作業系統
(Operating System)

記憶體管理

(Memory Management)

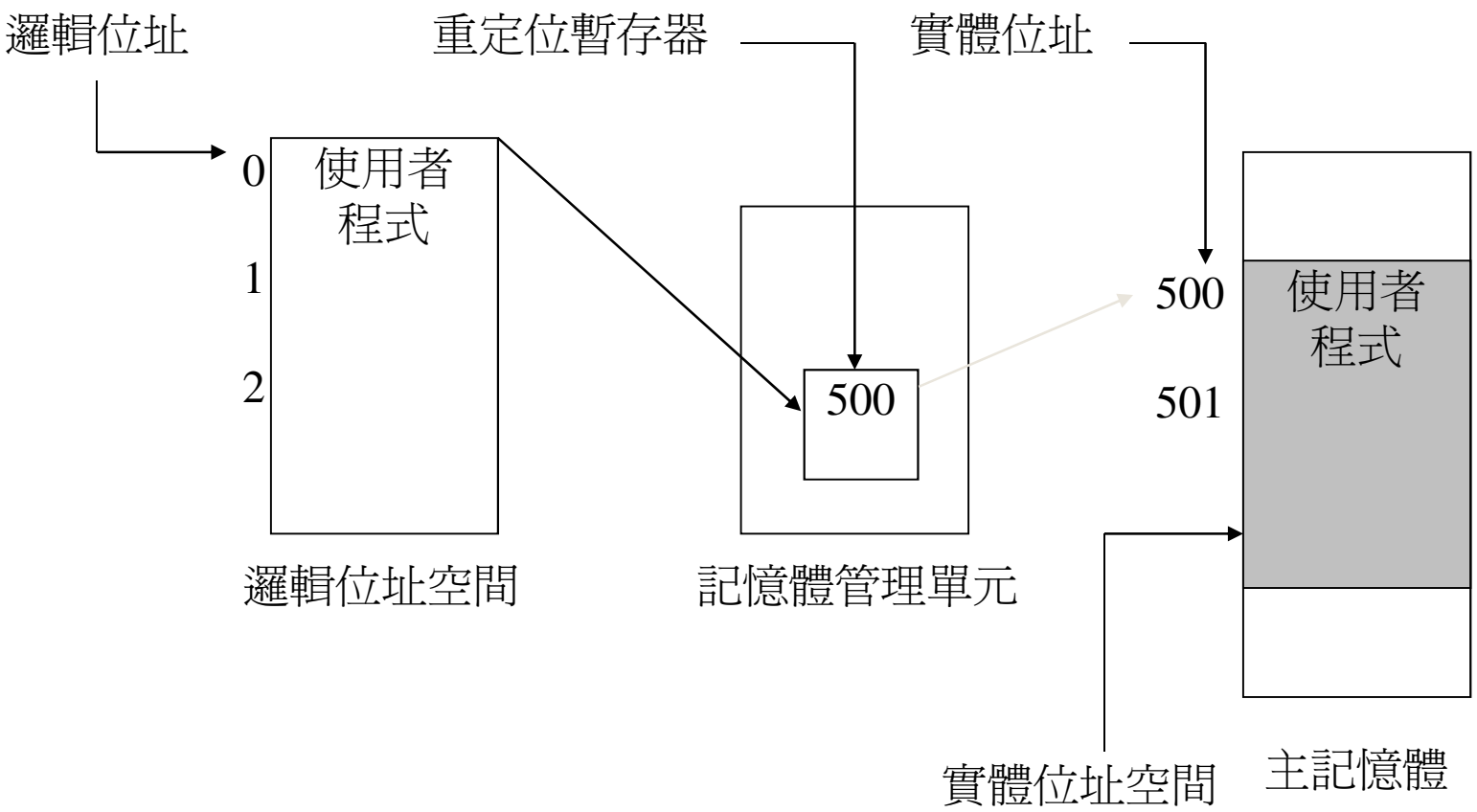
目標與趨勢

- 目標
  - ***追蹤***記憶體空間使用與否
  - ***配置***記憶體給需要的行程
  - ***回收***行程釋放出的記憶體
  - ***有效率的置換***（swapping）方法
- 趨勢
  - ***程式成長的速度***快於記憶體成長的速度
  - ***多媒體***應用環境，使用更多的記憶體

位址空間

- 記憶體位址
  - ***邏輯位址***，邏輯位址空間
  - ***實體位址***，實體位址空間
- 執行程式時，邏輯與實體空間的位址轉換
  - ***載入器（loader）***：在主記憶體中尋找一塊可供使用的記憶體空間來載入程式
  - ***基底暫存器（base register）***：又名***重定址暫存器***，存放邏輯位址轉換成實體位址的基底值
  - ***記憶體管理單元（memory management unit,MMU）***：負責將邏輯位址加上基底值，以轉換成實體位址

邏輯位址空間到實體位址空間的轉換



位址連結（１）

- 當許多行程都要求將程式載入記憶體時：
  - 行程均進入***輸入佇列***
  - 依據排程器的排程結果選擇一個行程載入
  - 行程執行時，從記憶體取得***指令***與***資料***；執行結束後，會釋放所佔有的記憶體空間
- 位址轉換的步驟：
  - ***原始程式中的位址***：以***符號***表示
  - ***編譯器或組譯器***：將符號所指之位址連結（binding）到一可重新定址的相對位址
  - ***鏈結編譯器或載入器***：將可重新定址的位址連結到記憶體中的絕對實體位址

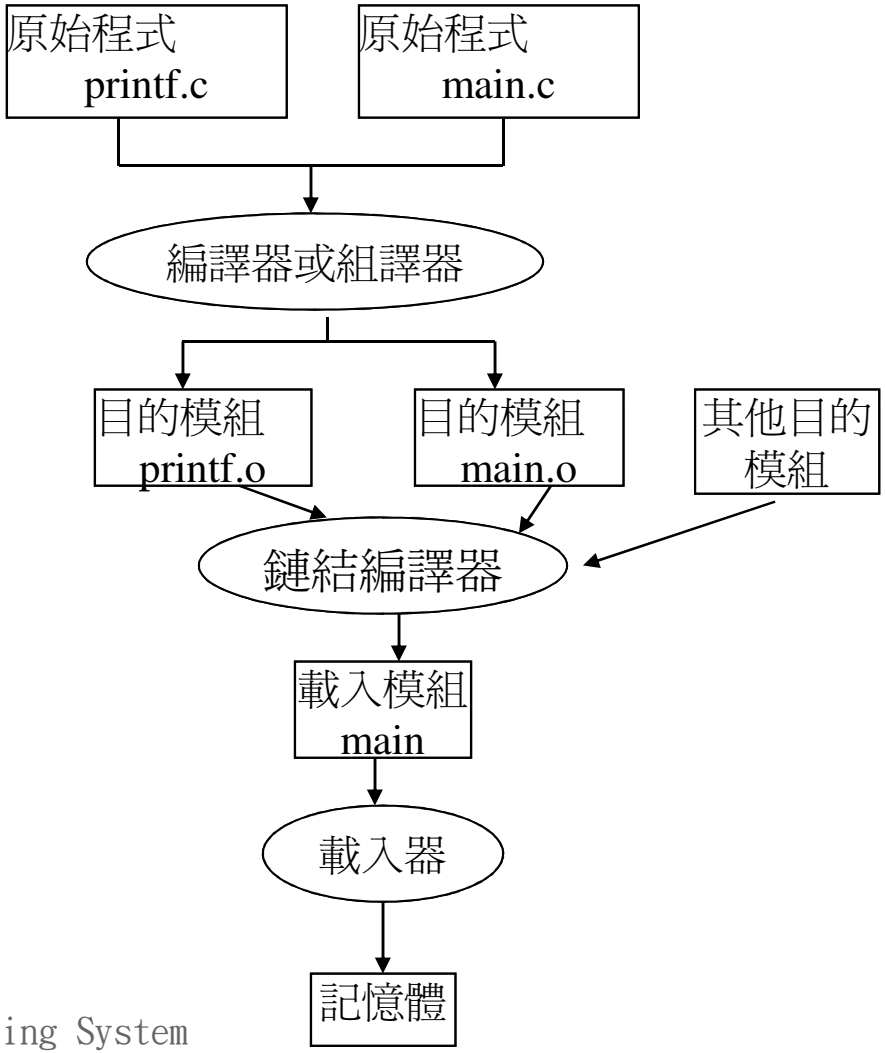
位址連結（２）

- 對於同一份資料或指令而言，所謂「位址」是***隨時間而變***的
- 資料或指令連結到實體位址的動作可在下列任一階段完成
  - ***編譯階段***
    - 已確定程式要在記憶體的某個位址執行
    - 當起始位址改變，程式必須***重新編譯***，以產生新的絕對位址的程式碼

位址連結（３）

- ***載入階段***
  - 不知道程式將在記憶體何處執行
  - 程式需編譯成可重新定址的程式碼，當起始位址改變，程式碼只需重新載入
  - 例：***動態鏈結程式庫***
- ***執行階段***
  - 若在執行時，行程會從一記憶體區塊移動到另一區塊；或是內含執行時才能確定的資料型態
  - 例：***大部分現代的作業系統中，動態產生行程或執行緒***

程式執行前的處理過程


原始程式
printf.c
原始程式
main.c
編譯器或組譯器
目的模組
printf.o
目的模組
main.o
其他目的
模組
鏈結編譯器
載入模組
main
載入器
記憶體

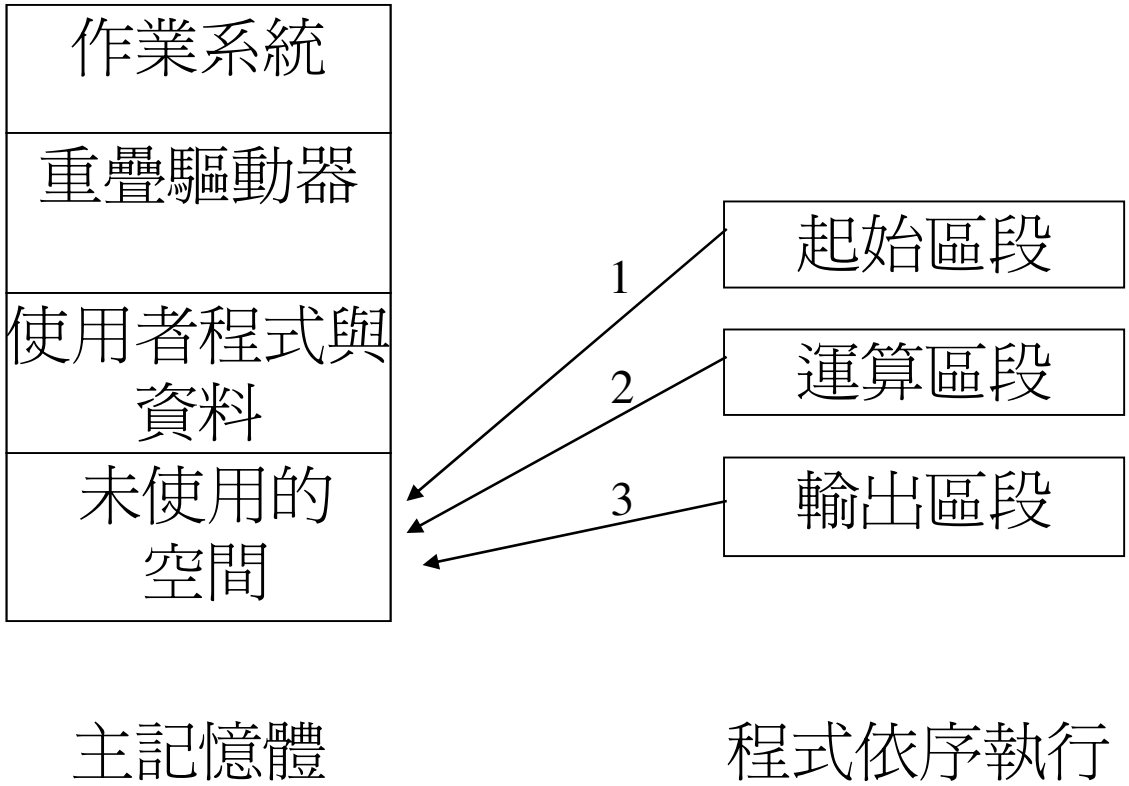
重疊（1）

- 目的：解決記憶體容量的限制
- 做法：在***編譯***時，將程式與資料分割成多個***獨立區域***；在執行時，記憶體中只保留有需要的區段
- ***重疊驅動器***：載入目前要用的區段
- 若有區段可共用記憶體，新載入的區段會覆蓋舊區段

重疊（２）

- ***多重重疊***：造成程式設計師的負擔，一般會***避免使用***，因爲：
  - 區段分割太多→置換次數過多→降低程式執行效能
  - 區段分割太少→可重疊的程式部分過少→記憶體可能不夠，系統效能降低
- 除外：***嵌入式系統***（記憶體有限，沒有虛擬記憶體）

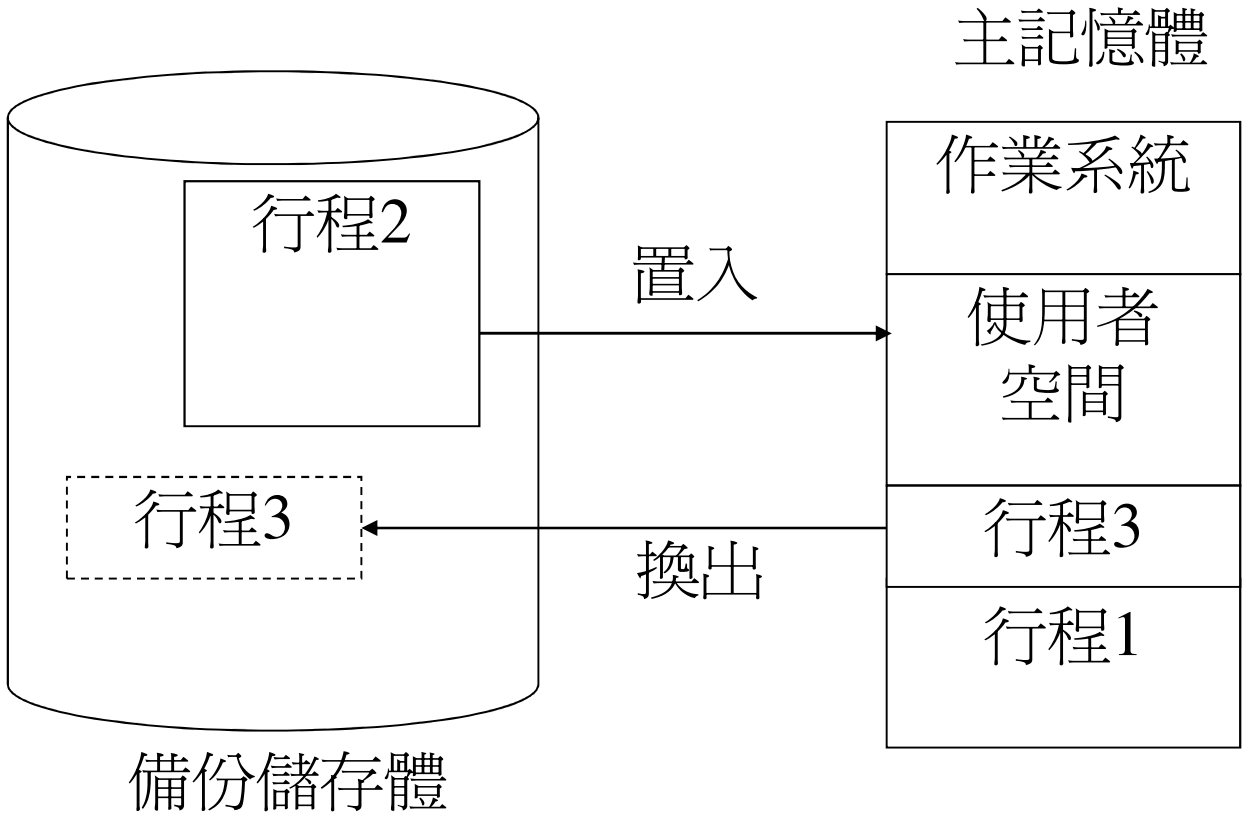
重疊（3）



置換（１）

- 時機：系統無足夠空間容納所有行程
  - 非執行中的行程暫時移到備份儲存體，要執行時再搬回記憶體中
  - 備份儲存體：一般而言指***磁碟***
- 換出、置入過程
  - CPU排程器決定下一個執行的行程
  - 分派程式到記憶體中尋找該行程
    - 若不存在且無足夠記憶體空間→先換出某些行程
    - 在置入該行程時，需重新載入暫存器內容，將控制權交給該行程

置換兩個行程



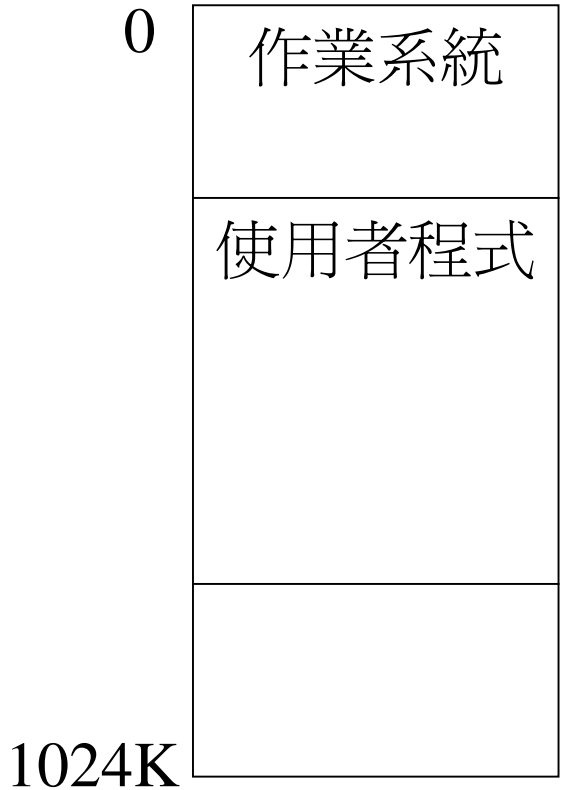
置換（２）

- 產生*內文切換*的額外負擔
  - 時間浪費在資料傳遞上
- 爲了提高效率
  - 行程必須隨時告知作業系統行程對記憶體需求的變化
  - 以便作業系統*只置換實際所需的記憶體空間*，節省置換所消耗的時間

置換（３）

- 發生記憶體存取的錯誤
  - 原因：置換出不處於閒置狀態的行程（例：置換出的行程正在等待非同步的 I/O 操作）
- 解決方式：
  1. 任何企圖作***I/O操作***的行程不會被置換
  2. 只有進入***作業系統緩衝區***的行程才可作I/O操作，而作業系統與行程記憶體間的資料傳遞，只有在行程被置入時才可以進行

記憶體分割



- 一部分供作業系統常駐使用
  - 放置位址常考量***中斷向量的位址***，因此作業系統常位於***低位址***
- 另一部分供使用者行程使用
- 配置方式：
  - 單一分割配置（單一使用者）
  - 多重分割配置（多元程式概念）

單一分割配置

- 單一使用者
- 記憶體分隔成兩部分，一用來***常駐作業系統***，剩餘僅供***一個使用者行程***執行
- 缺點：
  - 僅讓一個行程執行，造成記憶體空間的浪費
  - 若行程執行 I/O 操作， CPU 閒置，使整體系統效能降低
  - 若行程大小超過可用的記憶體空間，將導致程式無法執行（可能解決方法：***重疊***）

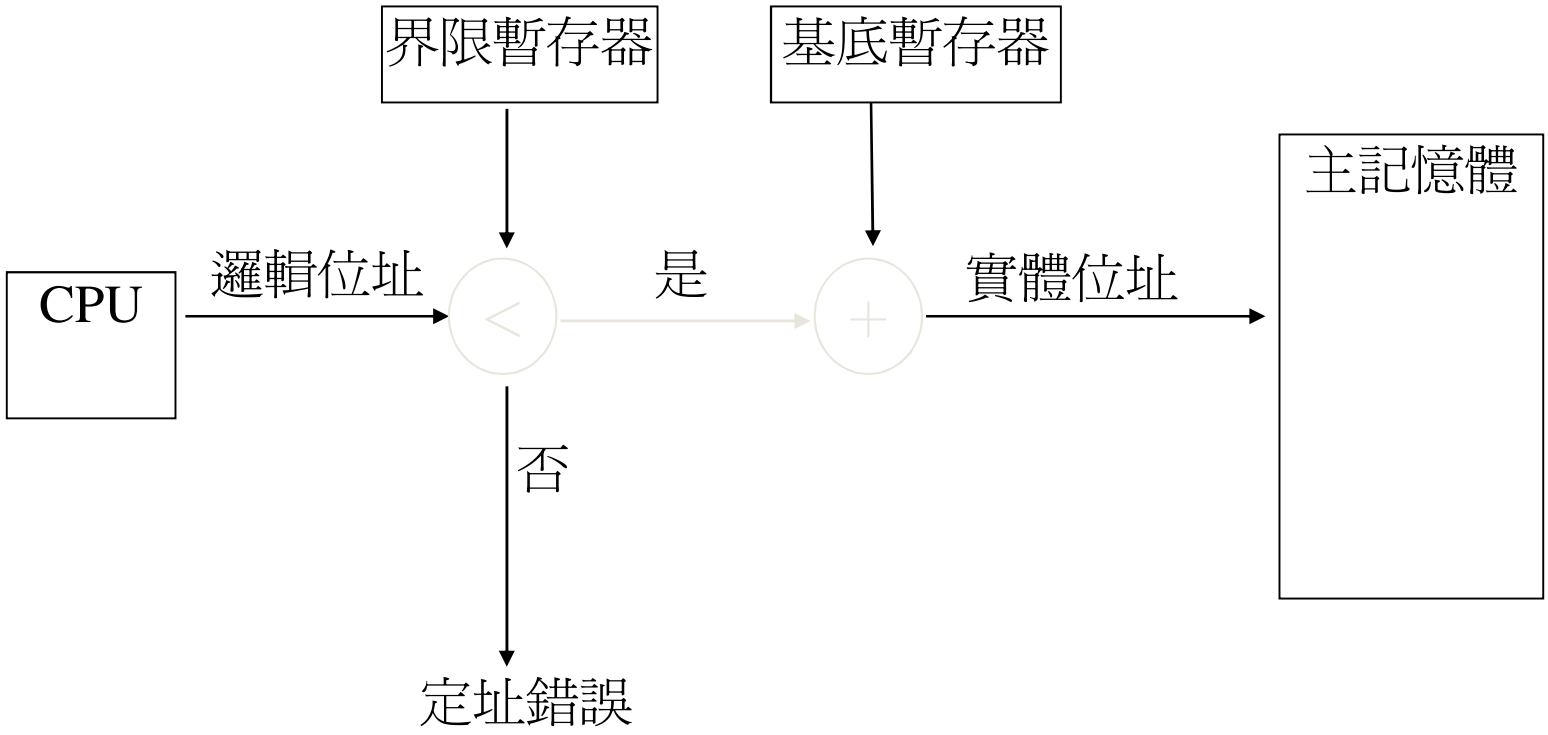
單一分割配置（２）

- 如何保護作業系統與使用者程式不會遭到對方不當修改？
  - 利用***基底暫存器***與***界限暫存器***的輔助
    - ***基底暫存器***：存放最小的記憶體實體位址
    - ***界限暫存器***：存放邏輯位址範圍
    - 每一個邏輯位址都須小於界限暫存器的邏輯位址範圍
    - 邏輯位址＋基底暫存器內的數值→記憶體的實體位址
  - CPU排程器選定一行程

    →分派程式將正確的值載入到基底和界限兩暫存器中

    →CPU每存取一次邏輯位址都經由兩暫存器的核對轉 換，確保作業系統與使用者程式不會互相影響

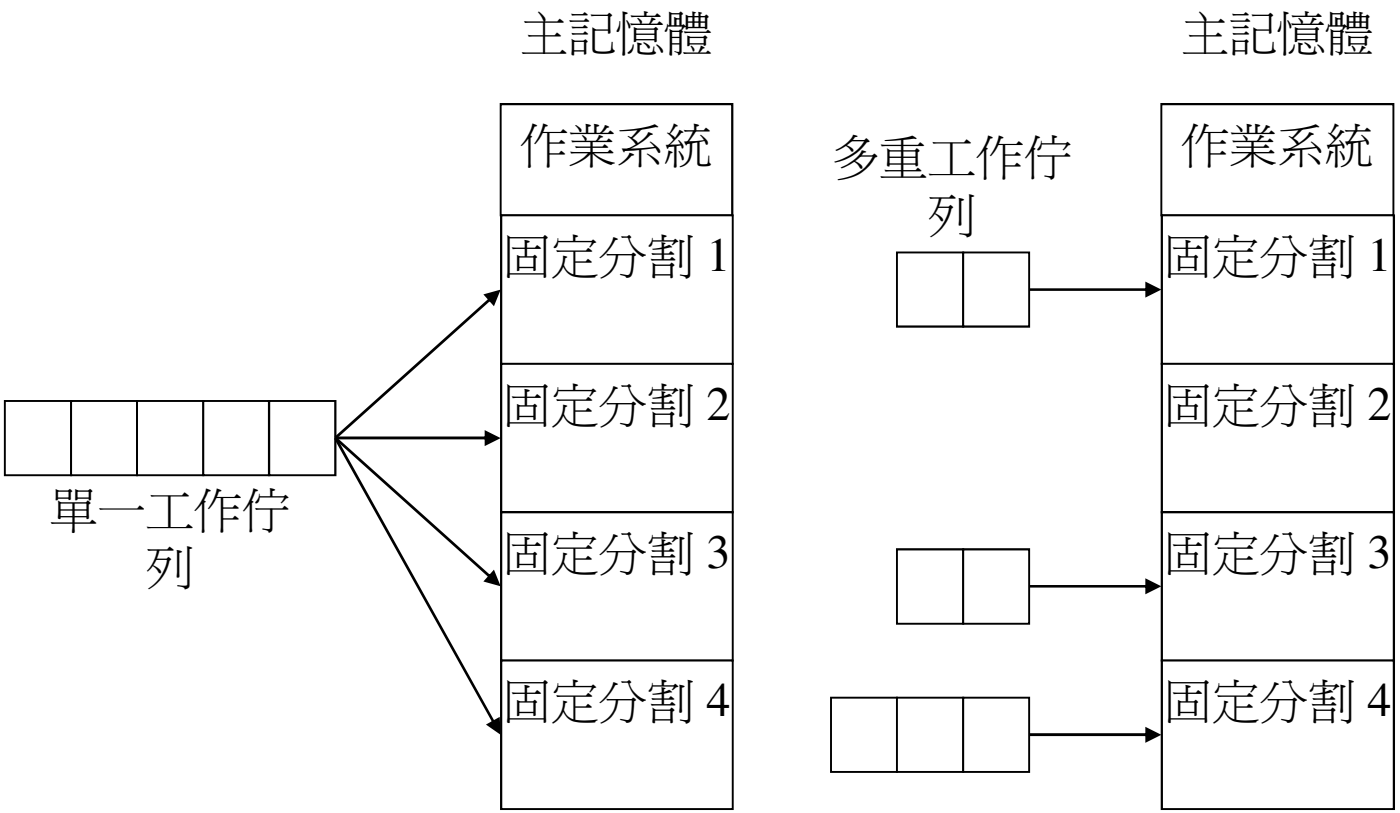
位址保護機制



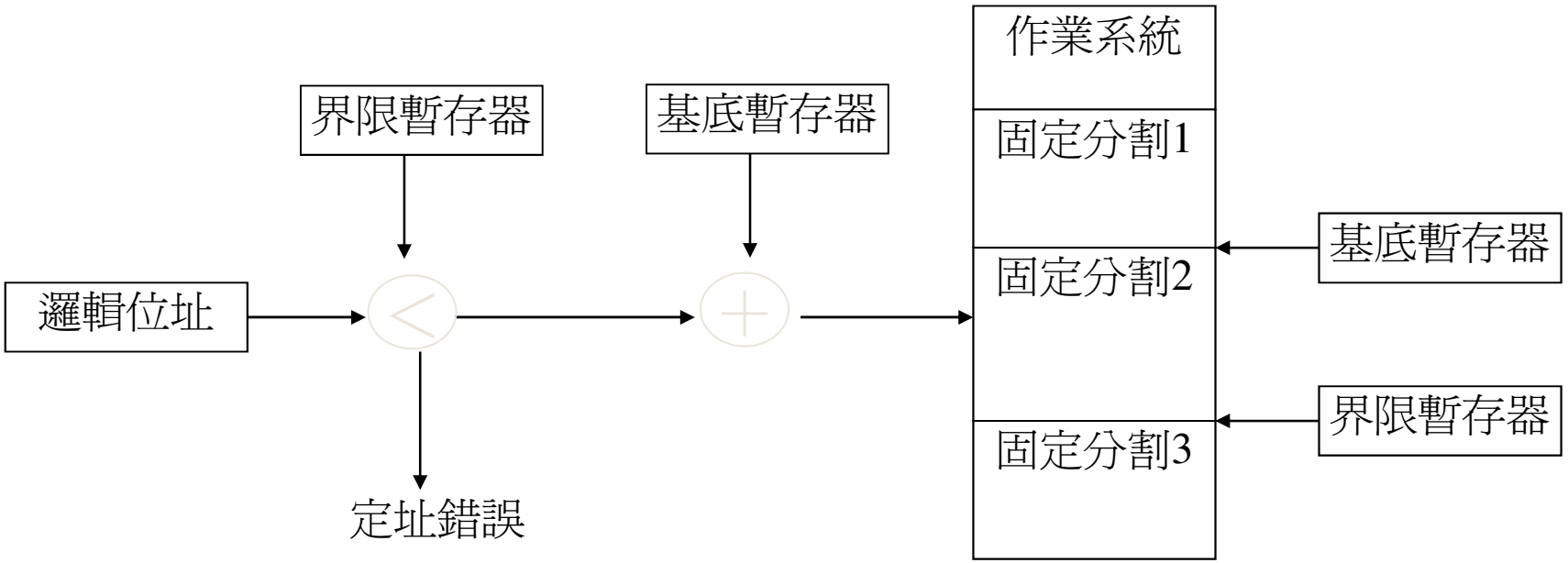
多重分割配置 (1)

- 多元程式概念
- 作業系統如何將可用的記憶體空間配置給正在輸入佇列中等待的多個使用者程式？
  - 將記憶體劃分成許多***固定大小***的***區域***或***分割***
  - 每個分割***只能容納一個行程***
  - 分割數目的多寡會影響到同時放置的***使用者行程數量***
- 設計輸入佇列兩方法：
  - ***單一工作佇列***：所有的分割都**對應到同一個輸入佇列**
    - 記憶體使用率較高；程式的等待時間較一致
  - ***多重工作佇列***：每個分割均有一個對應的輸入佇列
    - **缺點**：作業系統須針對每個程式找到適合的輸入佇列，會造成記憶體中雖有可用空間卻無法使用的情況

單一與多重佇列


主記憶體
作業系統
固定分割 1
固定分割 2
固定分割 3
固定分割 4
單一工作佇列
主記憶體
多重工作佇列
作業系統
固定分割 1
固定分割 2
固定分割 3
固定分割 4

多重分割下的系統保護



多重分割配置（２）

- 記憶體劃分成固定大小的分割之問題：
  - 若程式***小於***分割大小的行程，會造成**記憶體空間的浪費**
  - 而程式***大於***分割大小的行程，則因爲需要跨越分割，增加系統額外的***負擔***
- 解決方法：***動態分割法***
  - 依照程式執行時的大小，在記憶體中找到夠大的可用區塊給此行程使用
  - 需在作業系統維護一個表格，隨時記錄記憶體中哪些區塊使用中、哪些區塊空閒

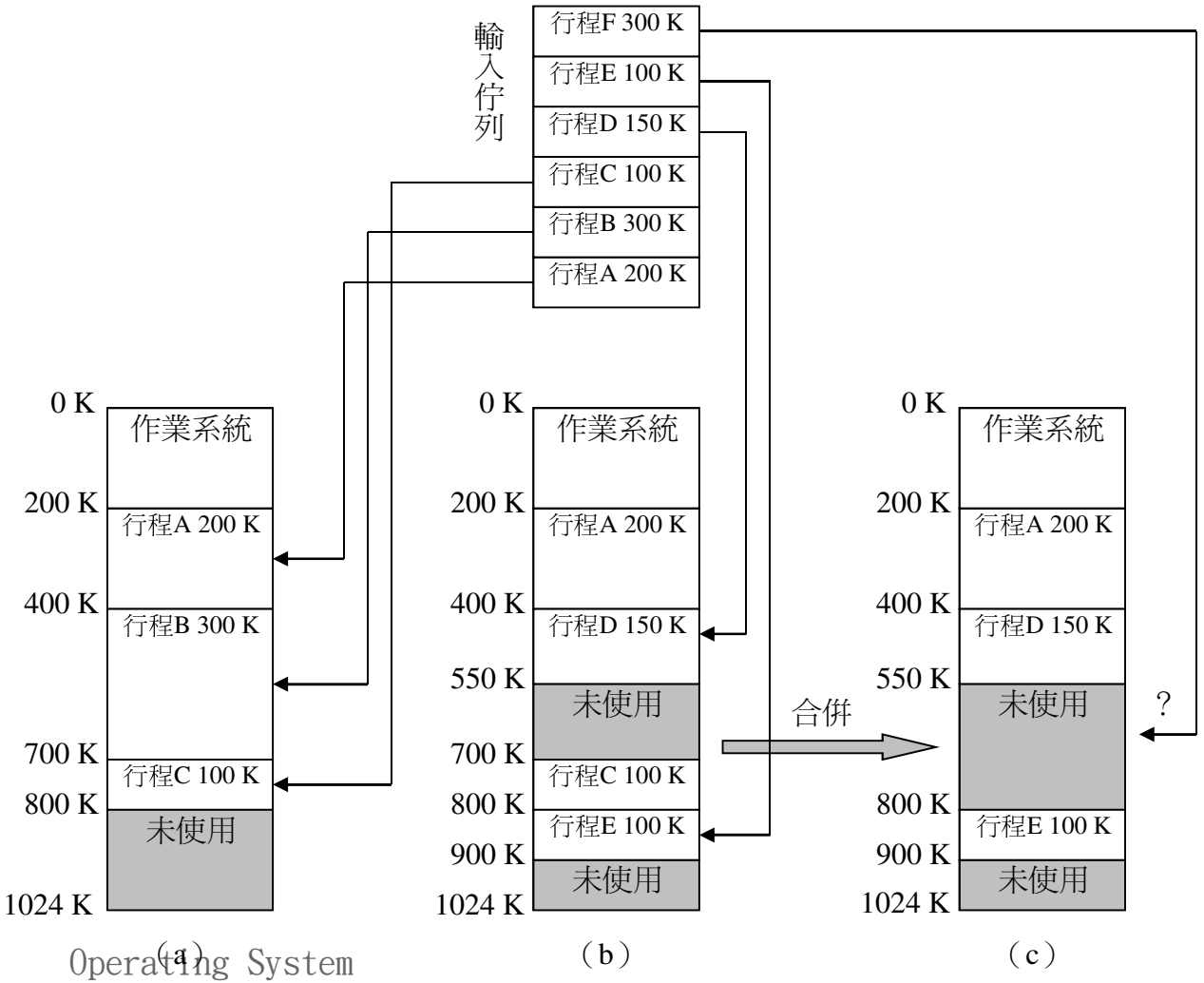
多重分割配置（3）

- 作業系統以動態分割法載入程式時，依據 3 種策略：
  - ***最先符合法（First Fit）***：從第一個可用的區塊開始循序找起，只要找到夠大的空間，就把程式載入
  - ***最佳符合法（Best Fit）***：找到一個與載入程式大小最爲接近的區塊，再把程式載入。
  - ***最差符合法（Worst Fit）***：找到最大的區塊，再把程式載入。

多重分割配置（4）

- 根據電腦模擬分析
  - ***最先符合法***與***最佳符合法***在搜尋時間和記憶體空間的使用率都優於***最差符合法***
  - 最先符合法的執行速度通常比最佳符合法與最差符合法要快
  - 若不考慮搜尋時間，行程大小變化較大的適合最佳符合法；反之，則適合最差符合法。
- 行程執行結束後，作業系統將會
  - ***釋放***此行程所佔用的記憶體區塊
  - 檢查此被釋放區塊是否可與可用的***相鄰區塊合併***
  - 同時作業系統檢查輸入佇列中是否有程式正等待配置記憶體→如果有，則檢查此新合併的區塊大小是否夠該行程所用。

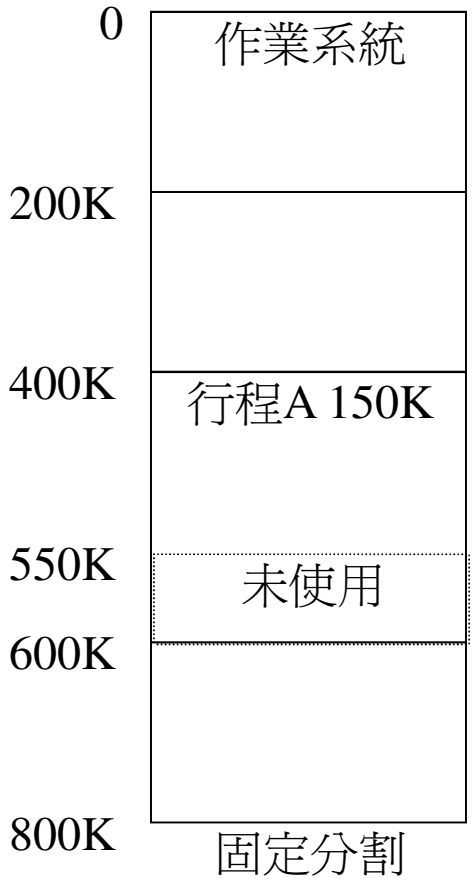
斷裂


輸入佇列
行程F 300 K
行程E 100 K
行程D 150 K
行程C 100 K
行程B 300 K
行程A 200 K
0 K
200 K
400 K
700 K
800 K
1024 K
作業系統
行程A 200 K
行程B 300 K
行程C 100 K
未使用
(a)
0 K
200 K
400 K
550 K
700 K
800 K
900 K
1024 K
作業系統
行程A 200 K
行程D 150 K
未使用
行程C 100 K
行程E 100 K
未使用
(b)
合併
0 K
200 K
400 K
550 K
800 K
900 K
1024 K
作業系統
行程A 200 K
行程D 150 K
未使用
?
行程E 100 K
未使用
(c)

斷裂

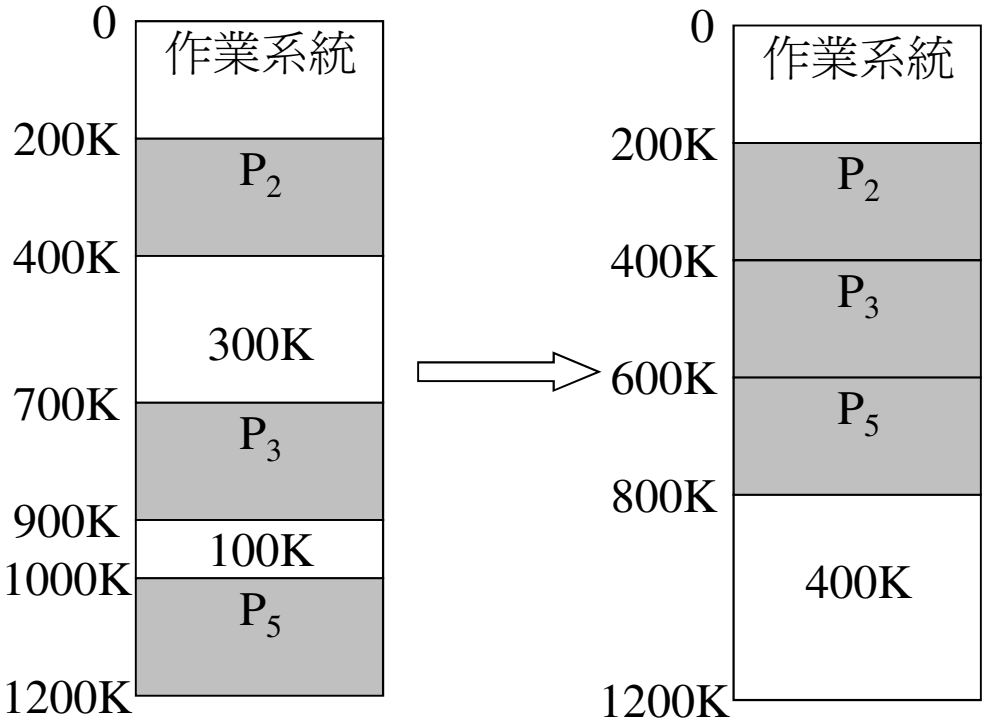
- 有記憶體空間卻無法用（記憶體區塊太小）
- ***外部斷裂***
  - 因爲行程持續地被載入與置換，使得可用的記憶體空間被分割成許多不連續的區塊
  - 雖然記憶體所剩空間總和足夠讓此行程執行，卻因爲空間不連續，導致程式無法載入執行
- ***內部斷裂***
  - 發生在以***固定分割方式配置***的記憶體
  - 當一個程式載入到固定大小的分割，假如程式小於此分割，則此分割剩餘空間無法被使用
  - 不能使用的空間散佈在各個分割內，造成輸入佇列中的程式無法順利載入執行，造成浪費

內部斷裂



聚集

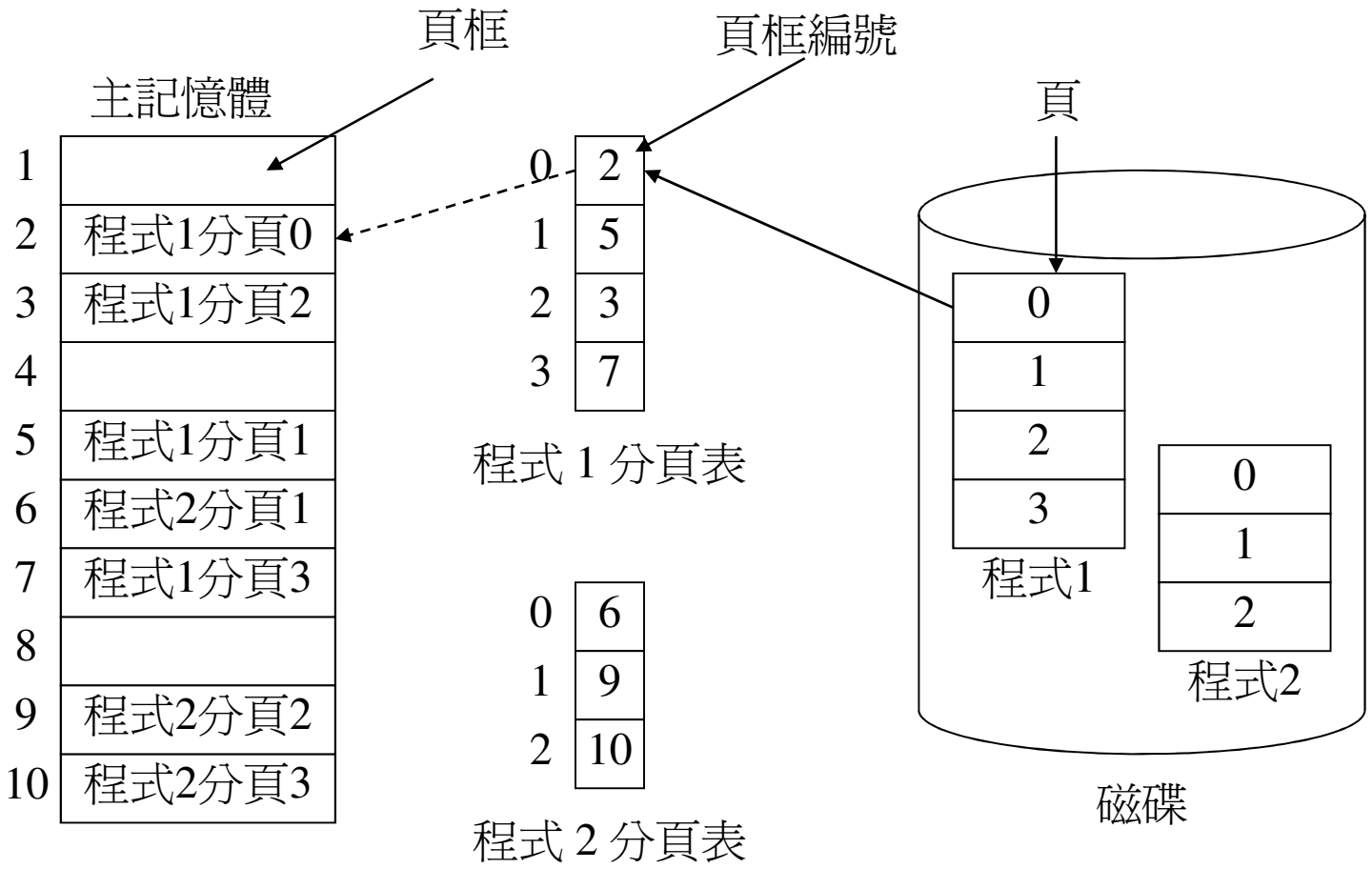
- ***解決***外部斷裂的問題
- 將記憶體中可用的不連續空間聚集成一大塊連續空間
  - 程式必須都可重新定位
  - ***代價高***，因爲要搬動許多行程的實體記憶體空間



基本方法（1）

- 程式可被不連續放置，沒有外部斷裂的問題
  - 將載入的程式分割成固定大小的***分頁***
  - 主記憶體也分割成固定大小的***頁框***，大小與分頁相同
  - 執行程式時，把程式所有的分頁載入記憶體任何可用的頁框中
- 每個程式有一個***分頁表***，存有每分頁在記憶體中的***起始位址***。當程式的分頁被載入到主記憶體時，程式分頁表中記錄該分頁被載入至主記憶體的哪一個頁框中
- ***頁框表***：作業系統須知道主記憶體中頁框使用與否、系統中總頁框數目有多少等資訊

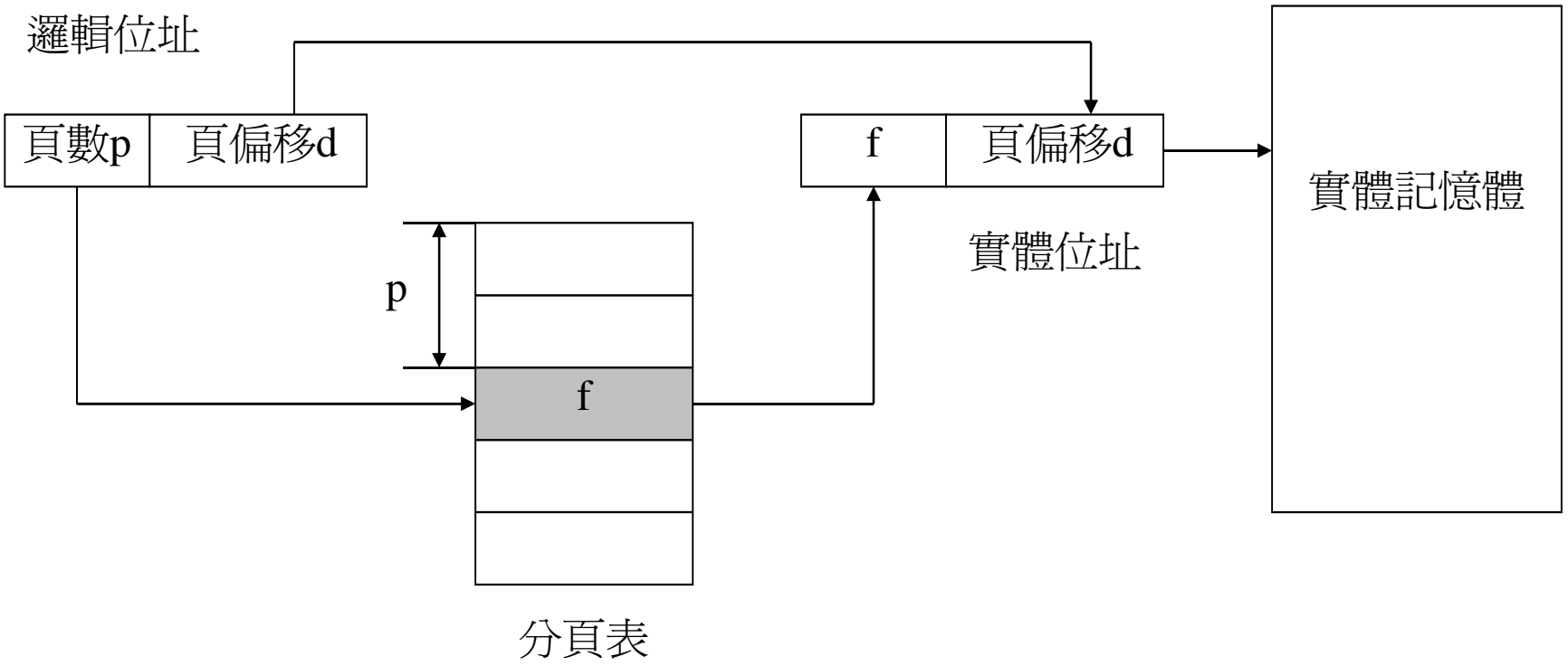
分頁法


頁框
頁框編號
主記憶體
頁
1
2 程式1分頁0
3 程式1分頁2
4
5 程式1分頁1
6 程式2分頁1
7 程式1分頁3
8
9 程式2分頁2
10 程式2分頁3
0 2
1 5
2 3
3 7
程式 1 分頁表
0 6
1 9
2 10
程式 2 分頁表
0
1
2
3
程式1
0
1
2
程式2
磁碟

基本方法（２）

- CPU 產生的邏輯位址分成兩部分：
  - ***分頁碼***爲指向分頁表的索引
  - ***頁偏移***表示與***該分頁起始位址的距離***
- 在分頁表中找到此分頁在實體記憶體中對應的基底位址後，再和***頁偏移***組合定義出***實體記憶體位址***
- 分頁法仍有***內部斷裂***的問題
  - 如果一個程式所需要的記憶體大小不能被頁框大小整除，最後一個頁框就不會全部被使用而形成內部斷裂
  - 若使分頁大小縮小，就可節省因內部斷裂而產生記憶體空間浪費；但***須更大空間儲存分頁表***（分頁的數量增加）
  - 每個程式平均會有 ***1/2 分頁的內部斷裂***

分頁邏輯位址空間與實體位址空間的轉換



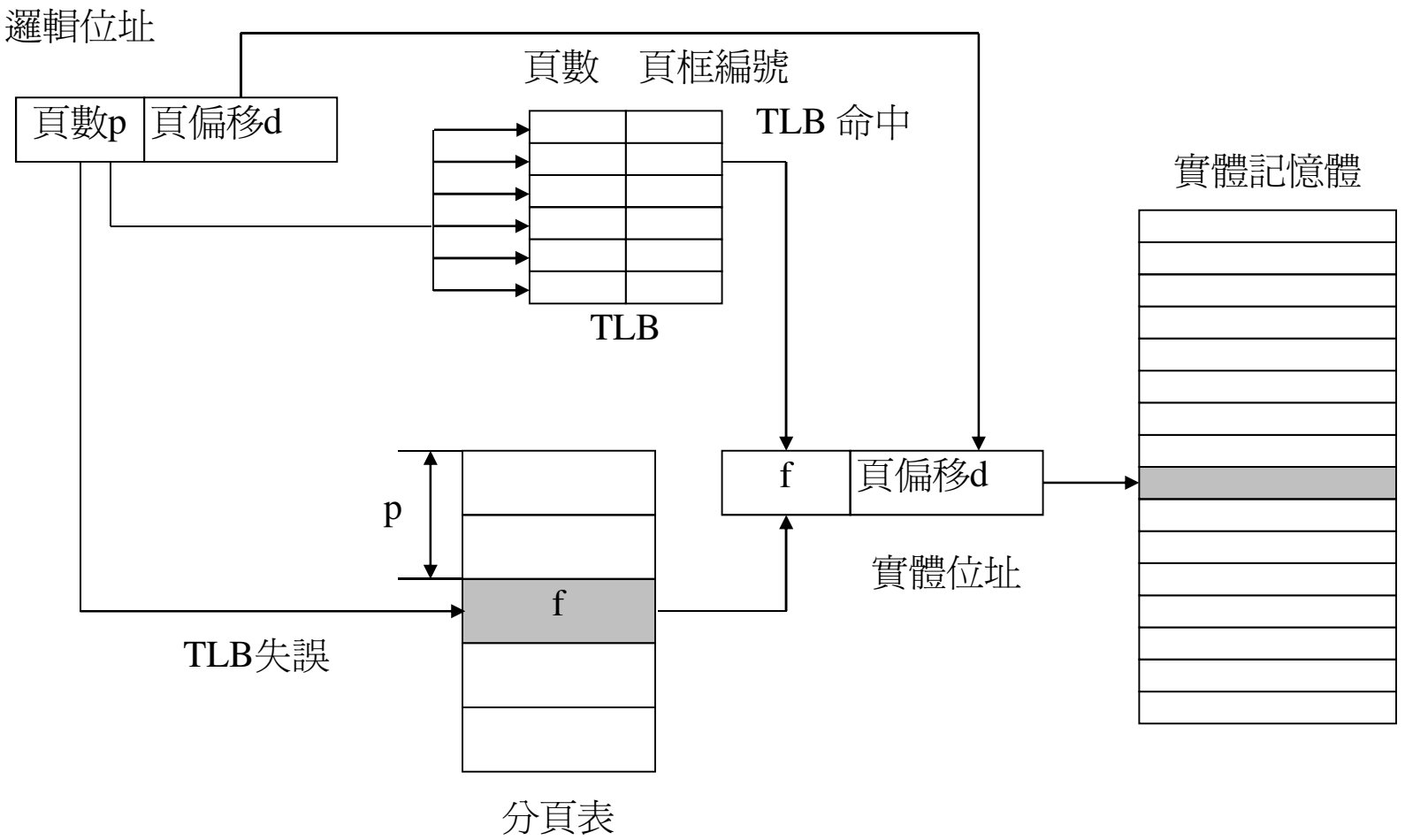
分頁表的結構（1）

- ***儲存分頁表***：若儲存在主記憶體中，系統需作兩次記憶體存取，效率低
- ***解決方法***：將分頁表放在***關聯式記憶體***(*Associative memory*)（也稱***位址查閱緩衝***(*Translation Look-aside Buffer, TLB*)）中
  - 每筆資料有***分頁碼***，***頁框編號***兩欄位。關聯式記憶體以分頁碼爲索引，平行地在相對的頁框中找尋
  - 很短的時間就可找尋到，時間複雜度爲 ***O(1)***

分頁表的結構（２）

- 實際上的做法（因關聯式記憶體昂貴）：使用***主記憶體建立分頁表***，將關聯式記憶體當成快取記憶體，只保存***分頁表的部分內容***
- 若頁框編號在關聯式記憶體中找得到，就直接與頁偏移相加得到實體記憶體位址，否則再到分頁表中找尋
- 在更新關聯式記憶體時，如果關聯式記憶體已經存滿，則系統必須置換掉其中一筆記錄（參考第９章的***分頁置換規則***）

使用 TLB 硬體支援的分頁法



分頁表的結構（3）

- 使用關聯式記憶體來儲存分頁表的效率：
  - 如果要尋找的分頁碼已經在關聯式記憶體中，則稱爲***命中***，否則稱爲***失誤***
  - ***命中率***的定義爲〔命中次數 /（ 命中次數 + 失誤次數）〕× 100%
  - 例：假設存取關聯式記憶體的時間爲 20 奈秒，直接存取主記憶體的時間爲 100 奈秒。假設在關聯式記憶體內命中率爲 95%，則所需的時間爲：

分頁表的結構（4）

- 0.95 × 120 ns + 0.05 × 220 ns = 125 ns
- 其中 120 奈秒的 20 奈秒花費在關聯式記憶體中找尋（命中），100 奈秒花費在存取主記憶體的資料；而 220 奈秒中的 20 奈秒花費在關聯式記憶體中找尋（失誤），100 奈秒花費在主記憶體中找尋，另外 100 奈秒花費在存取主記憶體的資料

- 適當地使用關聯式記憶體，可以降低主記憶體的存取時間，也可以節省成本

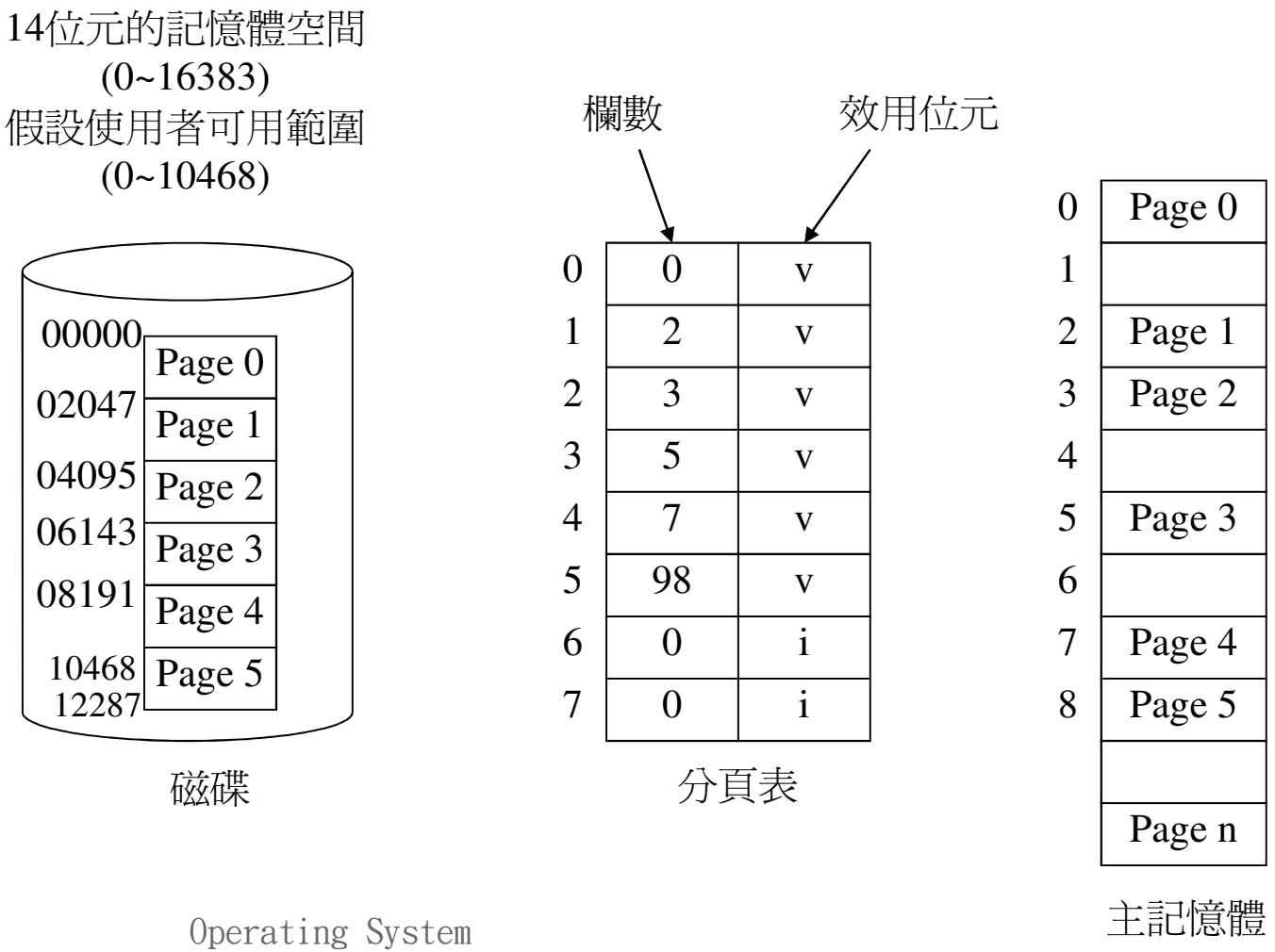
分頁表的結構（５）

- 分頁的環境中，記憶體的保護可以使用分頁上面的***保護位元***來完成（通常保存在***分頁表***）
- 保護位元可以定義某一分頁是***可讀***、***可寫***或者***兩者皆可***
- 每次的位址轉換均會去參考相對應的保護位元；企圖以非保護位元所提供的動作來存取此分頁，將會引發***例外中斷***

分頁表的結構（6）

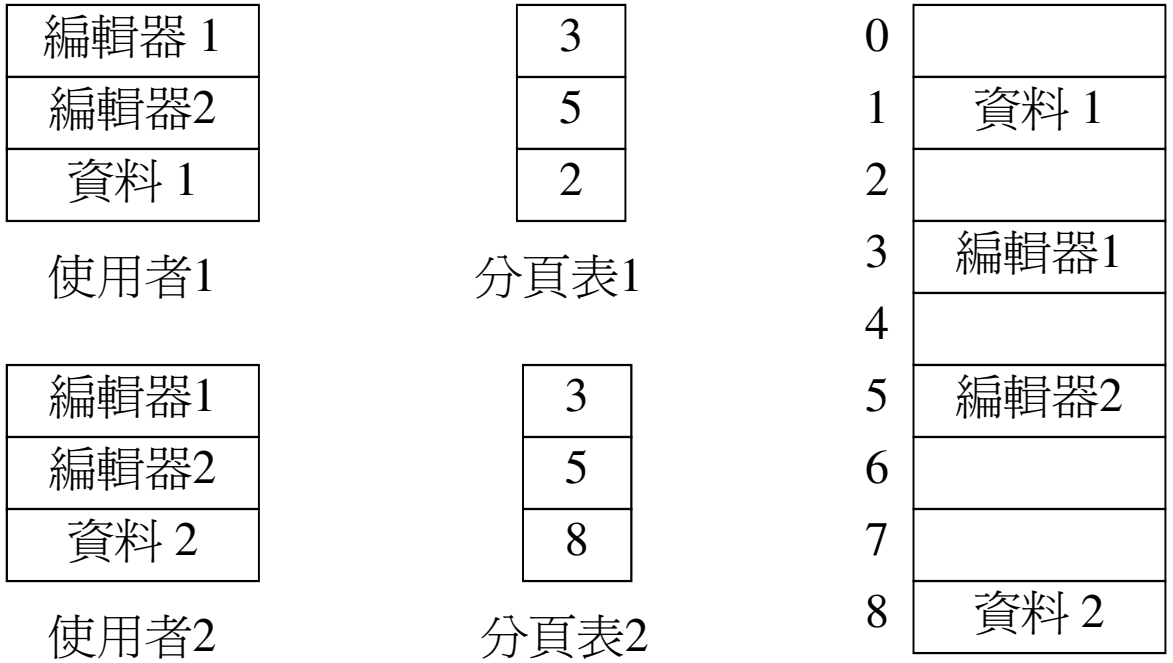
- ***效用位元(Valid-invalid bit)***：
  - 當此位元被設定爲有效（v）時，表示所對應的分頁目前在***主記憶體中***；如果此位元被設定成無效（i），則表示所對應的分頁在***輔助記憶體中***
  - 作業系統會設定每分頁的有效位元，以核對該分頁的存取動作；當系統進行位址轉換時，也會去檢查此位元，若企圖存取無效的分頁也會引發***例外中斷***

分頁表中的效用位元


14位元的記憶體空間
(0~16383)
假設使用者可用範圍
(0~10468)
00000
02047
04095
06143
08191
10468
12287
Page 0
Page 1
Page 2
Page 3
Page 4
Page 5
磁碟
欄數
效用位元
0 0 v
1 2 v
2 3 v
3 5 v
4 7 v
5 98 v
6 0 i
7 0 i
分頁表
0 Page 0
1
2 Page 1
3 Page 2
4
5 Page 3
6
7 Page 4
8 Page 5
Page n
主記憶體

兩個使用者共用頁框

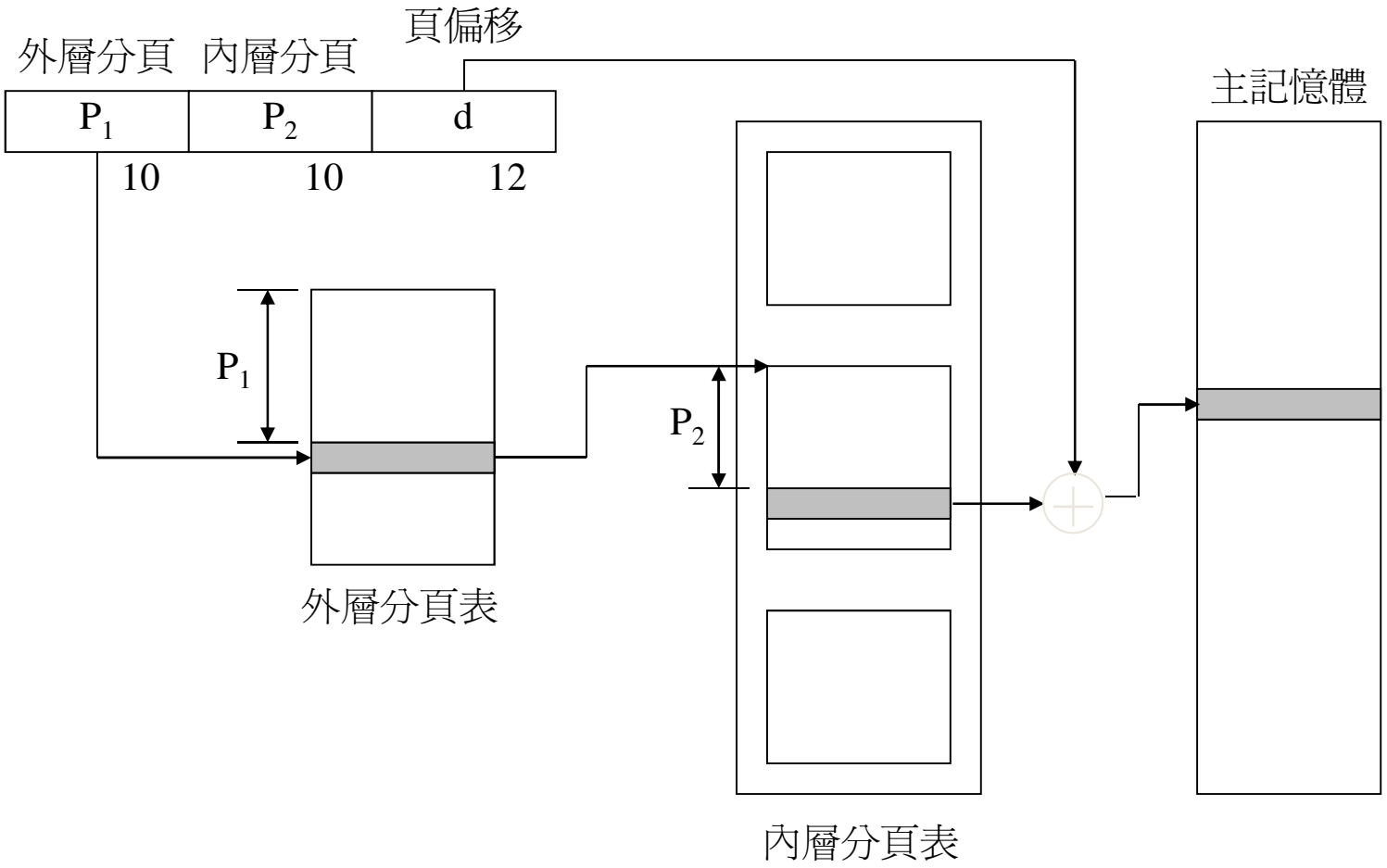
- 使用分頁法另一項好處
  - 簡單達到*程式碼共用*



多層分頁法（１）

- 不希望分頁表佔用連續的記憶體空間→想把分頁表分成較小的單位儲存
- 多層式分頁將分頁表也進行***分頁***
  - 以兩層式的分頁來說，在由邏輯位址轉換成實體位址時；
  - 先以 P_1 爲索引到***外層分頁表***中找尋，所找到的記錄會指向一個相對應的***內層分頁表***
  - 然後再以 P_2 爲索引到該***內層分頁表***中找尋，所找到的記錄會指向一個***頁框***
  - 最後和***頁偏移*** d 相加成爲***實體記憶體位址***

兩層式分頁法



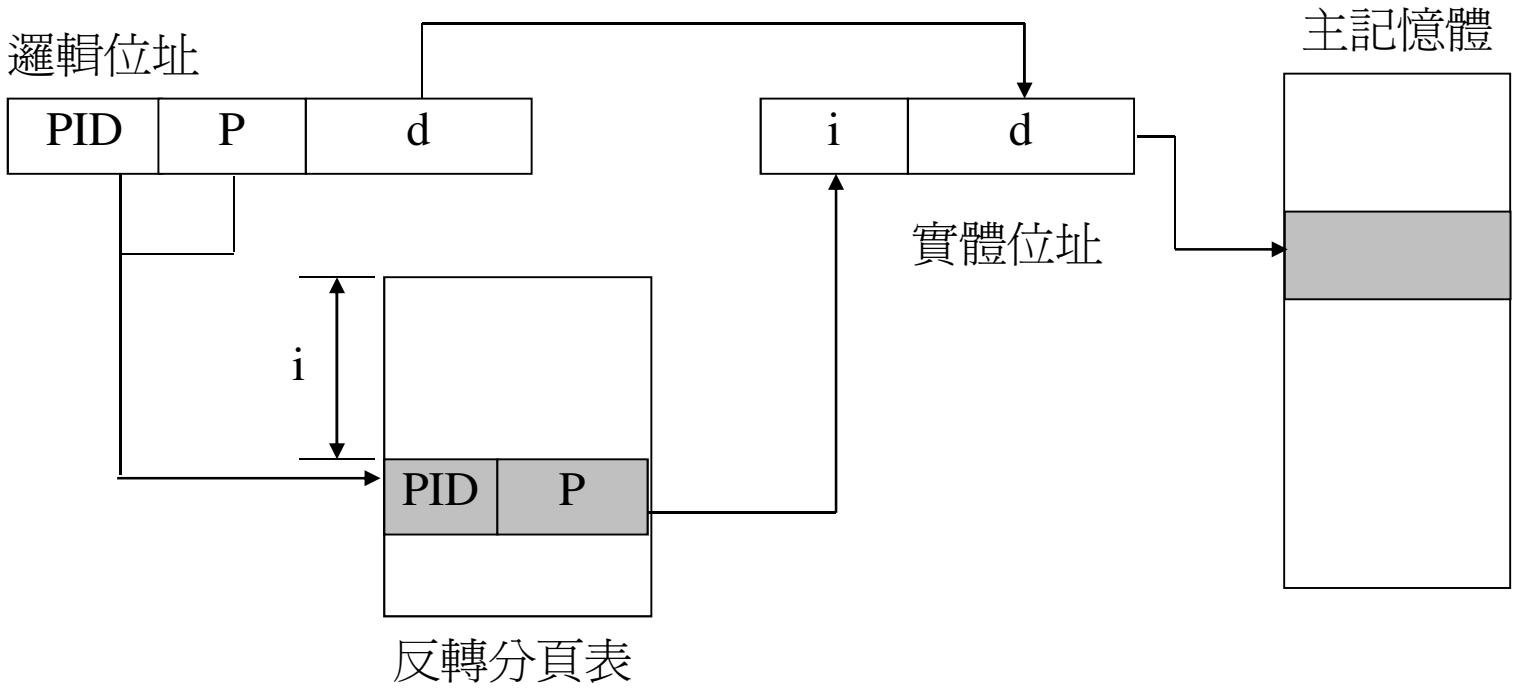
多層分頁法（2）

- 64 位元的架構*不適合階層式的分頁方式*（外層分頁表將佔用太大的連續記憶體空間）
- 對於邏輯位址*大於 32 位元*的*硬體架構*，可使用*雜湊分頁表*
  - 利用雜湊函數將邏輯位址對應到實體位址
  - 可能發生*雜湊碰撞*而*降低效率*，但是可有效*減少分頁表空間*

反轉分頁表 (Inverted Page Table)（1）

- 解決分頁表空間太大的問題：整個系統僅用一個分頁表
  - ***行程辨識碼***加上***邏輯分頁***與***實體頁框***是一對一的對應
  - 反轉分頁表的***記錄數目***要與***頁框***一致
- 反轉分頁表的架構下
  - 一個邏輯位址包含三個欄位：***行程辨識碼（PID）***、***分頁碼***、與***頁偏移***
  - 反轉分頁表中每個記錄包含：***行程辨識碼***與***分頁碼***兩欄位
  - 當一個行程要進行位址轉換，必須以 PID 與分頁碼當索引，到反轉分頁表中找尋所屬的頁框編號，再與頁偏移相加就可以得到實體位址

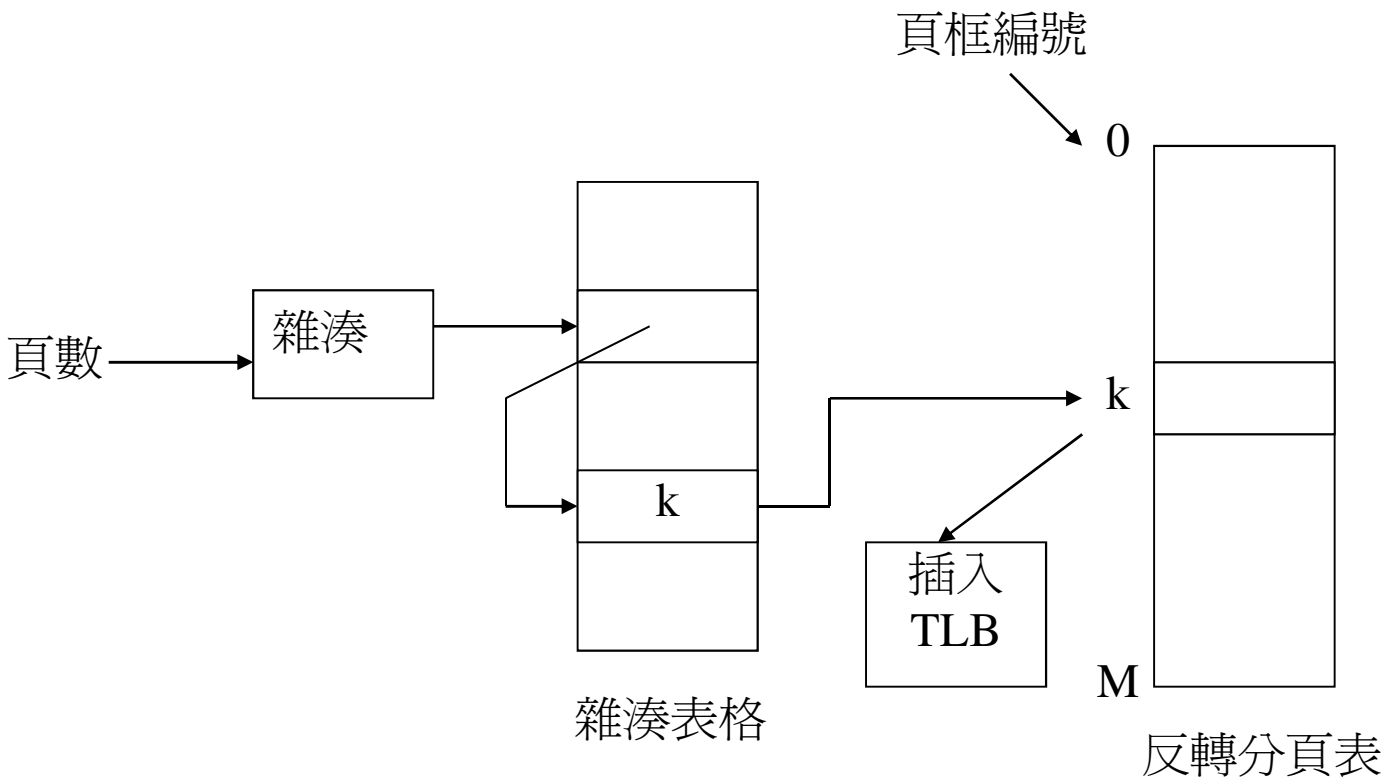
反轉分頁表（２）



反轉分頁表（3）

- 反轉分頁表可以降低儲存每個分頁表所需要的空間，但是***搜尋分頁表所用的時間卻大量增加***
  - 減輕此問題：使用***雜湊表格***
- 從取得邏輯位址開始到在記憶體中存得資料，至少需要兩次的記憶體讀取：一次是***雜湊表格***，另一次是***反轉分頁表***
  - 解決的辦法是利用***類似關聯式記憶體***的方式來加快搜尋
- 在一個使用反轉分頁表的系統中，***無法***達到***分頁共用***的目的

反轉分頁表 - 雜湊表格



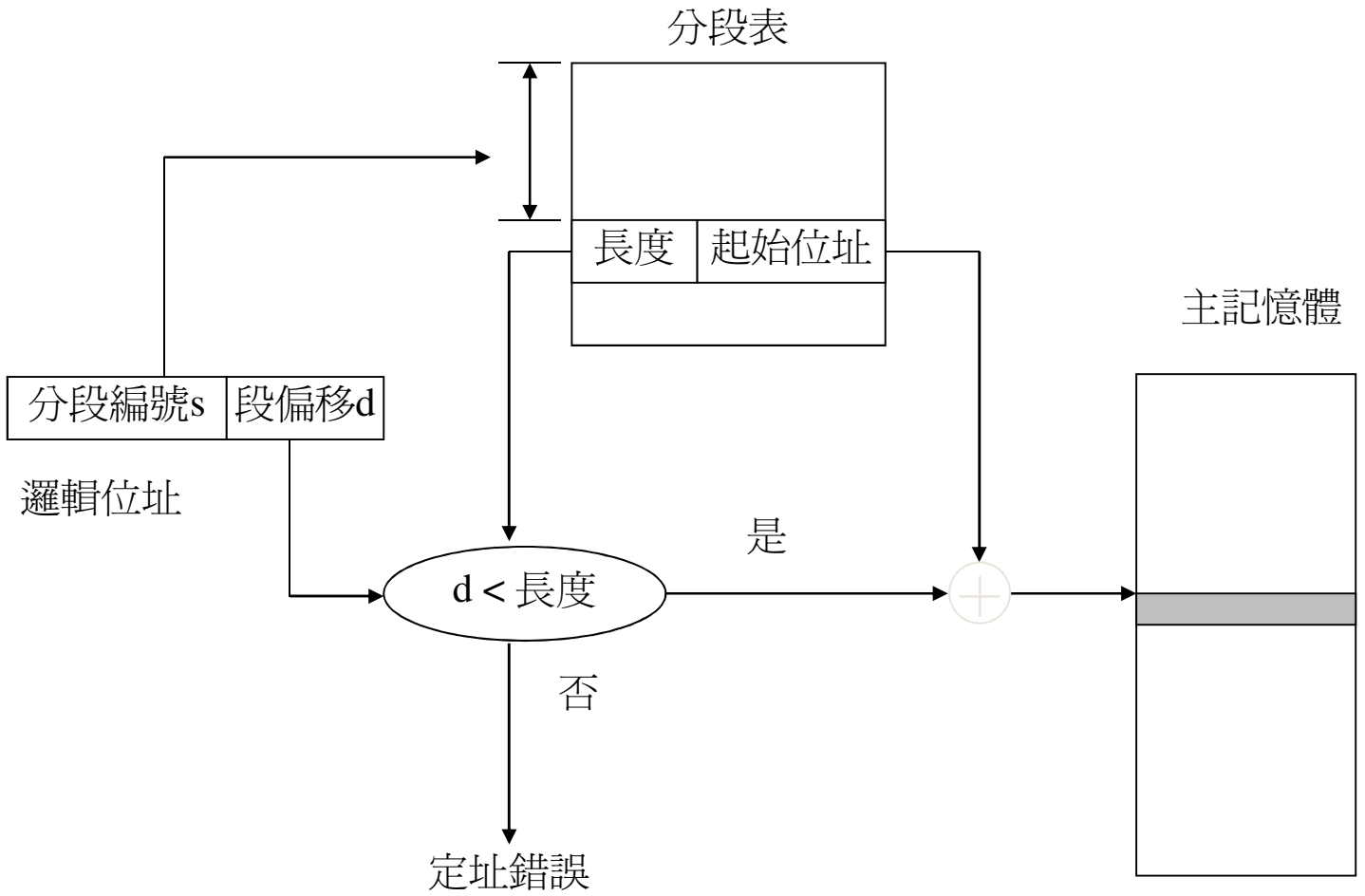
分段 (Segmentation)

- 依照程式的邏輯功能將邏輯位址空間*切割成許多分段*
- 每個分段的長度都*不盡相同*，長度是由分段中的程式大小來決定
- 若要參考分段中某一位元組，可以藉由此分段的*起始位址*配合*段偏移*來決定

基本方法（1）

- 一個行程的邏輯位址空間被分成許多分段，每個分段都有一個***名稱***和***長度***，也有一個***分段編號***
  - 一個行程的邏輯位址被分成兩個欄位：***分段編號***（當作分段表的索引）與***段偏移***
- 每個程式均有一個***分段表***，其中每一記錄都有：
  - ***分段基底值***：記錄分段在主記憶體中實際***開始的位址***
  - ***分段界限值***：記錄該***分段的大小***，以避免程式執行時超過該分段界限

分段法的使用

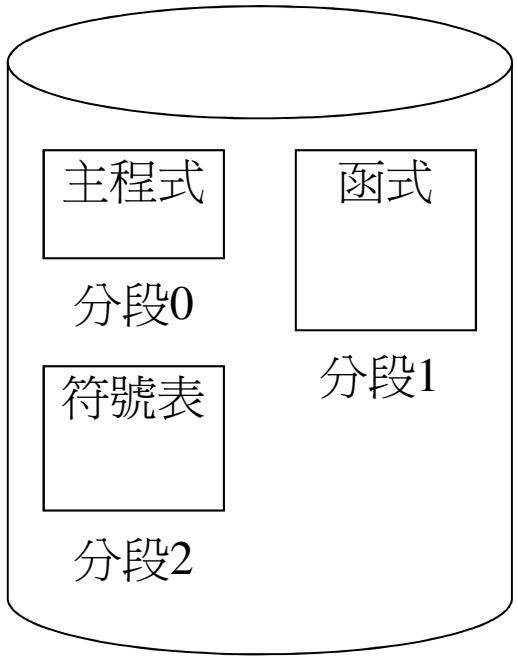


分段法

主記憶體

	位址	長度
0	5100	200
1	1500	300
2	3200	300

分段表

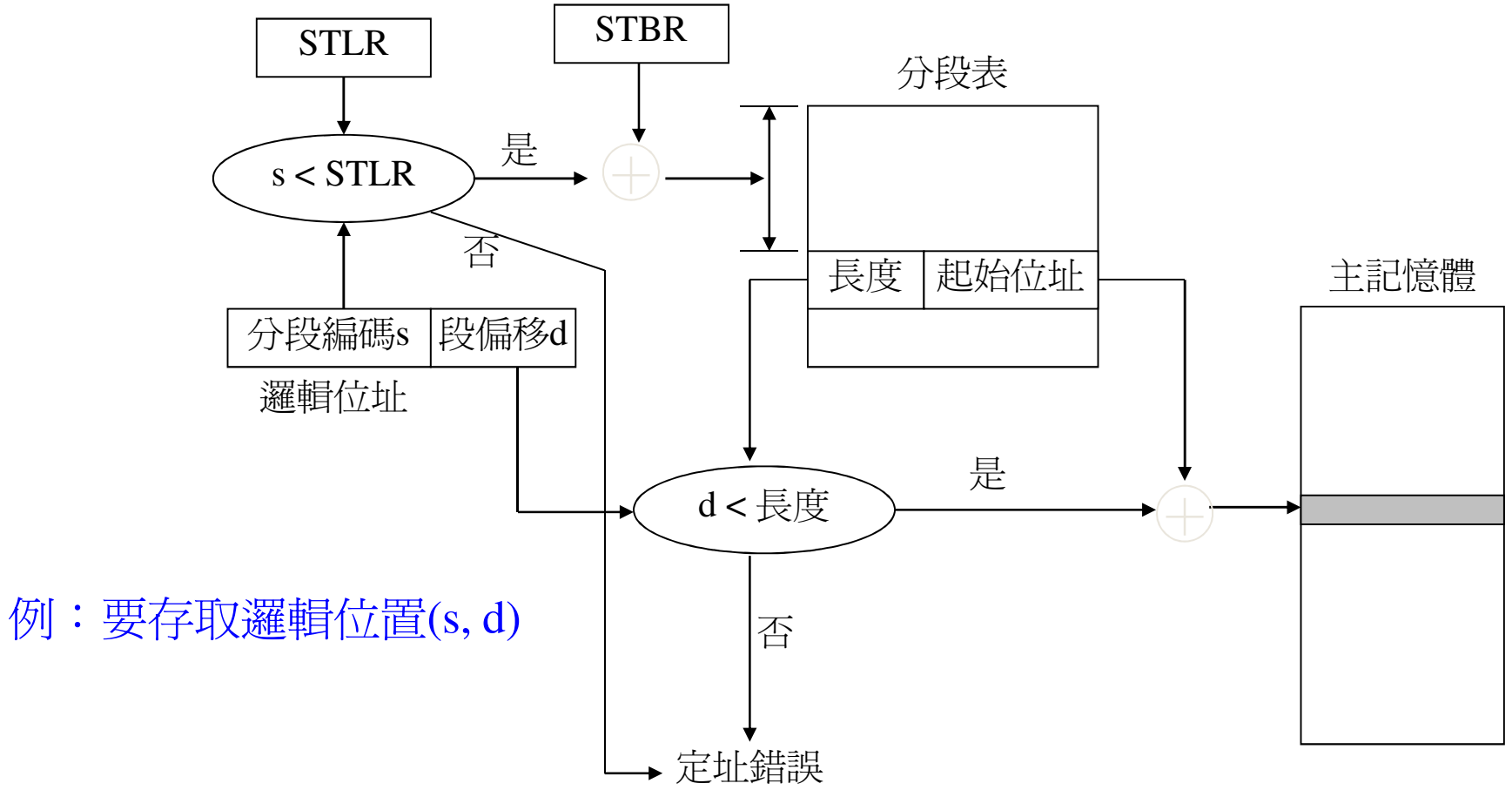


邏輯位址空間

基本方法（2）

- 分段表可以存放在
  - ***快速的暫存器***：節省對界限暫存器相加與基底暫存器比較的時間
  - ***主記憶體***：對於有***大量分段***的程式而言，不宜放在快速暫存器中
    - 利用一個***分段表基底暫存器***(Segment Table Base Register, STBR)來指向***分段表***；使用一個***分段表長度基底暫存器***(Segment Table Length Register, STLR)記錄***分段大小***
    - 將***分段基底值***與***段偏移***相加，取得要存取位元組的實體位址
    - ***缺點***：每個邏輯位址需要對***實際記憶體存取兩次***，所以***速度比較慢***
    - 解決方法：也可利用***關聯式暫存器***儲存經常會被使用的分段表記錄

配合 STLR 與 STBR 的分段法

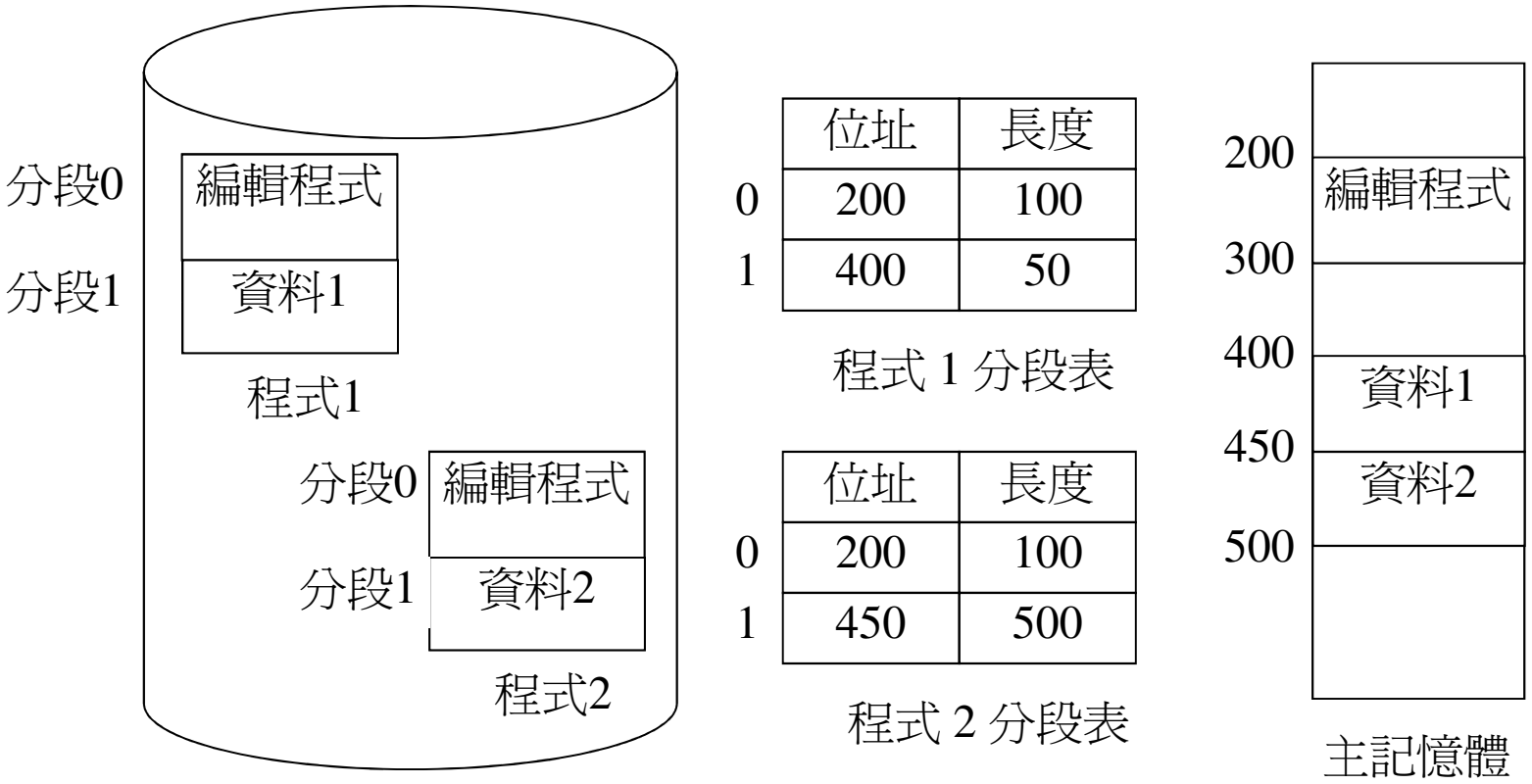


例：要存取邏輯位置(s, d)

基本方法（3）

- 分段法可達到***分段保護***的功能
  - 可以指定指令分段是***唯讀***或是***唯執行***的方式
  - 分段表中每個項目有一組***保護位元***，以避免不正當的存取
- 分段法可達到***資料***或***程式碼共用***的功能
  - 兩個分段表中有共同的項目指向同一個***實體位址***
- 問題：行程可能跳到本身以外的記憶體位址執行程式碼
  - 如果此記憶體位址的程式碼共用，勢必所有共用行程所屬共用分段的分段編號都要相同
  - 該編輯程式如何參考本身的程式？且共用的行程增多後，想要找到一個***可共用的編號***將會更加地困難

分段共用


分段0
分段1
編輯程式
資料1
程式1
分段0
分段1
編輯程式
資料2
程式2
位址
長度
0
200
100
1
400
50
程式 1 分段表
位址
長度
0
200
100
1
450
500
程式 2 分段表
200
300
400
450
500
編輯程式
資料1
資料2
主記憶體

基本方法（4）

- 分段法中分段長度是變動的，若有***外部斷裂***，便可能導致記憶體中根本找不到足夠空間供一個分段載入
  - 該行程可以等待別的***行程釋放記憶體空間***，或是利用***聚集法***取得夠大記憶體空間以供載入
  - 若選擇等待其他行程釋放空間，CPU 排程程式可以考慮讓***優先權較高***、***需要記憶體空間較小***的行程先行，以增進效率

分頁式分段

- 在分頁式分段的方式下
  - 每個行程會依據邏輯功能切成***分段***，而每個分段會再被分割成***分頁***，所以在這種架構下***外層***爲***分段表***，***內層***爲***分頁表***
  - 一個邏輯位址包含了分段編號 s、分頁碼 p 與頁偏移 d
  - 系統會先透過 s 在分段表中找到分頁表的起始位址，再透過 p 在分頁表中找到分頁的頁框位置，最後再與 d 相加成爲實體記憶體位址
- 平均每個分段也會有 1/2 分頁的內部斷裂
  - 所以分頁式分段中每個行程***會有較多的內部斷裂***
  - 著名的 OS/2 作業系統便是使用分頁式分段法，使用的分頁大小爲 4 KB

分頁式分段（２）