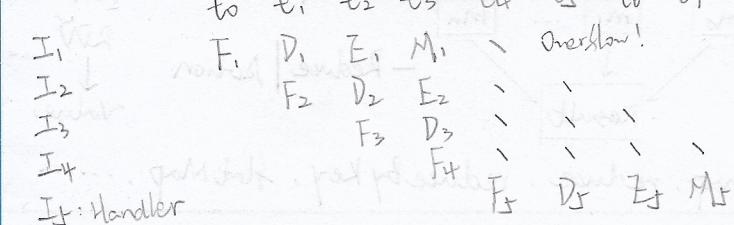
Polling: Loop to read Control Register
 Cost = % process time to poll = $\frac{\# \text{ of polls}}{\text{sec}} \times \frac{\# \text{ of cycles}}{\text{poll}} / \frac{\# \text{ of cycles}}{\text{sec}}$



#2: Bootstrap: same start address stored in ROM (BIOS)

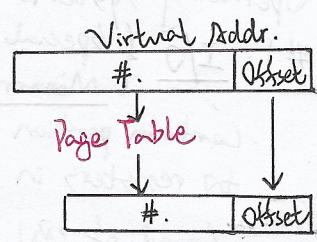


- Commit Point Only at M stage
- Ex of earlier instructions override later instructions
- Ex of earlier stages override later stages for one to t₁, t₂, t₃, t₄, t₅, t₆



Virtual Memory

- + Add disks into hierarchy
- + Simplify memory management for Apps
- + Protection { User \leftrightarrow Kernel } User1 \leftrightarrow User2



Page Tables reside in Main memory - per process.

- Linear: Set active proc's base Reg + #. as index
 $\text{Size} = \frac{\text{Size of User Space}}{\text{Size of a page}} \times \text{Size of an PTE}$

- Hierarchical: Index of L1 | Index of L2 | Offset

Translation Lookaside Buffer (TLB)

- + As a cache for page translations.

- On context switch, MUST flush away.

TLB Reach = Total size of VM space that can be mapped
 $= \# \text{ of TLB entries} \times \text{Size of a page}$

Page Fault: Found that a page not in memory now.
 ↳ A precise trap to bring in from disk

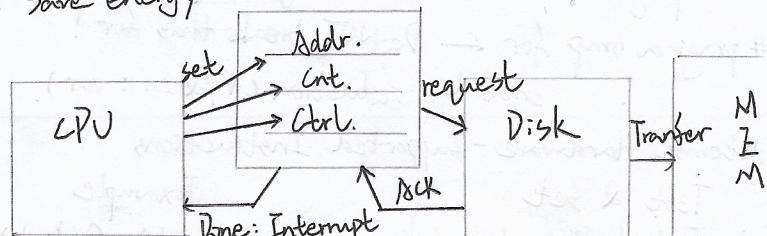
Virtual Machine: User - Kernel - VM mode

- Direct Memory Address (DMA) v.s. programmed I/O

- + General Purpose CPU can leave data sending tasks to DMA device

- + CPU utilizes spare time computing

- + Save energy



- Burst Mode: CPU out

- Cycle Stealing Mode

- Transparent Mode: Only when CPU not accessing

Magnetic Disks (HDD)

Access time = Seek + Rotation + Transfer + Ctrl overhead

$$\text{Average: } \frac{\# \text{ of Tracks}}{3}; \frac{\text{Cycle}}{2}$$

Spatial Redundancy = Replicated Copies

Temporal Redundancy = Retry in time

Mean time to Fail + Mean time to repair = --- between (MTTF) 取前 1/3 平均 (MTTR) CMTBF

$$\text{Availability} = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}} \times 100\% : 5^{\text{th}} = 99.999\%$$

$$\text{Annualized Failure Rate (AFR)} = \frac{\# \text{ of Failed}}{\text{Total} \#} = \# \text{ of Hours in a year (8760)} / \text{MTTF in hours}$$

Redundant Arrays of Inexpensive Disks (RAID)

RAID 1: Disk fully "mirroring" / "shadowing", expensive

RAID 3: Single Parity Disk, 1.5x faster

RAID 4: High I/O Rate Parity, 1.5x faster

RAID 5: 原理基本同 RAID 4, 但

Parity Disk 分散在不同 Disk 中 ✓

EC, Hamming Codes on cheatsheet 1 ;)

MIPS Reference Data



CORE INSTRUCTION SET

NAME, MNEMONIC	FOR-MAT	OPERATION (in Verilog)	OPCODE / FUNCT (Hex)
Add	add	R[Rd] = R[rs] + R[rt]	(1) 0 / 20 _{hex}
Add Immediate	addi	I R[rt] = R[rs] + SignExtImm	(1,2) 8 _{hex}
Add Imm. Unsigned	addiu	I R[rt] = R[rs] + SignExtImm	(2) 9 _{hex}
Add Unsigned	addu	R R[Rd] = R[rs] + R[rt]	0 / 21 _{hex}
And	and	R R[Rd] = R[rs] & R[rt]	0 / 24 _{hex}
And Immediate	andi	I R[rt] = R[rs] & ZeroExtImm	(3) C _{hex}
Branch On Equal	beq	I if(R[rs]==R[rt]) PC=PC+4+BranchAddr	(4) 4 _{hex}
Branch On Not Equal	bne	I if(R[rs]!=R[rt]) PC=PC+4+BranchAddr	(4) 5 _{hex}
Jump	j	J PC=JumpAddr	(5) 2 _{hex}
Jump And Link	jal	J R[31]=PC+8; PC=JumpAddr	(5) 3 _{hex}
Jump Register	jr	R PC=R[rs]	0 / 08 _{hex}
Load Byte Unsigned	lbu	I R[rt]={24'b0,M[R[rs] +SignExtImm](7:0)}	(2) 24 _{hex}
Load Halfword Unsigned	lhu	I R[rt]={16'b0,M[R[rs] +SignExtImm](15:0)}	(2) 25 _{hex}
Load Linked	ll	I R[rt] = M[R[rs]+SignExtImm]	(2,7) 30 _{hex}
Load Upper Imm.	lui	I R[rt] = {imm, 16'b0}	f _{hex}
Load Word	lw	I R[rt] = M[R[rs]+SignExtImm]	(2) 23 _{hex}
Nor	nor	R R[rd] = ~{R[rs] R[rt]}	0 / 27 _{hex}
Or	or	R R[rd] = R[rs] R[rt]	0 / 25 _{hex}
Or Immediate	ori	I R[rt] = R[rs] ZeroExtImm	(3) d _{hex}
Set Less Than	slt	R R[rd] = (R[rs] < R[rt]) ? 1 : 0	0 / 2a _{hex}
Set Less Than Imm.	slti	I R[rt] = (R[rs] < SignExtImm) ? 1 : 0 (2)	a _{hex}
Set Less Than Imm. Unsigned	sltiu	I R[rt] = (R[rs] < SignExtImm) ? 1 : 0 (2,6)	b _{hex}
Set Less Than Unsigned	sltu	R R[rd] = (R[rs] < R[rt]) ? 1 : 0 (6)	0 / 2b _{hex}
Shift Left Logical	sll	R R[rd] = R[rt] << shampt	0 / 00 _{hex}
Shift Right Logical	srl	R R[rd] = R[rt] >> shampt	0 / 02 _{hex}
Store Byte	sb	I M[R[rs]+SignExtImm](7:0) = R[rt](7:0) (2)	28 _{hex}
Store Conditional	sc	I M[R[rs]+SignExtImm] = R[rt]; R[rt] = (atomic) ? 1 : 0 (2,7)	38 _{hex}
Store Halfword	sh	I M[R[rs]+SignExtImm](15:0) = R[rt](15:0) (2)	29 _{hex}
Store Word	sw	I M[R[rs]+SignExtImm] = R[rt] (2)	2b _{hex}
Subtract	sub	R R[Rd] = R[rs] - R[rt] (1)	0 / 22 _{hex}
Subtract Unsigned	subu	R R[Rd] = R[rs] - R[rt] (0)	0 / 23 _{hex}

- (1) May cause overflow exception
- (2) SignExtImm = { 16{immediate[15]}, immediate }
- (3) ZeroExtImm = { 16{1b'0}, immediate }
- (4) BranchAddr = { 14{immediate[15]}, immediate, 2'b0 }
- (5) JumpAddr = { PC+4[31:28], address, 2'b0 }
- (6) Operands considered unsigned numbers (vs. 2's comp.)
- (7) Atomic test&set pair; R[rt] = 1 if pair atomic, 0 if not atomic

BASIC INSTRUCTION FORMATS

R	opcode	rs	rt	rd	shamt	funct	
	31	26 25	21 20	16 15	11 10	6 5	0
I	opcode	rs	rt		immediate		
	31	26 25	21 20	16 15			0
J	opcode			address			
	31	26 25					0

Copyright 2009 by Elsevier, Inc., All rights reserved. From Patterson and Hennessy, Computer Organization and Design, 4th ed.

(1)

ARITHMETIC CORE INSTRUCTION SET

NAME, MNEMONIC	FOR-MAT	OPERATION	OPCODE / FMT / FT / FUNCT (Hex)
Branch On FP True	bclt	FI if(FPcond)PC=PC+4+BranchAddr (4)	11/8/1/-
Branch On FP False	bcif	FI if(!FPcond)PC=PC+4+BranchAddr(4)	11/8/0/-
Divide	div	R Lo=R[rs]/R[rt]; Hi=R[rs]%R[rt]	0/-/-/1a
Divide Unsigned	divu	R Lo=R[rs]/R[rt]; Hi=R[rs]%R[rt] (6)	0/-/-/1b
FP Add Single	add.s	FR F[fd] = F[fs] + F[ft]	11/10/-/0
FP Add		F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} + {F[ft],F[ft+1]}	11/11/-/0
Double	add.d	FR F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} + {F[ft],F[ft+1]}	11/11/-/y
FP Compare Single	c.x.s*	FR FPcond = {F[fs] op F[ft]} ? 1 : 0	11/10/-/y
FP Compare	c.x.d*	FR FPcond = {F[fs],F[fs+1]} op {F[ft],F[ft+1]} ? 1 : 0	11/11/-/y
Double		* (x is eq, lt, or le) (op is ==, <, or <=) (y is 32, 3c, or 3e)	
FP Divide Single	div.s	FR F[fd] = F[fs] / F[ft]	11/10/-/3
FP Divide		F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} / {F[ft],F[ft+1]}	11/11/-/3
Double	div.d	FR F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} / {F[ft],F[ft+1]}	11/11/-/2
FP Multiply Single	mul.s	FR F[F[fd]] = F[fs] * F[ft]	11/10/-/2
FP Multiply		F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} * {F[ft],F[ft+1]}	11/11/-/2
Double	mul.d	FR F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} * {F[ft],F[ft+1]}	11/11/-/1
FP Subtract Single	sub.s	FR F[fd]=F[fs] - F[ft]	11/10/-/1
FP Subtract		F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} - {F[ft],F[ft+1]}	11/11/-/1
Double	sub.d	FR F[F[fd],F[fd+1]] = {F[fs],F[fs+1]} - {F[ft],F[ft+1]}	11/11/-/1
Load FP Single	lwc1	I F[rt]=M[R[rs]+SignExtImm]	(2) 31/-/-/-
Load FP	ldc1	I F[rt]=M[R[rs]+SignExtImm]; F[rt+1]=M[R[rs]+SignExtImm+4]	(2) 35/-/-/-
Double			
Move From Hi	mfhi	R R[rd] = Hi	0 / -/-/10
Move From Lo	mflo	R R[rd] = Lo	0 / -/-/12
Move From Control	mfc0	R R[rd] = CR[rs]	10 / 0/-/0
Multiply	mult	R {Hi,Lo} = R[rs] * R[rt]	0 / -/-/18
Multiply Unsigned	multu	R {Hi,Lo} = R[rs] * R[rt] (6)	0 / -/-/19
Shift Right Arith.	sra	R R[rd] = R[rt] >> shampt	0 / -/-/3
Store FP Single	swc1	I M[R[rs]+SignExtImm] = F[rt]	(2) 39/-/-/-
Store FP	sdc1	I M[R[rs]+SignExtImm] = F[rt]; M[R[rs]+SignExtImm+4] = F[rt+1]	(2) 3d/-/-/-
Double			

(2)

FLOATING-POINT INSTRUCTION FORMATS

FR	opcode	fmt	ft	fs	fd	funct	
	31	26 25	21 20	16 15	11 10	6 5	0
FI	opcode	fmt	ft		immediate		0

PSEUDOINSTRUCTION SET

NAME	MNEMONIC	OPERATION
Branch Less Than	blt	if(R[rs]<R[rt]) PC = Label
Branch Greater Than	bgt	if(R[rs]>R[rt]) PC = Label
Branch Less Than or Equal	ble	if(R[rs]<=R[rt]) PC = Label
Branch Greater Than or Equal	bge	if(R[rs]>=R[rt]) PC = Label
Load Immediate	li	R[rd] = immediate
Move	move	R[rd] = R[rs]

REGISTER NAME, NUMBER, USE, CALL CONVENTION

NAME	NUMBER	USE	PRESERVED ACROSS A CALL?
\$zero	0	The Constant Value 0	N.A.
\$at	1	Assembler Temporary	No
\$v0-\$v1	2-3	Values for Function Results and Expression Evaluation	No
\$a0-\$a3	4-7	Arguments	No
\$t0-\$t7	8-15	Temporaries	No
\$s0-\$s7	16-23	Saved Temporaries	Yes
\$t8-\$t9	24-25	Temporaries	No
\$k0-\$k1	26-27	Reserved for OS Kernel	No
\$gp	28	Global Pointer	Yes
\$sp	29	Stack Pointer	Yes
\$fp	30	Frame Pointer	Yes
\$ra	31	Return Address	No

OPCODES, BASE CONVERSION, ASCII SYMBOLS

MIPS (1)	MIPS (2)	Binary	Deci-	Hexa-	ASCII	Deci-	Hexa-	ASCII
opcode	funct		mal	mal	Char-	mal	mal	acter
(31:26)	(5:0)	(5:0)			mal			acter
(1)	sll	add.f	00 0000	0	0 NUL	64	40 @	
		sub.f	00 0001	1	1 SOH	65	41 A	
j	srl	mul.f	00 0010	2	2 STX	66	42 B	
jal	sra	div.f	00 0011	3	3 ETX	67	43 C	
beq	sllv	sqrt.f	00 0100	4	4 EOT	68	44 D	
bne		abs.f	00 0101	5	5 ENQ	69	45 E	
blez	srlv	mov.f	00 0110	6	6 ACK	70	46 F	
bgtz	sraev	neg.f	00 0111	7	7 BEL	71	47 G	
addi	jr		00 1000	8	8 BS	72	48 H	
addiu	jalr		00 1001	9	9 HT	73	49 I	
slti	movz		00 1010	10	a LF	74	4a J	
sltiu	movn		00 1011	11	b VT	75	4b K	
andi	syscall	round.wf	00 1100	12	c FF	76	4c L	
ori	break	trunc.wf	00 1101	13	d CR	77	4d M	
xori		ceil.wf	00 1110	14	e SO	78	4e N	
lui	sync	floor.wf	00 1111	15	f SI	79	4f O	
(2)	mfhi		01 0000	16	10 DLE	80	50 P	
mthi			01 0001	17	11 DC1	81	51 Q	
mflo	movzf		01 0010	18	12 DC2	82	52 R	
mtlo	movnf		01 0011	19	13 DC3	83	53 S	
			01 0100	20	14 DC4	84	54 T	
			01 0101	21	15 NAK	85	55 U	
			01 0110	22	16 SYN	86	56 V	
			01 0111	23	17 ETB	87	57 W	
mult			01 1000	24	18 CAN	88	58 X	
multu			01 1001	25	19 EM	89	59 Y	
div			01 1010	26	1a SUB	90	5a Z	
divu			01 1011	27	1b ESC	91	5b [
			01 1100	28	1c FS	92	5c \	
			01 1101	29	1d GS	93	5d]	
			01 1110	30	1e RS	94	5e ^	
			01 1111	31	1f US	95	5f -	
lb	add	cvt.s.f	10 0000	32	20 Space	96	60 .	
lh	addu	cvt.d.f	10 0001	33	21 !	97	61 a	
lw	sub		10 0010	34	22 "	98	62 b	
lw	subu		10 0011	35	23 #	99	63 c	
lbu	and	cvt.w.f	10 0100	36	24 \$	100	64 d	
lhu	or		10 0101	37	25 %	101	65 e	
lwr	xor		10 0110	38	26 &	102	66 f	
	nor		10 0111	39	27 ,	103	67 g	
sb			10 1000	40	28 (104	68 h	
sh			10 1001	41	29)	105	69 i	
swl	slt		10 1010	42	2a *	106	6a j	
sw	sltu		10 1011	43	2b +	107	6b k	
			10 1100	44	2c ,	108	6c l	
			10 1101	45	2d -	109	6d m	
swr			10 1110	46	2e .	110	6e n	
cache			10 1111	47	2f /	111	6f o	
ll	tge	c.f.f	11 0000	48	30 0	112	70 p	
lwc1	tgeu	c.un.f	11 0001	49	31 1	113	71 q	
lwc2	tlr	c.eqf	11 0010	50	32 2	114	72 r	
pref	tlru	c.ueqf	11 0011	51	33 3	115	73 s	
	teq	c.col.f	11 0100	52	34 4	116	74 t	
ldc1		c.ultr.f	11 0101	53	35 5	117	75 u	
ldc2	tne	c.ole.f	11 0110	54	36 6	118	76 v	
		c.cule.f	11 0111	55	37 7	119	77 w	
sc		c.sff.f	11 1000	56	38 8	120	78 x	
swc1		c.nglef.f	11 1001	57	39 9	121	79 y	
swc2		c.seqf.f	11 1010	58	3a :	122	7a z	
		c.nglf.f	11 1011	59	3b ;	123	7b {	
		c.lt.f	11 1100	60	3c <	124	7c	
sdc1		c.ngef.f	11 1101	61	3d =	125	7d }	
sdc2		c.lef.f	11 1110	62	3e >	126	7e ~	
		c.ngtf.f	11 1111	63	3f ?	127	7f DEL	

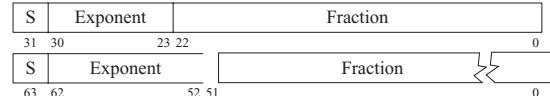
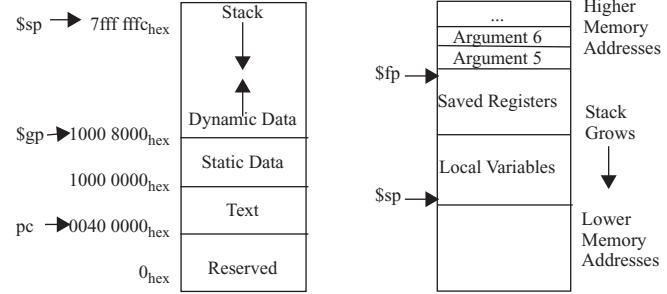
(1) opcode(31:26) == 0

 (2) opcode(31:26) == 17₁₀ (11_{hex}); if fmt(25:21) == 16₁₀ (10_{hex}) f == s (single); if fmt(25:21) == 17₁₀ (11_{hex}) f == d (double)

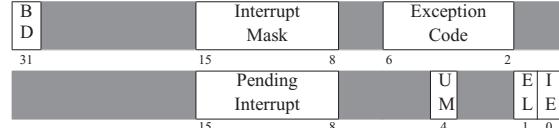
IEEE 754 FLOATING-POINT STANDARD

$$(-1)^S \times (1 + \text{Fraction}) \times 2^{(\text{Exponent} - \text{Bias})}$$

where Single Precision Bias = 127, Double Precision Bias = 1023.

IEEE Single Precision and Double Precision Formats:

MEMORY ALLOCATION

DATA ALIGNMENT

Double Word							
Word				Word			
Halfword		Halfword		Halfword		Halfword	
Byte	Byte	Byte	Byte	Byte	Byte	Byte	Byte
0	1	2	3	4	5	6	7
Value of three least significant bits of byte address (Big Endian)							

EXCEPTION CONTROL REGISTERS: CAUSE AND STATUS


BD = Branch Delay, UM = User Mode, EL = Exception Level, IE = Interrupt Enable

EXCEPTION CODES

Number	Name	Cause of Exception	Number	Name	Cause of Exception
0	Int	Interrupt (hardware)	9	Bp	Breakpoint Exception
4	AdEL	Address Error Exception (load or instruction fetch)	10	RI	Reserved Instruction Exception
5	AdES	Address Error Exception (store)	11	CpU	Coprocessor Unimplemented
6	IBE	Bus Error on Instruction Fetch	12	Ov	Arithmetic Overflow Exception
7	DBE	Bus Error on Load or Store	13	Tr	Trap
8	Sys	Syscall Exception	15	FPE	Floating Point Exception

SIZE PREFIXES (10^x for Disk, Communication; 2^x for Memory)

PRE-SIZE	FIX	PRE-SIZE	FIX	PRE-SIZE	FIX
10 ³ , 2 ¹⁰	Kilo-	10 ¹⁵ , 2 ⁵⁰	Peta-	10 ⁻³	milli-
10 ⁶ , 2 ²⁰	Mega-	10 ¹⁸ , 2 ⁶⁰	Exa-	10 ⁻⁶	micro-
10 ⁹ , 2 ³⁰	Giga-	10 ²¹ , 2 ⁷⁰	Zetta-	10 ⁻⁹	nano-
10 ¹² , 2 ⁴⁰	Tera-	10 ²⁴ , 2 ⁸⁰	Yotta-	10 ⁻¹²	pico-

The symbol for each prefix is just its first letter, except μ is used for micro.

Base Integer Instructions: RV32I, RV64I, and RV128I							RV Privileged Instructions									
Category	Name	Fmt	RV32I Base			+RV{64,128}			Category	Name	RV mnemonic					
Loads	Load Byte	I	LB	rd,rs1,imm		L{D Q} rd,rs1,imm			CSR Access	Atomic R/W	CSRRW	rd,csr,rs1				
	Load Halfword	I	LH	rd,rs1,imm						Atomic Read & Set Bit	CSRRS	rd,csr,rs1				
	Load Word	I	LW	rd,rs1,imm						Atomic Read & Clear Bit	CSRRC	rd,csr,rs1				
	Load Byte Unsigned	I	LBU	rd,rs1,imm						Atomic R/W Imm	CSRRWI	rd,csr,imm				
	Load Half Unsigned	I	LHU	rd,rs1,imm						Atomic Read & Set Bit Imm	CSRRSI	rd,csr,imm				
Stores	Store Byte	S	SB	rs1,rs2,imm		S{D Q} rs1,rs2,imm			Change Level	Env. Call	ECALL					
	Store Halfword	S	SH	rs1,rs2,imm						Environment Breakpoint	EBREAK					
	Store Word	S	SW	rs1,rs2,imm						Environment Return	ERET					
Shifts	Shift Left	R	SLL	rd,rs1,rs2		SLL{W D} rd,rs1,rs2			Trap Redirect to Supervisor	MRTS						
	Shift Left Immediate	I	SLLI	rd,rs1,shamt		SLLI{W D} rd,rs1,shamt				Redirect Trap to Hypervisor	MRTH					
	Shift Right	R	SRL	rd,rs1,rs2		SRL{W D} rd,rs1,rs2				Hypervisor Trap to Supervisor	HRTS					
	Shift Right Immediate	I	SRLI	rd,rs1,shamt		SRLI{W D} rd,rs1,shamt				Interrupt	WFI					
	Shift Right Arithmetic	R	SRA	rd,rs1,rs2		SRA{W D} rd,rs1,rs2				MMU	Supervisor FENCE	SFENCE.VM rs1				
	Shift Right Arith Imm	I	SRAI	rd,rs1,shamt		SRAI{W D} rd,rs1,shamt										
Arithmetic	ADD	R	ADD	rd,rs1,rs2		ADD{W D} rd,rs1,rs2			Optional Compressed (16-bit) Instruction Extension: RVC							
	ADD Immediate	I	ADDI	rd,rs1,imm		ADDI{W D} rd,rs1,imm										
	SUBtract	R	SUB	rd,rs1,rs2		SUB{W D} rd,rs1,rs2										
	Load Upper Imm	U	LUI	rd,imm		C.LW rd',rs1',imm										
	Add Upper Imm to PC	U	AUIPC	rd,imm						C.LWSP	rd,imm	LW rd',sp,imm*4				
Logical	XOR	R	XOR	rd,rs1,rs2		C.LD rd',rs1',imm				C.LD	rd',rs1',imm	LD rd',sp,imm*8				
	XOR Immediate	I	XORI	rd,rs1,imm						C.LDSP	rd,imm	LD rd,sp,imm*8				
	OR	R	OR	rd,rs1,rs2						C.LQ	rd',rs1',imm	LQ rd',sp,imm*16				
	OR Immediate	I	ORI	rd,rs1,imm		C.LQSP rd,imm				C.LQSP	rd,imm	LQ rd,sp,imm*16				
	AND	R	AND	rd,rs1,rs2												
Compare	AND Immediate	I	ANDI	rd,rs1,imm		C.SW rs1',rs2',imm				C.SW	rs1',rs2',imm	SW rs1',sp,imm*4				
	Set <	R	SLT	rd,rs1,rs2						C.SWSWP	rs2,imm	SW rs2,sp,imm*4				
	Set < Immediate	I	SLTI	rd,rs1,imm						C.SD	rs1',rs2',imm	SD rs1',sp,imm*8				
	Set < Unsigned	R	SLTU	rd,rs1,rs2		C.SDSP rs2,imm				C.SDSP	rs2,imm	SD rs2,sp,imm*8				
	Set < Imm Unsigned	I	SLTIU	rd,rs1,imm						C.SQ	rs1',rs2',imm	SQ rs1',sp,imm*16				
Branches	Branch =	SB	BEQ	rs1,rs2,imm		C.SQSP rs2,imm				C.SQSP	rs2,imm	SQ rs2,sp,imm*16				
	Branch ≠	SB	BNE	rs1,rs2,imm												
	Branch <	SB	BLT	rs1,rs2,imm		CR C.ADD rd,rs1				C.ADD	rd,rs1	ADD rd,rd,rs1				
	Branch ≥	SB	BGE	rs1,rs2,imm						CR C.ADDW rd,rs1	rd,rs1	ADDW rd,rd,imm				
	Branch < Unsigned	SB	BLTU	rs1,rs2,imm		CR C.ADDI rd,imm				C.ADDI	rd,imm	ADDI rd,rd,imm				
Jump & Link	Branch ≥ Unsigned	SB	BGEU	rs1,rs2,imm						CR C.ADDIW rd,imm	rd,imm	ADDIW rd,rd,imm				
	Jump & Link Register	UJ	JAL	rd,imm		CR C.ADDI16SP x0,imm				C.ADDI16SP	x0,imm	ADDI sp,sp,imm*16				
	Jump & Link Register	UJ	JALR	rd,rs1,imm						CR C.ADDI4SPN rd',imm	rd',imm	ADDI rd',sp,imm*4				
	Synch thread	I	FENCE			CI C.LI rd,imm				C.LI	rd,imm	ADDI rd,x0,imm				
	Synch Instr & Data	I	FENCE.I							CI C.LUI rd,imm	rd,imm	LUI rd,imm				
System	System CALL	I	SCALL			CR C.MV rd,rs1				C.MV	rd,rs1	ADD rd,rs1,x0				
	System BREAK	I	SBREAK							CR C.SUB rd,rs1	rd,rs1	SUB rd,rd,rs1				
Counters	ReaD CYCLE	I	RDCYCLE	rd		CI C.SLLI rd,imm				C.SLLI	rd,imm	SLLI rd,rd,imm				
	ReaD CYCLE upper Half	I	RDCYCLES	rd						CB C.BEQZ rs1',imm	rs1',imm	BEQ rs1',x0,imm				
	ReaD TIME	I	RDTIME	rd		CB C.BNEZ rs1',imm				C.BNEZ	rs1',imm	BNE rs1',x0,imm				
	ReaD TIME upper Half	I	RDTIMES	rd						C.J imm		JAL x0,imm				
	ReaD INSTR RETired	I	RDINSTRET	rd		CR C.JR rd,rs1				C.JR	rd,rs1	JALR x0,rs1,0				
Shifts	ReaD INSTR upper Half	I	RDINSTRETH	rd						C.J imm		JAL ra,imm				
	Shift Left Imm	CI	SLLI	rd,imm						C.JALR rs1		JALR ra,rs1,0				
	Branch=0	CB	BEQZ	rs1',imm		CJ imm										
	Branch≠0	CB	BNEZ	rs1',imm												
	Jump	CJ	J imm			CJ imm				C.J	imm	JAL x0,imm				
Jump & Link	Jump Register	CR	JR imm							C.JR	rd,rs1	JALR x0,rs1,0				
	J&L	CJ	J imm			CR C.JAL imm				C.JAL	imm	JAL ra,imm				
	Jump & Link Register	CR	JALR imm							C.JALR rs1		JALR ra,rs1,0				
	Env. BREAK	CI	EBREAK			CR C.EBREAK						EBREAK				

32-bit Instruction Formats

R	31	30	25 24	21 20	19	15 14	12 11	8	7	6	0
	funct7		rs2		rs1	funct3		rd		opcode	
I		imm[11:0]			rs1	funct3		rd		opcode	
S		imm[11:5]			rs2	rs1	funct3	imm[4:0]		opcode	
SB	imm[12]	imm[10:5]			rs2	rs1	funct3	imm[4:1]	imm[11]	opcode	
U			imm[31:12]					rd		opcode	
UJ	imm[20]	imm[10:1]	imm[11]		imm[19:12]			rd		opcode	

CR	15 14	13	12	11 10	9 8	7	6	5	4	3	2	1	0
	funct4		rd/rs1			rs2		op					
CI	funct3	imm	rd/rs1			imm		op					
CSS				imm		rs2		op					
CIW	funct3		imm			rd'		op					
CL	funct3	imm	rs1'		imm	rd'		op					
CS	funct3	imm	rs1'	imm	rs2'			op					
CB	funct3	offset	rs1'		offset			op					
CJ	funct3		jump target					op					

RISC-V Integer Base (RV32I/64I/128I), privileged, and optional compressed extension (RVC). Registers x1-x31 and the pc are 32 bits wide in RV32I, 64 in RV64I, and 128 in RV128I (x0=0). RV64I/128I add 10 instructions for the wider formats. The RVI base of <50 classic integer RISC instructions is required. Every 16-bit RVC instruction matches an existing 32-bit RVI instruction. See risc.org.

Optional Multiply-Divide Instruction Extension: RVM				
Category	Name	Fmt	RV32M (Multiply-Divide)	+RV{64,128}
Multiply	MULTiply	R	MUL rd,rs1,rs2	MUL{W D} rd,rs1,rs2
	MULTiply upper Half	R	MULH rd,rs1,rs2	
	MULTiply Half Sign/Uns	R	MULHSU rd,rs1,rs2	
	MULTiply upper Half Uns	R	MULHU rd,rs1,rs2	
Divide	DIVide	R	DIV rd,rs1,rs2	DIV{W D} rd,rs1,rs2
	DIVide Unsigned	R	DIVU rd,rs1,rs2	
Remainder	REMAinder	R	REM rd,rs1,rs2	REM{W D} rd,rs1,rs2
	REMAinder Unsigned	R	REMU rd,rs1,rs2	REMU{W D} rd,rs1,rs2

Optional Atomic Instruction Extension: RVA				
Category	Name	Fmt	RV32A (Atomic)	+RV{64,128}
Load	Load Reserved	R	LR.W rd,rs1	LR.{D Q} rd,rs1
Store	Store Conditional	R	SC.W rd,rs1,rs2	SC.{D Q} rd,rs1,rs2
Swap	SWAP	R	AMOSWAP.W rd,rs1,rs2	AMOSWAP.{D Q} rd,rs1,rs2
Add	ADD	R	AMOADD.W rd,rs1,rs2	AMOADD.{D Q} rd,rs1,rs2
Logical	XOR	R	AMOXOR.W rd,rs1,rs2	AMOXOR.{D Q} rd,rs1,rs2
	AND	R	AMOAND.W rd,rs1,rs2	AMOAND.{D Q} rd,rs1,rs2
	OR	R	AMOOR.W rd,rs1,rs2	AMOOR.{D Q} rd,rs1,rs2
Min/Max	MINimum	R	AMOMIN.W rd,rs1,rs2	AMOMIN.{D Q} rd,rs1,rs2
	MAXimum	R	AMOMAX.W rd,rs1,rs2	AMOMAX.{D Q} rd,rs1,rs2
	MINimum Unsigned	R	AMOMINU.W rd,rs1,rs2	AMOMINU.{D Q} rd,rs1,rs2
	MAXimum Unsigned	R	AMOMAXU.W rd,rs1,rs2	AMOMAXU.{D Q} rd,rs1,rs2

Three Optional Floating-Point Instruction Extensions: RVF, RVD, & RVQ				
Category	Name	Fmt	RV32{F D Q} (HP/SP,DP,QP,F1 Pt)	+RV{64,128}
Move	Move from Integer	R	FMV.{H S}.X rd,rs1	FMV.{D Q}.X rd,rs1
	Move to Integer	R	FMV.X.{H S} rd,rs1	FMV.X.{D Q} rd,rs1
Convert	Convert from Int	R	FCVT.{H S D Q}.W rd,rs1	FCVT.{H S D Q}.{L T} rd,rs1
	Convert from Int Unsigned	R	FCVT.{H S D Q}.WU rd,rs1	FCVT.{H S D Q}.{L T}U rd,rs1
	Convert to Int	R	FCVT.W.{H S D Q} rd,rs1	FCVT.{L T}.{H S D Q} rd,rs1
	Convert to Int Unsigned	R	FCVT.WU.{H S D Q} rd,rs1	FCVT.{L T}U.{H S D Q} rd,rs1

RISC-V Calling Convention				
Register	ABI Name	Saver	Description	
x0	zero	---	Hard-wired zero	
x1	ra	Caller	Return address	
x2	sp	Callee	Stack pointer	
x3	gp	---	Global pointer	
x4	tp	---	Thread pointer	
x5-x7	t0-2	Caller	Temporaries	
x8	s0/fp	Callee	Saved register/frame pointer	
x9	s1	Callee	Saved register	
x10-x11	a0-1	Caller	Function arguments/return values	
x12-x17	a2-7	Caller	Function arguments	
x18-x27	s2-11	Callee	Saved registers	
x28-x31	t3-t6	Caller	Temporaries	
f0-7	ft0-7	Caller	FP temporaries	
f8-9	fs0-1	Callee	FP saved registers	
f10-11	fa0-1	Caller	FP arguments/return values	
f12-17	fa2-7	Caller	FP arguments	
f18-27	fs2-11	Callee	FP saved registers	
f28-31	ft8-11	Caller	FP temporaries	

RISC-V calling convention and five optional extensions: 10 multiply-divide instructions (RVM); 11 optional atomic instructions (RV32A); and 25 floating-point instructions each for single-, double-, and quadruple-precision (RV32F, RV32D, RV32Q). The latter add registers f0-f31, whose width matches the widest precision, and a floating-point control and status register fcsr. Each larger address adds some instructions: 4 for RVM, 11 for RVA, and 6 each for RVF/D/Q. Using regex notation, {} means set, so L{D|Q} is both LD and LQ. See riscv.org. (8/21/15 revision)