

Procesadores para cómputo de altas prestaciones

Tema 2 Procesadores Segmentados

Contenidos

- ✓ Introducción: MPIS-DLX
- Excepciones y control
- ✓ Segmentación
- Riesgos: Estructurales, de datos y de control
- ✓ Segmentación del procesador. Diseño del control
- ✓ Diseño del control con riesgos
- Excepciones: una segunda mirada
- ✓ Operaciones multi-ciclo
- ✓ Un Ejemplo: MIPS R4000
- ✓ Bibliografía
 - Apéndice A [HePa07]
 - o Capítulos 4 y 5 de [SiFK97]
 - o Simulador WinDLX

Recordatorio

☐ Arquitectura MIPS (DLX)

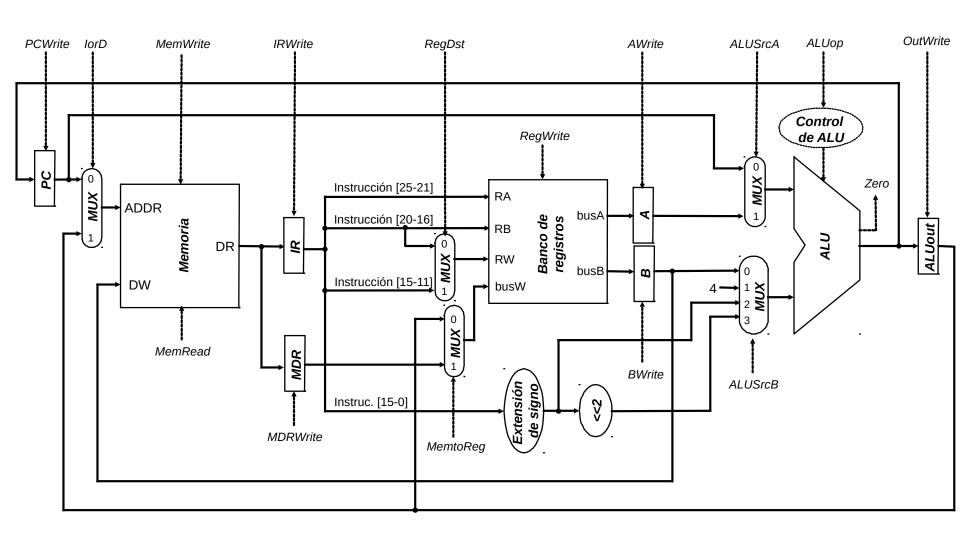
Todas las instrucciones del repertorio del MIPS tienen 32 bits de anchura, repartidas en 3 formatos de instrucción diferentes: 31 26 21 16 6 11 Tipo R: rd shamt funct op rt rs aritmético-lógicas 6 bits 5 bits 5 bits 6 bits 5 bits 5 bits 26 31 21 16 Tipo I: inmediato con memoria rt op rs salto condicional 6 bits 5 bits 5 bits 16 bits 31 26 0 Tipo J: dirección op salto incondicional 6 bits 26 bits

✓ El significado de los campos es:

- op: identificador de instrucción
- rs, rt, rd: identificadores de los registros fuentes y destino
- shamt: cantidad a desplazar (en operaciones de desplazamiento)
- funct: selecciona la operación aritmética a realizar
- inmediato: operando inmediato o desplazamiento en direccionamiento a registro-base
- dirección: dirección destino del salto

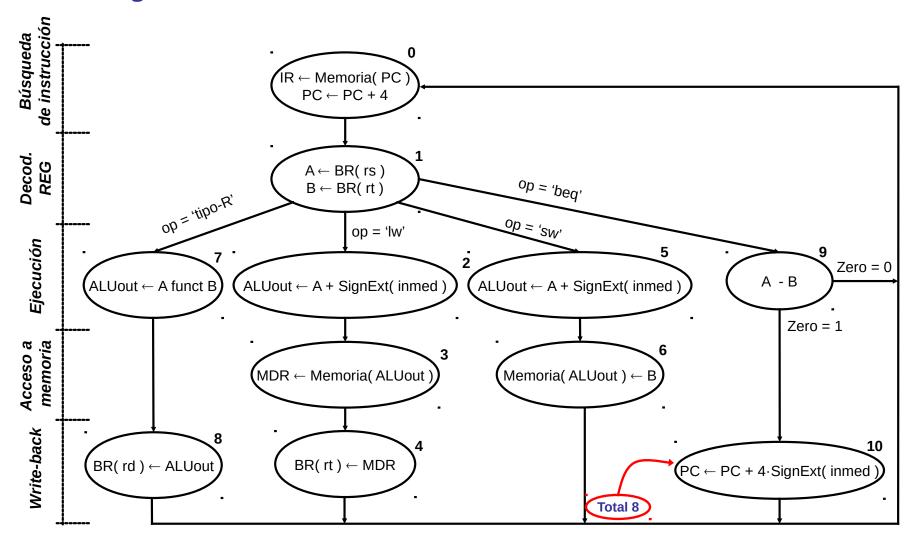
Recordatorio

☐ Ruta de datos (multiciclo)



Recordatorio

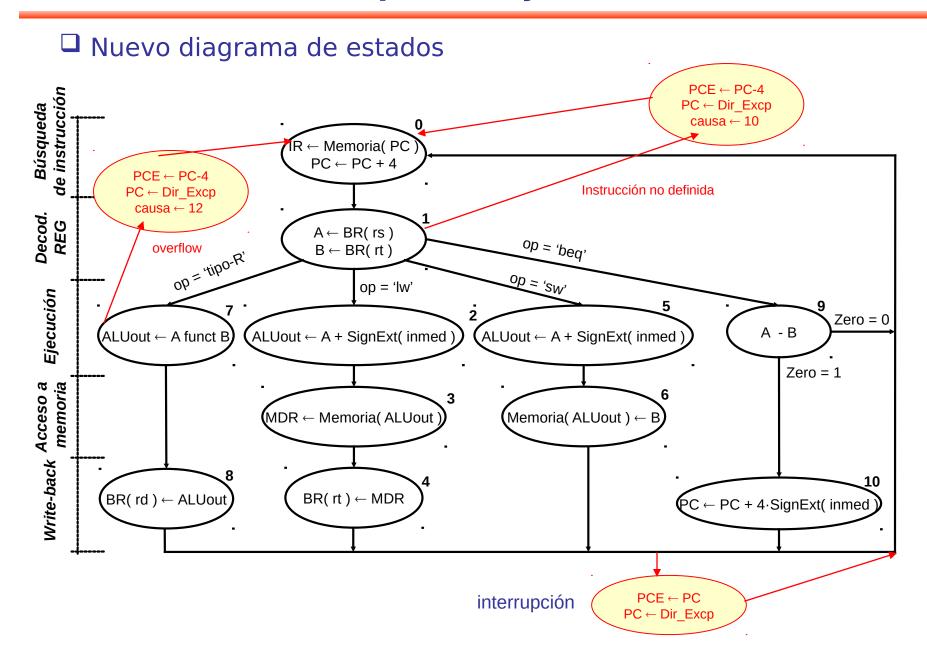
Diagrama de estados del controlador multiciclo



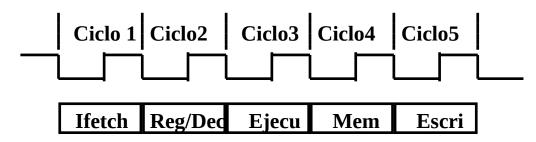
Excepciones y control

- Dos tipos
 - ✓ Interrupciones
 - Se producen a causa de sucesos externos al procesador.
 - Son asíncronas a la ejecución del programa.
 - Se pueden tratar entre instrucciones
 - ✓ Traps
 - Se producen por causas internas. Overflow, errores, fallos de pagina...
 - Son síncronas con la ejecución del programa
 - La condiciones deben ser almacenadas.
 - El programa debe ser abortado o continuado desde esa instrucción
- Requerimientos
 - ✓ Tratamiento dentro del control
 - ✓ Salvar el estado (CP y la causa)
 - Stack 68k, x86
 - Registros específicos CPE, Causa

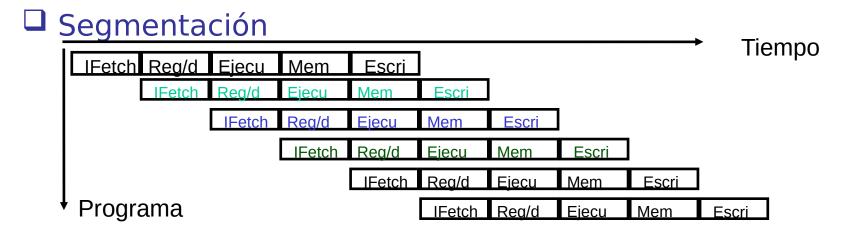
Excepciones y control

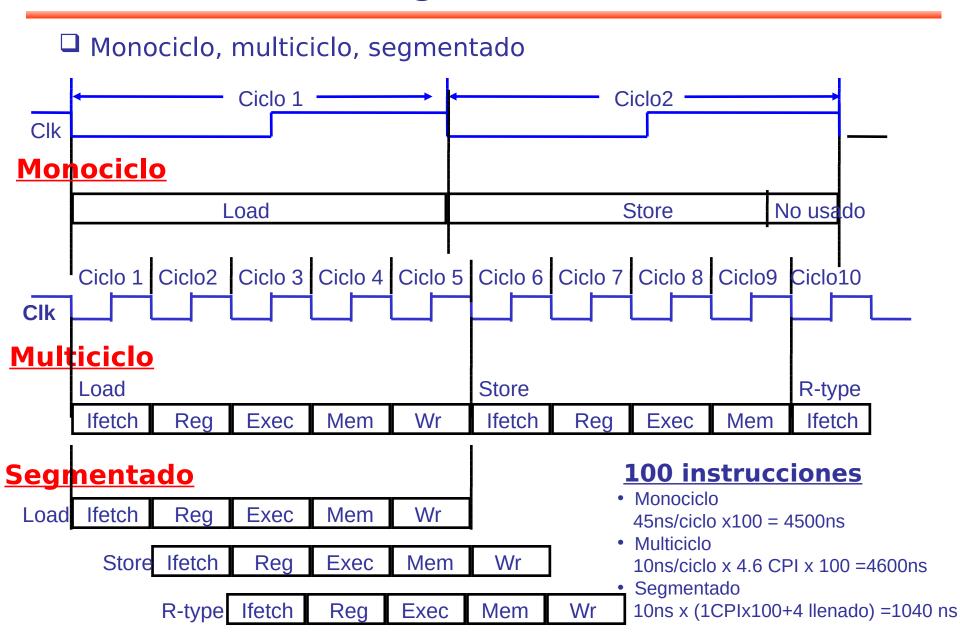


Etapas de una instrucción (Load)



- ° lfetch (búsqueda). Lee una instrucción desde la memoria
- ° Reg/dec. Decodifica la instrucción y lee registros
- ° Ejecuta. Calcula dirección de memoria
- ° Mem. Lee el dato de la memoria
- ° Escri. Escribe el dato en el registro





Riesgos

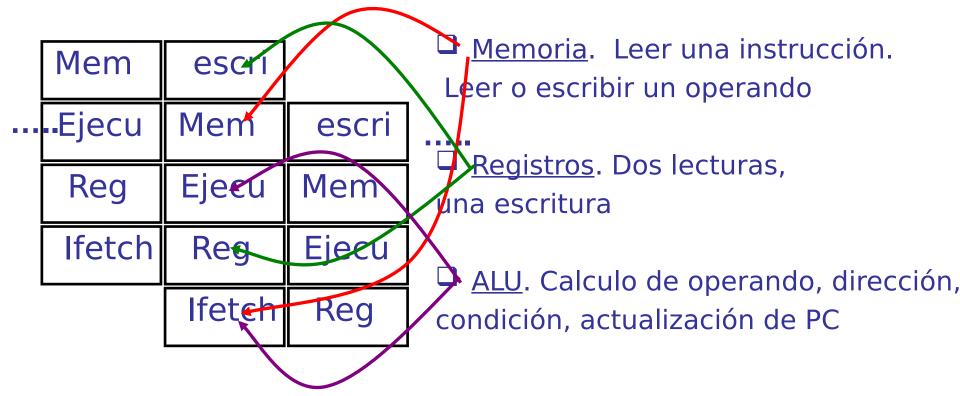
✓ Situaciones que impiden que cada ciclo se inicie la ejecución de una nueva instrucción

✓ Tipos:

- ✓ Estructurales. Se producen cuando dos instrucciones tratan de utilizar el mismo recurso en el mismo ciclo.
- ✓ De datos. Se intenta utilizar un dato antes de que este preparado. Mantenimiento del orden estricto de lecturas y escrituras.
- ✓ De control. Intentar tomar una decisión sobre una condición todavía no evaluada.
- ✓ Los riesgos se deben <u>detectar y resolver</u>

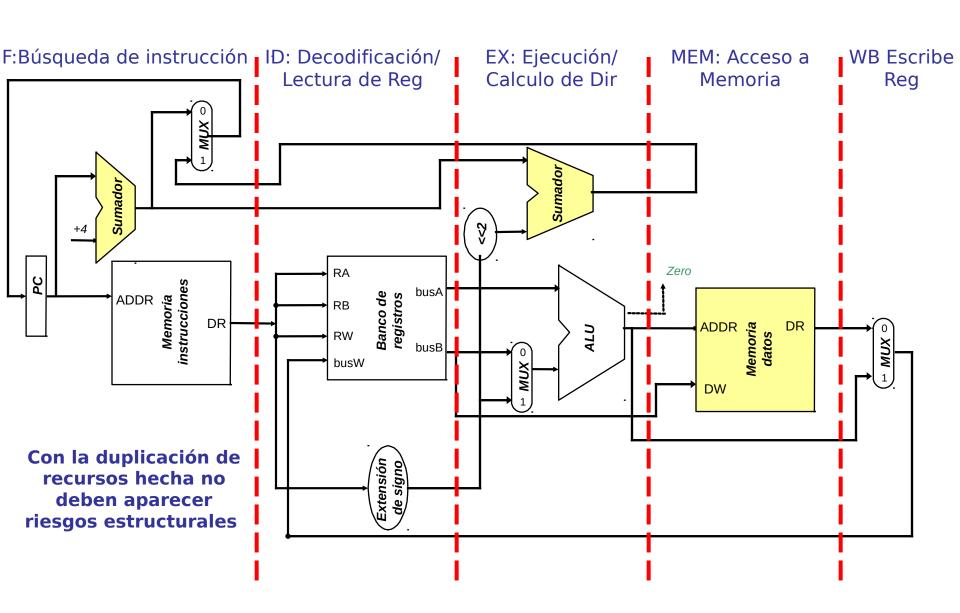
☐ Riesgos estructurales.

Objetivo: Ejecutar sin conflicto cualquier combinación de instrucciones



Segmentación: Riesgos estructurales

☐ Nueva Ruta de datos del DLX para ejecución segmentada (idea inicial)



Segmentación: Riesgos de datos

- Se produce cuando por la segmentación, el orden de LECTURA de los operandos y la ESCRITURA de resultados se modifica respecto al especificado por el programa.
- ☐ Se produce un riesgo si existe dependencia entre instrucciones que se ejecutan concurrentemente.

Si
$$a = b + c$$

Si $a = b + c$

Sk $a = d + e$

Dominio: operandos de la instrucción **Rango**: resultado de la instrucción

Situaciones: i precede a j

```
D(i) \cap D(j) \neq \phi no RIESGO

D(i) \cap R(j) \neq \phi riesgo EDL (WAR)

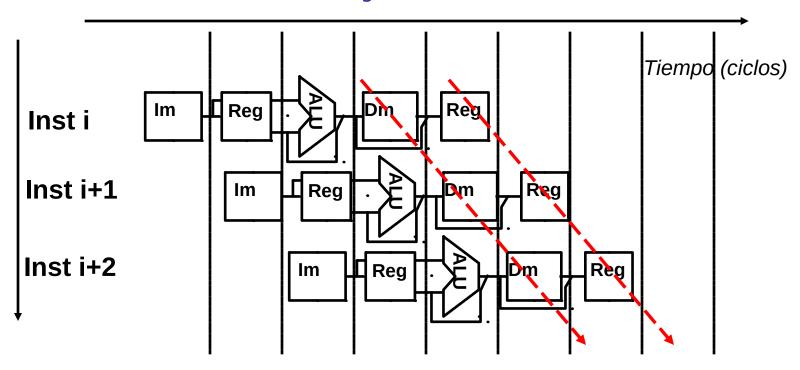
R(i) \cap D(j) \neq \phi riesgo LDE (RAW)

R(i) \cap R(j) \neq \phi riesgo EDE (WAW)
```

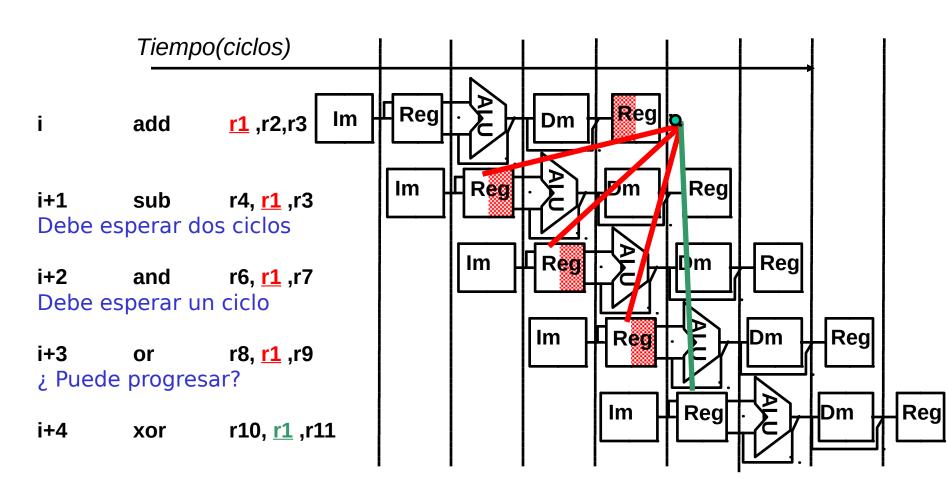
Riesgos de datos

```
Riesgos. D(i) \cap R(i+?) \neq \phi EDL (WAR)
R(i) \cap R(i+?) \neq \phi EDE (WAW)
```

- o Se leen los registros en el final de la segunda etapa
- o Todos las instrucciones escriben en la ultima etapa
- o Todas las instrucciones tienen igual duración



□ Riesgo R(i) \cap D(i+?) $\neq \phi$ LDE (RAW)

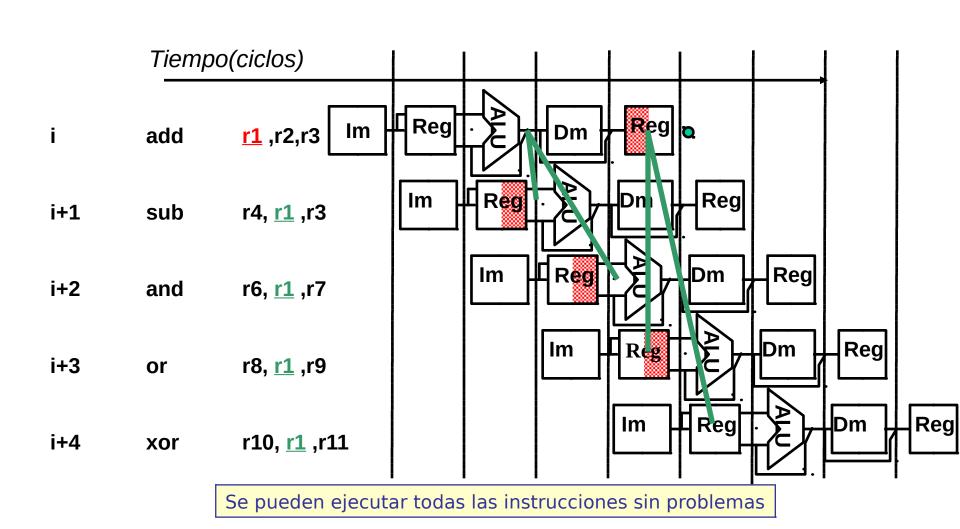


¿ Cuando esta listo el operando ?

Riesgos de datos

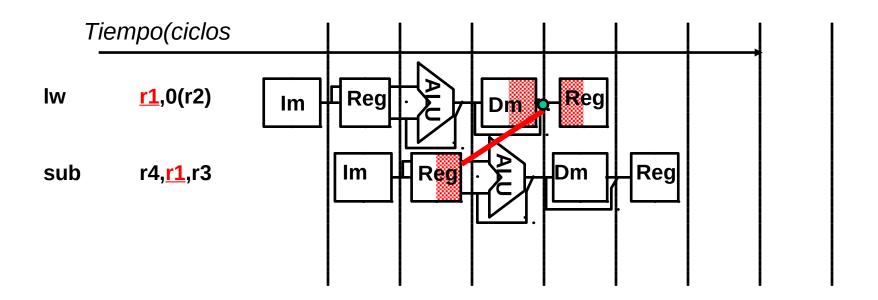
□ Riesgo R(i) \cap D(i+?) $\neq \phi$ LDE (RAW).

✓ Cortocircuito. Enviar el dato a las etapas que lo necesitan, cuanto esta listo



Riesgos de datos

□ Riesgo R(i) \cap D(i+?) $\neq \phi$ LDE (RAW). \checkmark Caso del Load

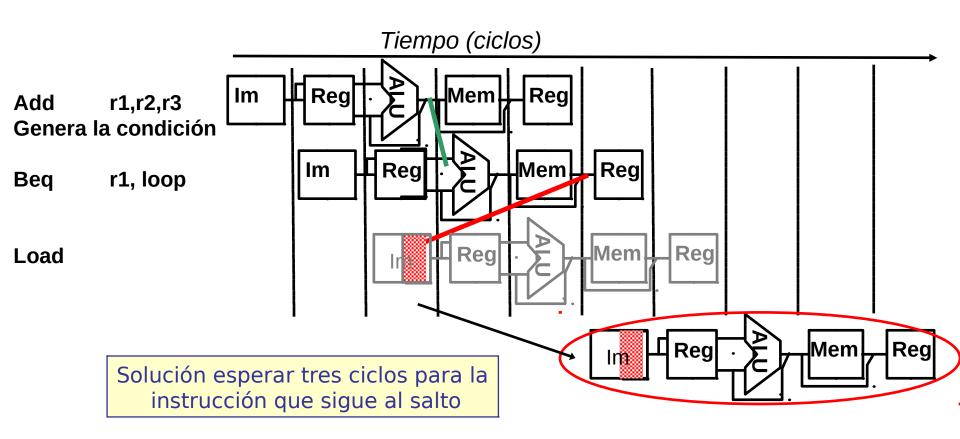


No se puede resolver con el cortocircuito La instrucción dependiente del Load debe esperar un ciclo

Riesgos de control

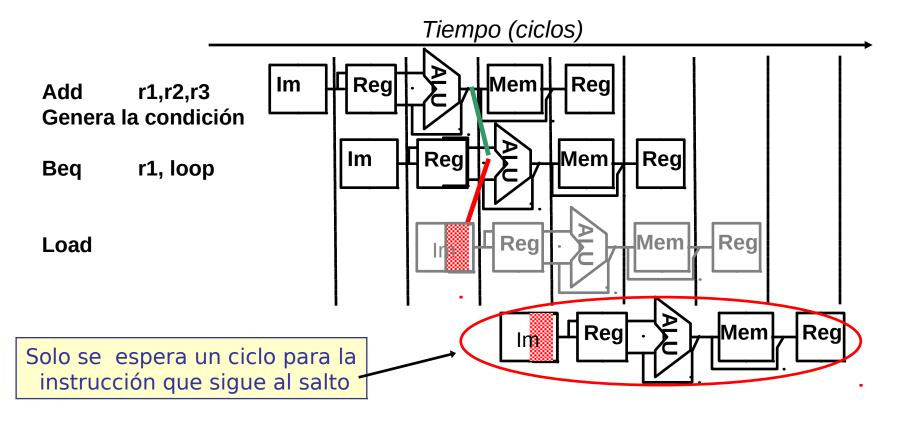
Se debe evaluar la condición y obtener el nuevo valor del PC

✓ La evaluación de la condición y el calculo del nuevo PC se realizan en la etapa de ejecución



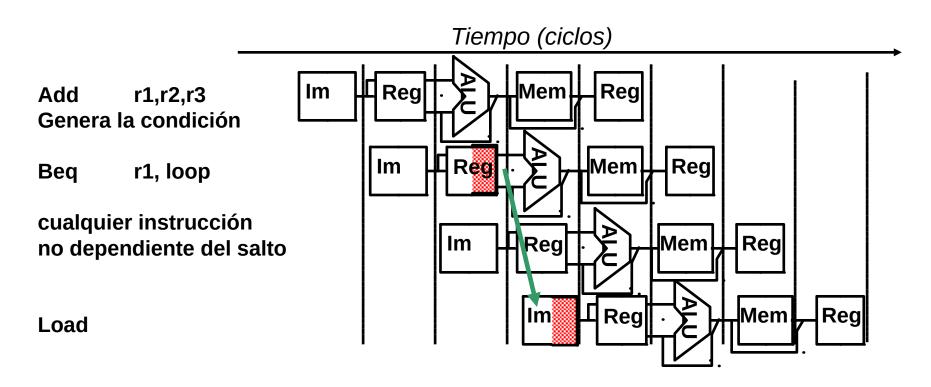
Riesgos de control

☐ Solución: Desplazar el calculo de la dirección y la evaluación de la condición a la etapa ID



Riesgos de control

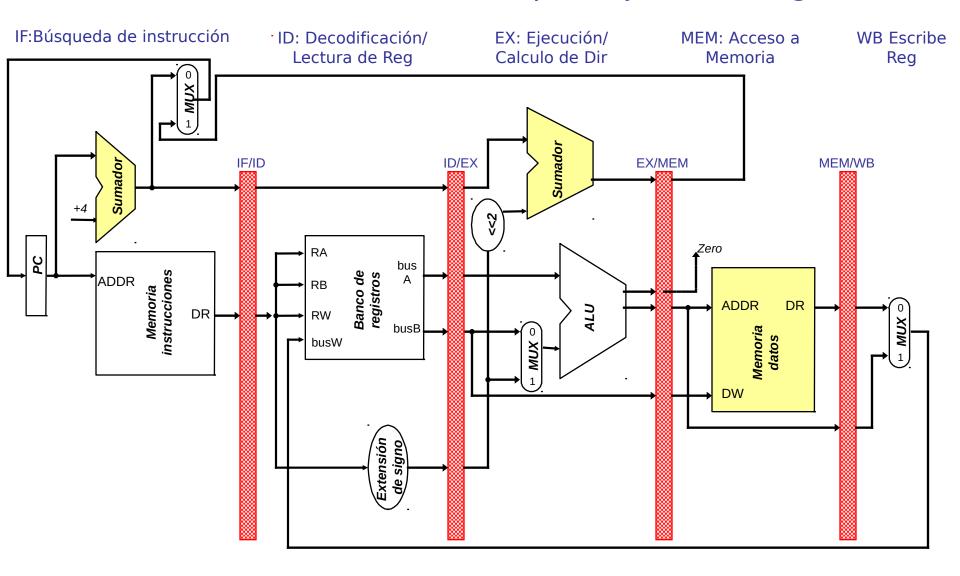
- Reinterpretar el comportamiento de los saltos: Salto retardado
 - √ Se efectúa después de la siguiente instrucción
 - √ La instrucción siguiente al salto se ejecuta siempre



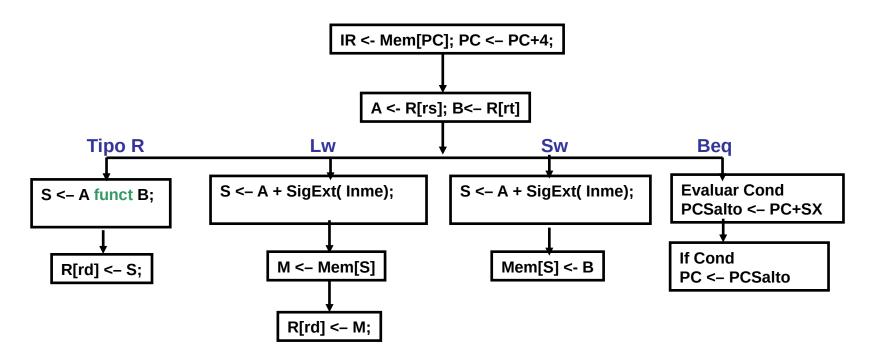
<u>Si es posible encontrar una instrucción ⇒ 0 ciclos de</u> <u>penalización</u>

Segmentación del procesador

Nueva Ruta de datos del DLX para ejecución segmentada

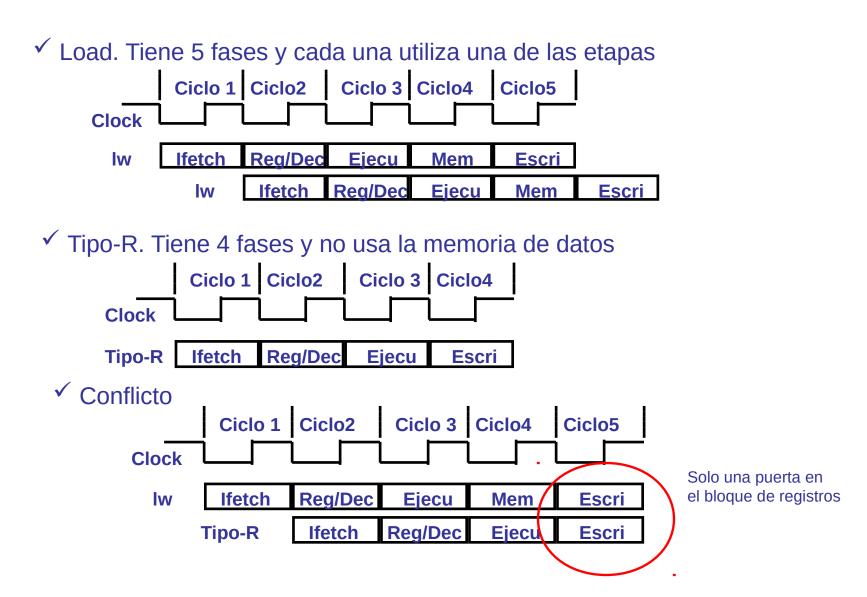


☐ Ejecución de instrucciones: diagrama RT

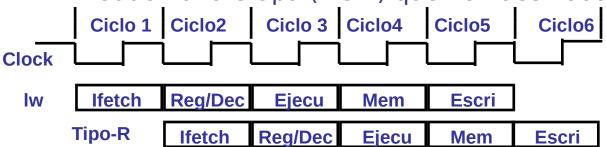


- ☐ Recordatorio
 - $\sqrt{rs} = Ins[25-21]$
 - $\sqrt{\text{rt} = lns[20-16]}$
 - $\sqrt{\text{rd}} = \ln[15-11]$
 - \checkmark Inme = Ins[15-0]

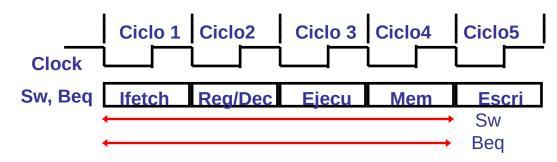
☐ Las instruciones: Load (lw) y tipo-R



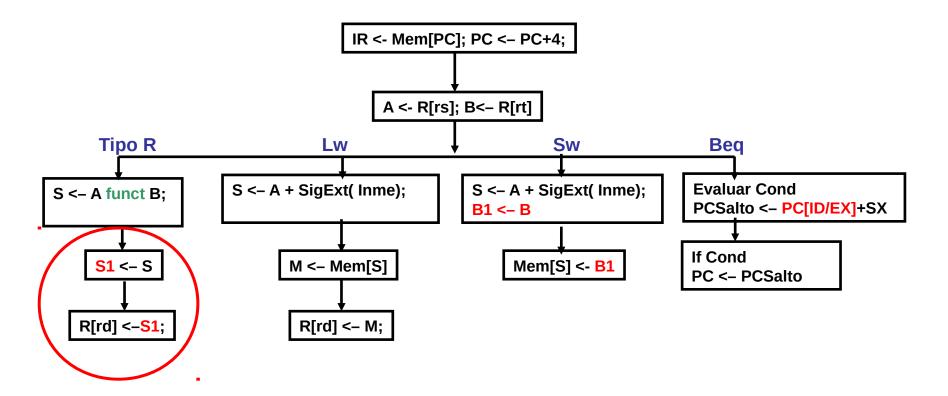
- ☐ Importante. ¿ como evitarlo ?
 - ✓ Cada etapa del procesador puede usarse en cada ciclo por una sola instrucción.
 - ✓ Cada instrucción debe usar cada etapa del procesador en la misma fase de ejecución.
- ☐ Solución.
 - ✓ Introducir una etapa (Mem) que no hace nada.



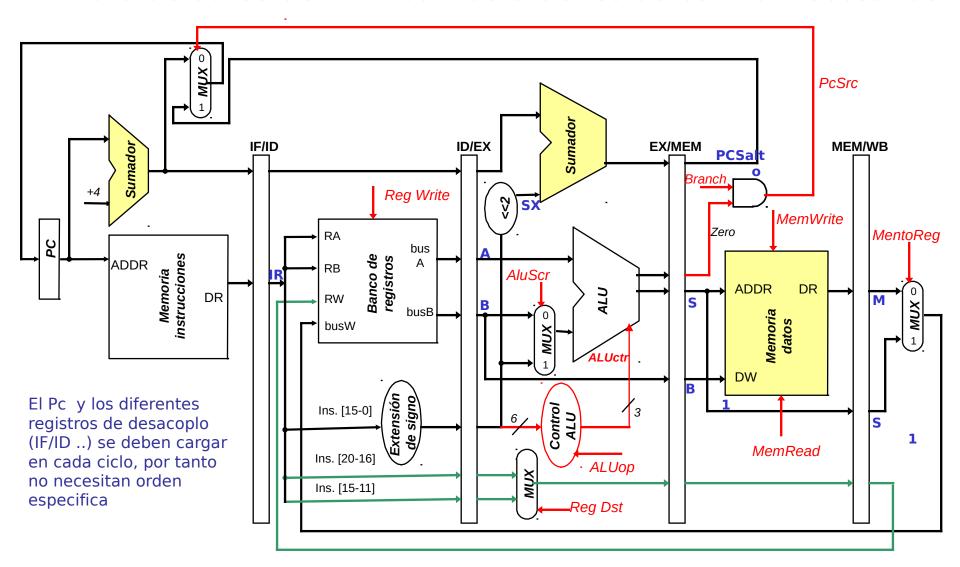
☐ Store 4 etapas con actividad. Branch cuatro etapas activas



☐ Ejecución de instrucciones: Diagrama RT modificado



Ruta de datos del DLX con las ordenes de control necesarias



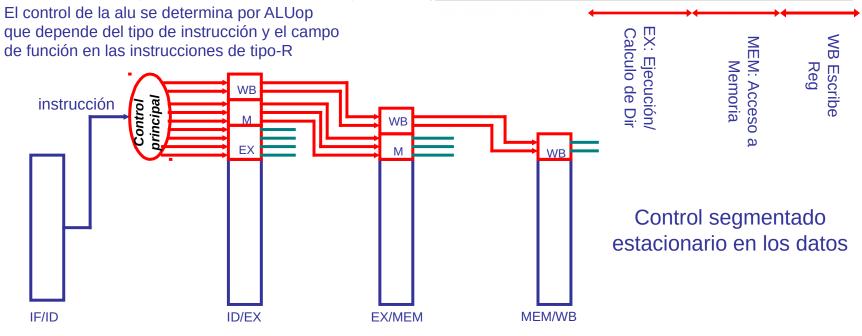
Señales de control

Control de la ALU

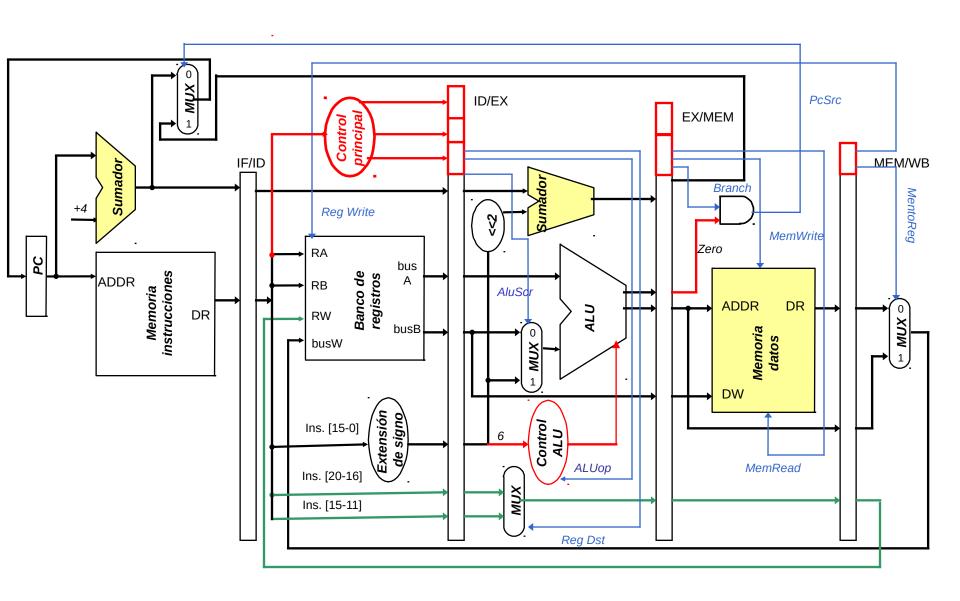
Control principal

op	funct	ALUop	ALUctr			T
100011 (lw)		00	010		ــا	l
101011 (sw)	XXXXXX	00	010	ор	Ost	l
000100 (beq)		01	110		RegDst	
000000 (tipo-R)	100000 (add)	10	010		8	
	100010 (sub)	10	110	100011 (lw)	0	
	100100 (and)	10	000	101011 (sw)	X	I
	100101 (or)	10	001	000100 (beq)	X	
	101010 (slt)	10	111	000000 (tipo-R)	1	

ор		ALUSrc	ALUop	MemRead	MemWrite	Branch	RegWrite	MemtoReg				
100011 (lw)	0	1	00	1	0	0	1	0				
101011 (sw)	X	1	00	0	1	0	0	X				
000100 (beq)		0	01	0	0	1	0	X				
000000 (tipo-R)	1	0	10	0	0	0	1	1				

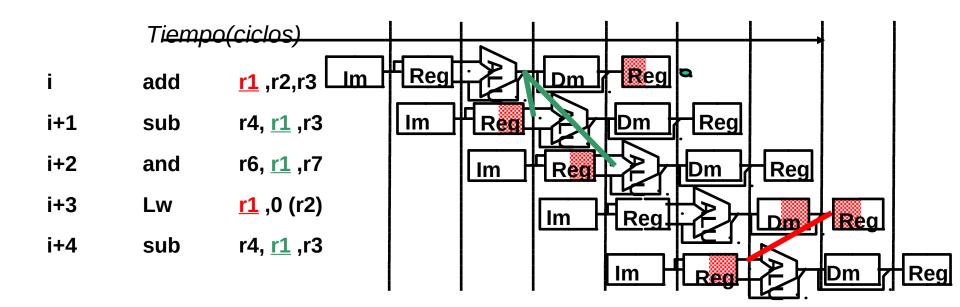


☐ Ruta de datos del DLX con las ordenes de control y control estacionario en los datos



- ☐ ¿Qué facilita el control segmentado?
 - √ Todas las instrucciones con igual duración
 - ✓ Pocos formatos diferente de instrucción
 - ✓ Accesos a memoria solo en load y stores
- ☐ ¿Qué dificulta el control segmentado?
 - ✓ Riesgos estructurales. Conflictos de recursos
 - ✓ Riesgos de datos. Solo LDE
 - ✓ Riesgos de control
- El diseño anterior no tenia en cuenta los riesgos de datos y los riesgos de control los eliminaba con saltos retardados
 - ✓ Implementación del cortocircuito
 - ✓ Caso del load
 - ✓ Mejora en el comportamiento de los saltos.

- □ Riesgos de datos $R(i) \cap D(i+?) \neq \phi$ LDE (RAW).
 - ✓ Cortocircuito. Enviar el dato a las etapas que lo necesitan, cuanto esta listo



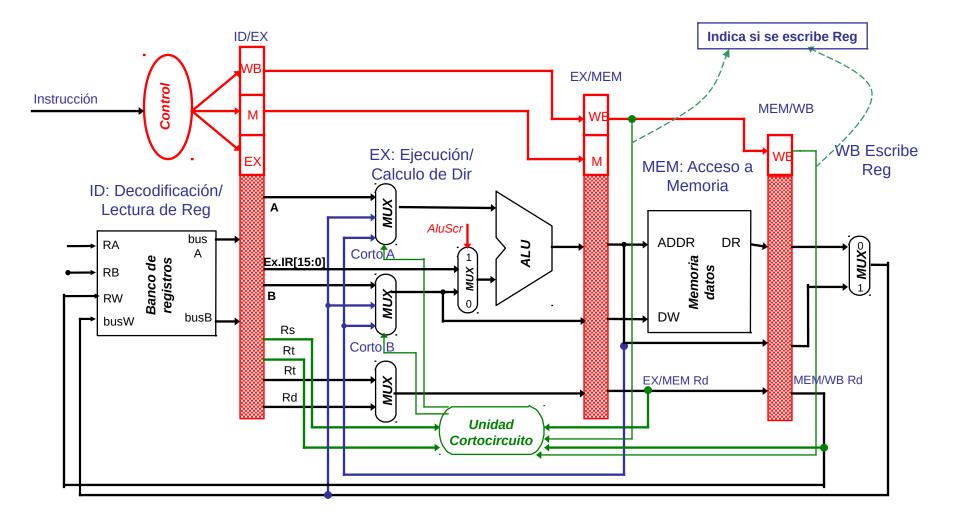
Dos caminos de datos: Desde salida ALU (EX/MEM) a entrada ALU

Desde la salida de la memoria (MEM/WB) a entrada ALU

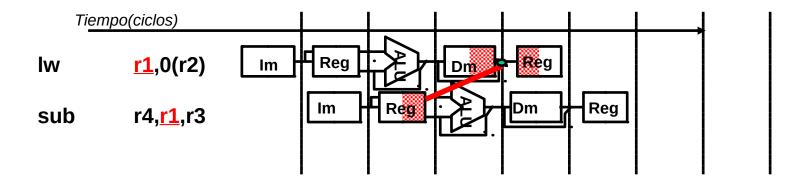
Información necesaria: Registro a escribir en ultima etapa (Rd en Tipo-R y Rt en Lw)

Registros que se leen en segunda etapa (Rs y Rt)

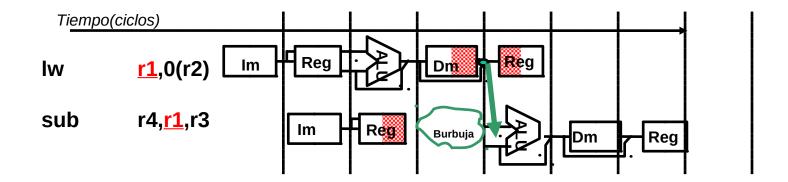
- ☐ Riesgos de datos LDE: Implementación del cortocircuito
 - ✓ Cortocircuitos de datos
 - ✓ Información de control del cortocircuito



☐ Riesgos de datos LDE: Caso del load

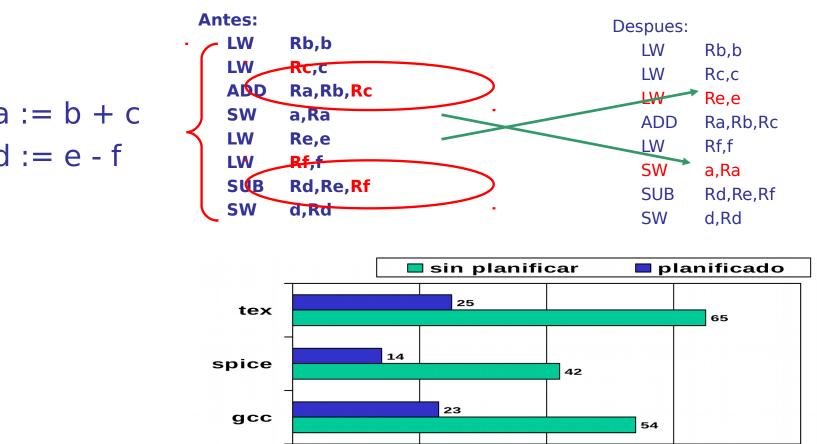


Se debe esperar un ciclo a pesar del cortocircuito



☐ Riesgos de datos LDE: Caso del load

Solución SW: Anticipar el Load en la planificación de instrucciones que hace el compilador



20

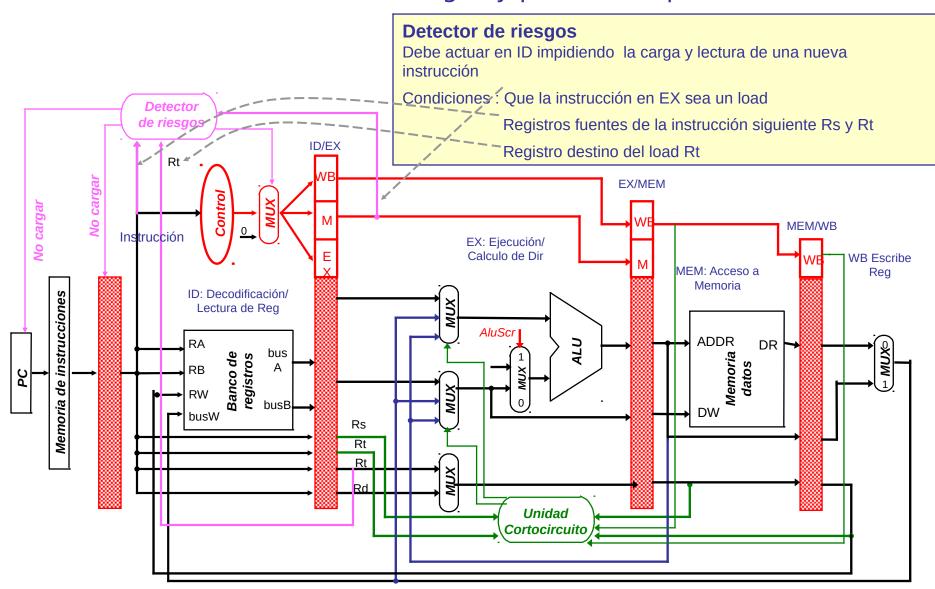
O

60

80

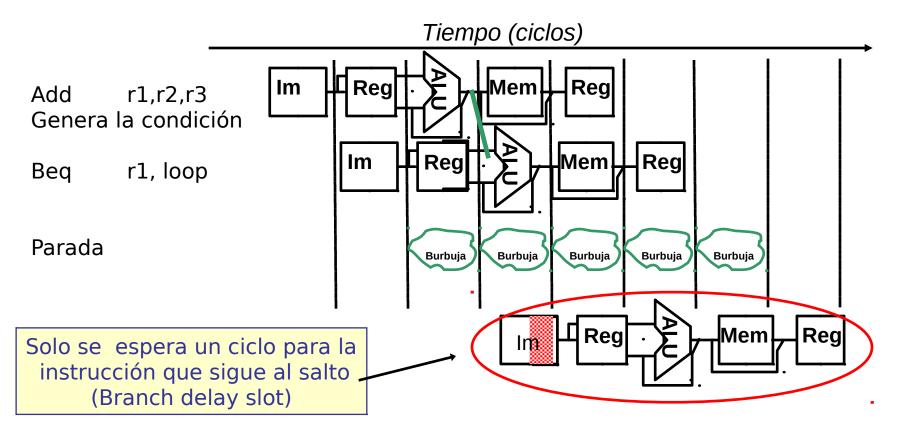
40

Solución HW: Detección de riesgos y parada del procesador un ciclo

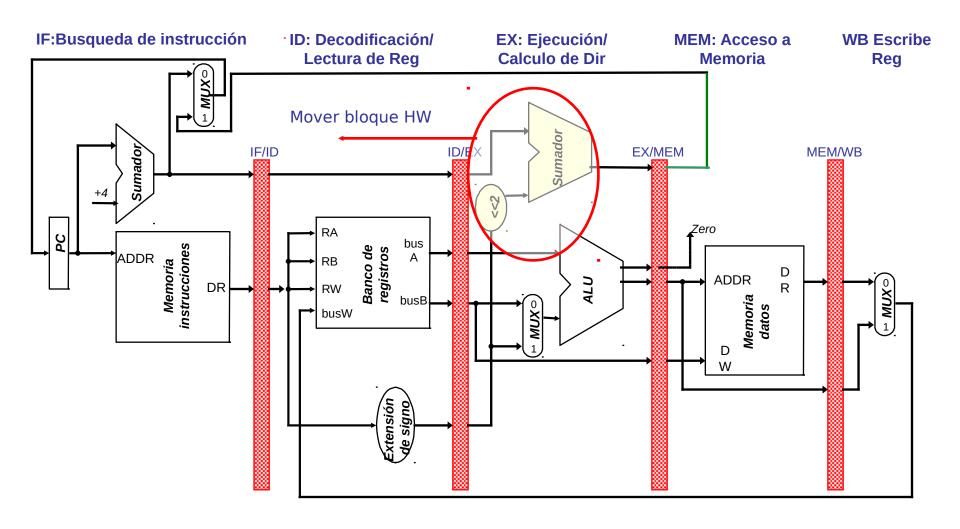


☐ Riesgos de control

✓ Solución: Desplazar el calculo de la dirección y la evaluación de la condición a la etapa anterior

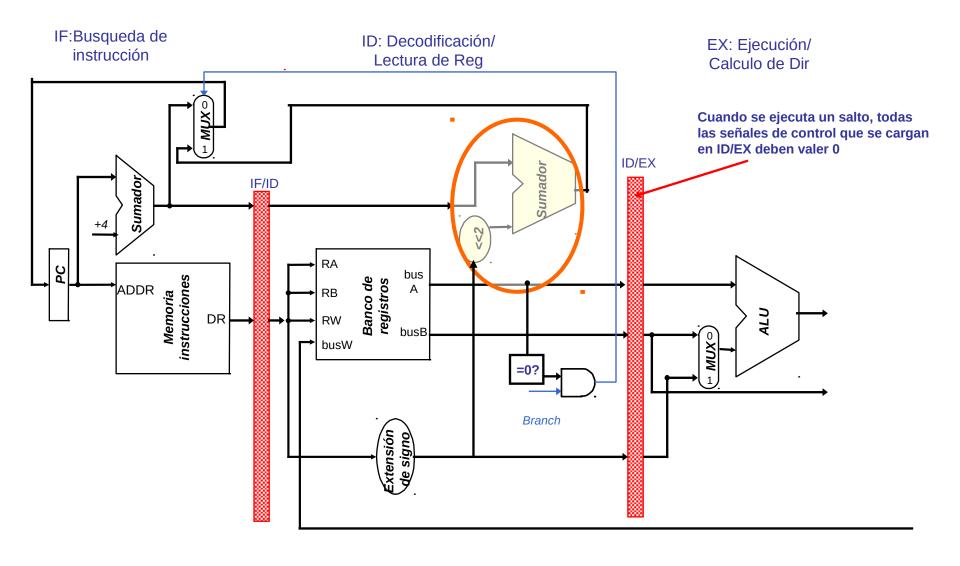


- Mejora del comportamiento de los saltos. Solo un ciclo de parada
 - ✓ Calculo de la dirección. Operandos disponibles (Pc y desplazamiento)
 - ✓ Calculo de la condición. Unidad de detección de cero



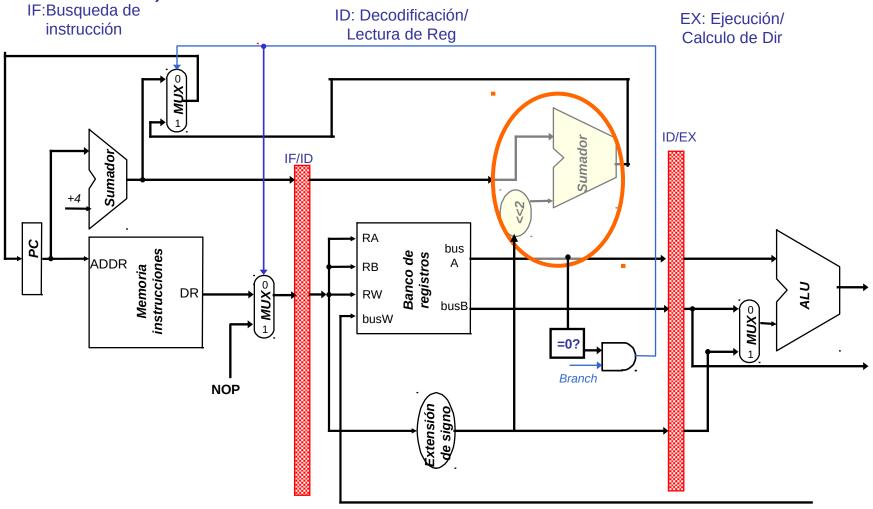
Saltos retardados

☐ Delay slot = 1 ciclo de reloj. La instrucción siguiente al salto se ejecuta siempre



Predicción estática

 Predecir que el salto no se toma: Si la predicción es falsa, la instrucción siguiente se aborta (se cambia por NOP)



Pérdida de rendimiento

- ☐ EL CPI ideal es 1
- ☐ Hay perdidas de rendimiento por las paradas del pipe

CPI _{real} = CPI _{ideal} + Penaliz. media por instrucción +
$$\sum_{i=1}^{\text{#tipos de instr}} Penaliz_i \times Frec_i$$

☐ Caso de los saltos. Un programa tipico 30% de saltos o $CPI = 1 + (1 \times 0.3) = 1.3$

$$N^{\circ}$$
 Instrucciones x n° de etapas 5

Speedup = = 3.84

 N° instrucciones x CPI 1.3

o Se pierde un 24 % respecto al caso ideal: Speedup $_{real}$ / Speedup $_{ideal}$ = 3.84 / 5 = 0.76

Excepciones

- ☐ Interrupciones, Excepciones, Fallos
 - O Síncronas asincronas
 - o Solicitadas por el programa generadas por el programa
 - **O Dentro de instrucciones entre instrucciones**
 - o Continuar terminar
- Problema: El solapamiento en la ejecución de las instrucciones dificulta el saber si una instrucción puede cambiar el estado de la maquina sin peligro
- Cualquier instrucción en el pipeline puede provocar una excepción
- ☐ El sistema debe resolver la excepción y recomenzar la ejecución. El sistema debe recuperar el estado previo a la excepción
- Excepciones problemáticas: (un ejemplo fallo de página)
 - O Ocurren en el medio de una instrucción
 - Deben ser recomenzables
- Interrupciones externas (I/O)
 - Vaciar el pipeline y entrar no operaciones (NOPs)
 - o Almacenar el PC con la dirección de interrupción

Excepciones

☐ Excepciones en el DLX

IF Fallo de pagina de instrucción; Acceso no alineado;

Violación de protección

ID Instrucción ilegal

Ex Excepción aritmética

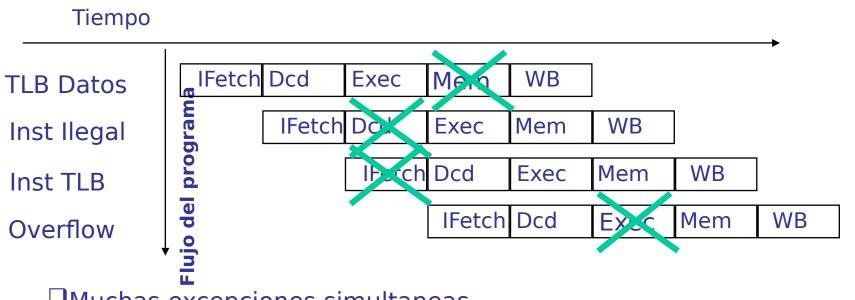
MEM Fallo de pagina de datos; Acceso no alineado; Violación

de protección

WB Ninguna

- Excepciones (fallo de pagina). Reinicio
 - o Introducir una instrucción de trap en IF
 - o Anular las escrituras de la instrucción que produce la excepción y las posteriores
 - o Salvar el PC(se debe conservar hasta MEM) de la instrucción que produjo la excepción y pasar a la rutina de gestión
- Excepciones precisas:
 - Todas las instrucciones anteriores completas
 - La que produce la interrupción y las siguientes como si no hubieran empezado

Excepciones



- ☐ Muchas excepciones simultaneas
- □ Fallo de pagina en MEM y otras excepciones en el resto. Se atiende el fallo de pagina y se recomienza la siguiente instrucción.
- □Problema:

Fallo de pagina en MEM , instrucción ilegal en <u>reg./De</u> y fallo de pagina en <u>IF</u>.

¡La excepción de la segunda y tercera aparece antes!

□Solución:

Vector de estado, un bit por cada etapa donde es posible excepción Cada excepción se marca en un vector de estado asociado con la instrucción,

y se impide la escritura. El vector se chequea al final de MEM inicio de WB)

Saltos retardados: Compilador

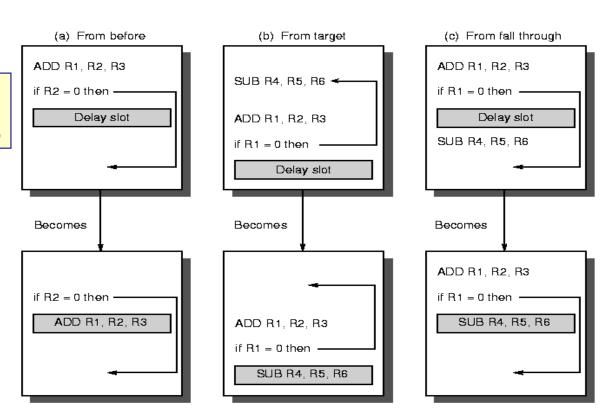
Se puede mejorar el rendimiento introduciendo una instrucción no dependiente del salto en la burbuja

fundamento de los saltos retardados cutar instrucciones independientes del salto durante los ciclos de retardo

Estas instrucciones se ejecutarán **siempre** (tanto si el salto es tomado como si no)

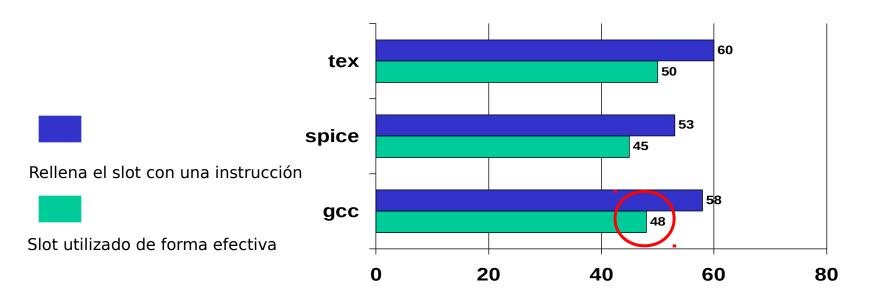
El compilador debe encargarse de elegir adecuadamente las instrucciones que se planifican en los ciclos de retardo

- (a) mejor solución siempre trabajo útil
- (b) y (c) la instrucción elegida no debe modificar la semántica (aunque se ejecute indebidamente).



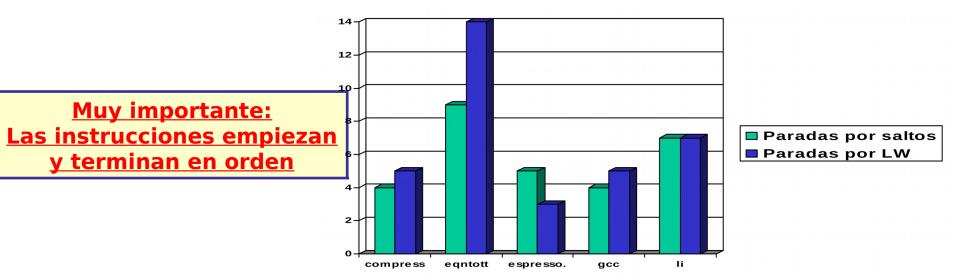
Saltos retardados: benchmarks

- ☐ Tex, Spice, Gcc: En media el 80 % de las posiciones rellenas son útiles (barra verde / barra azul).
 - ✓ El resto son NOP
 - ✓ En Gcc el 22% de las instrucciones son saltos CPI = $1+(1-0.48) \times 0.22 = 1.11$



Resumen

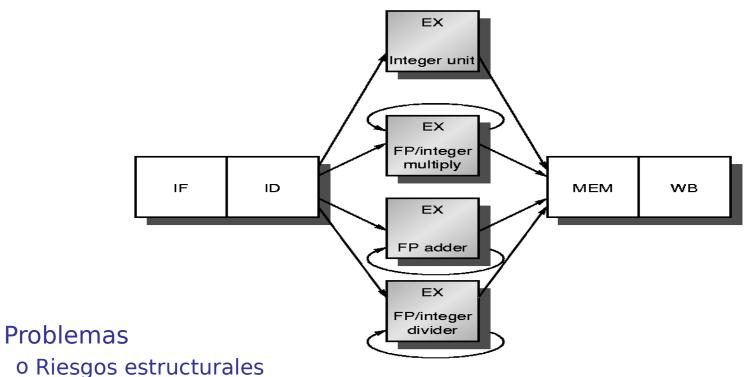
- Procesador segmentado
 - ✓ Todas las instrucciones tienen igual duración
 - ✓ Rendimiento ideal, una instrucción por ciclo CPI=1
 - ✓ Riesgos estructurales y de datos EDE y EDL se resuelven por construcción
 - ✓ Riesgso LDE en instrucciones tipo-R se solucionan con el cortocircuito.
 - ✓ Riesgos LDE en instrucciones de *load* implican paradas del procesador. Ayuda del compilador planificando las instrucciones.
 - ✓ Riesgos de control. Paradas y saltos retardados con ayuda del compilador.



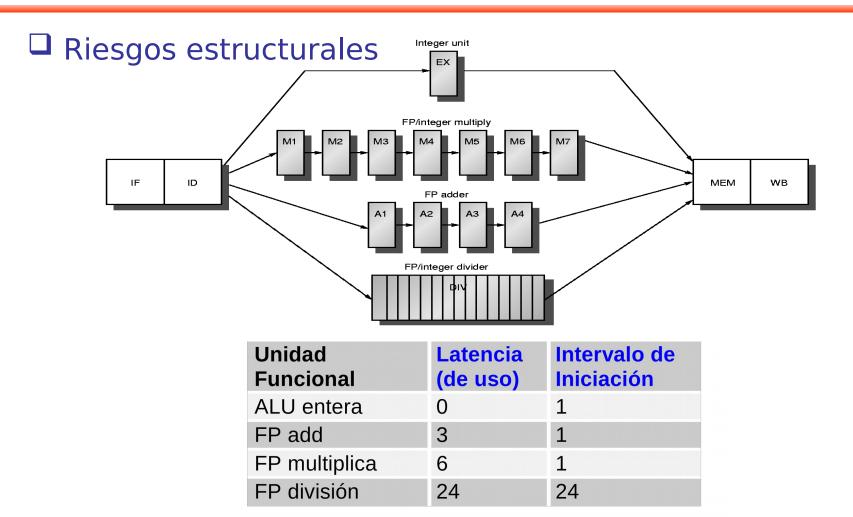
<u>Que ocurre si las instrucciones tienen diferentes duración?</u> <u>Instrucciones de aritmética en punto flotante</u>

☐ Es el caso típico de las operaciones en punto flotante

Estructura de etapas de procesador: Es necesario HW adicional para la fase de ejecución

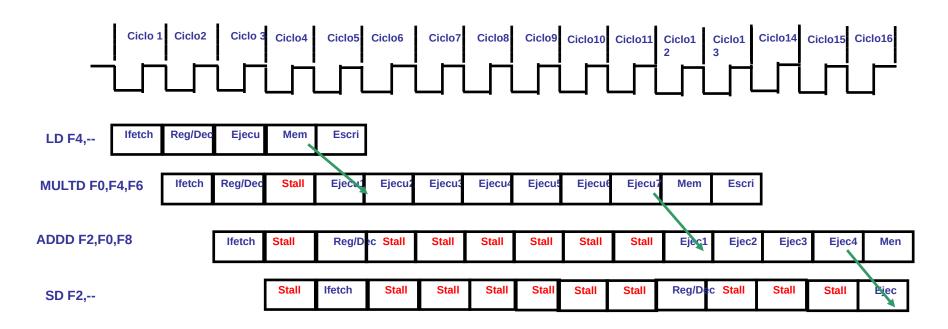


- o Mayor popalización do los riosgos
- o Mayor penalización de los riesgos LDE
- o Problemas con la finalización fuera de orden
- Solapamiento operaciones enteras y PF
 - O No hay problemas operandos independientes



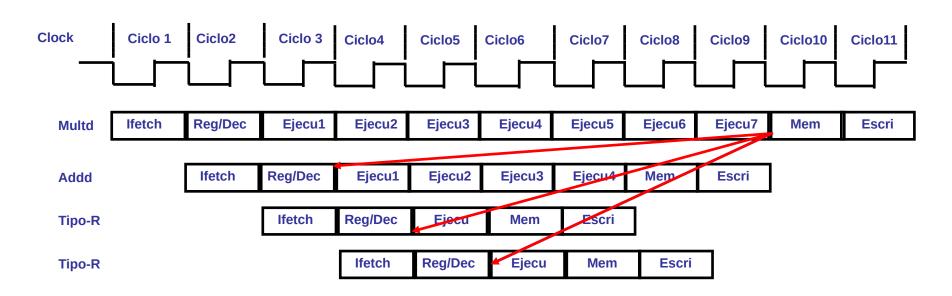
- Replicación o segmentación de las unidades de PF
- ☐ La división no suele estar segmentada. Detección del riesgo
 - y parada de procesador

☐ Riesgos LDE (RAW)



- Mayor impacto en el rendimiento
 - o La gran duración de las instrucciones implica más ciclos de detención
 - o Es necesaria una planificación más cuidadosa de las instrucciones

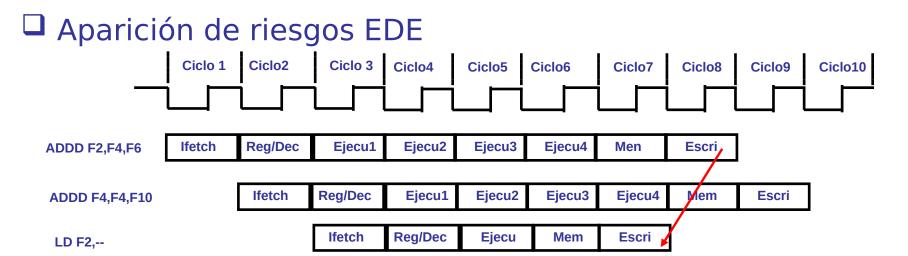
☐ Finalización fuera de orden



- Las instrucciones acaban en orden distinto al lanzamiento
- Problemas
 - o Conflictos por escritura simultanea en registros
 - o Aparición de riesgos EDE
 - o Problema con las excepciones

Muy importante Las instrucciones empiezan en orden y terminan fuera de orden

- Conflictos por escritura simultánea en registros
 - □ Problemas si el bloque de registros tiene un único puerto de escritura→ riesgo
 - ☐ Solución:
 - o Detener, en la etapa Reg/Dec, las instrucciones que produzcan conflicto:
 - Si la instrucción en Reg/Dec necesita escribir en registros en el mismo ciclo que una instrucción ya emitida, la primera se detiene un ciclo.
 - Se puede usar un registro de desplazamiento para indicar cuándo usarán el bloque de registros las instrucciones ya lanzadas.
 - o Detener las instrucciones conflictivas al final de Ejecu
 - Necesidad de establecer prioridades de acceso: dar mayor prioridad a la unidad con mayor latencia
 - Lógica de chequeo de detenciones en dos puntos

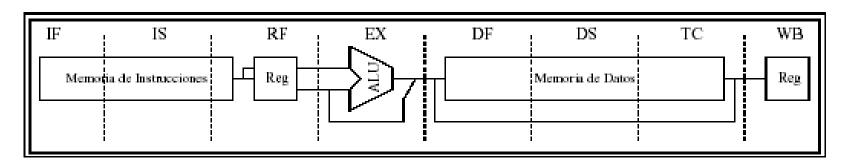


- LD F2,0(R2) escribe en F2 antes que ADDD F2,F4,F6 → error
 - o Situación poco común: Instrucción ADDD F2,F4,F6 inútil. Puede ocurrir con instrucciones que ocupen un hueco de retardo
 - o Si la segunda ADDD lee F2 → riesgo LDE que elimina el EDE
 - Riesgos EDL(WAR) No aparecen porque <u>las lecturas de registros son en</u> <u>orden</u> etapa Reg/Dec)
- Tratamiento de riesgos EDE WAW:
 - o Detener la instrucción que produce el riesgo (LD en el ejemplo)
 - o Eliminar la escritura de la primera instrucción (ADDD en el ejemplo)
- Situaciones poco comunes. No implica demasiado problema cualquiera de las aproximaciones

Más segmentación 8 etapas

- IF: Primera mitad de búsqueda de instrucción. Inicia el acceso a cache de instrucciones. Actualización del PC.
- o IS:{ Segunda mitad de búsqueda de instrucción. Completa el acceso a cache.
- Reg/Dec: Comprobación de acierto en cache. Decodificación y búsqueda de registros. Chequeo de riesgos.
- O EX: Ejecución; operación de la ALU, calculo de dirección efectiva, evaluación de condición de salto y cálculo de destino (puede durar varios ciclos)
- O DF: Inicio de la búsqueda de datos
- O DS: Segunda mitad de la búsqueda de datos
- o TC: Chequeo de etiquetas, determinación de acierto en la cache
- o WB: Postescritura

Los datos se encuentran disponibles al final de DS, aunque no se ha comprobado el acierto en la cache



Más penalización en cargas y saltos

Dos ciclos de latencia en los Load con cortocircuito	IF	IS IF	RF IS IF	EX RF IS IF	DF EX RF IS IF	DS DF EX RF IS IF	TC DS DF EX RF IS	WB TC DS DF EX RF IS
Tres ciclos de latencia en saltos (la condición se evalú EX) Salto retardado de un 2 ciclos de "predict-no	ciclo <u>y</u>		RF IS IF	EX RF IS IF	DF EX RF IS	DS DF EX RF IS IF	TC DS DF EX RF IS	WB TC(*) DS(**) DF(**) EX RF IS

(*) Continúa la ejecución en cq caso

(**) Se abortan en caso de que el salto se tome

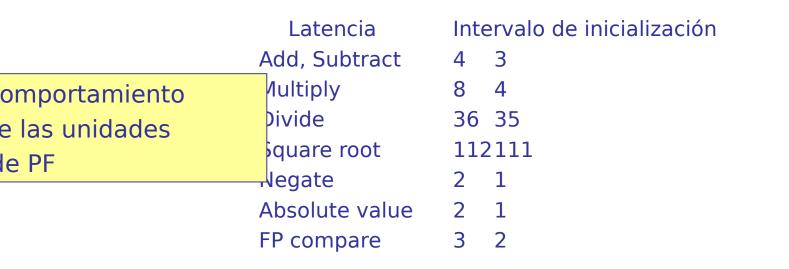
Operaciones en PF

- Tres unidades funcionales: divisor, multiplicador y sumador en punto flotante
- O Diferentes unidades son usadas para completar una operación
- Operaciones con duración entre 2 ciclos (para una negación) y 112 ciclos (para una raíz cuadrada)
- Segmentadas con un total de 8 estados
- Descripción de la etapas de las unidades en PF

Etapas	Unidad	Descripción
Α	FP adder	Etapa de suma de mantisas
D	FP divider	Etapa de división
Е	FP multiplier	Etapa de test de excepciones
M	FP multiplier	1ª etapa del multiplicador
N	FP multiplier	2ª etapa del multiplicador
R	FP adder	Etapa de redondeo
S	FP adder	Etapa de desplazamiento
U		Etapa de desempaquetado

Operaciones en PF: Implementación

```
Instr PF
                          3
                                     5
                                                     8
                                4
                                           6
Add, Subtract
                 U S+A
                          A+R R+S
Multiply
                 U = E + M
                           M
                                M
                                           Ν
                                                N+AR
                                     М
                                           D+A D+R, D+R, D+A, D+R, A, R
Divide
                 U
                    Α
                           R
                                D28
Square root
                 U
                   Ε
                          (A+R)^{108}
                                          Α
                                                R
                 U S
Negate
                 U S
Absolute value
FP compare
                 U
                    Α
                           R
```



- Rendimiento
 - \square No se alcanza el rendimiento ideal CPI= 1:
 - Load stalls (1 o 2 ciclos)
 - Saltos (2 ciclos + slots no rellenos)
 - FP stalls (resultados): riesgo de LDE (latencia)
 - o FP stalls (estructurales): Hw FP muy escaso (paralelismo)

