

Theory Part

Thursday, February 17, 2022 10:35 AM

CDB:
• all Func Units write their result here as soon as an instruction is completely processed
• all reservation stations waiting for operand, register file, ROB read from here
• having multiple CDBs will allow multiple writes in one cycle, hence improving performance (but cost increased)

Loop Unrolling:
• static technique -> reduce # of iterations
• loop body modified by compiler
adv:
• # of branches reduced
• increased loop size -> increased chance of ILP identification
disadv:
• increased code size

Example:
for (x=0;x<MAX_X+1;) {
 y[0] = x+1;
 for (y=0;y<MAX_Y-4;) {
 y[0] = x+1;
 y[1] = x+1;
 y[2] = x+2+1;
 y[3] = x+3+1;
 }
}

structural resource conflicts
• pipeline unit can exec all op in given cc
• is 1-reg-file write port, req 2 reg writes
• is single-port mem, two instr access mem together
data
int depending on res of prev instr
• dependence b/w instr and they are close
• fixed by fwding and stall
control
dep on branches and other instr that may change pc after the instr is fetched already
• stall at decode, decide early if branch taken or not, compute new PC value
• delayed branch

BTB:
1. Architecture
• n entries
• 2 x 32-bit fields i.e Address and Target
• one additional bit (optional)
2. Behaviour
• accessed on each Fetch
• entry selected $i > \log(n)$ bits (log is base 2)
• address field of entry & fetched instr compared; if matched, target field uploaded in PC
• btb possibly if branch taken updated when instr is completed

BHT
1. Architecture
• N elements of M bits each (1 or 2)
2. Behaviour
• Accessed at decode stage of a conditional branch instr
• using log (n) bits while ignoring alignment bits
• entry predicts taken or not taken
• based on previous execution of the same instr
• if res of branch is known, BHT is possibly updated

ROB
• ROB is crucial to allow in-order instruction commitment while execution is performed out-of-order
• committed buffer are composed of the following fields:
• instruction type
• value
• branch
Warning:
• The ROB is allocated to an instruction when it is successfully issued.
• If a commit is written into the ROB while the execution finishes its execution, then the ROB is corrupted.
• Committing a PBO is a circular buffer. When an instruction becomes the oldest in the ROB, it is overwritten (i.e., it is moved to the destination and the entry is被淘汰).
• If the ROB is full, then the processor has to implement a mechanism to evict the oldest instruction.
Advantage
• The ROB allows the processor to implement dynamic scheduling with speculation. It also allows precise exception management.

Feature	ARM LPC1768	Intel 8086
Architecture	RISC	CISC
Data bus width	32 bits	16 bits
Endianness	Little-endian	Little-endian
Clock speed	Up to 100 MHz	Up to 10 MHz
Memory addressing	Up to 64 MB	Up to 1 MB
Register file size	16 general-purpose registers	8 general-purpose registers
Interrupt handling	Uses vectored interrupts	Uses interrupt vector table
I/O interface	Uses GPIO pins for I/O	Uses separate I/O address space

Here are some key points about the address bus:

- Memory Addressing:** The primary function of the address bus is to facilitate memory addressing. It allows the CPU (Central Processing Unit) to specify the exact location in the computer's memory where data or instructions are to be read from or written to.
- Size:** The width of the address bus determines the maximum number of unique memory locations or addresses that the computer can access. For example, a 16-bit address bus can address 2^{16} (65,536) different memory locations, while a 32-bit address bus can address 2^{32} (approximately 4.3 billion) different locations.
- Binary Representation:** Addresses on the address bus are represented in binary format. Each wire or line in the address bus carries a binary digit (bit), and the combination of bits on all the lines forms the binary address.
- Addressable Memory:** The size of the address bus also dictates the maximum amount of addressable memory. For example, a 32-bit address bus can address up to 4 GB (gigabytes) of memory because it can specify 2^{32} different memory locations.
- Expansion:** In systems with larger memory requirements, the width of the address bus may be expanded to accommodate more memory. For instance, 64-bit processors have a 64-bit address bus, allowing them to address much larger memory spaces.
- Address Decoding:** The computer's memory management hardware uses the address bus to decode memory addresses and route data requests to the appropriate memory cells or storage locations.
- Input/Output (I/O) Addressing:** In addition to memory addressing, the address bus can also be used for addressing I/O devices and ports. This allows the CPU to communicate with peripherals and external hardware components.
- Control Signals:** The address bus is often accompanied by control signals, which include read and write signals, to indicate the direction of data transfer and control memory access.
- Compatibility:** The width of the address bus is an important factor in determining the compatibility of hardware components. For example, a CPU with a 32-bit address bus is typically compatible with memory modules designed for 32-bit addressing.

- 1. Associative Memory:**
• Associative memory, also known as content-addressable memory (CAM), works by searching for data based on its content rather than its address.

Feature	CISC	RISC
Philosophy	Complex instruction set	Reduced instruction set
Instruction length	Variable	Fixed
Number of instructions	Large	Small
Execution cycle	One cycle per instruction	Multiple cycles per instruction
Execution time	Longer execution time per instruction	Shorter execution time per instruction
Register usage	Memory-to-memory operations	Register-to-register operations
Memory access	Direct memory access	Load/store architecture
Pipeline architecture	More difficult to pipeline	Easier to pipeline
Code density	High	Low
Power consumption	Higher	Lower
Instruction set	Rich and complex	Simplified and limited
Applications	Desktops, servers, mainframes	Embedded systems, mobile devices

Aspect	Monolithic Architecture	Microkernel Architecture
Architecture	Single tightly integrated unit	Modular design with cores services in kernel
Philosophy	More complex	Less complex
Reliability and Stability	Prone to system-wide failures	More resilient to failures
Performance	Lower communication overhead	Potentially higher communication overhead
Flexibility and Extensibility	Challenging to extend and modify	Easier to add/replace components

Aspect	Superscalar Architecture	Multi-core Architecture
Parallel Execution	Executes multiple instructions within a single core in parallel during a clock cycle.	Has multiple independent processor cores that execute instructions simultaneously.
Instruction Dispatch	Dispatches multiple instructions to different execution units within a single core.	Dispatches different threads or tasks to separate cores, each with its own resources.
Independence	Parallel execution limited to a single core; instructions must be independent within that core.	True parallelism achieved by having multiple independent cores, allowing for concurrent execution.
Usage Scenario	Effective for improving single-threaded performance and instruction throughput within a single core.	Effective for handling multiple threads or tasks simultaneously, benefiting multi-threaded applications and multitasking.

Data Bus:

cc:	Instruction	j	k	Issue	Execution complete	Write Result
	LD F6	34+	R2	1	3	4
add 2	LD F2	45+	R3	2	5	6
mul 10	MULT F0	F2	F4	3	16	17
div 40	SUBD F8	F6	F2	4	8	9
ld 2	DIVD F10	F0	F6	5	58	59
	ADDD F6	F8	F2	6	12	13

1	L_D	F6, 34(R2)
2	L_D	F2, 45(R3)
3	MUL_D	F0, F2, F4
4	SUB_D	F8, F2, F6
5	DIV_D	F10, F0, F6
6	ADD_D	F6, F8, F2

- Assumptions:**
• Add takes 2 clock cycles, Multiply 10, Divide 40
• Loads requires 2 ccs

#iteration	ISSUE	EXE	MEM	CDB	COMMIT
1	I.d f1.v1(r1)	2	3	4	5
1	I.d f2.v2(r1)	1	3	4	5
1	I.d f3.v3(r1)	2	2	X	5
1	divd.f5, f3, f11	3	6	X	8
1	subd.f4, f1, f2	3	16	X	18
1	add.d f6, f4, f5	3	16	X	18
1	divd.f7,f6,f12	4	23	X	31
1	s.d f7,v4(r1)	4	5m		
1	daddui.r1,r1,8	5	6i		
1	addir.r2,r2,-1	5	7i		
1	bnez r2,loop	6	10j		
2	I.d f1.v1(r1)	7	8m	9	10
2	I.d f2.v2(r1)	7	9m	10	11
2	I.d f3.v3(r1)	8	10m	11	12
2	div.d f5, f3, f11	8	15d		
2	sub.d f4, f1, f2	9	12a		
2	add.d f6, f4, f5	9	24a		
2	div.d f7,f6,f12	10	31d		
2	s.d f7,v4(r1)	10	11m		
2	daddui.r1,r1,8	11	12i		
2	addir.r2,r2,-1	11	13i		
2	bnez r2,loop	12	17j		

→ issue 9 first now
→ exe of b2 & cd
→ sequence in column
→ greater than issue
→ Mem 'X' : ALU, Store, Branch
→ CDB → EXE 1 in MEM & MEM is in CDB
→ F.P ALU : EXE & CDB
→ other ALU & Branch.

#iteration	ISSUE	EXE	MEM	CDB	COMMIT	Notes
1	I.d f1.v1(r1)	1	2m	3	4	
1	I.d f2.v2(r1)	1	3m	4	5	
1	I.d f3.v3(r1)	2	4m	5	6	
1	div.d f5, f3, f11	2	7d		15	Wait for f3
1	sub.d f4, f1, f2	3	6a		8	Wait for f2
1	add.d f6, f4, f5	3	16a		18	Wait for f5
1	div.d f7,f6,f12	4	23d		31	Wait for f6, then for div unit
1	s.d f7,v4(r1)	4	5m			
1	daddui.r1,r1,8	5	6i			
1	addir.r2,r2,-1	5	7i			
1	bnez r2,loop	6	10j			
2	I.d f1.v1(r1)	7	8m	9	10	
2	I.d f2.v2(r1)	7	9m	10	11	
2	I.d f3.v3(r1)	8	10m	11	12	
2	div.d f5, f3, f11	8	15d		23	Wait for f3, then for div unit
2	sub.d f4, f1, f2	9	12a		14	Wait for f2
2	add.d f6, f4, f5	9	24a		26	Wait for f5
2	div.d f7,f6,f12	10	31d		39	Wait for f6, then for div unit
2	s.d f7,v4(r1)	10	11m			
2	daddui.r1,r1,8	11	12i			
2	addir.r2,r2,-1	11	13i			
2	bnez r2,loop	12	17j			

Address	Code	BHT	Iteration #1	Iteration #2	Iteration #3	Iteration #4
0x0000	dadui r1, r0, 0					
0x0004	dadui r2, r0, -1					
0x0008	dadui r3, r0, 0					
0x000c	sbz r3, r0, 0					
0x0010	cyc	dadui r2, r2, 1				
0x0014	lb r3, r0, 0					
0x0018	dadui r4, r4, 0					
0x0020	divd r5, r5, 1					
0x0024	divd r6, r6, 1					
0x0028	j nx					
0x002c	j nx					
0x0030	eq0					
0x0034	eq1					
0x0038	nxt					
0x0042	beq r5, r5, 0					
0x0046	beq r6, r6, 1					
0x004a	beq r7, r7, 0					
0x004e	term -ib r5, r0, 0					
0x004f	term -ib r6, r0, 0					
0x0054	halt					

compatibility of hardware components. For example, a CPU with a 32-bit address bus is typically compatible with memory modules designed for 32-bit addressing.

1. Associative Memory:

- Associative memory, also known as content-addressable memory (CAM), works by searching for data based on its content rather than its address.

- If stores data along with associated tags or keywords.

- When you query the memory with a specific content or keyword, it searches through the stored data to find a match, returning the associated data.

- This type of memory is often used in applications like caching and hardware routing tables to quickly locate information based on its content.

2. Cross Compiler:

- A cross compiler is a software tool that generates executable code for a platform different from the one on which the compiler itself runs.

- It is commonly used in embedded systems and cross-development environments.

- Cross compilers are necessary when you want to develop software for a target platform (e.g., an embedded system or a different architecture) that is not compatible with the host system where you write the code.

- The cross compiler translates the source code into machine code that can run on the target platform, allowing developers to create software for various environments without needing a dedicated compiler for each.

Design Scenario	Directive for improving single-threaded performance and instruction throughput within a single core.	Directive for handling multiple threads or tasks simultaneously, benefiting multi-threaded applications and multitasking.
-----------------	--	---

Data Bus:

- The data bus is another critical component of a computer's system bus alongside the address bus and control bus.

- It is a set of parallel lines or wires that allow data to be transferred between various parts of the computer, such as the CPU, memory, and peripherals.

- The width of the data bus determines how many bits of data can be transferred simultaneously. For example, a 32-bit data bus can carry 32 bits of data at a time.

- The data bus is bidirectional, meaning it can carry data in both read and write directions, facilitating the flow of information within the computer.

- Data buses play a central role in tasks like reading data from memory, writing data to memory, exchanging information with peripherals, and performing arithmetic and logic operations.

Control Bus:

- The control bus, like the data bus, is a set of parallel lines or wires within the computer's system bus.

- Its purpose is to carry control signals and commands that govern the operation of various components within the computer system.

- Control signals on the control bus include signals like "read," "write," "interrupt request," "clock," and various control codes that coordinate activities between the CPU, memory, and peripherals.

- The control bus plays a crucial role in activities such as fetching instructions from memory, signaling when data is ready to be read from or written to memory, and managing input/output operations.

- Control signals on the control bus help ensure that the computer's components work together seamlessly, allowing for orderly and synchronized data flow and operation.

0x001c	beq r6, r0, eq0					
0x0020	beq r6, r1, eq1					
0x0024	j r0					
0x0028	eq0:	daddiu r3, r3, 1				
0x002c	lwt					
0x0030	cgt:	daddiu r4, r4, 1				
0x0034	nxt:	daddiu r2, r2, 1				
0x0038	j cyc					
0x003c	term:	sb r5, res0(2)				
0x0040		sb r4, res1(2)				
0x0044	halt					

Address	Code	BHT	0		0		1		Iteration #8	
			entry #	prediction	result	prediction	result	prediction	result	prediction
0x0000	daddiu r1, r0, 1									
0x0004	daddiu r2, r0, 0									
0x0008	daddiu r3, r0, 0									
0x000c	daddiu r4, r0, 0									
0x0010	daddiu r5, r0, 6									
0x0014	cyc:	beq r2, r5, term								
0x0018		lb r6, vec(r2)								
0x001c		beq r6, r0, eq0								
0x0020		beq r6, r1, eq1								
0x0024	j r0									
0x0028	eq0:	daddiu r3, r3, 1								
0x002c	lwt									
0x0030	cgt1:	daddiu r1, r4, 1								
0x0034	nxt:	daddiu r2, r2, 1								
0x0038	j cyc									
0x003c	term:	sb r5, res0(2)								
0x0040		sb r4, res1(2)								
0x0044	halt									

Address	Code	BHT	0		0		1		Iteration #8	
			entry #	prediction	result	prediction	result	prediction	result	prediction
0x0000	daddiu r1, r0, 1									
0x0004	daddiu r2, r0, 0									
0x0008	daddiu r3, r0, 0									
0x000c	daddiu r4, r0, 0									
0x0010	daddiu r5, r0, 6									
0x0014	cyc:	beq r2, r5, term								
0x0018		lb r6, vec(r2)								
0x001c		beq r6, r0, eq0								
0x0020		beq r6, r1, eq1								
0x0024	j r0									
0x0028	eq0:	daddiu r3, r3, 1								
0x002c	lwt									
0x0030	cgt1:	daddiu r1, r4, 1								
0x0034	nxt:	daddiu r2, r2, 1								
0x0038	j cyc									
0x003c	term:	sb r5, res0(2)								
0x0040		sb r4, res1(2)								
0x0044	halt									

0x002c	beq r6, r0, eq0					
0x0030	daddiu r1, r5, 1					
0x0034	beq r6, r1, eq1					
0x0038	addiu r2, r0, 0					
0x0040	addiu r3, r0, 0					
0x0044	term:	sb r5, res0(0)				
0x0048		sb r6, res0(0)				
0x004c	halt					

Address	Code	BHT	0		0		1		Iteration #8	
			entry #	prediction	result	prediction	result	prediction	result	prediction
0x0000	daddiu r1, r0, 1									
0x0004	daddiu r2, r0, 0									
0x0008	daddiu r3, r0, 0									
0x000c	daddiu r4, r0, 0									
0x0010	daddiu r5, r0, 6									
0x0014	cyc:	beq r2, r5, term								
0x0018		lb r6, vec(r2)								
0x002c		beq r6, r0, eq0								
0x0030		beq r6, r1, eq1								
0x0034	daddiu r2, r0, 0									
0x0038	j cyc									
0x003c	term:	sb r5, res0(0)								
0x0040		sb r6, res0(0)								
0x0044	halt									

final content
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15

mis-predicted

main: daddiu r1,r0,0	f d e m w									
0x0000	daddiu r2,r0,30	f d e m w								
loop: 1d f1,v1(r1)	f d e m w									
0x0004	daddiu r3,r1,r0,0	f d e m w								
0x0008	daddiu r4,r1,r0,0	f d e m w								
0x000c	daddiu r5,r1,r0,0	f d e m w								
0x0010	daddiu r6,r1,r0,0	f d e m w								
0x0014	daddiu r7,r1,r0,0	f d e m w								
0x0018	daddiu r8,r1,r0,0	f d e m w								
0x002c	daddiu r9,r1,r0,0	f d e m w								
0x0030	daddiu r10,r1,r0,0	f d e m w								
0x0034	daddiu r11,r1,r0,0	f d e m w								
0x0038	daddiu r12,r1,r0,0	f d e m w								
0x0040	daddiu r13,r1,r0,0	f d e m w								
0x0044	daddiu r14,r1,r0,0	f d e m w								
0x0048	daddiu r15,r1,r0,0	f d e m w								
0x004c	daddiu r16,r1,r0,0	f d e m w								
0x0050	daddiu r17,r1,r0,0	f d e m w								
0x0054	daddiu r18,r1,r0,0	f d e m w								
0x0058	daddiu r19,r1,r0,0	f d e m w								
0x005c	daddiu r20,r1,r0,0	f d e m w								
0x0060	daddiu r21,r1,r0,0	f d e m w								
0x0064	daddiu r22,r1,r0,0	f d e m w								
0x0068	daddiu r23,r1,r0,0	f d e m w								
0x0072	daddiu r24,r1,r0,0	f d e m w								
0x0076	daddiu r25,r1,r0,0	f d e m w								
0x0080	daddiu r26,r1,r0,0	f d e m w								
0x0084	daddiu r27,r1,r0,0	f d e m w								
0x0088	daddiu r28,r1,r0,0	f d e m w								
0x0092	daddiu r29,r1,r0,0	f d e m w								
0x0096	daddiu r30,r1,r0,0	f d e m w								
0x00a0	daddiu r31,r1,r0,0	f d e m w								
0x00a4	daddiu r32,r1,r0,0	f d e m w								
0x00a8	daddiu r33,r1,r0,0	f d e m w</td								

Iteration	issue	exe	mem	cdb	commit
1	l.d f1,v1(r1)	1	2	3	4
1	l.d f2,v2(r1)	1	3	4	5
1	l.d f3,v3(r1)	2	4	5	6
1	l.divd f5,r3,r1	2	7	x	15
1	l.addd f4,r1,r2	3	8	x	8
1	l.addd f4,r1,r5	3	16	x	18
1	l.divd f7,f6,r2	4	23	x	31
1	s.d f7,v4(r1)	4	5	x	33
1	daddui r1,r1,8	5	6	x	7
1	daddi r2,r2,1	5	7	x	9
1	bnez r2,loop	6	10	x	x
2	l.d f1,v1(r1)	7	8	9	10
2	l.d f2,v2(r1)	7	9	10	11
2	l.d f3,v3(r1)	6	10	11	38
2	l.divd f5,r2,r1	6	12	12	39
2	l.divd f5,f2,r1	8	15	x	23
2	s.add d f4,f1,f2	9	12	x	14
2	s.add d f6,f4,f5	9	24	x	26
2	l.divd f7,f6,r2	10	31	x	39
2	s.d f7,v4(r1)	10	11	x	44
2	daddui r1,r1,8	11	12	x	13
2	daddi r2,r2,1	11	13	x	16
2	bnez r2,loop	12	x	x	47

First name, Last name, ID.....									
		c	e3	o	75	d	64	l	69
Address	Code	BHT	Iteration #5	prediction	result	Iteration #6	prediction	result	Iteration #4
0x0000	daddui r1,r0,1								
0x0004	daddi r2,r0,0								
0x0008	daddi r3,r0,0								
0x000c	daddi r4,r0,0								
0x0010	daddi r5,r0,6								
0x0014	cyc:								
0x0018	lb r5, vec(r2)								
0x0020	beq r6,r1,eq0	8	NT	T	T	NT	NT	NT	NT
0x0024	1 nxz								
0x0028	eq0:								
0x002c	1 nxz								
0x0030	eq1:								
0x0034	nxt:								
0x0038	1 cyc								
0x003c	term:								
0x0040	sb r3, res0(r2)								
0x0044	sb r4, res1(r2)								
halt									

The number of mispredicted branches during the execution of the code is: 6

BHT - Final content

Entry 0	0	Entry 4	0
Entry 1	0	Entry 5	0
Entry 2	0	Entry 6	1
Entry 3	1	Entry 7	0

Address	Code	BHT	Iteration #1		Iteration #2		Iteration #3		Iteration #4	
			entry #	prediction	result	prediction	result	prediction	result	prediction
0x0000	daddui r1,r0,1									
0x0004	daddi r2,r0,0									
0x0008	daddi r3,r0,0									
0x000c	daddi r4,r0,0									
0x0010	daddi r5,r0,6									
0x0014	cyc:									
0x0018	lb r5, vec(r2)									
0x0020	beq r6,r1,eq0	8	NT	T	T	T	NT	NT	NT	T
0x0024	1 nxz									
0x0028	eq0:									
0x002c	1 nxz									
0x0030	eq1:									
0x0034	nxt:									
0x0038	1 cyc									
0x003c	term:									
0x0040	sb r3, res0(r2)									
0x0044	sb r4, res1(r2)									
halt										

The number of mispredicted branches during the execution of the code is: 6

BHT - Final content

Entry 0	0	Entry 4	0
Entry 1	0	Entry 5	1
Entry 2	0	Entry 6	0
Entry 3	0	Entry 7	0
Entry 8	1	Entry 9	0
Entry 10	0	Entry 11	0
Entry 12	0	Entry 13	0
Entry 14	0	Entry 15	0