

计算机组织与体系结构

第十七讲

计算机科学与技术学院

舒燕君

Recap

- 相关性
 - ✓ 数据相关（RAW）
 - ✓ 名相关（反相关WAR，输出相关WAW）
 - ✓ 控制相关
- 指令动态调度
 - ✓ 动态调度的目的、优点及缺点
 - ✓ 译码段再划分为流出和读操作数两段
 - ✓ 乱序执行
- 记分牌
 - ✓ 基本原理（集中控制指令的流出和执行）
 - ✓ 执行过程（流出、读操作数、执行和写结果）
 - ✓ 记分牌的结构（指令状态表、功能部件状态表、结果寄存器状态表）

7.2 指令的动态调度

7.2.1 动态调度的原理

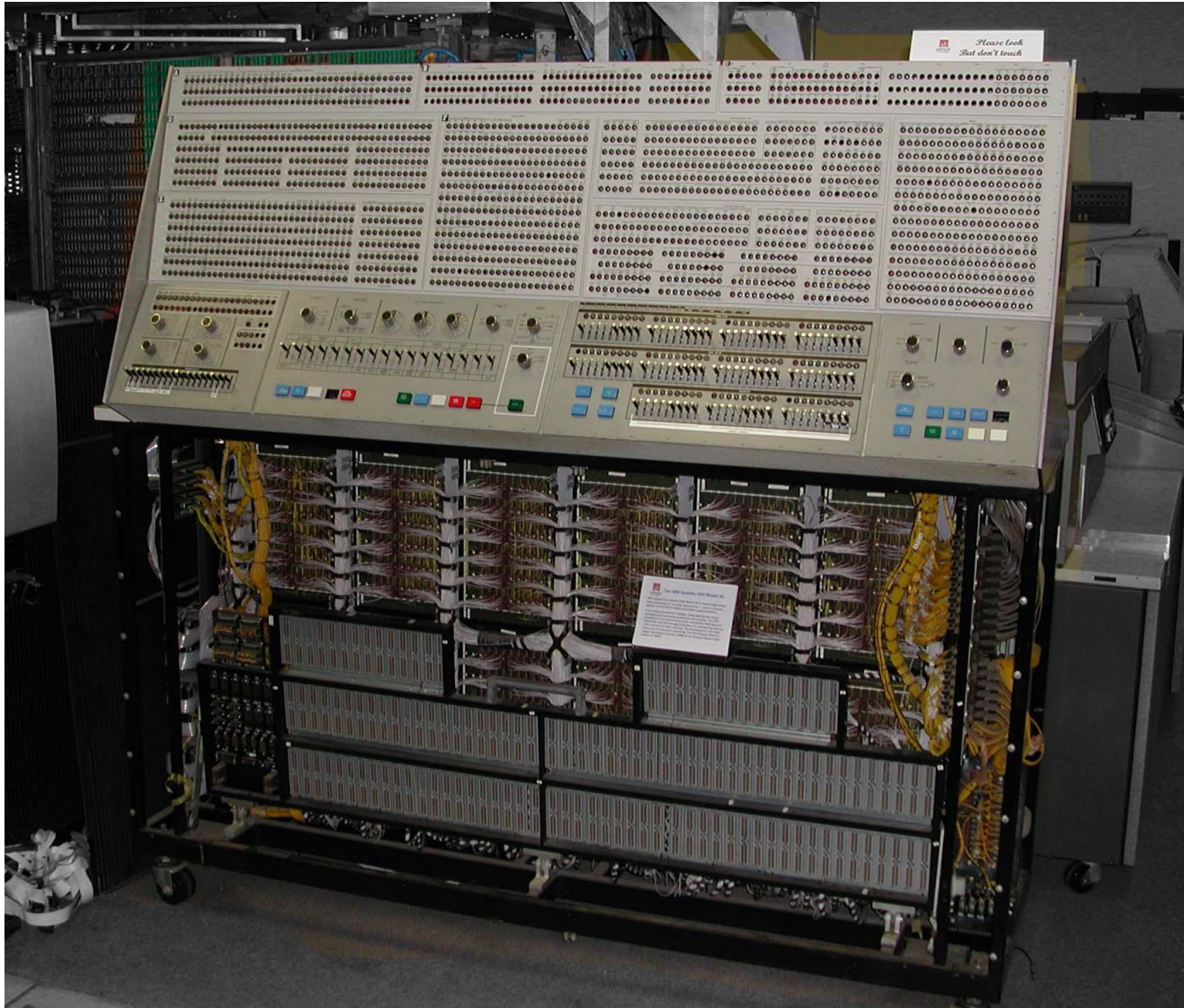
7.2.2 动态调度算法之一：记分牌

7.2.3 动态调度算法之二：Tomasulo算法

7.2.2 动态调度算法之二：Tomasulo算法

- IBM 360/91比CDC6600晚三年推出
 - 商业计算机使用Cache技术之前
- 整个360系列仅一个指令系统和一个编译器
 - 要求具有很高的浮点性能，但不是通过高端机器的专用的编译器实现
 - 只有四个双精度浮点寄存器，编译器调度的有效性受到很大限制
 - 访存时间和浮点计算时间都很长
 - 可支持循环的多次迭代重叠执行

IBM 360/91前面板



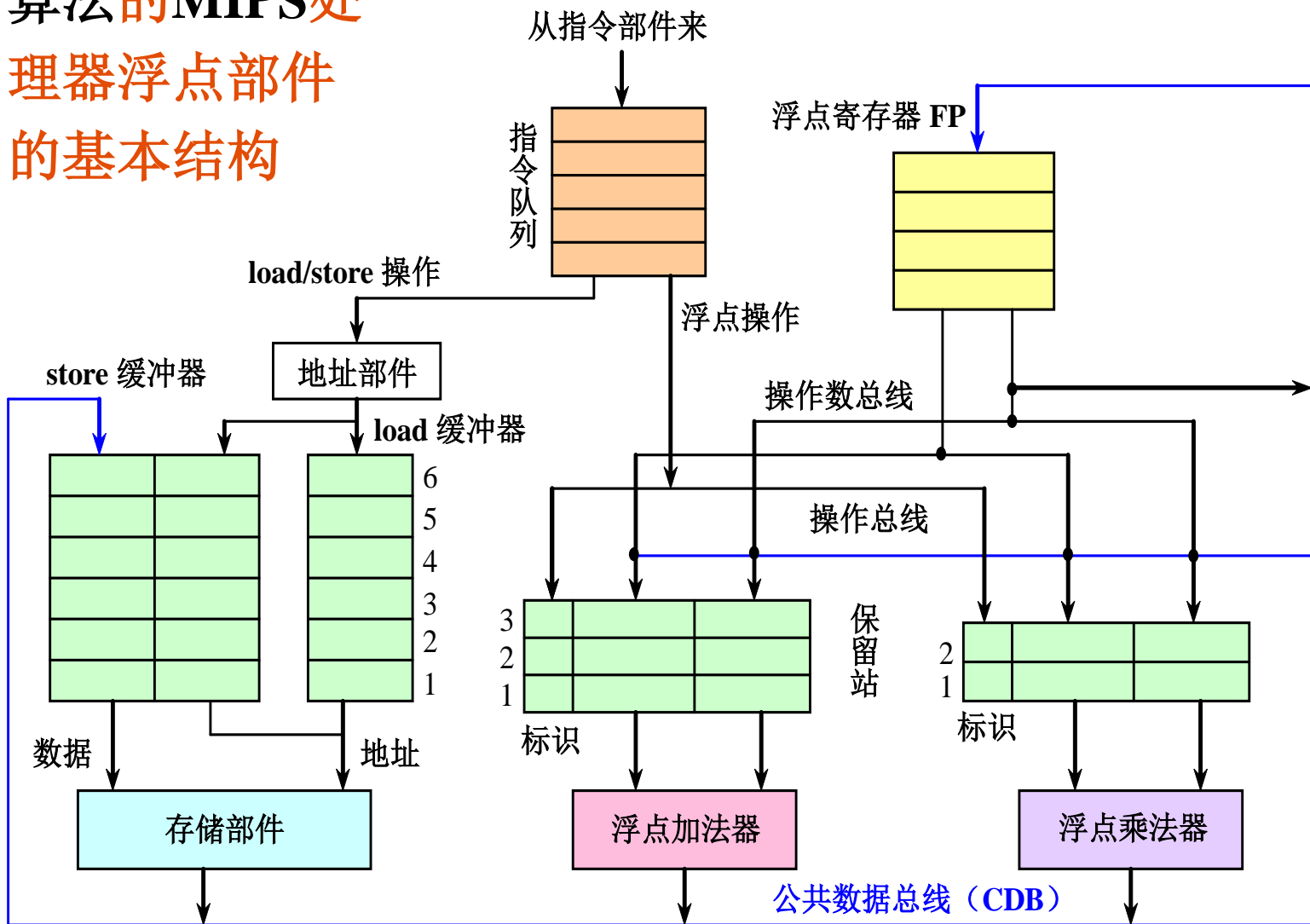
Tomasulo算法与记分牌

- 采用了许多记分牌中的理念
- 两个较大的差异
 - Tomasulo算法中，冲突检测和执行控制是分布的，利用保留站实现。
 - Tomasulo算法不检查WAR和WAW相关，通过算法本身消除。计算结果通过专用通道直接从功能部件进入到保留站进行缓冲，而不是写到寄存器。

Tomasulo算法核心思想

- 记录并检测指令的**RAW**相关，操作数一旦准备就绪就立即执行，把发生**RAW**冲突的可能性降到最低。
- 不检查**WAR**和**WAW**相关，通过换名（**Renaming**）技术来消除**WAR**和**WAW**冲突（注意：Tomasulo算法利用保留站缓冲实现换名）。

➤ 基于Tomasulo
算法的MIPS处
理器浮点部件
的基本结构



MIPS五阶段的流水线的改造

- ID和EX阶段被以下三个阶段代替：
 - 1、流出（Issue）
 - 2、执行（Execute）
 - 3、结果写回（Write result）

Tomasulo 算法的三阶段

1. Issue: 从指令队列中取指令

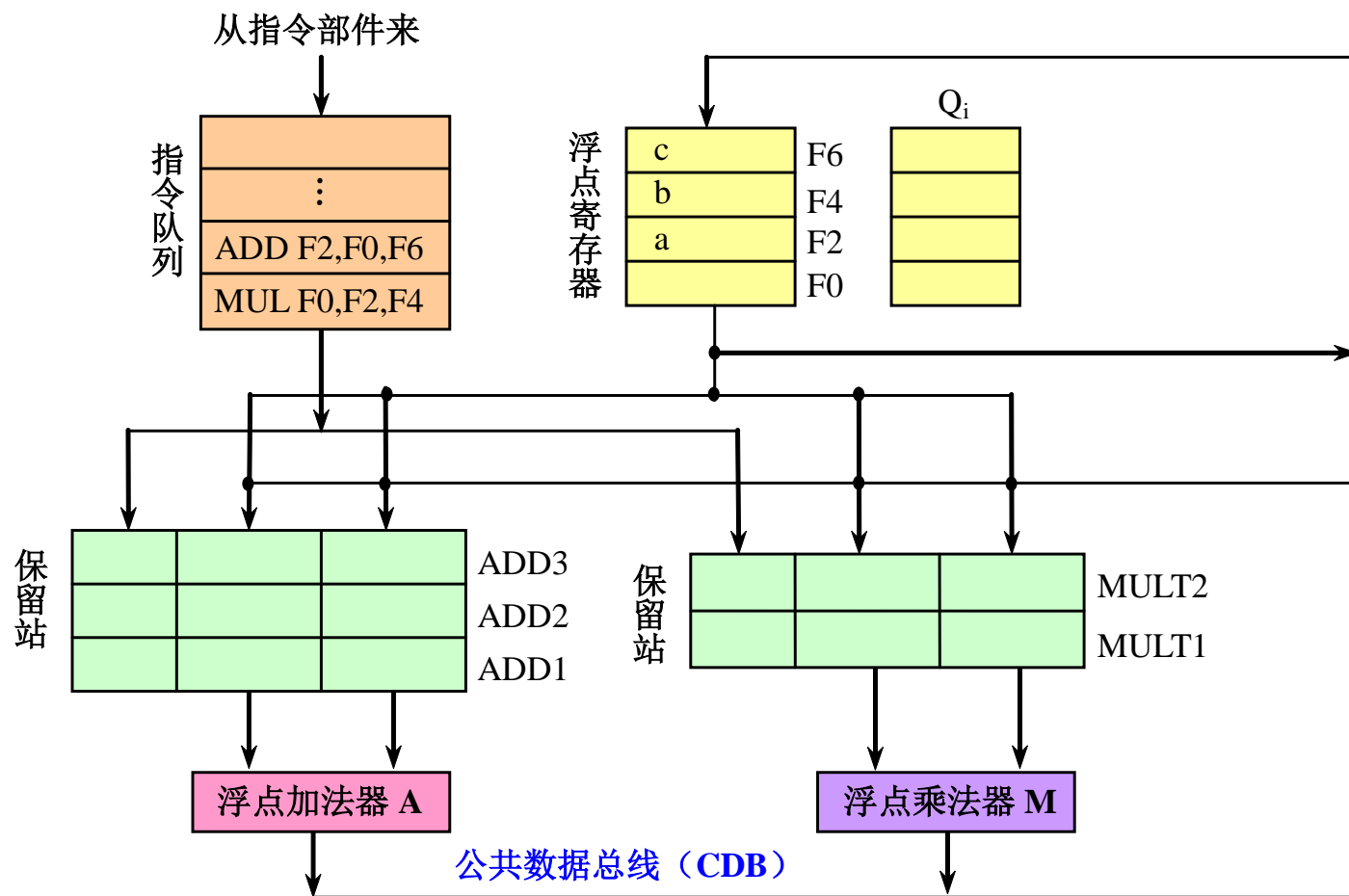
如果存在一个空闲的保留站 (no structural hazard), 则控制发射指令和操作数 (renames registers), 消除WAR, WAW相关。

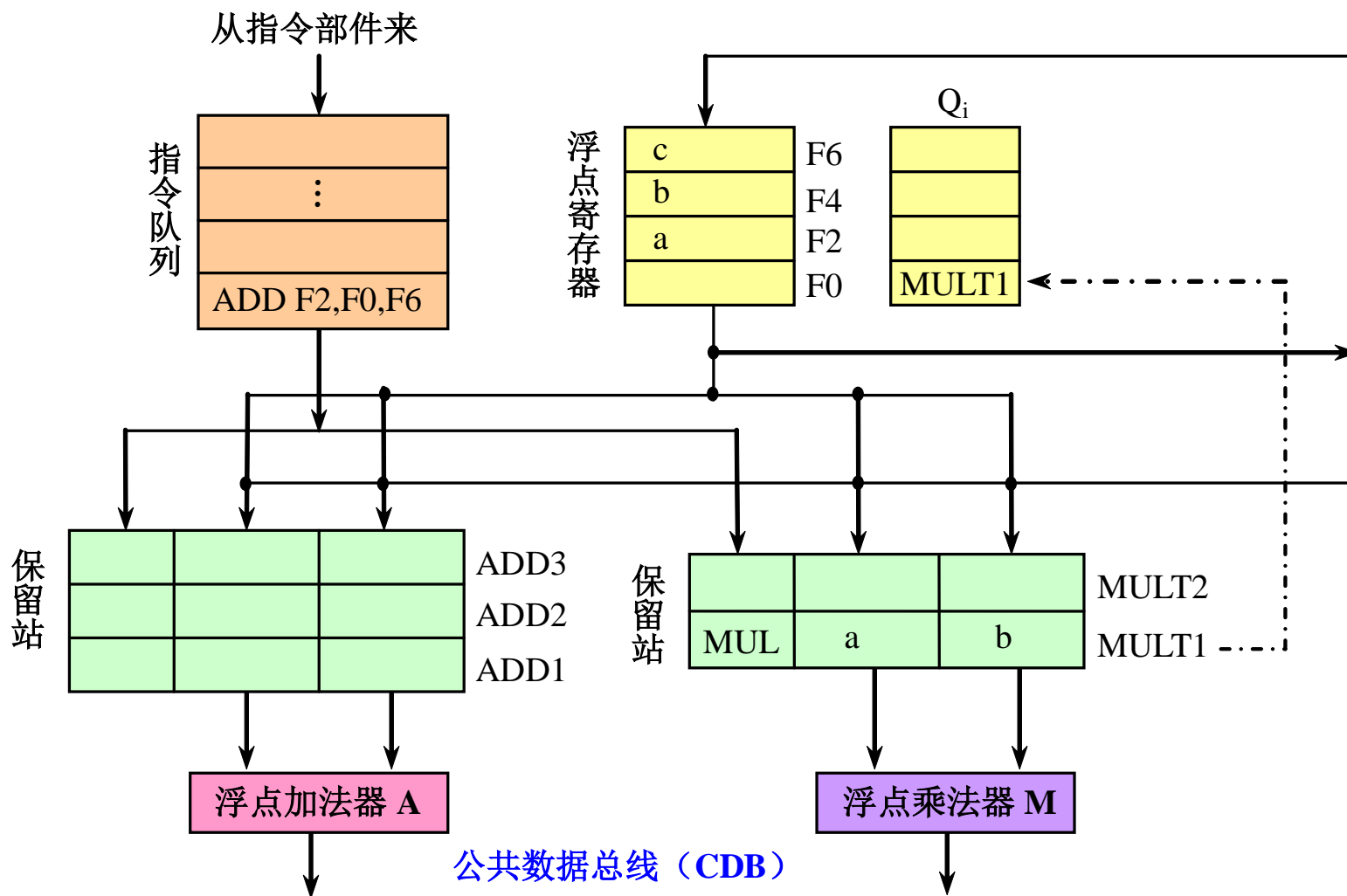
2. Execution: operate on operands (EX)

当两操作数就绪后, 就可以执行;
如果没有准备好, 则监测Common Data Bus 以获取结果;
通过推迟指令执行避免RAW相关。

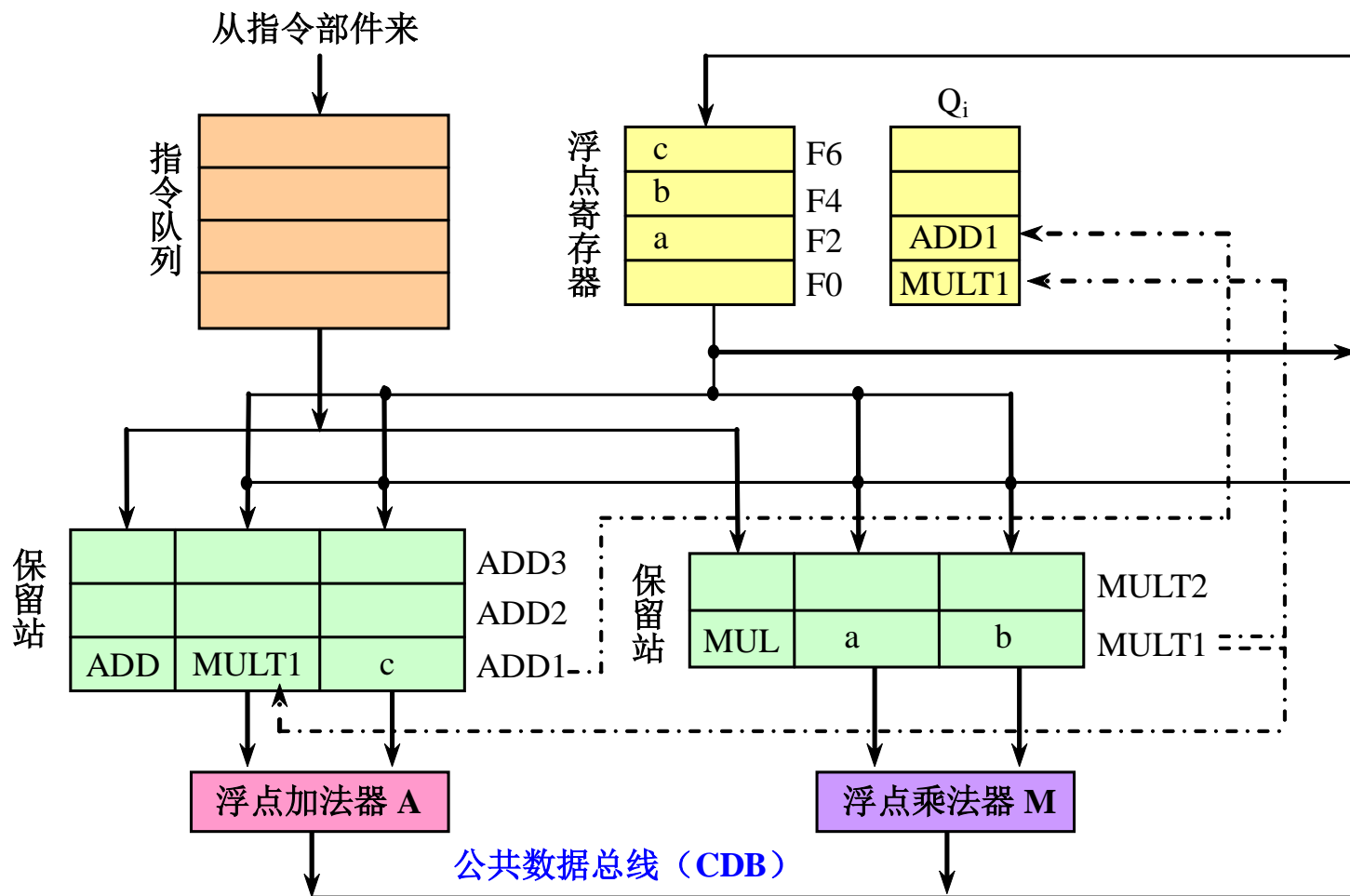
3. Write result: finish execution (WB)

将结果通过Common Data Bus传给所有等待该结果的部件。

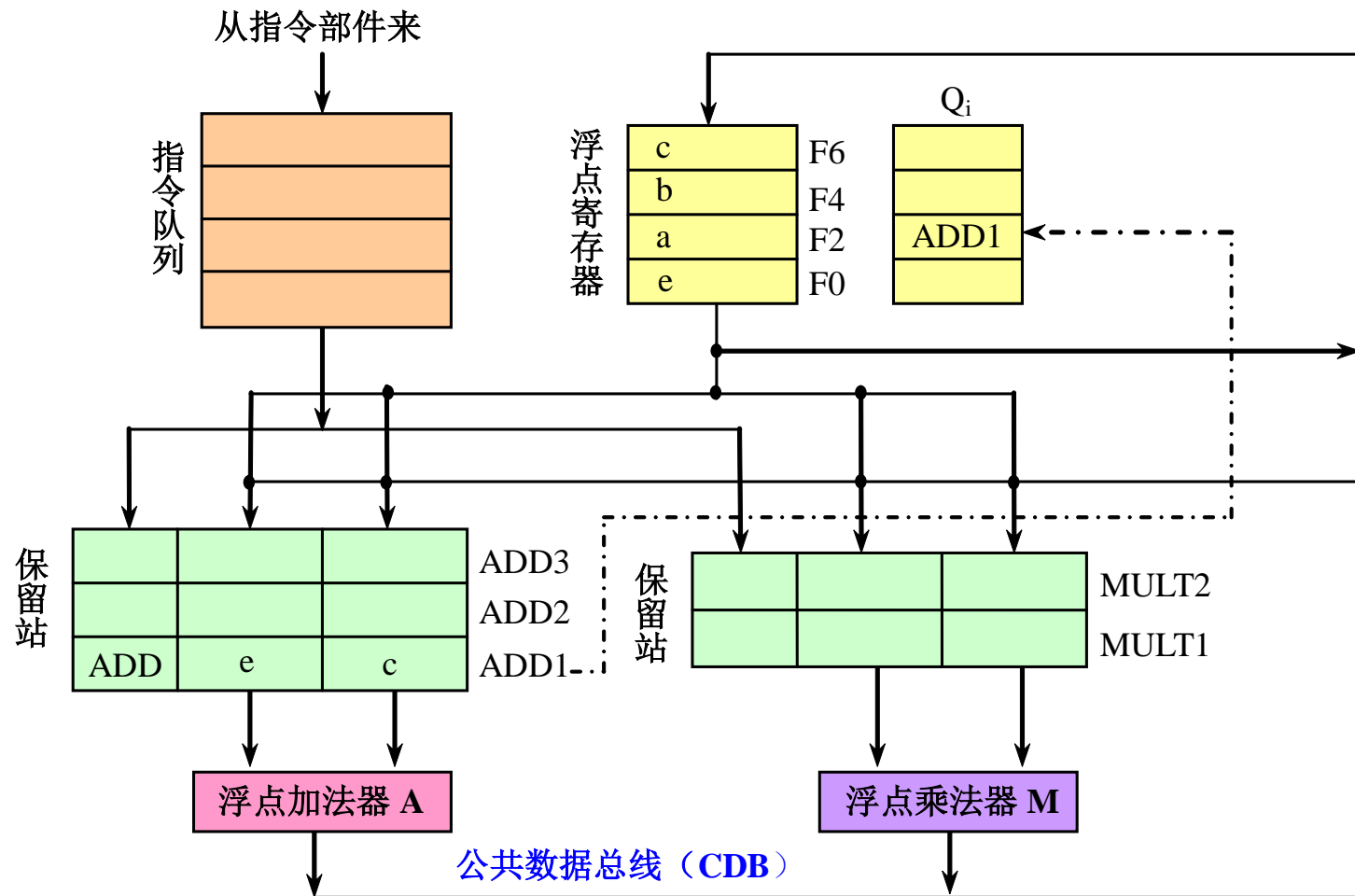




指令队列里的第一条指令流出



指令队列里的第一条指令执行，第二条指令流出



指令队列里的第一条指令写结果，结果同时送给其他需要的保留站

与记分牌的不同

- 无须任何操作来检查数据的**写后写（WAW）**和**先读后写（WAR）**冲突，在指令流出的过程中，操作数寄存器号换成操作数本身（若已就绪）或者相应的保留站标识。
- 通过公共数据总线来广播结果，将计算结果直接从产生它的保留站传送到所有需要它的功能部件，而不用经过寄存器（动态解决了**RAW**）。
- 存储器存和取都作为基本的功能部件。
- 由于保留站技术能够有效地解决先写后读，记分牌中的用于判断**RAW**的“取操作数”段被取消。

Reservation Station 结构

- **Op**: 部件所进行的操作
- **Vj, Vk**: 源操作数的值。Store 缓冲区有Vk域，用于存放要写入存储器的值
- **A**: 存放存储器地址。开始存立即数，计算出有效地址后，存放有效地址
- **Qj, Qk**: 产生源操作数的保留站号
 - ✓ 没有记分牌中的准备就绪标志， $Qj, Qk=0 \Rightarrow \text{ready}$
 - ✓ Store 缓存区只有Qk表示产生结果的保留站号
- **Busy**: 标识保留站RS或相应的功能部件FU是否空闲

Register result status—如果存在对寄存器的写操作，指示对该寄存器进行写操作的部件。

Qi: 保留站的编号

Tomasulo实例

Instruction status:

<i>Instruction status:</i>				<i>Exec</i>	<i>Write</i>		
Instruction	<i>j</i>	<i>k</i>	<i>Issue</i>	<i>Comp</i>	<i>Result</i>		Busy Address
LD	F6	34+	R2			Load1	No
LD	F2	45+	R3			Load2	No
MULTD	F0	F2	F4			Load3	No
SUBD	F8	F6	F2				
DIVD	F10	F0	F6				
ADDD	F6	F8	F2				

Reservation Stations:

<i>on Stations:</i>				$S1$	$S2$	RS	RS
$Time$	$Name$	$Busy$	Op	Vj	Vk	Qj	Qk
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	No					
	Mult2	No					

Register result status:

Clock		$F0$	$F2$	$F4$	$F6$	$F8$	$F10$	$F12$...	$F30$
0	FU									

指令序列在Tomasulo算法下如何工作

- 用以下来说明Tomasulo算法原理

1. **LD** **F6, 34(R2)**
2. **LD** **F2, 45(R3)**
3. **MULTD** **F0, F2, F4**
4. **SUBD** **F8, F6, F2**
5. **DIVD** **F10, F0, F6**
6. **ADDD** **F6, F8, F2**

- 假设各种操作的延迟为:

加法: **2**个时钟周期

乘法: **10**个时钟周期

除法: **40**个时钟周期

Tomasulo实例: Cycle 1

Instruction status:

Instruction	<i>j</i>	<i>k</i>	Issue	Exec	Write	Comp	Result
LD	F6	34+	R2	1			
LD	F2	45+	R3				
MULTD	F0	F2	F4				
SUBD	F8	F6	F2				
DIVD	F10	F0	F6				
ADDD	F6	F8	F2				

	Busy	Address
Load1	Yes	34+R2
Load2	No	
Load3	No	

Reservation Stations:

Time	Name	Busy	Op	<i>S1</i> <i>Vj</i>	<i>S2</i> <i>Vk</i>	<i>RS</i> <i>Qj</i>	<i>RS</i> <i>Qk</i>
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	No					
	Mult2	No					

Register result status:

Clock	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
1				Load1					

Tomasulo实例: Cycle 2

Instruction status:

Instruction		<i>j</i>	<i>k</i>	Issue	Exec	Write
LD	F6	34+	R2	1		
LD	F2	45+	R3	2		
MULTD	F0	F2	F4			
SUBD	F8	F6	F2			
DIVD	F10	F0	F6			
ADDD	F6	F8	F2			

	Busy	Address
Load1	Yes	34+R2
Load2	Yes	45+R3
Load3	No	

Reservation Stations:

on Stations:

				$S1$	$S2$	RS	RS
Time	Name	Busy	Op	Vj	Vk	Qj	Qk
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	No					
	Mult2	No					

Register result status:

Clock		<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
2	FU		Load2		Load1					

Note: Unlike 6600, can have multiple loads outstanding

Tomasulo实例: Cycle 3

Instruction status:

				Exec	Write		
Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy	Address
LD	F6	34+	R2	1	3	Load1	Yes 34+R2
LD	F2	45+	R3	2		Load2	Yes 45+R3
MULTD	F0	F2	F4	3		Load3	No
SUBD	F8	F6	F2				
DIVD	F10	F0	F6				
ADDD	F6	F8	F2				

Reservation Stations:

				<i>S1</i>	<i>S2</i>	<i>RS</i>	<i>RS</i>
Time	Name	Busy	Op	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	Yes	MULTD		R(F4)	Load2	
	Mult2	No					

Register result status:

Clock	F0	F2	F4	F6	F8	F10	F12	...	F30
3	FU	Mult1	Load2		Load1				

- Note: registers names are removed ("renamed") in Reservation Stations; MULT issued vs. scoreboard
- Load1 completing; what is waiting for Load1?

Tomasulo实例: Cycle 4

Instruction status:

				Exec		Write		
Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result		Busy	Address
LD	F6	34+	R2	1	3	4	Load1	No
LD	F2	45+	R3	2	4		Load2	Yes 45+R3
MULTD	F0	F2	F4	3			Load3	No
SUBD	F8	F6	F2	4				
DIVD	F10	F0	F6					
ADDD	F6	F8	F2					

Reservation Stations:

				S1	S2	RS	RS
Time	Name	Busy	Op	Vj	Vk	Qj	Qk
Add1	Yes	SUBD	M(A1)				Load2
Add2	No						
Add3	No						
Mult1	Yes	MULTD			R(F4)	Load2	
Mult2	No						

Register result status:

Clock		F0	F2	F4	F6	F8	F10	F12	...	F30
4	FU	Mult1	Load2		M(A1)	Add1				

- Load2 completing; what is waiting for Load2?

Tomasulo实例: Cycle 5

Instruction status:

Instruction	<i>j</i>	<i>k</i>	Issue	Exec Comp	Write Result	Load	Busy	Address
LD	F6	34+	R2	1	3	4	Load1	No
LD	F2	45+	R3	2	4	5	Load2	No
MULTD	F0	F2	F4	3			Load3	No
SUBD	F8	F6	F2	4				
DIVD	F10	F0	F6	5				
ADDD	F6	F8	F2					

Reservation Stations:

Time	Name	Busy	Op	S1 Vj	S2 Vk	RS Qj	RS Qk
2	Add1	Yes	SUBD	M(A1)	M(A2)		
	Add2	No					
	Add3	No					
10	Mult1	Yes	MULTD	M(A2)	R(F4)		
	Mult2	Yes	DIVD		M(A1)	Mult1	

Register result status:

Clock	F0	F2	F4	F6	F8	F10	F12	...	F30
5	Mult1	M(A2)		M(A1)	Add1	Mult2			

- Timer starts down for Add1, Mult1?

Tomasulo实例: Cycle 6

Instruction status:

				Exec Write			
Instruction		<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy Address
LD	F6	34+	R2	1	3	4	<div> Load1 No Load2 No Load3 No </div>
LD	F2	45+	R3	2	4	5	
MULTD	F0	F2	F4	3			
SUBD	F8	F6	F2	4			
DIVD	F10	F0	F6	5			
ADDD	F6	F8	F2	6			

Reservation Stations:

				<i>S1</i>	<i>S2</i>	<i>RS</i>	<i>RS</i>
Time	Name	Busy	Op	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>
1	Add1	Yes	SUBD	M(A1)	M(A2)		
	Add2	Yes	ADDD		M(A2)	Add1	
	Add3	No					
9	Mult1	Yes	MULTD	M(A2)	R(F4)		
	Mult2	Yes	DIVD		M(A1)	Mult1	

Register result status:

Clock	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
6	<div> FU Mult1 M(A2) Add2 Add1 Mult2 </div>								

- Issue ADDD here vs. scoreboard?

Tomasulo实例: Cycle 7

Instruction status:

				Exec	Write		
Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy	Address
LD	F6	34+	R2	1	3	4	Load1
LD	F2	45+	R3	2	4	5	Load2
MULTD	F0	F2	F4	3			Load3
SUBD	F8	F6	F2	4	7		
DIVD	F10	F0	F6	5			
ADDD	F6	F8	F2	6			

Reservation Stations:

				<i>S1</i>	<i>S2</i>	<i>RS</i>	<i>RS</i>
Time	Name	Busy	Op	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>
0	Add1	Yes	SUBD	M(A1)	M(A2)		
	Add2	Yes	ADDD		M(A2)	Add1	
	Add3	No					
8	Mult1	Yes	MULTD	M(A2)	R(F4)		
	Mult2	Yes	DIVD		M(A1)	Mult1	

Register result status:

Clock	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
7	FU	Mult1	M(A2)		Add2	Add1	Mult2		

- Add1 completing; what is waiting for it?

Tomasulo实例: Cycle 9

Instruction status:

<i>Instruction status:</i>				<i>Exec</i>	<i>Write</i>		
Instruction	<i>j</i>	<i>k</i>	<i>Issue</i>	<i>Comp</i>	<i>Result</i>		Busy Address
LD	F6	34+	R2	1	3	4	Load1 No
LD	F2	45+	R3	2	4	5	Load2 No
MULTD	F0	F2	F4	3			Load3 No
SUBD	F8	F6	F2	4	7	8	
DIVD	F10	F0	F6	5			
ADDD	F6	F8	F2	6			

Reservation Stations:

<i>on Stations:</i>				<i>S1</i>	<i>S2</i>	<i>RS</i>	<i>RS</i>
<i>Time</i>	<i>Name</i>	<i>Busy</i>	<i>Op</i>	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>
	Add1	No					
1	Add2	Yes	ADDD	(M-M)	M(A2)		
	Add3	No					
6	Mult1	Yes	MULTD	M(A2)	R(F4)		
	Mult2	Yes	DIVD		M(A1)	Mult1	

Register result status:

Clock		$F0$	$F2$	$F4$	$F6$	$F8$	$F10$	$F12$...	$F30$
9	FU	Mult1	M(A2)		Add2	(M-M)	Mult2			

Tomasulo实例: Cycle 15

Instruction status:

<i>Instruction status:</i>				<i>Exec</i>	<i>Write</i>		
Instruction	<i>j</i>	<i>k</i>	<i>Issue</i>	<i>Comp</i>	<i>Result</i>		Busy Address
LD	F6	34+	R2	1	3	4	Load1 No
LD	F2	45+	R3	2	4	5	Load2 No
MULTD	F0	F2	F4	3	15		Load3 No
SUBD	F8	F6	F2	4	7	8	
DIVD	F10	F0	F6	5			
ADDD	F6	F8	F2	6	10	11	

Reservation Stations:

<i>on Stations:</i>				<i>S1</i>	<i>S2</i>	<i>RS</i>	<i>RS</i>
<i>Time</i>	<i>Name</i>	<i>Busy</i>	<i>Op</i>	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>
	Add1	No					
	Add2	No					
	Add3	No					
0	Mult1	Yes	MULTD	M(A2)	R(F4)		
	Mult2	Yes	DIVD		M(A1)	Mult1	

Register result status:

Clock		$F0$	$F2$	$F4$	$F6$	$F8$	$F10$	$F12$...	$F30$
15	FU	Mult1	M(A2)		(M-M+M	(M-M)	Mult2			

Tomasulo实例: Cycle 55

(跳过一些节拍)

Instruction status:

				<i>Exec</i>		<i>Write</i>		
Instruction	<i>j</i>	<i>k</i>		<i>Issue</i>	<i>Comp</i>	<i>Result</i>	Busy	Address
LD	F6	34+	R2	1	3	4	Load1	No
LD	F2	45+	R3	2	4	5	Load2	No
MULTD	F0	F2	F4	3	15	16	Load3	No
SUBD	F8	F6	F2	4	7	8		
DIVD	F10	F0	F6	5				
ADDD	F6	F8	F2	6	10	11		

Reservation Stations:

on Stations:

				<i>S1</i>	<i>S2</i>	<i>RS</i>	<i>RS</i>
<i>Time</i>	<i>Name</i>	<i>Busy</i>	<i>Op</i>	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	No					
1	Mult2	Yes	DIVD	M*F4	M(A1)		

Register result status:

Clock	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
55	<i>FU</i>	M*F4	M(A2)		(M-M+N	(M-M)	Mult2		

Tomasulo实例: Cycle 56

Instruction status:

				Exec	Write		
Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy	Address
LD	F6	34+	R2	1	3	4	Load1
LD	F2	45+	R3	2	4	5	Load2
MULTD	F0	F2	F4	3	15	16	Load3
SUBD	F8	F6	F2	4	7	8	
DIVD	F10	F0	F6	5	56		
ADDD	F6	F8	F2	6	10	11	

Reservation Stations:

Time	Name	Busy	Op	<i>S1</i> <i>Vj</i>	<i>S2</i> <i>Vk</i>	<i>RS</i> <i>Qj</i>	<i>RS</i> <i>Qk</i>
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	No					
0	Mult2	Yes	DIVD	M*F4	M(A1)		

Register result status:

Clock	F0	F2	F4	F6	F8	F10	F12	...	F30
56	FU	M*F4	M(A2)		(M-M+M)	(M-M)	Mult2		

- Mult2 is completing; what is waiting for it?

Tomasulo实例: Cycle 57

Instruction status:

Instruction	<i>j</i>	<i>k</i>	Issue	Exec Comp	Write Result	Load	Busy	Address
LD	F6	34+	R2	1	3	4	Load1	No
LD	F2	45+	R3	2	4	5	Load2	No
MULTD	F0	F2	F4	3	15	16	Load3	No
SUBD	F8	F6	F2	4	7	8		
DIVD	F10	F0	F6	5	56	57		
ADDD	F6	F8	F2	6	10	11		

Reservation Stations:

Time	Name	Busy	Op	S1 Vj	S2 Vk	RS Qj	RS Qk
	Add1	No					
	Add2	No					
	Add3	No					
	Mult1	No					
0	Mult2	Yes	DIVD	M*F4	M(A1)		

Register result status:

Clock	F0	F2	F4	F6	F8	F10	F12	...	F30
56	FU	M*F4	M(A2)		(M-M+N	(M-M)	Mult2		

- Once again: In-order issue, out-of-order execution and completion.

基于Tomasulo算法的动态循环展开

- 用以下指令序列来说明Tomasulo算法的动态循环展开

```
Loop: LD      F0, 0(R1)
      MULTD   F4, F0, F2
      SD      0(R1), F4
      SUBI    R1, R1, #8
      BNEZ    R1, Loop
```

- ✓ 为了保证正确的异常行为，Tomasulo算法对指令的执行有一个限制：
一旦有一条分支指令还没有执行完，其后的指令是不允许进入执行段。
- ✓ 动态存储器地址判别技术，解决存储器访问时的RAW、WAW、WAR冲突。
- ✓ 第一次取数LD，由于Cache miss，需要8个时钟周期延迟，之后的取数LD，只需1个时钟周期。存数SD需要3个时钟周期。
- ✓ 乘法需要4个时钟周期

Tomasulo 循环实例

Instruction status:

Exec Write

<i>ITER</i>	Instruction	<i>j</i>	<i>k</i>	<i>Issue CompResult</i>	<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	Load1	No	
1	MULTD	F4	F0	F2	Load2	No	
1	SD	F4	0	R1	Load3	No	
2	LD	F0	0	R1	Store1	No	
2	MULTD	F4	F0	F2	Store2	No	
2	SD	F4	0	R1	Store3	No	

Reservation Stations:

Reservation Stations:					$S1$	$S2$	RS	
$Time$	$Name$	$Busy$	Op	Vj	Vk	Qj	Qk	$Code:$
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	No						SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

<i>Clock</i>	R1		<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	<i>...</i>	<i>F30</i>
0	80	<i>Fu</i>									

Tomasulo 循环实例: Cycle 1

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue Comp</i>	<i>Result</i>	<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1	Load1	Yes	80	
1	MULTD	F4	F0	F2		Load2	No		
1	SD	F4	0	R1		Load3	No		
2	LD	F0	0	R1		Store1	No		
2	MULTD	F4	F0	F2		Store2	No		
2	SD	F4	0	R1		Store3	No		

Reservation Stations:

Reservation Stations:				$S1$	$S2$	RS			
Time	Name	Busy	Op	V_j	V_k	Q_j	Q_k	Code:	
	Add1	No						LD	F0 0 R1 ←
	Add2	No						MULTD	F4 F0 F2
	Add3	No						SD	F4 0 R1
	Mult1	No						SUBI	R1 R1 #8
	Mult2	No						BNEZ	R1 Loop

Register result status

<i>Clock</i>	R1		<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	<i>...</i>	<i>F30</i>
1	80	<i>Fu</i>	Load1								

Tomasulo 循环实例: Cycle 2

Instruction status:

Exec Write

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1		Load1	Yes	80
1	MULTD	F4	F0	F2	2		Load2	No	
1	SD	F4	0	R1			Load3	No	
2	LD	F0	0	R1			Store1	No	
2	MULTD	F4	F0	F2			Store2	No	
2	SD	F4	0	R1			Store3	No	

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load1		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
2	80	Fu	Load1	Mult1						

Tomasulo 循环实例: Cycle 3

Instruction status:

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1		Load1	Yes	80
1	MULTD	F4	F0	F2	2		Load2	No	
1	SD	F4	0	R1	3		Load3	No	
2	LD	F0	0	R1			Store1	Yes	80
2	MULTD	F4	F0	F2			Store2	No	Mult1
2	SD	F4	0	R1			Store3	No	

Reservation Stations:

Time	Name	Busy	Op	Vj	Vk	Qj	Rs	Code:
Add1	No							LD
Add2	No							MULTD
Add3	No							SD
Mult1	Yes	Multd			R(F2)	Load1		SUBI
Mult2	No							BNEZ

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
3	80	Fu	Load1	Mult1						

- Implicit renaming sets up “DataFlow” graph

Tomasulo 循环实例: Cycle 4

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue Comp</i>	<i>Result</i>	<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1 2 3	Load1	Yes	80	Mult1
1	MULTD	F4	F0	F2		Load2	No		
1	SD	F4	0	R1		Load3	No		
2	LD	F0	0	R1		Store1	Yes	80	
2	MULTD	F4	F0	F2		Store2	No		
2	SD	F4	0	R1		Store3	No		

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load1		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1		F0	F2	F4	F6	F8	F10	F12	...	F30
4	80	Fu	Load1		Mult1						

- Dispatching SUBI Instruction

Tomasulo 循环实例: Cycle 5

Instruction status:

Exec Write

<i>ITER</i>	Instruction	<i>j</i>	<i>k</i>	<i>Issue</i>	<i>CompResult</i>		<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1 2 3	Load1	Yes	80	Mult1
1	MULTD	F4	F0	F2		Load2	No		
1	SD	F4	0	R1		Load3	No		
2	LD	F0	0	R1	Store1	Yes	80		
2	MULTD	F4	F0	F2	Store2	No			
2	SD	F4	0	R1	Store3	No			

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load1		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
5	72	Fu	Load1	Mult1						

- And, BNEZ instruction

Tomasulo 循环实例: Cycle 6

Instruction status:

Exec Write

<i>ITER</i>	Instruction	<i>j</i>	<i>k</i>	<i>Issue</i>	<i>Comp</i>	<i>Result</i>	<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1 2 3 6	Load1	Yes	80	Mult1
1	MULTD	F4	F0	F2		Load2	Yes	72	
1	SD	F4	0	R1		Load3	No		
2	LD	F0	0	R1		Store1	Yes	80	
2	MULTD	F4	F0	F2		Store2	No		
2	SD	F4	0	R1		Store3	No		

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load1		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
6	72	Fu	Load2	Mult1						

- Notice that F0 never sees Load from location 80

Tomasulo 循环实例: Cycle 7

Instruction status:

Exec Write

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1		Load1	Yes 80	
1	MULTD	F4	F0	F2	2		Load2	Yes 72	
1	SD	F4	0	R1	3		Load3	No	
2	LD	F0	0	R1	6		Store1	Yes 80	Mult1
2	MULTD	F4	F0	F2	7		Store2	No	
2	SD	F4	0	R1			Store3	No	

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load1		SUBI R1 R1 #8
	Mult2	Yes	Multd		R(F2)	Load2		BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
7	72	Fu	Load2	Mult2						

- Register file completely detached from iteration 1

Tomasulo 循环实例: Cycle 8

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue Comp</i>	<i>Result</i>	<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1 2 3 6 7 8	Load1	Yes	80	
1	MULTD	F4	F0	F2		Load2	Yes	72	
1	SD	F4	0	R1		Load3	No		
2	LD	F0	0	R1		Store1	Yes	80	Mult1
2	MULTD	F4	F0	F2		Store2	Yes	72	Mult2
2	SD	F4	0	R1		Store3	No		

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:				
	Add1	No						LD	F0	0	R1	
	Add2	No						MULTD	F4	F0	F2	
	Add3	No						SD	F4	0	R1	←
	Mult1	Yes	Multd		R(F2)	Load1		SUBI	R1	R1	#8	
	Mult2	Yes	Multd		R(F2)	Load2		BNEZ	R1	Loop		

Register result status

Clock	R1		F0	F2	F4	F6	F8	F10	F12	...	F30
8	72	Fu	Load2		Mult2						

- First and Second iteration completely overlapped

Tomasulo 循环实例: Cycle 9

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue CompResult</i>			<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1	9	Load1	Yes	80	
1	MULTD	F4	F0	F2	2		Load2	Yes	72	
1	SD	F4	0	R1	3		Load3	No		
2	LD	F0	0	R1	6		Store1	Yes	80	Mult1
2	MULTD	F4	F0	F2	7		Store2	Yes	72	Mult2
2	SD	F4	0	R1	8		Store3	No		

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>	Code:				
	Add1	No						LD	F0	0	R1	
	Add2	No						MULTD	F4	F0	F2	
	Add3	No						SD	F4	0	R1	
	Mult1	Yes	Multd		R(F2)	Load1		SUBI	R1	R1	#8	←
	Mult2	Yes	Multd		R(F2)	Load2		BNEZ	R1	Loop		

Register result status

Clock	R1	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
9	72	<i>Fu</i>	Load2		Mult2					

- Load1 completing: who is waiting?
- Note: Dispatching SUBI

Tomasulo 循环实例: Cycle 10

Instruction status:

Exec Write

ITER	Instruction		<i>j</i>	<i>k</i>	<i>Issue CompResult</i>				<i>Busy</i>	<i>Addr</i>	<i>Fu</i>	
1	LD	F0	0	R1	1	9	10	Load1	No			
1	MULTD	F4	F0	F2				2	Load2	Yes	72	
1	SD	F4	0	R1				3	Load3	No		
2	LD	F0	0	R1	6	10		Store1	Yes	80	Mult1	
2	MULTD	F4	F0	F2	7			Store2	Yes	72	Mult2	
2	SD	F4	0	R1	8			Store3	No			

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
4	Mult1	Yes	Multd	M[80]	R(F2)			SUBI R1 R1 #8
	Mult2	Yes	Multd		R(F2)	Load2		BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
10	64	Fu	Load2	Mult2						

- Load2 completing: who is waiting?
- Note: Dispatching BNEZ

Tomasulo 循环实例: Cycle 11

Instruction status:

Exec Write

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1	9	10	Load1	No
1	MULTD	F4	F0	F2	2			Load2	No
1	SD	F4	0	R1	3			Load3	Yes 64
2	LD	F0	0	R1	6	10	11	Store1	Yes 80 Mult1
2	MULTD	F4	F0	F2	7			Store2	Yes 72 Mult2
2	SD	F4	0	R1	8			Store3	No

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
3	Mult1	Yes	Multd	M[80]	R(F2)			SUBI R1 R1 #8
4	Mult2	Yes	Multd	M[72]	R(F2)			BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
11	64	Fu	Load3	Mult2						

- Next load in sequence

Tomasulo 循环实例: Cycle 12

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue CompResult</i>				<i>Busy</i>	<i>Addr</i>	<i>Fu</i>	
1	LD	F0	0	R1	1	9	10	Load1	No			
1	MULTD	F4	F0	F2				2	Load2	No		
1	SD	F4	0	R1				3	Load3	Yes	64	
2	LD	F0	0	R1	6	10	11	Store1	Yes	80	Mult1	
2	MULTD	F4	F0	F2				7	Store2	Yes	72	Mult2
2	SD	F4	0	R1				8	Store3	No		

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
2	Mult1	Yes	Multd	M[80]	R(F2)			SUBI R1 R1 #8
3	Mult2	Yes	Multd	M[72]	R(F2)			BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
12	64	Fu	Load3	Mult2						

- Why not issue third multiply?

Tomasulo 循环实例: Cycle 13

Instruction status:

Exec Write

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1	9	10	Load1	No
1	MULTD	F4	F0	F2	2			Load2	No
1	SD	F4	0	R1	3			Load3	Yes 64
2	LD	F0	0	R1	6	10	11	Store1	Yes 80 Mult1
2	MULTD	F4	F0	F2	7			Store2	Yes 72 Mult2
2	SD	F4	0	R1	8			Store3	No

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
1	Mult1	Yes	Multd	M[80]	R(F2)			SUBI R1 R1 #8
2	Mult2	Yes	Multd	M[72]	R(F2)			BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
13	64	Fu	Load3	Mult2						

Tomasulo 循环实例: Cycle 14

Instruction status:

Exec Write

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1	9	10	Load1	No
1	MULTD	F4	F0	F2	2	14		Load2	No
1	SD	F4	0	R1	3			Load3	Yes 64
2	LD	F0	0	R1	6	10	11	Store1	Yes 80 Mult1
2	MULTD	F4	F0	F2	7			Store2	Yes 72 Mult2
2	SD	F4	0	R1	8			Store3	No

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
0	Mult1	Yes	Multd	M[80]	R(F2)			SUBI R1 R1 #8
1	Mult2	Yes	Multd	M[72]	R(F2)			BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
14	64	Fu	Load3	Mult2						

- Mult1 completing. Who is waiting?

Tomasulo 循环实例: Cycle 15

Instruction status:

Exec Write

ITER	Instruction	j	k	Issue	Comp	Result	Busy	Addr	Fu
1	LD	F0	0	R1	1	9	10	Load1	No
1	MULTD	F4	F0	F2	2	14	15	Load2	No
1	SD	F4	0	R1	3			Load3	Yes 64
2	LD	F0	0	R1	6	10	11	Store1	Yes 80 [80]*R2
2	MULTD	F4	F0	F2	7	15		Store2	Yes 72 Mult2
2	SD	F4	0	R1	8			Store3	No

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
	Mult1	No						SUBI R1 R1 #8
0	Mult2	Yes	Multd	M[72]	R(F2)			BNEZ R1 Loop

Register result status

Clock	R1	F0	F2	F4	F6	F8	F10	F12	...	F30
15	64	Fu	Load3	Mult2						

- Mult2 completing. Who is waiting?

Tomasulo 循环实例: Cycle 16

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue CompResult</i>				<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1	9	10	Load1	No		
1	MULTD	F4	F0	F2	2	14	15	Load2	No		
1	SD	F4	0	R1	3			Load3	Yes	64	
2	LD	F0	0	R1	6	10	11	Store1	Yes	80	[80]*R2
2	MULTD	F4	F0	F2	7	15	16	Store2	Yes	72	[72]*R2
2	SD	F4	0	R1	8			Store3	No		

Reservation Stations:

Reservation Stations:					<i>S1</i>	<i>S2</i>	<i>RS</i>	
<i>Time</i>	<i>Name</i>	<i>Busy</i>	<i>Op</i>	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>	<i>Code:</i>
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2 ←
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load3		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

<i>Clock</i>	R1		<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	<i>...</i>	<i>F30</i>
16	64	<i>Fu</i>	Load3		Mult1						

Tomasulo 循环实例: Cycle 17

Instruction status:

Exec Write

ITER	Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy	Addr	<i>Fu</i>
1	LD	F0	0	R1	1	9	10	Load1	No
1	MULTD	F4	F0	F2	2	14	15	Load2	No
1	SD	F4	0	R1	3			Load3	Yes 64
2	LD	F0	0	R1	6	10	11	Store1	Yes 80 [80]*R2
2	MULTD	F4	F0	F2	7	15	16	Store2	Yes 72 [72]*R2
2	SD	F4	0	R1	8			Store3	Yes 64 Mult1

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load3		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
17	64	<i>Fu</i>	Load3		Mult1					

Tomasulo 循环实例: Cycle 18

Instruction status:

Exec Write

ITER	Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy	Addr	<i>Fu</i>
1	LD	F0	0	R1	1	9	10	Load1	No
1	MULTD	F4	F0	F2	2	14	15	Load2	No
1	SD	F4	0	R1	3	18		Load3	Yes 64
2	LD	F0	0	R1	6	10	11	Store1	Yes 80 [80]*R2
2	MULTD	F4	F0	F2	7	15	16	Store2	Yes 72 [72]*R2
2	SD	F4	0	R1	8			Store3	Yes 64 Mult1

Reservation Stations:

S1 S2 RS

Time	Name	Busy	Op	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load3		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1	<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	...	<i>F30</i>
18	64	<i>Fu</i>	Load3		Mult1					

Tomasulo 循环实例: Cycle 19

Instruction status:

Exec Write

<i>ITER</i>	Instruction		<i>j</i>	<i>k</i>	<i>Issue CompResult</i>				<i>Busy</i>	<i>Addr</i>	<i>Fu</i>
1	LD	F0	0	R1	1	9	10	Load1	No		
1	MULTD	F4	F0	F2	2	14	15	Load2	No		
1	SD	F4	0	R1	3	18	19	Load3	Yes	64	
2	LD	F0	0	R1	6	10	11	Store1	No		
2	MULTD	F4	F0	F2	7	15	16	Store2	Yes	72	[72]*R2
2	SD	F4	0	R1	8	19		Store3	Yes	64	Mult1

Reservation Stations:

Reservation Stations:					$S1$	$S2$	RS	
Time	Name	Busy	Op	Vj	Vk	Qj	Qk	Code:
	Add1	No						LD F0 0 R1
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load3		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

<i>Clock</i>	R1		<i>F0</i>	<i>F2</i>	<i>F4</i>	<i>F6</i>	<i>F8</i>	<i>F10</i>	<i>F12</i>	<i>...</i>	<i>F30</i>
19	64	<i>Fu</i>	Load3		Mult1						

Tomasulo 循环实例: Cycle 20

Instruction status:

Exec Write

ITER	Instruction	<i>j</i>	<i>k</i>	Issue	Comp	Result	Busy	Addr	Fu		
1	LD	F0	0	R1	1	9	10	Load1	Yes	56	
1	MULTD	F4	F0	F2	2	14	15	Load2	No		
1	SD	F4	0	R1	3	18	19	Load3	Yes	64	
2	LD	F0	0	R1	6	10	11	Store1	No		
2	MULTD	F4	F0	F2	7	15	16	Store2	No		
2	SD	F4	0	R1	8	19	20	Store3	Yes	64	Mult1

Reservation Stations:

Reservation Stations:					<i>S1</i>	<i>S2</i>	<i>RS</i>	
<i>Time</i>	<i>Name</i>	<i>Busy</i>	<i>Op</i>	<i>Vj</i>	<i>Vk</i>	<i>Qj</i>	<i>Qk</i>	<i>Code:</i>
	Add1	No						LD F0 0 R1 ←
	Add2	No						MULTD F4 F0 F2
	Add3	No						SD F4 0 R1
	Mult1	Yes	Multd		R(F2)	Load3		SUBI R1 R1 #8
	Mult2	No						BNEZ R1 Loop

Register result status

Clock	R1		F0	F2	F4	F6	F8	F10	F12	...	F30
20	56	Fu	Load1		Mult1						

Tomasulo算法的优点

- 分布式硬件冲突检测。
- 利用寄存器换名，彻底消除WAW和WAR这两种名相关
- 如果多个保留站等待同一个操作数，当操作数在CDB上广播时，他们可以同时获得所需的数据
- 对于存储器访问，动态存储器地址判别技术可解决RAW冲突（取操作数时判断）、WAR和WAW冲突（存操作数时判断）。
- 能够达到很高的性能。

Tomasulo算法的缺点

- 高复杂性：需要大量硬件
- 存在瓶颈：单个公共数据总线（CDB）引发竞争
 - 额外的CDB：在每个保留站上需要为每条CDB设置重复的硬件接口
- 为了保证正确的异常行为，对指令的执行有一个限制：
一旦有一条分支指令还没有执行完，其后的指令是不允许进入执行段

动态调度方法中的异常行为处理

- 指令乱序大大增加了异常处理的复杂度。
- 不精确异常（**Imprecise Exception**）：当指令 i 导致发生异常时，处理机的现场（状态）与严格按程序顺序执行不同。
- 精确异常（**Precise Exception**）：处理机现场跟严格按程序顺序执行时指令 i 的现场相同。
- 不精确异常产生的原因：
 - ✓ 流水线可能已经执行完按程序顺序是位于指令 i 之后的指令
 - ✓ 流水线可能还没完成按程序顺序是指令 i 之前的指令

动态调度方法中的异常行为处理

- 例子:

DADDU R2, R3, R4

BEQZ R2, L1

LW R1, 0(R2)

L1: ...

- **BEQZ指令和LW指令没有数据相关**
- **但是LW指令可能会引起存储器的异常（如访存缺页、访存越界等）**

第七章 指令级并行

7.1 指令级并行的概念

7.2 指令的动态调度

7.3 控制相关的动态解决技术

7.4 多指令流出技术

7.3 控制相关的动态解决技术

7.3.1 分支预测缓冲

7.3.2 分支目标缓冲

7.3.3 基于硬件的前瞻执行

7.3.1 分支预测缓冲

- 动态分支预测的两个理由
 - n 流出的处理器加速上限为 n 倍
 - Amdahl定律提示：在较低CPI机器上，控制相关导致的空转对机器性能影响大
- 前面解决控制相关的静态策略
 - 需要编译器将一条或多条指令移动到流水线产生的分支延迟槽中
- 关于分支预测策略的两部分工作
 - 预测的分支是否成功
 - 执行分支目标指令

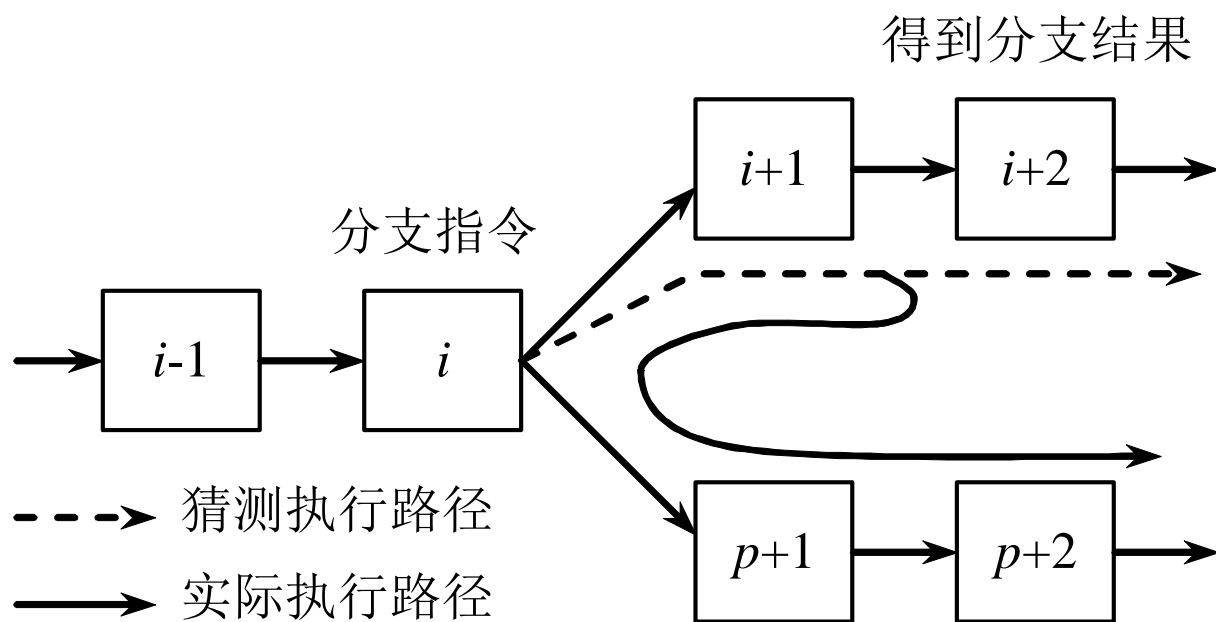
分支预测的效果

- 预测的准确率
- 分支的开销
 - 预测正确的开销
 - 预测错误的开销
 - 决定分支开销的因素：流水线的结构、预测的方法、预测错误时的恢复策略等

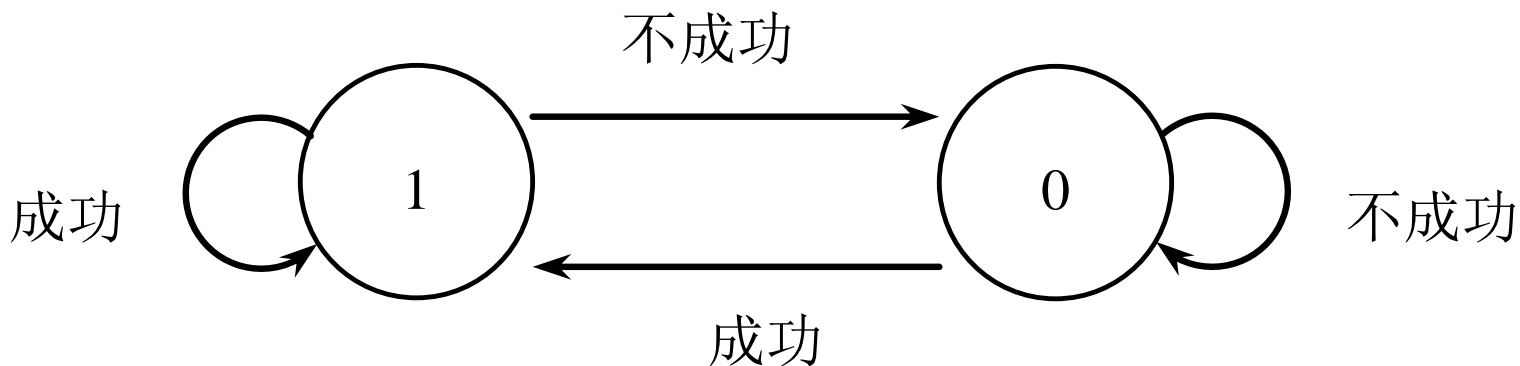
分支预测缓冲(BPB): 原理

- 最简单的分支预测策略
- **BPB**也被称为**BHT** (**Branch History Table**, **BHT**)
- 分支预测缓冲是一个小的存储器阵列
 - 每个单元最小可以只有1位，记录最近一次分支是否成功的信息
 - 预测位为1时，表示预测分支成功，并从目标位置开始取指令
 - 在预测错误时，要作废已经预取和分析的指令，恢复现场，并从另一条分支路径重新取指令。

分支预测执行不成功和重新执行过程



1位BPB状态图

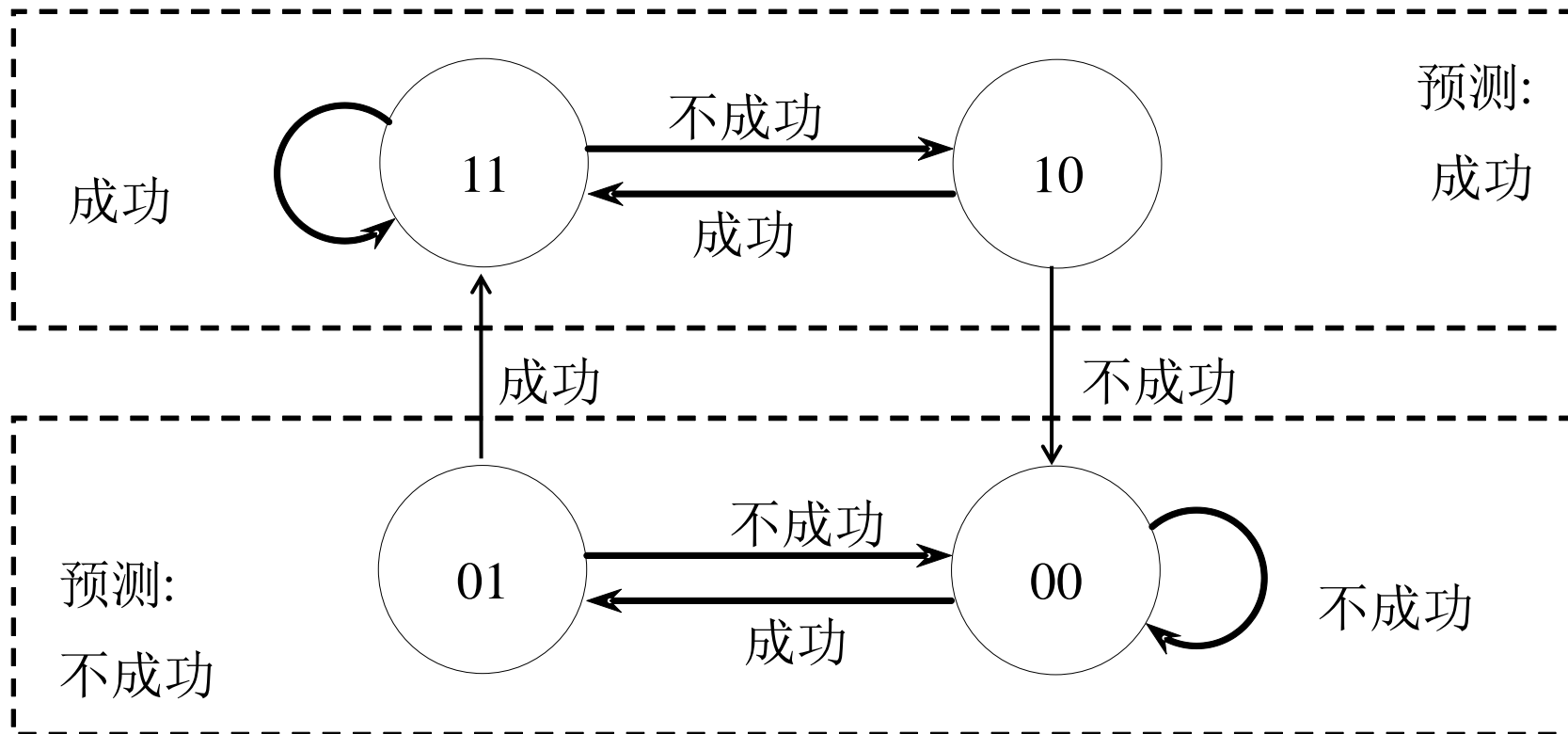


- 这种单位预测策略：
 - 当分支不成功时，将会发生连续两次预测出错
 - 2位预测策略能够改善这种情况

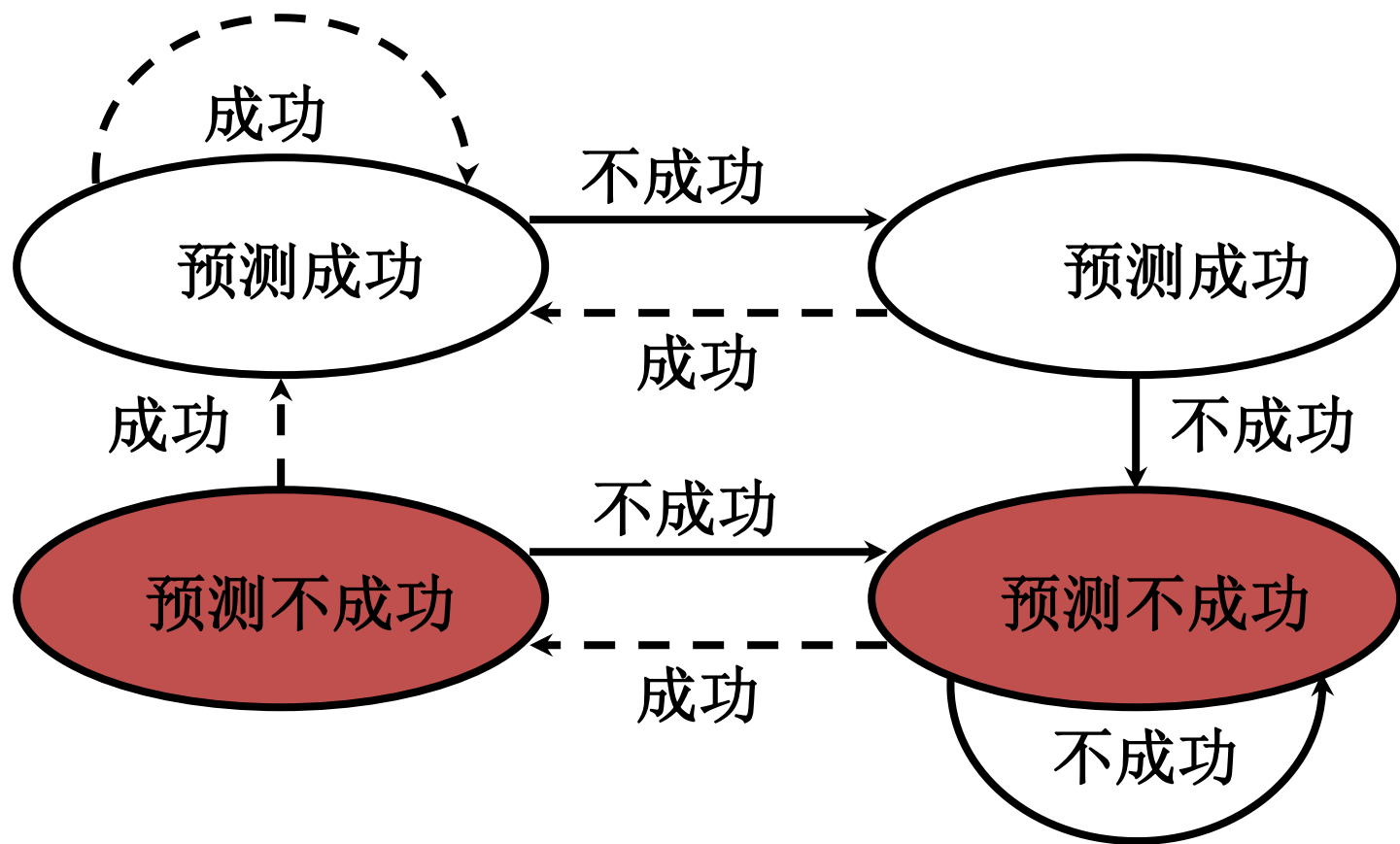
2位BPP工作原理

- 在2位预测策略中，一个预测必须错误两次才会改变。
- 对于一个4096条记录的BPP，利用2位预测策略，用SPEC89测试，命中率为82%到99%
 - 准确率最高的测试程序一般包含大量循环
 - 没有循环的线性代码一般准确率最差

2位BPP状态转换图



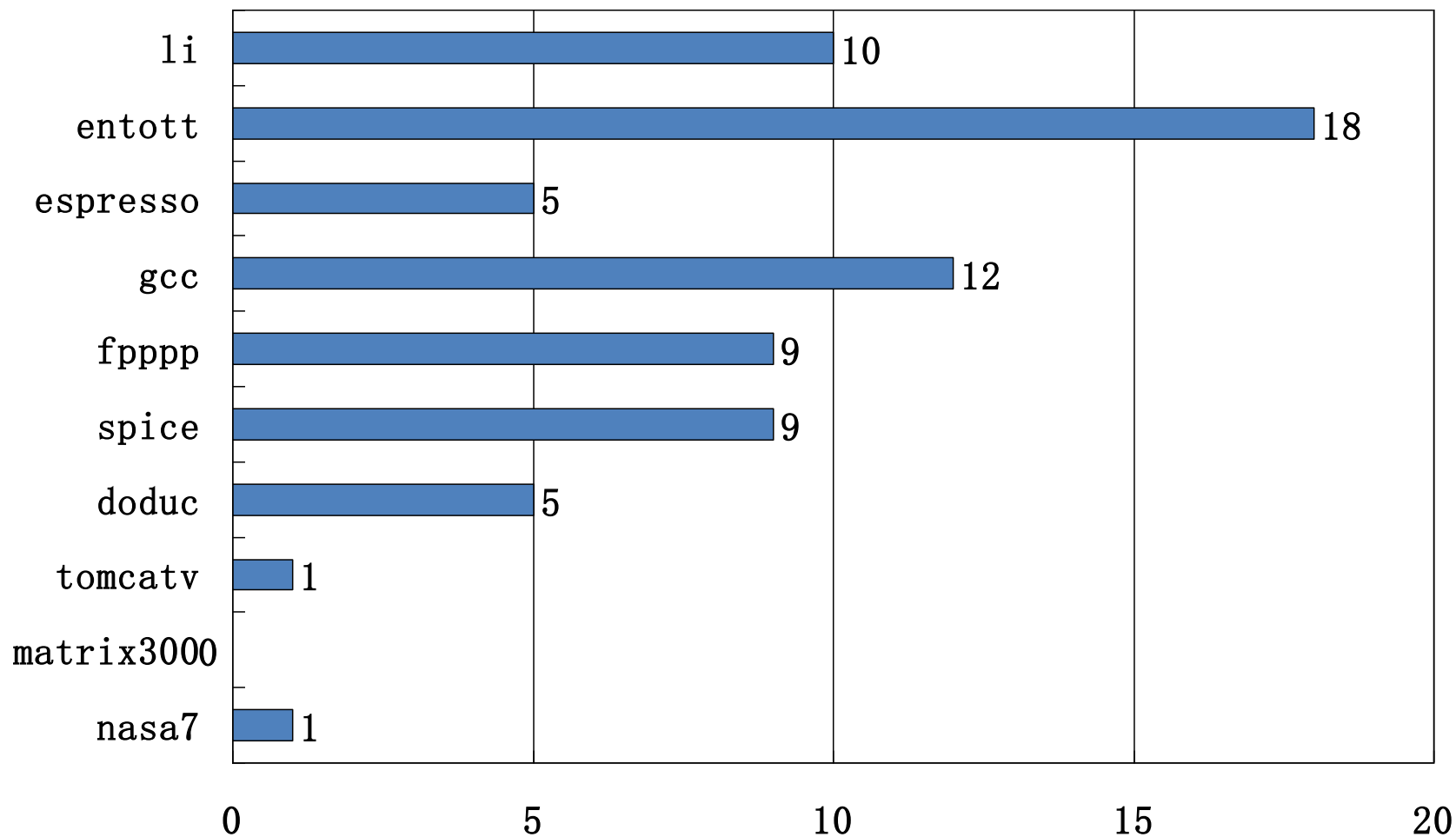
2位BPP另一种状态转换图



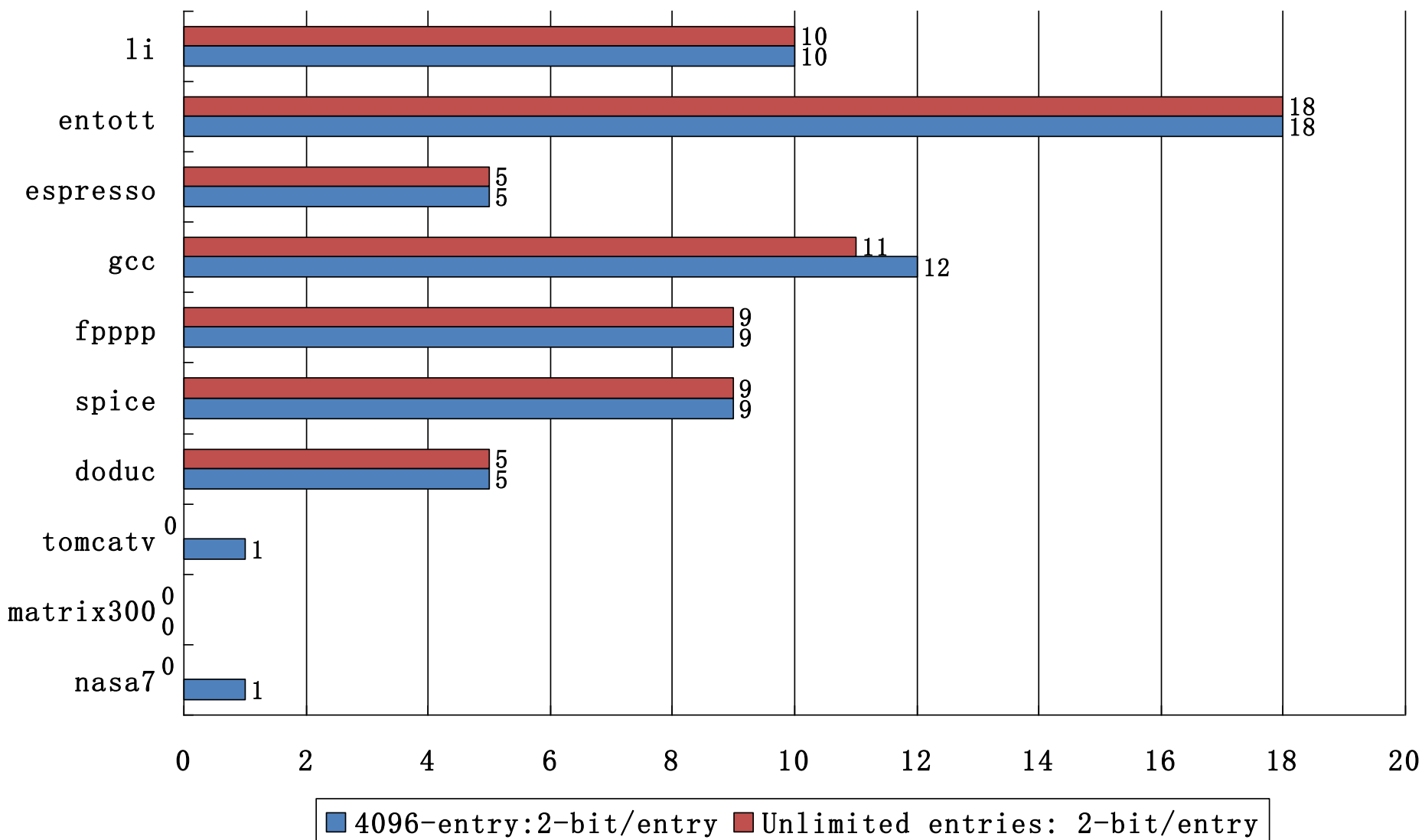
4096单元2位BPB的预测错误率

- 测试程序为SPEC89
- 整数测试程序: 平均 11%
 - gcc, espresso, eqntott, li
- 浮点测试程序: 平均 4%
 - nasa7, matrix300, tomcatv
- 为什么?

4096单元2位BPB: 预测错误率



2位BPB: 4K vs. 无穷多



BPB实现

- **BPB的实现方案**
 - 实现一个小而特殊的“**cache**”，利用指令地址进行索引，在**IF**流水段访问。
 - 或者，为指令**cache**中每一块增加附加位，与指令一起取出
- 若一个指令在**ID**段被译码为分支指令，且对应的**BPB**标志位预测其成功，则
 - 一旦**PC**已知，立刻从分支目标位置开始取指
- 对于改进**MIPS**，分支判断和计算分支目标地址都在**ID**段完成，**BPB**无效果

7.3 控制相关的动态解决技术

7.3.1 分支预测缓冲

7.3.2 分支目标缓冲

7.3.3 基于硬件的前瞻执行

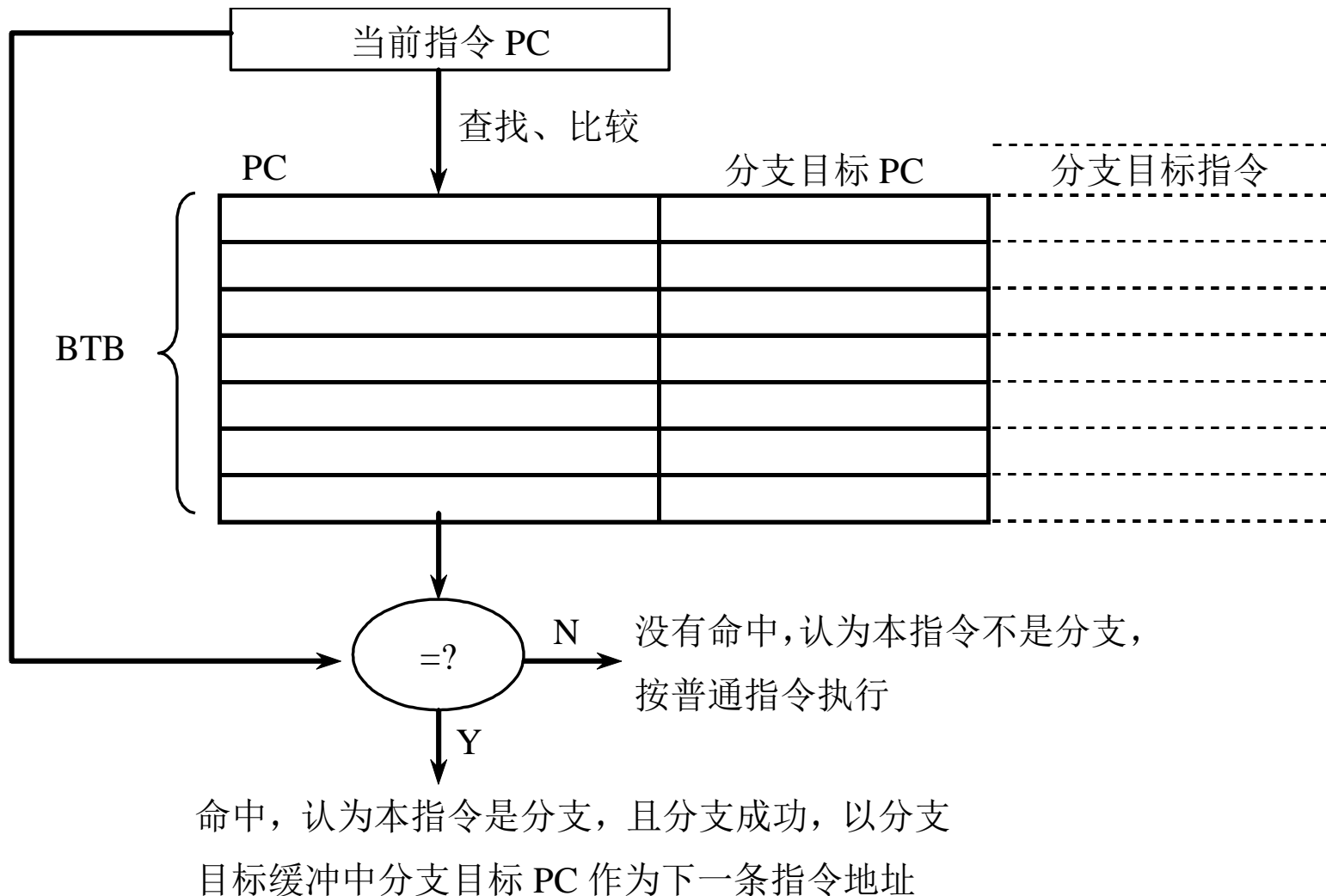
7.3.2 分支目标缓冲(BTB)

- 另一个动态分支预测方法：分支目标缓冲
 - **Branch Target Buffer, BTB**
 - 为了减小或消除流水线的分支开销，我们需要在IF段结束前知道从哪个地址开始取下一条指令
 - 换句话说，我们在IF段就需要知道这条未译码的指令是否为分支指令，并且如果它是分支指令，要尽快知道NPC值应当为多少

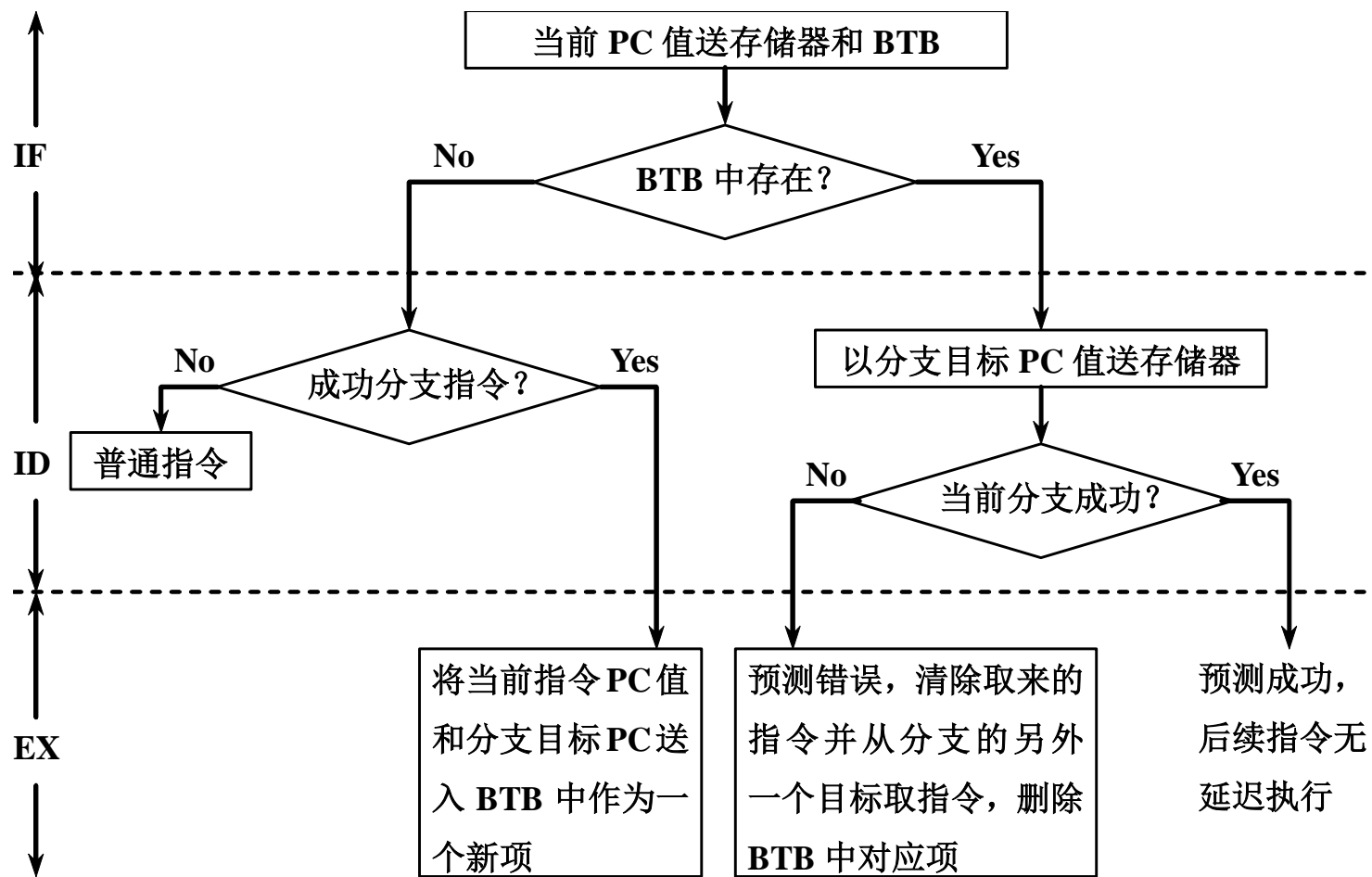
BTB实现

- 分支目标缓存**BTB**每个单元应该包括
 - 分支指令的地址
 - 分支目标的地址
 - 分支预测标识
- 取指阶段，所有指令地址都与**BTB**中保存的分支指令的地址做比较，一旦相同，就认为本指令是分支指令，并且分支成功
- 它的目标地址就是保存在缓冲区中的分支目标地址
 - 取出后直接送入**NPC**

BTB 结构



BTB 执行过程



采用BTB技术时指令在各种情况下的延迟

指令是否在BTB中	预测成功	实际的动作	延迟周期
是	成功	成功	0
是	成功	不成功	2
不是		成功	2
不是		不成功	0

按上表计算分支转移总的延迟，根据下面假设，计算分支目标缓冲的性能。

(1) 对于BTB中的指令，预测准确率**90%**

(2) 缓冲区命中率**90%**

(3) 不在BTB中的分支转移成功的比例为**60%**

解：在BTB中，预测成功且实际成功，延迟为**0**；

在BTB中，预测成功，实际不成功，延迟为：

$$90\% \times 10\% \times 2 = 0.18 \text{ 时钟周期}$$

不在BTB中，实际成功，延迟为：

$$10\% \times 60\% \times 2 = 0.12 \text{ 时钟周期}$$

不在BTB中，实际不成功，延迟为**0**；

综上，该分支目标缓冲产生总的时间延迟为**0.30** 是时钟周期

BTB的改进

- 分支预测技术受限于预测精度，以及预测失效后产生的开销
- 根据不同程序特点以及缓冲区的大小，典型的BTB可以实现80%到95%的预测精度
- 降低失效开销技术：在一个时钟周期内同时取出不同分支路径的指令
 - 会引入其他开销，比如存储系统的端口加倍

分支预测局限性总结

- 预测准确性：80%-90%
- 预测性能依赖于
 - 程序类型
 - 缓冲区大小
- 预测失效开销的优化
 - 预取不同分支路径指令
 - 存储端口数目加倍，交叉存取缓冲

Corei7分支预测浪费的指令

