## Arquitectura e Ingeniería de Computadores

# UT 2. Computadores segmentados Tema 2.1 Unidad de instrucción segmentada

J. Flich, P. López, V. Lorente, A. Pérez, S. Petit, J.C. Ruiz, S. Sáez, J. Sahuquillo

Departamento de Informática de Sistemas y Computadores Universitat Politècnica de València







#### Índice

- Concepto de segmentación
- El ciclo de instrucción
- Segmentación del ciclo de instrucción
- Riesgos
- Riesgos de datos
- Riesgos de control
- Riesgos estructurales
- Excepciones



## Bibliografía



John L. Hennessy and David A. Patterson.

Computer Architecture, Fifth Edition: A Quantitative Approach. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 5 edition, 2012.

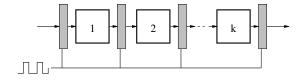
#### Índice

- 1 Concepto de segmentación
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



#### Segmentación

- Técnica que descompone un proceso en varios subprocesos
- Cada subproceso puede ejecutarse independientemente en un módulo autónomo
- Cada módulo opera concurrentemente con los demás.
- Un sistema se segmenta en *k* etapas:



→ Etapa: módulo que procesa un subproceso

registro de tipo latch

#### Segmentación (cont.)

- Registros inter-etapa:
  - Mantienen los datos estables durante el tiempo que necesita el módulo para hacer su función.
- Un reloj sincroniza el avance de los datos por las etapas. El reloj marca:
  - cuándo puede entrar un nuevo dato en la unidad segmentada.
  - el tiempo que las etapas disponen para efectuar su función.

### Segmentación (cont.)

#### Periodo de reloj

Caso ideal: mismo retardo en todos módulos.

$$\tau = \frac{D}{k}$$

- D: Retardo del circuito original.
- k: Número de etapas.
- Caso real: módulos con retardos distintos, registros inter-etapa y desfase del reloj.

$$au = \max_{i=1}^k ( au_i) + T_R + T_S \geq \frac{D}{k}$$

- $\blacksquare$   $T_R$ : Retardo del registro inter-etapa.
- $\blacksquare$   $T_S$ : Desfase del reloj.

### Mejora obtenida al segmentar

Aceleración (Speedup):

$$S = \frac{T_{ns}}{T_s} \approx \frac{nD}{n\tau} = \frac{D}{\tau}$$

- $T_{ns}$ : Tiempo para procesar *n* datos en unidad original.
- T<sub>s</sub>: Idem en la unidad segmentada.
- Caso ideal:  $\tau = \frac{D}{k}$ .

$$S = \frac{D}{\tau} = k$$

Caso real:  $\tau \geq \frac{D}{k}$ .

$$S = \frac{D}{\tau} \le k$$

#### Mejora obtenida al segmentar (cont.)

#### Productividad

Expresión general:

$$\chi = \frac{n}{T}$$
, siendo:

- n: datos procesados.
- T: Tiempo necesario para procesar n datos.
- Unidad no segmentada:

$$\chi_{ns} = \frac{n}{T_{ns}} = \frac{n}{nD} = \frac{1}{D}$$
 resultados/s

Unidad segmentada:

$$\chi_s = \frac{n}{T_s} \approx \frac{n}{n\tau} = \frac{1}{\tau}$$
 resultados/s

 $\rightarrow$  1 resultado cada  $\tau$  segundos = 1 resultado/ciclo de reloj.

#### Índice

- 1 Concepto de segmentación
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



#### 2. El ciclo de instrucción

#### Computador de ejemplo: MIPS reducido

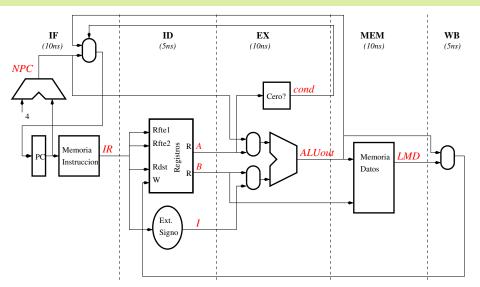
- Instrucciones aritméticas (reg—reg y reg—imm). Ejemplos: dadd Rdst,Rfte1,Rfte2 y daddi Rdst,Rfte1,Imm
- Cargas y almacenamientos. Ejemplos: ld Rdst, Desp (Rfte1) y sd Rfte2, Desp (Rfte1)
- Saltos condicionales: beqz Rfte1, Desp y bnez Rfte1, Desp

	0 5	6 10	11 15	16 20	21	31
R	Codop	R <sub>fte1</sub>	R <sub>fte2</sub>	R <sub>dst</sub>	Fui	nc (op)
	$\Rightarrow$ Instr.	UAL re	eg-reg: R <sub>d</sub>	$s_{t} \leftarrow R_{\mathit{fte}}$	op R <sub>fte2</sub>	
	0 5	6 10	11 15		16	31
I	Codop	R <sub>fte1</sub>	R <sub>fte2/dst</sub>		Imm/De	esp
	$\rightarrow$ Instr IIAI req-imm: $R_{++} \leftarrow R_{++}$ on Imm					

- $\Rightarrow$  Instr. UAL reg-imm:  $R_{dst} \leftarrow R_{fte1}$  op Imm
- $\Rightarrow$  Instr. Carga:  $R_{dst} \leftarrow M[R_{fte1} + Desp]$
- $\Rightarrow$  Instr. Almac.: M[R<sub>fte1</sub>+Desp]  $\leftarrow$  R<sub>fte2</sub>
- $\Rightarrow$  Instr. Salto cond.: if (R<sub>fte1</sub> =, $\neq$  0) then PC  $\leftarrow$  PC+4+Desp

#### 2. El ciclo de instrucción

#### Ruta de datos del MIPS



#### 2. El ciclo de instrucción

#### Ciclo de instrucción del MIPS

		Instrucciones	ejemplo		
	dadd Rdst,Rfte1,Rfte2				
Fase	daddi Rdst,Rfte1,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rfte1,Desp	
IF		Leer instrucción en Memo			
(10ns)		Calcular NPC: NF	<b>PC</b> ←PC+4;		
	Siguiente instrucción: if cond then <b>PC</b> ←ALUout else <b>PC</b> ←NPC;				
ID	Decodificar instrucción				
(5ns)	Leer operandos en Banco de Registros: A←Regs[Rfte1]; B←Regs[Rfte2];				
	Extender signo del campo inmediato/desplazamiento: I←Ext.Signo(Imm/Desp);				
EX				Calcular	
(10ns)	Calcular resultado:	Calcular	dirección:	destino y condición:	
	$ALUout \leftarrow A+(B \circ I);$	ALUoi	ut←A+I;	ALUout←NPC+I;	
				cond←A=0?;	
MEM		Leer dato en Memoria:	Escribir dato en Memoria:		
(10ns)		<b>LMD</b> ←Mem[ALUout];	Mem[ALUout]←B;		
WB	Escribir resultado en Banco de Registros:				
(5ns)	Regs[Rdst]←A	LUout o LMD			

#### Tiempo de ciclo T:

El tiempo de ciclo debe ser lo suficientemente largo para poder soportar la ejecución de la instrucción más larga.

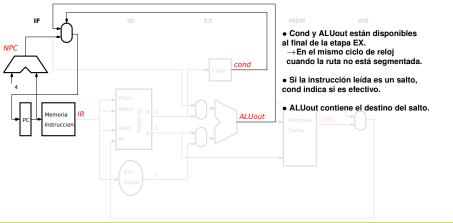
saltos ual almac. carga 
$$T = máx(25, 30, 35, 40) = 40 \text{ ns}$$

#### Índice

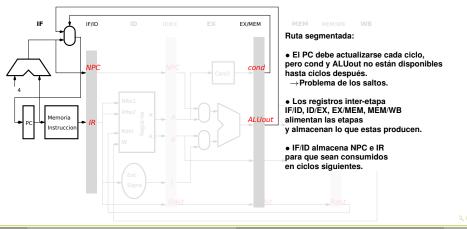
- 1 Concepto de segmentación
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



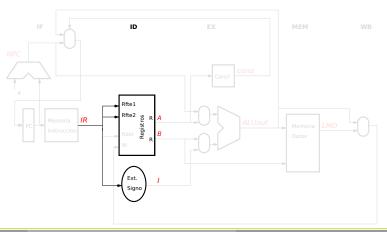
		Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2					
Etapa	daddi Rdst,Rfte1,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp		
IF	Leer instrucción en Memoria: IR←Mem[PC];					
	Calcular NPC: <b>NPC</b> ←PC+4;					
	Siguiente instrucción: if cond then <b>PC</b> ←ALUout else <b>PC</b> ←NPC;					



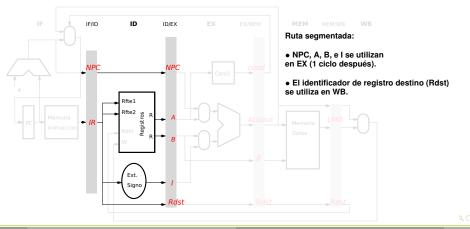
		Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2					
Etapa	daddi Rdst,Rftel,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rfte1,Desp		
IF	Leer instrucción en Memoria: IF/ID.IR←Mem[PC];					
İ	Calcular NPC: IF/ID.NPC←PC+4;					
	Siguiente instrucción: if EX/MEM.cond then PC←EX/MEM.ALUout else PC←NPC;					



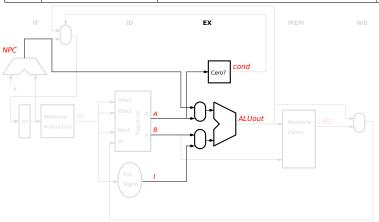
		Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2					
Etapa	daddi Rdst,Rfte1,Imm	ld Rdst,Desp(Rfte1)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp		
ID	Decodificar instrucción					
	Leer operandos en Banco de Registros: A←Regs[Rfte1]; B←Regs[Rfte2];					
	Extender sign	Extender signo del campo inmediato/desplazamiento: I←Ext.Signo(Imm/Desp);				



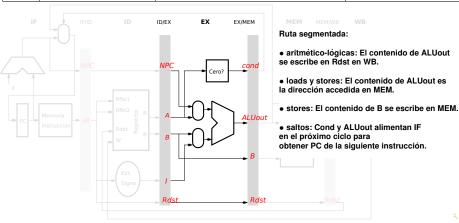
		Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2					
Etapa	daddi Rdst,Rfte1,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp		
ID	Decodificar instrucción					
	Leer operandos en Banco de Registros: ID/EX.A←Regs[IF/ID.IR.Rfte1]; ID/EX.B←Regs[IF/ID.IR.Rfte2];					
	Extender signo del c	ampo inmediato/desplazamier	nto: ID/EX.I←Ext.Signo(IF/ID.IR	I.lmm/Desp):		



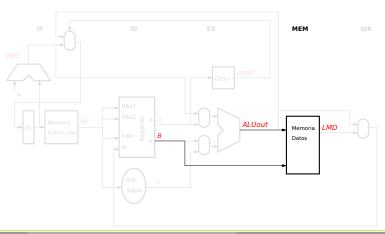
	Instrucciones ejemplo			
	dadd Rdst,Rfte1,Rfte2			
Etapa	daddi Rdst,Rfte1,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp
EX	Calcular resultado: <b>ALUout</b> ←A+(B o I);	Calcular dirección: ALUout←A+I;		Calcular destino y condición: ALUout←NPC+I; cond←A=0?;



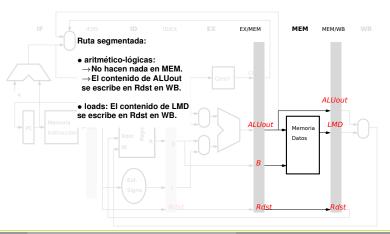
	Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2	ld	sd		
Etapa	daddi Rdst,Rfte1,Imm	Rdst, Desp (Rfte1)	Rfte2, Desp (Rfte1)	beqz Rfte1,Desp	
EX	Calcular resultado:	Calcular resultado: Calcular dirección:		Calcular destino y cond:	
	EX/MEM.ALUout	EX/MEM.ALUout←ID/EX.A+ID/EX.I;		EX/MEM.ALUout	
	←ID/EX.A+ID/EX.(B o I);		$\leftarrow$ ID/EX.NPC+ID/EX.I;		
				EX/MEM.cond←IF/ID.A=0?;	



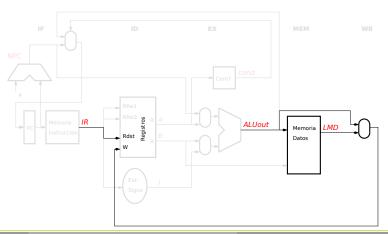
	Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2				
Fase	daddi Rdst,Rfte1,Imm	ld Rdst, Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp	
MEM		Leer dato en Memoria:	Escribir dato en Memoria:		
		<b>LMD</b> ←Mem[ALUout];	Mem[ALUout]←B;		



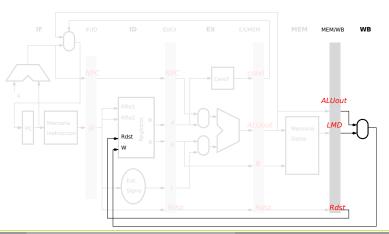
	Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2	ld	sd		
Etapa	daddi Rdst,Rftel,Imm	Rdst, Desp(Rfte1)	Rfte2,Desp(Rfte1)	beqz Rftel,Desp	
MEM		Leer dato en Mem.:	Escribir dato en Mem.:		
		MEM/WB.LMD←	Mem[ <b>EX/MEM.</b> ALUout]		
		Mem[EX/MEM.ALUout];	←EX/MEM.B;		



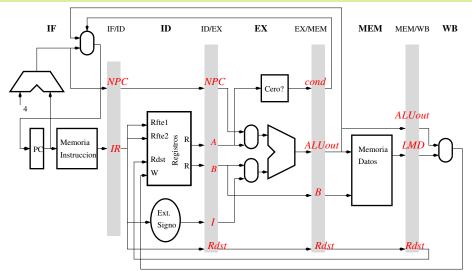
	Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2				
Fase	daddi Rdst,Rftel,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp	
WB	Escribir resultado en Banco de Registros:				
İ	Regs[Rdst]←ALUout o LMD				



	Instrucciones ejemplo				
	dadd Rdst,Rfte1,Rfte2				
Fase	daddi Rdst,Rfte1,Imm	ld Rdst,Desp(Rftel)	sd Rfte2,Desp(Rfte1)	beqz Rftel,Desp	
WB	Escribir resultado en Banco de Registros:				
	Regs[ <b>MEM/WB.</b> Rdst]← <b>MEM/WB.</b> (ALUout o LMD)				



# Ruta de datos segmentada



## Ciclo de instrucción segmentado

	Instrucciones ejemplo									
	dadd Rdst,Rfte1,Rfte2	ld	sd							
Fase	daddi Rdst,Rfte1,Imm	Rdst, Desp(Rfte1)	Rfte2, Desp (Rfte1)	beqz Rftel,Desp						
IF	Leer instrucción en Memoria: IF/ID.IR←Mem[PC];									
(10ns)	Calcular NPC: IF/ID.NPC←PC+4;									
	Siguiente instrucción: if <b>EX/MEM.</b> cond then PC← <b>EX/MEM.</b> ALUout else PC←NPC;									
ID	Decodificar instrucción									
(5ns)	Leer operandos en Band	Leer operandos en Banco de Registros: ID/EX.A   — Regs[IF/ID.IR.Rfte1]; ID/EX.B   — Regs[IF/ID.IR.Rfte2];								
	Extender signo del campo inmediato/desplazamiento: ID/EX.I←Ext.Signo(IF/ID.IR.Imm/Desp);									
EX	Calcular resultado:	Calcular o	Calcular destino y cond:							
(10ns)	EX/MEM.ALUout	EX/MEM.ALUout ←	EX/MEM.ALUout							
	←ID/EX.A+ID/EX.(B o I);		$\leftarrow$ ID/EX.NPC+ID/EX.I;							
			EX/MEM.cond←IF/ID.A=0?;							
MEM		Leer dato en Mem.:	Escribir dato en Mem.:							
(10ns)		MEM/WB.LMD←	Mem[EX/MEM.ALUout]							
		Mem[EX/MEM.ALUout];	←EX/MEM.B;							
WB	Escribir resultado en B									
(5ns)	Regs[ <b>MEM/WB.</b> Rdst]← <b>MEM/WB.</b> (ALUout o LMD)									

#### Tiempo de ciclo T:

- El tiempo de ciclo debe ser lo suficientemente largo para poder soportar la ejecución de la etapa más larga.
- T = máx(5, 10) = 10 ns

## Ejecución de las instrucciones

i	IF	ID	EX	MEM	WB				
i + 1		IF	ID	EX	MEM	WB			
i + 2			IF	ID	EX	MEM	WB		
<i>i</i> + 3				IF	ID	EX	MEM	WB	
<i>i</i> + 4					IF	ID	EX	MEM	WB

#### Aceleración:

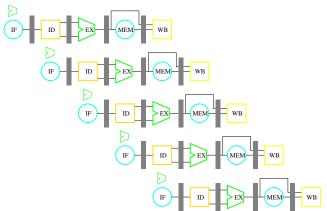
			Segmentado				
I	n						n
CPI				1			1
		saltos	ual	almac.	carga		
T	máx(	25,	30,	35,	40	) = 40  ns	máx(5, 10) = 10 ns
$T_{\rm ej} = I \cdot CPI \cdot T$			$n \cdot 1 \cdot 10 = 10n \text{ ns}$				

$$S = 4$$

El número de etapas k = 5 es la cota superior de la aceleración.

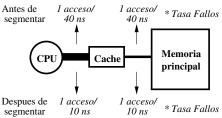
#### Requisitos hardware:

■ En un ciclo dado, hay 5 instrucciones en ejecución → hay que evitar conflictos en las unidades funcionales.

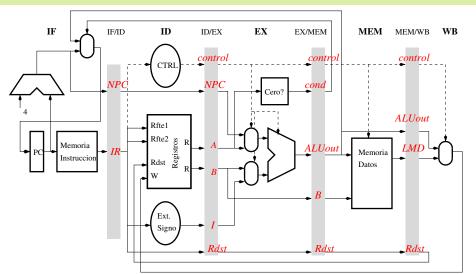


#### Requisitos *hardware:* (cont.)

- Caché de instrucciones (IF) y de datos (MEM) separadas.
- Banco de registros con dos puertos de lectura (ID) y uno de escritura (WB) simultáneos.
- El tiempo de acceso a la memoria caché no varía, pero el ancho de banda requerido es 4 veces superior:



#### Señales de control



#### Señales de control

IF	ID	EX	MEM	WB	
Mem_Read	Reg_Read	ALU_Op	Mem_Read	Reg_Write	
PC_Src		ALU_Op1	Mem_Write	Mem_to_Reg	
		ALU_Op2			
PC_Src  Salto Efectivo Salto no efectivo		ALU_op1 Ins. Salto Otras ins.  Otras ins.  ALU_op2	Mem_Read Mem_Write	Mem_to_Reg Otras ins Ins. Carga	

Las señales de control necesarias en las etapas EX, MEM y WB se generan en la etapa ID.

#### Índice

- 1 Concepto de segmentación
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



## 4. Riesgos

#### Concepto y clasificación

Riesgo (hazard): Situación que produce que la ejecución de algunas instrucciones genere resultados diferentes a los de la ruta de datos no segmentada. ⇒ Pérdida de compatibilidad binaria.
 Los riesgos se producen entre dos o más instrucciones que están presentes de forma simultánea en la ruta de datos segmentada.

# Tipos de riesgos:

#### Datos

El resultado de una instrucción se utiliza como dato en la(s) instrucción(es) siguiente(s).

#### Control

En el flujo de instrucciones aparece una instrucción de salto.

#### Estructurales

Dos o más instrucciones pretenden utilizar el mismo recurso.

#### 4. Riesgos

#### Soluciones a los riesgos

Inserción de ciclos de parada Impedir el avance de la instrucción que origina el conflicto y de todas las que le siguen, pero no de las que le preceden (si no, el conflicto jamás desaparecería).

- En estos ciclos no se buscan nuevas instrucciones (ciclos de parada o stalls).
- Se origina una pérdida de prestaciones:

$$\mathsf{CPI}_{\mathcal{S}} = \mathsf{CPI}_{\mathcal{S}_{ideal}} + \frac{\mathsf{Ciclos}\;\mathsf{de}\;\mathsf{parada}}{\mathsf{instruccion}} = 1 + \frac{\mathsf{stalls}}{\mathsf{instruccion}} > 1$$

### 4. Riesgos

#### Soluciones a los riesgos (cont.)

- Modificación de la ruta de datos para detectarlos y resolverlos dinámicamente. La modificación puede:
  - Reducir el número de ciclos de parada necesarios para resolverlo
  - Resolver el riesgo completamente (a veces no es posible).
- Modificación de la arquitectura del juego de instrucciones Impedir que aparezca el problema prohibiendo la generación de ciertas secuencias de instrucciones. Para evitarlas, el compilador puede insertar instrucciones NOP o reordenar instrucciones independientes.
  - Inconveniente: Se pierde la compatibilidad a nivel binario.

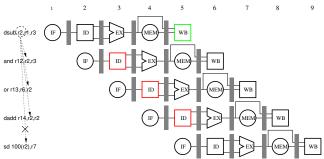
#### Índice

- 1 Concepto de segmentación
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



#### Causa

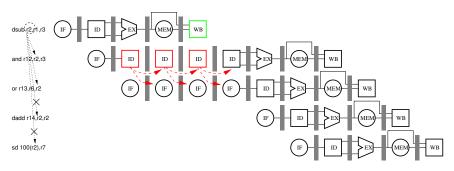
Una instrucción depende de los resultados producidos por otra anterior  $\rightarrow$  al segmentar se puede invertir el orden de acceso a los operandos.



Se puede dar hasta entre 4 instrucciones consecutivas.

#### Inserción de ciclos de parada (stalls)

Retrasar las operaciones que causan el conflicto.

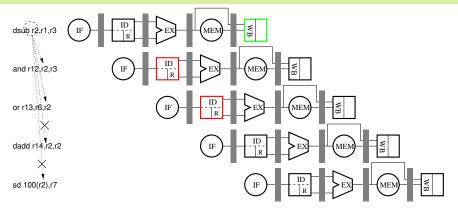


- ightarrow 3 stalls por cada dos instrucciones dependientes consecutivas.
- ightarrow 2 stalls por cada dos instr. dependientes separadas por 1 instr.
- → 1 stall por cada dos instr. dependientes separadas por 2 instr.
- ⇒ Notable pérdida de prestaciones.

#### Reducción del número de instrucciones involucradas

- ⇒ Modificación de la ruta de datos
  - Banco de registros multipuerto soportando lectura y escritura simultáneas.
  - Modificación del acceso al banco de registros: Lectura de registros → Segundo semiciclo. Escritura en registros → Primer semiciclo.
    - No hay problema si el tiempo de acceso al banco de registros es ≤ a la mitad del periodo de reloj.
    - Simplifica el diseño del banco de registros: No es necesario soportar lectura y escritura simultáneas.

### Reducción del número de instrucciones involucradas (cont.)



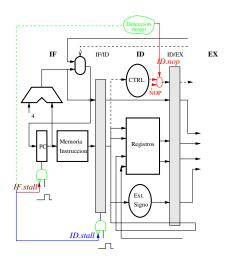
- ⇒ el número de instrucciones involucradas es 3: **dsub**, **and** y **or**.
- → Penalización por instrucciones dependientes consecutivas: 2 stalls

#### Inserción de 2 ciclos de parada

- Detectar las situaciones (en negrita) donde debe insertarse un ciclo de parada.
- Para insertar el ciclo de parada:
  - Poner las señales de control que circulan hacia la etapa EX como si de una instrucción NOP se tratara.
  - Conservar las instrucciones en IF y ID.

#### Inserción de 2 ciclos de parada (cont.)

# Lógica de control



#### Inserción de 2 ciclos de parada (cont.)

#### Primer ciclo:

```
if (ID/EX.IR.CODOP = "ALU") and
   (IF/ID.IR.CODOP = "ALU") and
   ((ID/EX.IR.Rdst = IF/ID.IR.Rfte1) or
    (ID/EX.IR.Rdst = IF/ID.IR.Rfte2))
t.hen
   IF.stall, ID.stall, ID.nop
```

#### Segundo ciclo:

```
if (EX/ME.IR.CODOP = "ALU") and
   (IF/ID.IR.CODOP = "ALU") and
   ((EX/MEM.IR.Rdst = IF/ID.IR.Rfte1) or
    (EX/MEM.IR.Rdst = IF/ID.IR.Rfte2))
t.hen
   IF.stall, ID.stall, ID.nop
```

#### Cortocircuitos

Problema: 2 *stalls* por riesgo de datos entre instrucciones consecutivas

Consideremos el riesgo entre las dos primeras instrucciones si no insertáramos ciclos de parada:

```
dsub r2,r1,r3 IF ID EX ME WB and r12,r2,r3 IF ID EX ME WB
```

- and necesita el dato al principio del ciclo 4 (su etapa EX)
- dsub tiene el resultado al principio del ciclo 4 (su etapa MEM)

### Cortocircuitos (cont.)

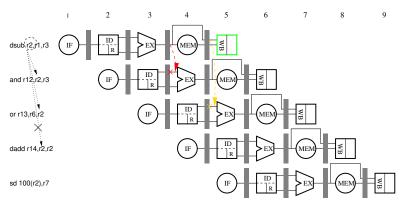
⇒ Modificación de la ruta de datos

Añadir un bus desde la salida de la UAL (etapa MEM) hacia su entrada (etapa EX) + lógica de control  $\Rightarrow$  "Cortocircuito" de MEM hacia EX.

Si consideramos la tercera instrucción, el riesgo se puede resolver de forma similar con un cortocircuito de WB hacia EX.

#### Cortocircuitos (cont.)

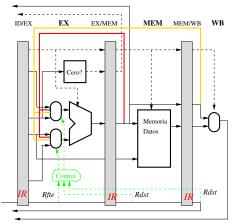
Instrucciones consumidoras posteriores ya no requieren cortocircuito.



### Cortocircuitos (cont.)

## Implementación

Multiplexores a la entrada de la UAL.



#### Cortocircuitos (cont.)

### Lógica de control

- Cortocircuito de MEM a EX:
  - Rfte1:

```
if (EX/MEM.IR.CODOP = "ALU") and
  (ID/EX.IR.CODOP = "ALU") and
  (EX/MEM.IR.Rdst = ID/EX.IR.Rfte1)
then
  Cortocircuito MEM-a-EX (entrada sup. UAL)
```

Rfte2:

```
if (EX/MEM.IR.CODOP = "ALU") and
  (ID/EX.IR.CODOP = "ALU") and
  (EX/MEM.IR.Rdst = ID/EX.IR.Rfte2)
then
  Cortocircuito MEM-a-EX (entrada inf. UAL)
```

## Cortocircuitos (cont.)

- Cortocircuito de WB a EX:
  - Rfte1:

```
if (MEM/WB.IR.CODOP = "ALU") and
  (ID/EX.IR.CODOP = "ALU") and
  (MEM/WB.IR.Rdst = ID/EX.IR.Rfte1)
then
  Cortocircuito WB-a-EX (entrada sup. UAL)
```

■ Rfte2:

```
if (MEM/WB.IR.CODOP = "ALU") and
  (ID/EX.IR.CODOP = "ALU") and
  (MEM/WB.IR.Rdst = ID/EX.IR.Rfte2)
then
```

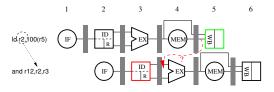
Cortocircuito WB-a-EX (entrada inf. UAL)

### Instrucción consecutiva y dependiente de una load

## Consideremos el riesgo entre las instrucciones:

```
ld r2,100(r5) and r12,r2,r5
```

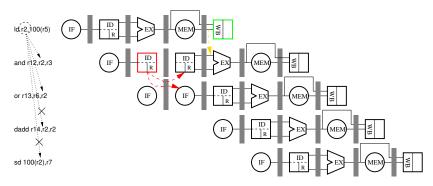
- ¿Cuándo necesita and el dato? En el ciclo 4 (etapa EX)
- ¿Cuándo tiene 1d el resultado? En el ciclo 5 (etapa WB)



⇒ No se puede resolver usando solamente la técnica del cortocircuito.

Instrucción consecutiva y dependiente de una *load* (cont.)

Utilizando el cortocircuito de WB a EX hace falta insertar 1 ciclo de parada.



#### Instrucción consecutiva y dependiente de una *load* (cont.)

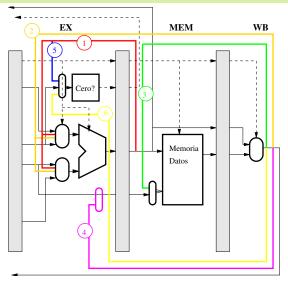
Lógica de control para insertar el ciclo de parada

```
if (ID/EX.IR.CODOP = "LOAD") and
   (IF/ID.IR.CODOP = "ALU") and
   ((ID/EX.IR.Rdst = IF/ID.IR.Rfte1) or
    (ID/EX.IR.Rdst = IF/ID.IR.Rfte2))
t.hen
   IF.stall, ID.stall, ID.nop
```

## Resumen de cortocircuitos y ciclos de parada

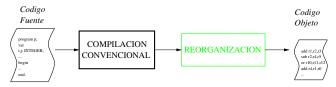
Tipos			Ciclos	
de instrucciones	Ejemplo	Cortocircuito	de parada	Fig.
UAL - UAL	DADD R1,R2,R3			
	DSUB R4,R1,R5	MEM a EX	0	1
	AND R7,R1,R6	WB a EX	0	2
Carga - UAL	LD R1,20(R2)			
	DADD R3,R1,R4	WB a EX	1	2
UAL - Carga/Almac.	DADD R1,R2,R3			
	LD R2,20(R1)	MEM a EX	0	1
	LD R3,40(R1)	WB a EX	0	2
UAL - Almac.	DADD R1,R2,R3			
	SD R1,20(R2)	WB a MEM	0	3
	SD R1,40(R2)	WB a EX	0	4
Carga - Almac.	LD R1,20(R3)			
	SD R1,20(R2)	WB a MEM	0	3
	SD R1,40(R2)	WB a EX	0	4
Carga - Carga/Almac.	LD R1,30(R3)			
	LD R2,20(R1)	WB a EX	1	2
UAL - Salto	DSLT R1,R2,R3			
	BEQZ R1,loop	MEM a EX	0	5
UAL - Salto	DSLT R1,R2,R3			
	BEQZ R1,loop	WB a EX	0	6
Carga - Salto:	LD R1,20(R2)			
	BEQZ R1,loop	WB a EX	1	6

## Resumen de cortocircuitos y ciclos de parada (cont.)



#### Reorganización del código

⇒ El compilador tiene una fase más: reorganización del código:



⇒ El objetivo es evitar ciclos de parada. Ejemplo:

Código convencional	Código reorganizado		
LD Rb,b	LD Rb,b		
LD Rc,c	LD Rc,c		
DADD Ra,Rb,Rc	LD Re,e		
LD Re,e	DADD Ra,Rb,Rc		
LD Rf,f	LD Rf,f		
DSUB Rd,Re,Rf	SD Ra,a		
SD Ra,a	DSUB Rd,Re,Rf		
SD Rd,d	SD Rd,d		
8 ins + 2 stalls = 10 ciclos	8 ins = 8 ciclos		

# Tema 2.1 Unidad de instrucción segmentada

#### Índice

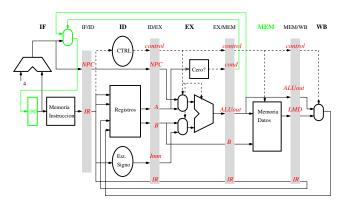
- 1 Concepto de segmentació
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de control
- 7 Riesgos estructurales
- 8 Excepciones



#### Causa

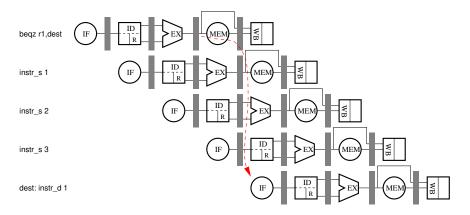
En el flujo de instrucciones aparecen instrucciones que modifican el contador de programa (PC).

Las instrucciones de salto modifican el PC en la etapa MEM.



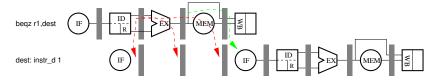
#### Causa (cont.)

En ese ciclo se han lanzado ya 3 instrucciones:



#### Inserción de ciclos de parada

Insertar ciclos de parada *siempre* que se decodifica una instrucción de salto:



3 ciclos de parada  $\rightarrow$  Pérdida de prestaciones.

#### Inserción de ciclos de parada (cont.)

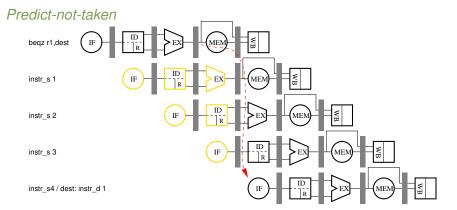
Lógica de control para la inserción de 3 ciclos de parada

Inhibir la etapa IF durante tres ciclos de reloj, pasando instrucciones NOP a la etapa ID:

#### Predicción (fija)

- ⇒ Modificación de la ruta de datos
- *Predict-not taken*: Suponer el salto no efectivo  $\rightarrow$  las instrucciones buscadas a continuación de la de salto son válidas.
  - Si, finalmente, el salto es efectivo, se abortan las tres instrucciones en curso.
  - IMPORTANTE: Estas instrucciones no deben haber modificado el estado.
- *Predict-taken*: Suponer el salto efectivo  $\rightarrow$  en cuanto se conoce la dirección destino se buscan las nuevas instrucciones.
  - Si, finalmente, el salto es no efectivo, se abortan estas instrucciones.
  - Solo es útil si se conoce la dirección destino antes que la condición de salto → inútil en MIPS.

## Predicción (fija) (cont.)



### Lógica de control

if EX/MEM.cond then
 IF.nop, ID.nop, EX.nop

#### Reducción de la latencia de salto

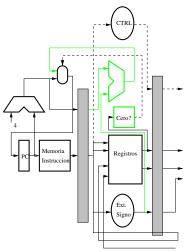
⇒ Modificación de la ruta de datos

Reducir el número de ciclos comprendidos entre la búsqueda de la instrucción de salto y la correspondiente al destino del salto.

- Actualizar el PC en la etapa  $EX \rightarrow N^o$  de ciclos = 2.
  - $\Rightarrow$  En la etapa EX del salto se conoce el nuevo PC para el ciclo siguiente.
- Actualizar el PC en la etapa ID → Nº de ciclos = 1.
  ⇒ En la etapa ID del salto se conoce el nuevo PC para el ciclo siguiente. Requiere:
  - $\blacksquare$  Trasladar el cálculo de la dirección efectiva de EX a ID  $\to$  hace falta un sumador adicional
  - Trasladar la evaluación de la condición de EX a ID.

### Reducción de la latencia de salto (cont.)

Implementación de la actualización del PC en ID



#### Reducción de la latencia de salto (cont.)

### Lógica de control para actualizar el PC en ID

- Asumiendo predict-not-taken, si el salto no es efectivo, se busca la siguiente instrucción.
- Pero si el salto es efectivo, la instrucción buscada en el mismo ciclo debe cancelarse.  $\rightarrow$  Se entrega una NOP a la etapa ID.

```
IF ID EX ME WB
 begz rl, dest
 <sqte>
                    IF X
 <dest>
                        TF TD EX ME WB
if (IF/ID.IR.CODOP = "Salto") and
   (Regs[IF/ID.IR.Rfte1] op 0)
t.hen
   IF.nop
   PC <- IF/ID.NPC + Ext.Signo(IF/ID.IR.Desp)</pre>
else
   PC \leftarrow PC + 4
```

#### Reducción de la latencia de salto (cont.)

Impacto en el periodo de reloj

#### Etapa ID

tiempo de acceso a los registros + retardo de la evaluación de la condición + selección + escritura sobre el PC

- escritura sobre eri C
- → ID puede convertirse en la etapa más lenta
- $\Rightarrow$  las condiciones de salto deben ser sencillas (= y  $\neq$ ).

Esta modificación puede ser incompatible con la lectura de registros en el segundo semiciclo.

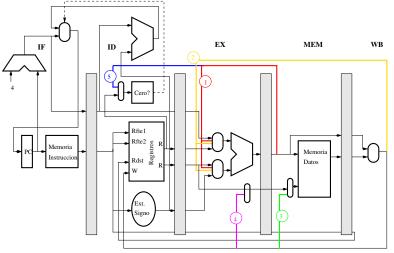
#### Reducción de la latencia de salto (cont.)

## Impacto en cortocircuitos y ciclos de parada

Instrucciones	Ejemplo	Cortocircuito	stalls	Fig.
UAL - UAL	DADD R1,R2,R3			
	DSUB R4,R1,R5	MEM a EX	0	1
	AND R7,R1,R6	WB a EX	0	2
Carga - UAL	LD R1,20(R2)			
	DADD R3,R1,R4	WB a EX	1	2
UAL - Carga/Almac.	DADD R1,R2,R3			
	LD R2,20(R1)	MEM a EX	0	1
	LD R3,40(R1)	WB a EX	0	2
UAL - Almac.	DADD R1,R2,R3			
	SD R1,20(R2)	WB a MEM	0	3
	SD R1,40(R2)	WB a EX	0	4
Carga - Almac.	LD R1,20(R3)			
	SD R1,20(R2)	WB a MEM	0	3
	SD R1,40(R2)	WB a EX	0	4
Carga - Carga/Almac.	LD R1,30(R3)			
	LD R2,20(R1)	WB a EX	1	2
UAL - Salto	DSLT R1,R2,R3	MEM a EX	0	
	BEQZ R1,loop	MEM a ID	1	5
UAL - Salto	DSLT R1,R2,R3			
		WB a EX		
	BEQZ R1,loop	MEM a ID	0	5
Carga - Salto	LD R1,20(R2)	WB a EX	4	
	BEQZ R1,loop	Por BR	2	-
Carga - Salto	LD R1,20(R2)			
			0	
	BEQZ R1,loop	Por BR	L 1 a	

### Reducción de la latencia de salto (cont.)

Impacto en cortocircuitos y ciclos de parada



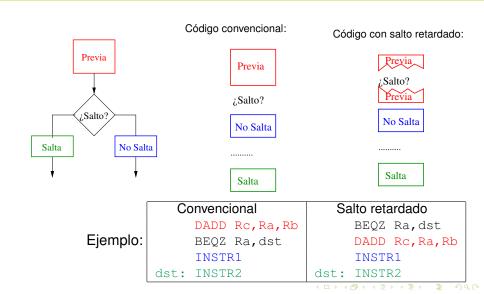
#### Salto retardado (delayed branch)

- → Modificación de la arquitectura del juego de instrucciones El compilador debe colocar tras los saltos instrucciones que se ejecutarán tanto si se salta como si no se salta.
  - Como estas instrucciones siempre se ejecutan, no es necesario introducir ciclos de parada o cancelarlas.
  - En general, se eligen instrucciones anteriores al salto que no dependan de la condición del salto.

## Branch delay slot

- Número de instrucciones tras el salto que se ejecutarán siempre independientemente de si es efectivo o no.
  - → Coincide con la latencia del salto.
  - $\rightarrow$  1 ciclo si se actualiza el PC en ID.

### Salto retardado (*delayed branch*) (cont.)



## Salto retardado (delayed branch) (cont.)

Salto retardado en los juegos de instrucciones actuales

El salto retardado deja de utilizarse aproximádamente a partir de los 90. Algunas razones:

- Si el branch delay slot es elevado (esto es, mayor que una instrucción), el compilador tiene más dificultad para encontrar instrucciones útiles para colocar tras los saltos. Esto ocurre si:
  - La latencia de salto es elevada (p.e. la ruta de datos tiene muchas etapas).
  - Se lanzan múltiples instrucciones por ciclo de reloj (Tema 2.5).
- Condiciona el juego de instrucciones a una implementación de la ruta de datos, lo cual es inflexible. Por ejemplo, un diseño de ruta de datos diferente podría variar la latencia de salto, perdiéndose la compatibilidad.
- Hay otras soluciones mucho mejores basadas en realizar una predicción dinámica de los saltos (Tema 2.3).

# Tema 2.1 Unidad de instrucción segmentada

#### Índice

- 1 Concepto de segmentación
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



#### Causas

- El hardware no permite todas las combinaciones posibles de las instrucciones presentes en la unidad.
  - → Un recurso no ha sido replicado suficientemente.
- **Ejemplo**: Procesador con caché única de instrucciones y datos.
  - → La etapa MEM de las instrucciones de Carga/Almac. colisiona con la etapa IF de la instrucción que ocupa el 3<sup>er</sup> puesto tras estas.

#### Soluciones

#### Modificaciones de la ruta de datos

Replicar el recurso para que sea posible esa combinación.

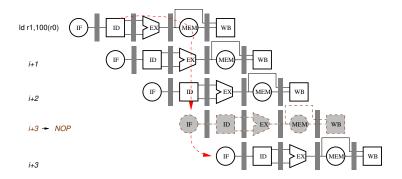
- → Ejemplo: Arquitectura *Harvard*: utiliza caché de instrucciones y datos separadas.
- → Aumento del coste.
- → No siempre es posible o tiene sentido replicar el recurso.

## Inserción de ciclos de parada

Retrasar una de las operaciones que causa el conflicto.

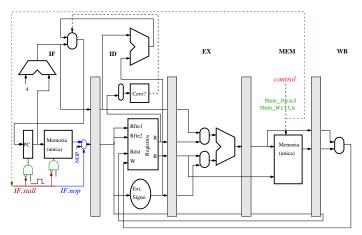
- $\rightarrow$  stalls  $\rightarrow$  Pérdida de prestaciones.
- → Las prestaciones dependen de la frecuencia de aparición de las combinaciones que originan los riesgos estructurales.

# Inserción de ciclos de parada



# Inserción de ciclos de parada (cont.)

# Implementación



### Inserción de ciclos de parada (cont.)

Lógica de control

Cuando una instrucción de Carga/Almac. está en la etapa MEM,

- No se accede a la memoria de instrucciones.
- Conservando la instrucción de la etapa IF.
- Entregando a la etapa ID una instrucción NOP.

```
if EX/MEM.Mem_Read or EX/MEM.Mem_Write then
    IF.stall, IF.nop
```

# Tema 2.1 Unidad de instrucción segmentada

## Índice

- 1 Concepto de segmentació
- 2 El ciclo de instrucción
- 3 Segmentación del ciclo de instrucción
- 4 Riesgos
- 5 Riesgos de datos
- 6 Riesgos de contro
- 7 Riesgos estructurales
- 8 Excepciones



# Concepto y clasificación

**Denominaciones:** Interrupción, **excepción** o falta.

# Tipos:

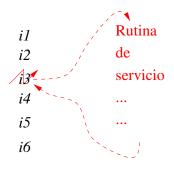
- Sincrona vs. asíncrona. Es síncrona si el evento ocurre en el mismo lugar cada vez que el programa se ejecuta.
- Solicitada por el usuario *vs.* lanzada hacia el usuario.
- Enmascarables por el usuario *vs.* no enmascarables.
- En medio de una instrucción vs. entre instrucciones.
- Continuar vs. terminar el programa.

Una excepción es un salto condicional implícito  $\to$  Si se produce la excepción, hay que saltar a la rutina de servicio.

# Excepciones posibles en el MIPS

Etapa	Excepciones
IF	Fallo de página de instrucción, Acceso desalineado
	Violación de protección, Petición E/S
ID	Instrucción ilegal, Petición E/S
EX	Excepción aritmética, Petición E/S
MEM	Fallo de página de datos, Acceso desalineado
	Violación de protección, Petición E/S
WB	Petición E/S

# Excepciones en computadores convencionales

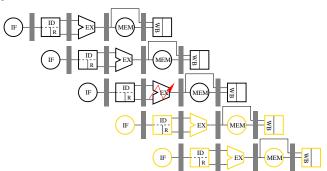


Secuencia correcta: ... i1, i2, i3 - Rutina de servicio - i3, i4, i5, i6 ...

# Excepciones en unidades de instrucción segmentadas

En el momento en que se produce la excepción, hay varias instrucciones en ejecución. Problema:

Hay instrucciones en ejecución posteriores a la que origina la excepción.



## Excepciones en unidades de instrucción segmentadas (cont.)

El funcionamiento es incorrecto:



Secuencia: ... *i*1, *i*2, *i*3, *i*4, *i*5 - Rutina de servicio - *i*3, *i*4, *i*5, *i*6 ...

El PC de las instrucciones solo se usa en IF.

Las primeras máquinas segmentadas finalizaban el programa, imprimiendo el PC de la instrucción que se encontraba en IF  $\Rightarrow$  indicaban *aproximadamente* la instrucción que había originado la excepción  $\rightarrow$  excepciones *imprecisas*.

## Excepciones precisas en unidades de instrucción segmentadas

Un computador soporta un comportamiento *preciso* frente a las excepciones si:

- Las instrucciones anteriores a la que origina la excepción terminan correctamente.
- La instrucción que origina la excepción y todas las posteriores son abortadas.
- Tras completar la rutina de servicio se puede relanzar el programa comenzando por la instrucción que originó la excepción.
- ⇒ Se puede identificar la instrucción causante de la excepción.
- ⇒ El comportamiento es idéntico al que tendría el mismo computador no segmentado.

# Implementación de excepciones precisas en el MIPS

# Requisitos:

- Se pueden producir no una excepción, sino hasta 5 excepciones.
- En el mismo o en distintos ciclos de reloj.
- Teniendo en cuenta cómo funciona el salto retardado.

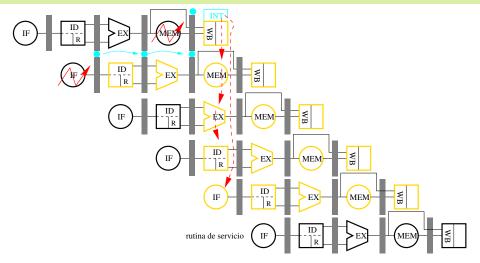
Idea: Asegurar que el orden de atención de las excepciones se realiza en el orden natural.

- ⇒ La llegada de las instrucciones a la última etapa se realiza en orden
  - Cada instrucción que entra en la unidad tiene asociado un registro con tantos bits como etapas en las que se pueden originar excepciones, el cual acompaña a la instrucción durante su recorrido por la misma.

## Implementación de excepciones precisas en el MIPS (cont.)

- Si se produce una excepción, se pone a "1" el bit de la etapa correspondiente, al tiempo que se convierte en NOP la instrucción implicada.
- En la última etapa del ciclo de instrucción se verifica si alguno de los bits está activado.
  - En caso afirmativo, convierte en NOP las instrucciones posteriores y escribe la dirección de la rutina de servicio en el PC.
- Se guarda la dirección de la instrucción que origina la excepción.
- La rutina de servicio toma el control.
- 6 Cuando finaliza la rutina de servicio, se restaura el PC con la dirección de la instrucción que originó la excepción, continuando la ejecución desde este punto.

# Implementación de excepciones precisas en el MIPS (cont.)



# Implementación de excepciones precisas en el MIPS (cont.)

