

Chapter 5 CPU Scheduling

5.1 Basic Concepts

- CPU Burst
 - 使用CPU的時間
- I/O Burst
 - 使用I/O的時間
- I/O Bound Process
 - 小的CPU burst. Spends more of its time doing I/O than computations.
- CPU Bound Process
 - 大的CPU burst Use more of its time doing computation than I/O.

Non-Preemptive and Preemptive Process(1)

- 有些排程是一旦處理元擁有中央處理器，則必須讓處理元執行完畢才允許分派別的處理元，有些則用不著這麼沒有彈性。
 - 不可奪取的 (Non Preemptive)
 - 可奪取的 (Preemptive)

Non-Preemptive and Preemptive Process(2)

- 當處理元生成之後，有些處理元優先等級非常的高，一旦它佔有中央處理器之後，就不允許別的處理元奪取中央處理器，我們稱此類處理元為不可奪取處理元（Non Preemptive Process）。
- 大部份處理元均依照排程策略，允許中央處理器被奪取，我們稱這類處理元為可奪取處理元（Preemptive Process）。

Non-preemptive and preemptive process(3)

- Non-preemptive process(不可奪取處理元)
 - After interrupt processing is complete, it gets CPU again.
 - 僅process結束或進入wait state可以讓別人使用。
- Preemptive process
 - After interrupt processing is complete, context switch to another process.
 - 除非ready queue內沒有其他process。

分派時間 (Dispatch Latency)

- 分派程式停止某一個處理元使用中央處理器，並分派中央處理器給另一個處理元所需的時間，稱為分派時間 (Dispatch Latency)。
- The time it takes for the dispatcher to stop one process and start another running.

中央處理器排程的選用標準(1)

5.2 Scheduling Criteria

- 儘量讓中央處理器不斷的忙碌，並有最高的中央處理器使用率（Maximum CPU Utilization）。
- 使得系統有最高的產量（Maximum Throughput）。
- 使得系統內最慢的反應時間是所有方法中最小的（Minimize the Maximum Response Time）。
- 使得往返時間是最小的（Minimize Turnaround Time）。
- 使得所有處理元之平均等待時間是最小的（Minimize Average Waiting Time）。此處的等待時間是指處理元在備妥佇列內的時間。

中央處理器排程的選用標準(2)

- 使得反應時間之變異數是最小的 (Minimize Variance in the Response Time) 。
- 使得交談使用者人數是最多的 (Maximum the Number of Interactive Users) 。
- 使得系統之負擔是最小的 (Minimize System Overhead) 。
- 使得資源被平衡的使用 (Balance Resources Used) 。
- 所有排程是公平的 (Be Fair) 。

先到先服務 (First Come First Served) 排程法(1)

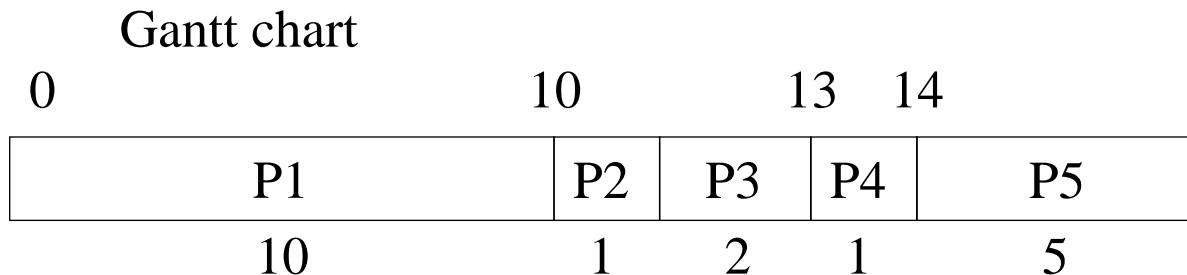
5.3 Scheduling Scheme

- 先到先服務 (First Come First Served) 排程法簡稱為FCFS。
 - 是一種不可奪取的排程法。
 - 不適合使用者進行交談，亦不適用於分時系統。

先到先服務 (First Come First Served) 排程法(2)

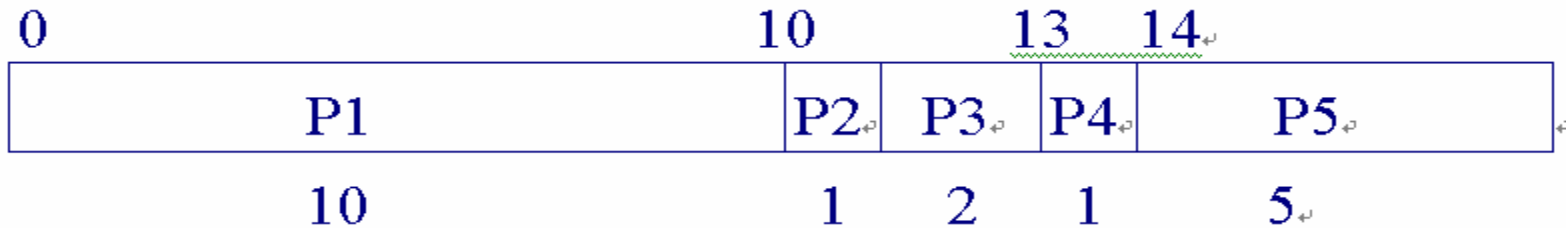
Process	CPU Burst Time
P1	10
P2	1
P3	2
P4	1
P5	5

圖 4.1 五個處理元使用中央處理器時間



先到先服務 (First Come First Served) 排程法(3)

Gantt Chart



Process	Turnaround Time	Waiting Time
P1	10	0
P2	11	10
P3	13	11
P4	14	13
P5	19	14

圖 4.2 先到先服務排程法之甘特圖、往返時間及等待時間

先到先服務 (First Come First Served) 排程法(4)

Process	CPU Burst Time
P1	10
P2	1
P3	2
P4	1
P5	5

圖 4.1 五個處理元使用中央處理器時間

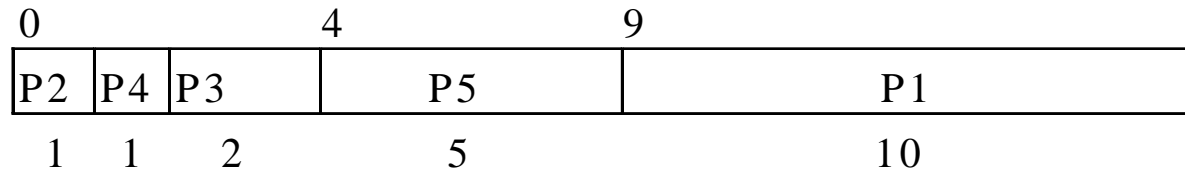


圖 4.3 最短工作優先排程法之甘特圖

先到先服務 (First Come First Served) 排程法(5)

Process	Turnaround Time	Waiting Time
P1	10	0
P2	11	10
P3	13	11
P4	14	13
P5	19	14

圖 4.2 先到先服務排程法之甘特圖、往返時間及等待時間

- 平均等待時間為9.6ms。

Process	Turnaround Time	Waiting Time
P1	19	9
P2	1	0
P3	4	2
P4	2	1
P5	9	4

圖 4.4 最短工作優先排程法之往返時間及等待時間

- 平均等待時間為3.2ms。

最短工作優先 (Shortest Job First) 排程法(1)

- 最短工作優先 (Shortest Job First) 排程法，簡稱為SJF。
- 依照使用中央處理器時間由小至大依序排列至備妥佇列內，以進行中央處理器排程。
 - 若處理元有相同中央處理器時間，則以FCFS排列。
- Smallest estimated run time to completion is run next.
- 可被區分為不可奪取最短工作優先排程法 (Non-Preemptive SJF) 及可奪取最短工作優先排程法 (Preemptive SJF)

最短工作優先 (Shortest Job First) 排程法(2)

- 不可奪取最短工作優先排程法是指，在此排程法下之處理元均為不可奪取處理元。
 - 不適合Time sharing
- 可奪取最短工作優先排程法則是，隨時注意是否有較小使用中央處理器時間之處理元出現，若有，則將中央處理器讓給此使用中央處理器時間較小的處理元。
 - 最短剩餘時間優先 (Shortest Remaining Time First) 排程法，簡稱為SRTF。
 - 適合Time sharing
- Running process may be preempted by a new process with a shortest estimated run time.

最短工作優先 (Shortest Job First) 排程法(3)

Process	CPU Burst Time
P1	10
P2	1
P3	2
P4	1
P5	5

圖 4.1 五個處理元使用中央處理器時間

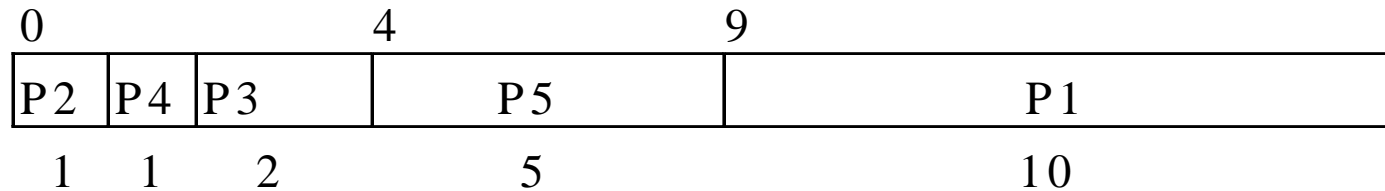


圖 4.3 最短工作優先排程法之甘特圖

最短工作優先 (Shortest Job First) 排程法(4)

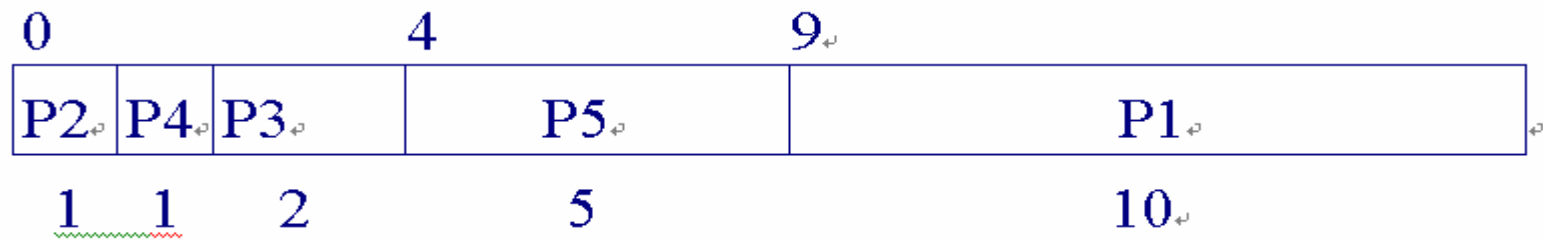


圖 4.3 最短工作優先排程法之甘特圖

Process	Turnaround Time	Waiting Time
P1	19	9
P2	1	0
P3	4	2
P4	2	1
P5	9	4

圖 4.4 最短工作優先排程法之往返時間及等待時間

最短工作優先 (Shortest Job First) 排程法(5)

- 最短工作優先排程法有最小的平均等待時間，所以若以等待時間來評估方法的好壞，最短工作優先排程法是一個最佳的演算法 (Optimal Algorithm)。
- It is difficult to implement, because no way to predict burst time.

最短工作優先 (Shortest Job First) 排程法(6)

假設 P_i 代表處理元 i 的使用中央處理器時間。

W_i 代表處理元 i 等待的時間。

依最短工作優先排程法 $P_0 < P_1 < P_2 < \dots < P_n$ 。

則 $W_0 = 0$

$$W_1 = P_0$$

$$W_2 = P_0 + P_1$$

$$W_3 = P_0 + P_1 + P_2$$

$$\vdots$$

$$\vdots$$

$$W_n = P_0 + P_1 + P_2 + \dots + P_{n-1}$$

將等式兩邊相加得到

$$W_0 + W_1 + W_2 + \dots + W_n = (0) + (P_0) + (P_0 + P_1) + (P_0 + P_1 + P_2) + \dots + (P_0 + P_1 + P_2 + \dots + P_{n-1})$$

所以

$$\sum_{i=0}^n W_i = \sum_{i=0}^{n-1} (n-i)P_i = \boxed{\sum_{i=0}^{n-1} nP_i} - \boxed{\sum_{i=0}^{n-1} iP_i}$$

項 A 項 B

最短工作優先 (Shortest Job First) 排程法(7)

Case1 :

假如我們依照使用中央處理器時間由小至大執行處理元，則執行次序為

$$P_0 < P_1 < P_2 < \dots < P_i < P_{i+1} < \dots < P_n$$

$$\sum_{i=0}^n W_i = \sum_{i=0}^{n-1} (n-i)P_i = \boxed{\sum_{i=0}^{n-1} nP_i} - \boxed{\sum_{i=0}^{n-1} iP_i}$$

Case 2 :

若我們任意對調二個處理元之執行次序，而依

$$P_0 < P_1 < P_2 < \dots < P_{i+1} > P_i < \dots < P_n \text{ 執行。}$$

固定

變動

不論 Case1 或 Case2 它們的項 A 均相同且固定不變。

比較變動項 B

Case1 之項 B 為 $P_1 + 2P_2 + 3P_3 + \dots + (iP_i + iP_{i+1} + P_{i+1}) + \dots + (n-1)P_{n-1}$

Case2 之項 B 為 $P_1 + 2P_2 + 3P_3 + \dots + (iP_{i+1} + iP_i + P_i) + \dots + (n-1)P_{n-1}$

由以上證明得知，若我們隨意改變處理元次序，而不依使用中央處理器時間由小至大執行處理元，則所得到之平均等待時間均較依規定執行者大，故最短工作優先排程法有最小的平均等待時間。

最短工作優先 (Shortest Job First) 排程法(8)

Process	Arrival Time	CPU Burst
P1	0	4
P2	1	2
P3	2	4

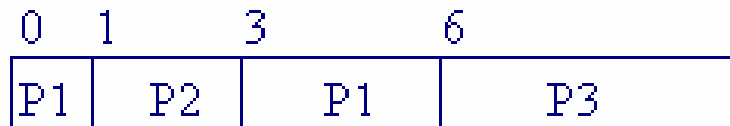


圖 4.5 可奪取最短工作優先排程之甘特圖

Process	Turnaround Time	Waiting Time
P1	6	2 (6 - 4 = 2)
P2	2	0 (2 - 2 = 0)
P3	8	4 (8 - 4 = 4)

圖 4.6 可奪取最短工作優先排程之往返時間及等待時間

最高反應時間比率優先排程法

- 反應時間比率（Response Ratio）與等待時間及中央處理器時間有關，它的公式為

$$\text{Response Ratio} = \frac{\text{Waiting Time} + \text{CPU Burst Time}}{\text{CPU Burst Time}}$$

- 計算出反應時間比率愈高的處理元優先處理。
- 以反應時間比率當作變動優先等級來排程。

優先等級(Priority)排程法(1)

- 先等級排程法是給予每一個處理元優先等級，並依照優先等級將處理元由高優先等級排至低優先等級，然後依序分派處理元擁有中央處理器。
- CPU is allocated to the process with the **highest priority**. Equal priority processes are scheduled in FCFS order.
 - 當處理元有相同的優先等級時，則採用先到先服務方式處理。
 - 若有一個較高優先等級的處理元後來才生成，它必須排在比它優先等級低的處理元之前，以進行排程。
 - 可分為可奪取優先等級排程及不可奪取優先等級排程

優先等級(Priority)排程法(2)

- 當處理元之優先等級很低，而系統內不斷有較高優先等級之處理元出現，則低優先等級之處理元可能造成永遠無法被執行的問題，我們稱此低優先等級處理元無限懸置 (Indefinite Blocking) 或餓死 (Starvation)。
- The major problem is Indefinite Blocking or Starvation for low priority process.
 - 利用時間升級 (Aging) 機制來解決。
 - 變動優先等級 (Dynamic Priority) 或浮動優先等級 (Floating Priority)。
- Aging: Gradually increasing the priority of processes that wait in the system for a long time.

優先等級(Priority)排程法(3)

Process	CPU Burst	Priority
P1	10	3
P2	1	1
P3	2	3
P4	1	4
P5	5	2



圖 4.7 優先等級排程法

0	1	6	16	18
P2	P5	P1	P3	P4

Process	Turnaround Time	Waiting Time
P1	16	6
P2	1	0
P3	18	16
P4	19	18
P5	6	1

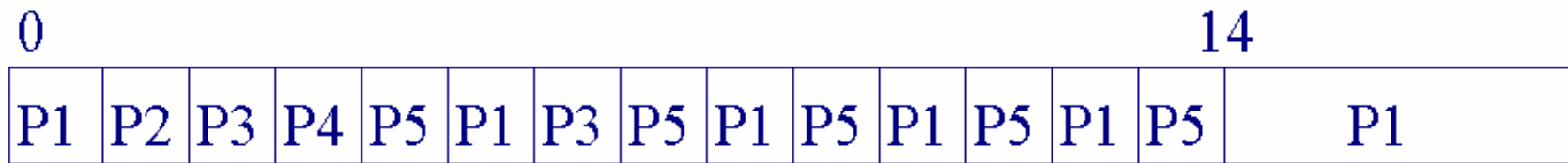
圖 4.8 優先等級排程法之甘特圖，往返時間、及等待時間

知更鳥式循環(Round Robin)排程法(1)

- 知更鳥式循環排程法就是分時系統所使用之排程法。
- Processes are dispatched FCFS but are given a **limited amount** of CPU time. (**time slice**)
- **Preemptive, 適合time sharing** 。
- 時間片段的選擇
 - 時間片段太大  先到先服務排程法
 - 時間片段太小  系統的效率太差
 - processor sharing: each of n processes has its own processor running are $1/n$ the speed of the real processor.

不能太小; context switch 時間被佔滿, process 根本無法 run
(或 run 極短時間)

知更鳥式循環(Round Robin)排程法(2)



Process	Turnaround Time	Waiting Time
P1	19	9
P2	2	1
P3	7	5
P4	4	3
P5	14	9

圖 4.9 知更鳥式循環排程法之甘特圖、往返時間、及等待時間

多階層佇列(Multilevel Queue)排程法(1)

- Partitions the ready queue into several separate queues.
- 當電腦系統內有許多不同優先等級編號之處理元要執行，且亦有許多相同優先等級編號之處理元要執行，則很顯然的最高優先等級之處理元會先被執行，且若同時有多個最高優先等級之處理元在系統內，則它們會優先在佇列內排隊等待被執行。

多階層佇列(Multilevel Queue)排程法(2)

- Preemptive for different priority.
- Fixed priority for each queue.

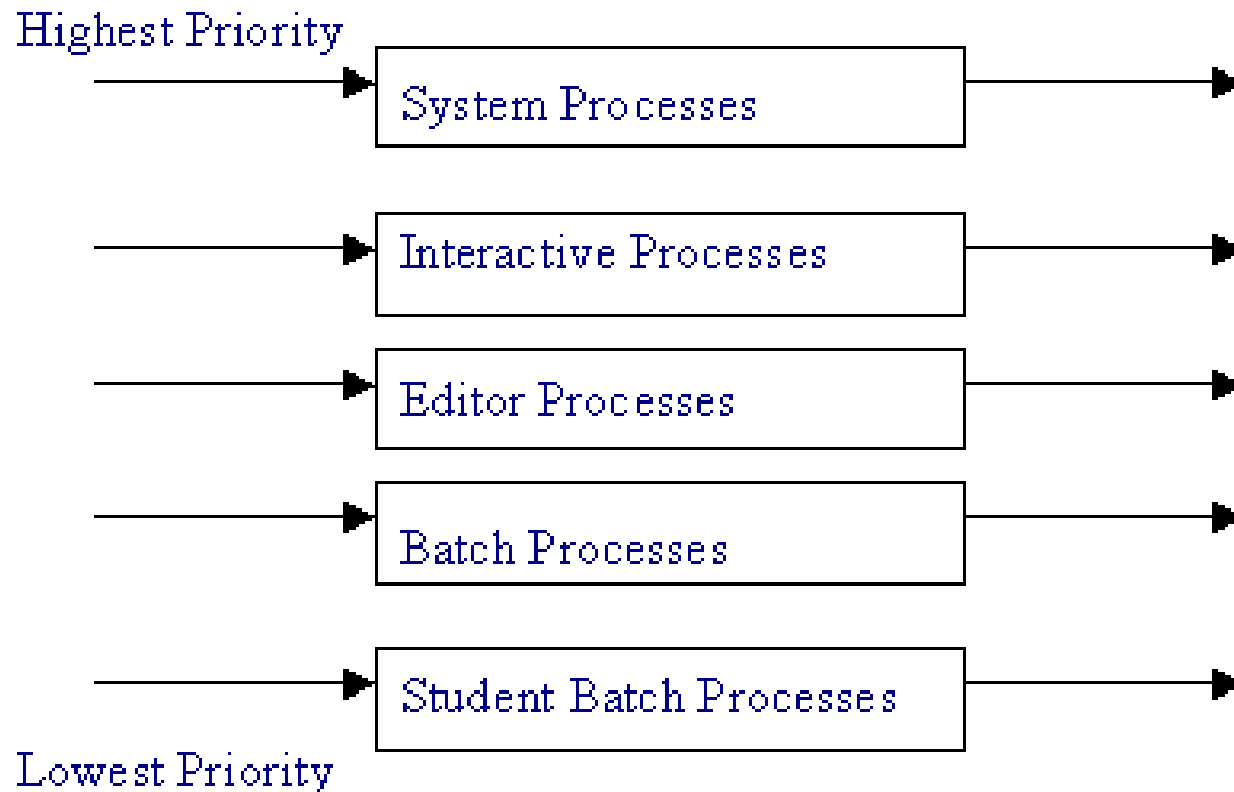
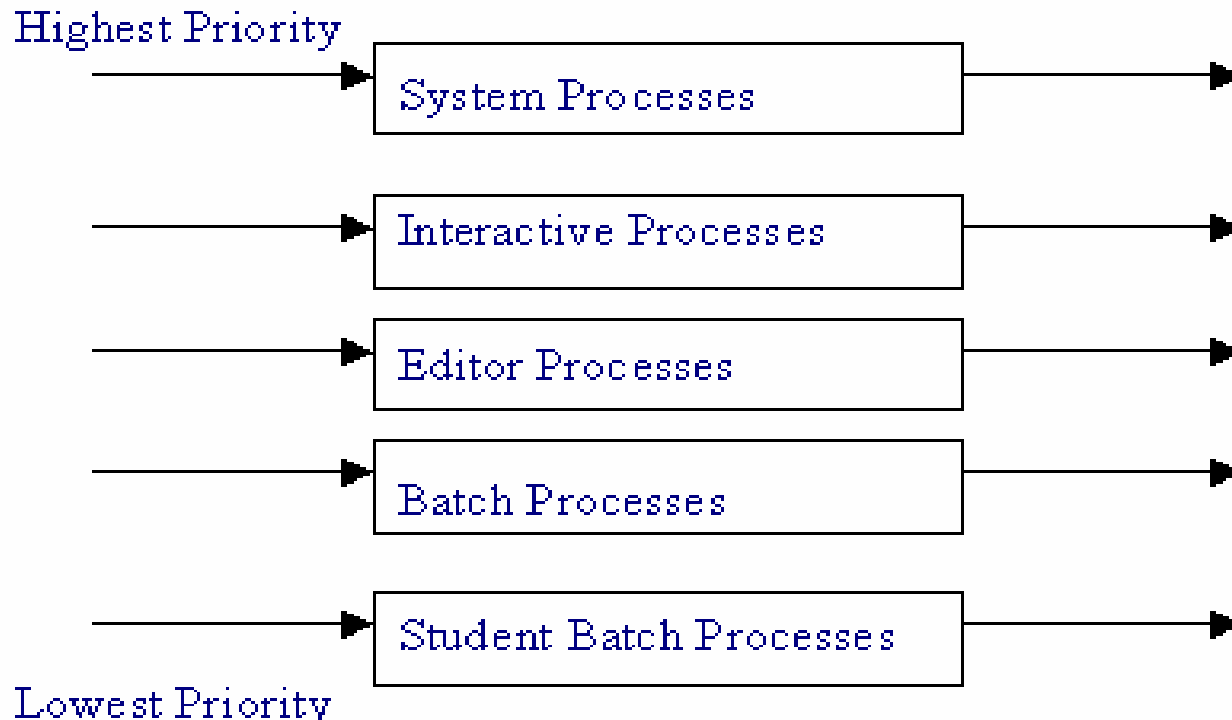


圖 4.10 多階層佇列排程法

多階層佇列(Multilevel Queue)排程法(3)

- Foreground process
 - for interactive, high priority, RR scheduling.
- Background process
 - for batch, lower priority, FCFS.



多階層佇列(Multilevel Queue)排程法(4)

- 由於在多階層佇列排程法內的處理元一旦生成之後，就永遠固定優先等級，以致有可能造成某些低優先等級之處理元餓死的情形，因此在實務上並不太適用。

多階層回饋佇列(Multilevel Feedback Queue)排程法(1)

- 此排程法中允許處理元依時間升級機制，**提升至較高優先等級或降低至較低等級佇列中**。
- The multilevel feedback queue scheduling allows a process to **move between queues**.
 - 原則上屬於系統之處理元可以固定於高優先等級，而批次作業處理元可以固定於低優先等級，且它們不必使用時間升級機制；其餘處理元則依時間升級機制游走於各個佇列中。

多階層回饋佇列(Multilevel Feedback Queue)排程法(2)

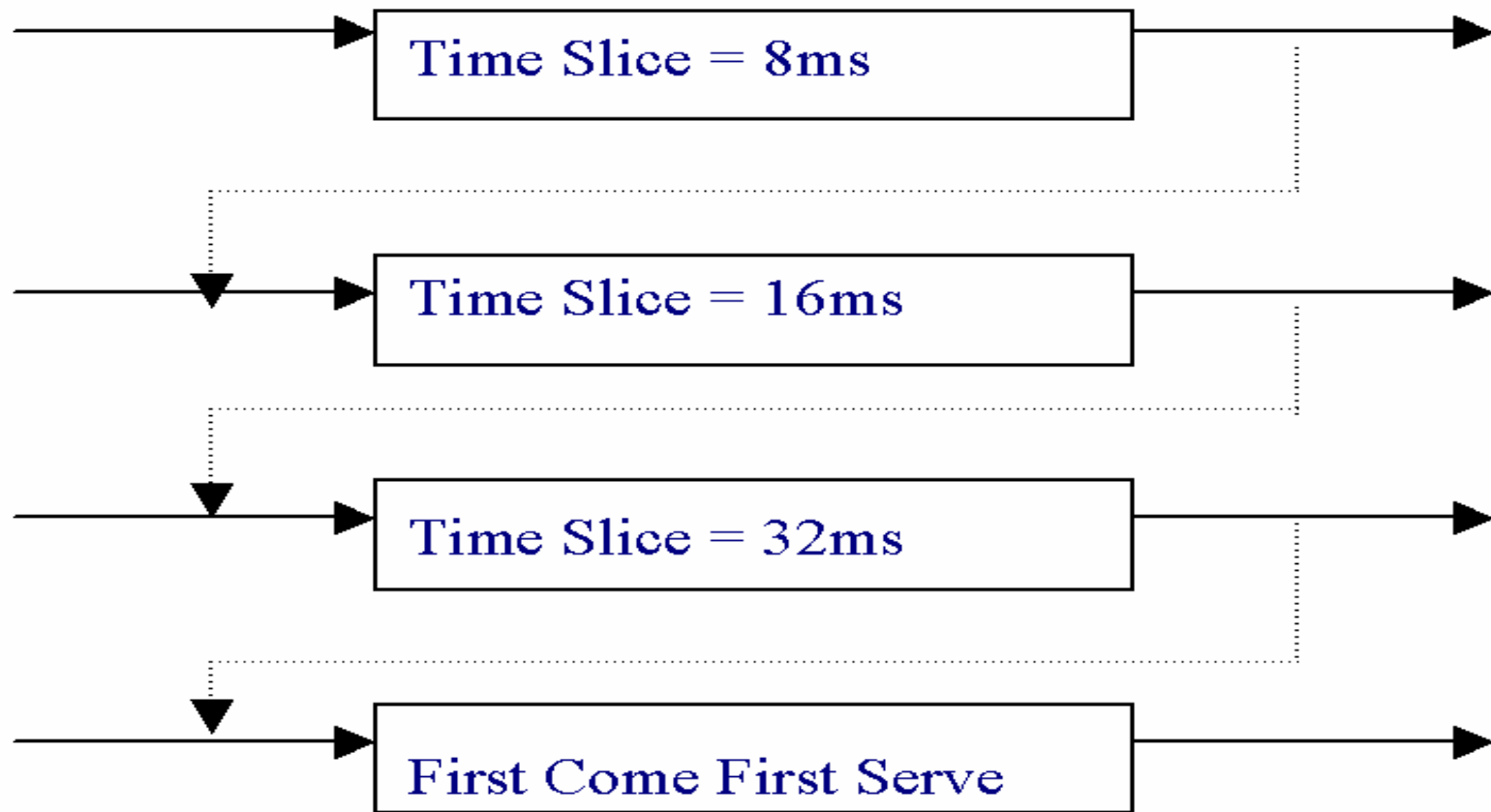


圖 4.11 多階層回饋佇列排程法

多階層回饋佇列(Multilevel Feedback Queue)排程法(3)

- 多階層回饋佇列排程法的特性
 - 是可奪取排程
 - 處理元不會有餓死的情形發生
 - 中央處理器限制處理元 (CPU Bound Process) (use too much CPU time process) 會隨著時間而降低優先等級, small time slice.
 - 輸出 / 輸入限制處理元 (I/O Bound Process) (interactive, wait too long process) 則隨著時間提升優先等級。Large time slice.

優先等級(Priority)

- 先到先服務排程法是以處理元進入備妥佇列的時間先後次序，當作優先等級。
- 不可奪取最短工作優先排程法是以預估處理元使用中央處理器時間大小，當作優先等級。
- 可奪取最短工作優先排程法是以處理元剩餘使用中央處理器時間大小，當作優先等級。
- 最高反應時間比率優先排程法是依其反應時間比率，當作優先等級。
- 知更鳥式循環排程法則是以處理元在備妥佇列的次序，當作優先等級。

多處理器 (Multiple Processor)

- Homogenous Processor
 - Tightly Coupled System
- Heterogeneous Processor
 - Loosely Coupled system

Real Time

- Preemptive
 - hard real time v.s. soft real time.

定量模式 (Deterministic Modeling)

5.4 系統效能評估

- 定量模式 (Deterministic Modeling)
 - 所謂定量模式評估法就是，事先設定好負載量的標準，並就現有環境提供確實的數據資料，然後利用這些數據資料評量各個演算法的效能，而得到對演算法效能的評比。
 - Taken a particular **predetermined workload** and defines the performance of each algorithm for that workload.
 - 不一定可以適用在各個系統內，但對特定環境內的特定行為大都可以反應出來。

佇列模式 (Queuing Modeling)

- 佇列模式 (Queuing Modeling)
 - There is no static set of processes (and time) to use for deterministic modeling.
 - 使用佇列理論 (Queuing Theory)
 - n 代表在佇列內排隊的平均長度
 - ω 代表每個處理元在佇列內的平均等待時間
 - λ 代表新的處理元平均出生率
 - 用 Little's formula 來解析評估
 - Little's formula 為 $n = \lambda * \omega$

模擬法 (Simulations)

- 模擬法 (Simulations)
 - 所謂模擬法就是撰寫模擬程式，並利用例如亂數產生器 (Random Number Generator) 產生所需資料，然後將此模擬資料送入模擬程式，以進行效能評估。