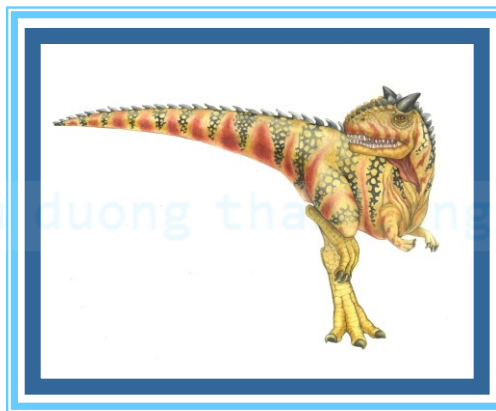


Chương 4: Định thời CPU - 1





Câu hỏi ôn tập

- Process control block (PCB) chứa những thông tin gì?
- Các tác vụ đối với tiến trình?
- Tại sao phải định thời, có mấy loại bộ định thời?

cuuduongthancong.com

cuuduongthancong.com





Mục tiêu

- Biết được các khái niệm cơ bản về định thời
- Biết được các tiêu chuẩn định thời CPU
- Hiểu được các giải thuật định thời
- Vận dụng các giải thuật định thời để làm bài tập và mô phỏng

[cuu duong than cong . com](http://cuuduongthancong.com)





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- Các giải thuật định thời
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling





Nội dung

- **Các khái niệm cơ bản về định thời**
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- Các giải thuật định thời
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling





Khái niệm cơ bản

- Trong các hệ thống multitasking
 - Thực thi nhiều chương trình đồng thời làm tăng hiệu suất hệ thống
 - Tại mỗi thời điểm, trong một hệ thống đơn bộ xử lý (single-processor system) chỉ có một process được thực thi; những process khác phải chờ cho đến khi bộ xử lý rảnh hoặc tái định thời lại.
 - Cần phải giải quyết vấn đề phân chia, lựa chọn process thực thi sao cho được hiệu quả nhất
 - Phải có chiến lược định thời CPU, tức như thế nào để đưa một process vào thực thi.





Nội dung

- Các khái niệm cơ bản về định thời
- **Các bộ định thời**
- Các tiêu chuẩn định thời CPU
- Các giải thuật định thời
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling





Các bộ định thời (tt)

- Long-term scheduling
- Medium-term scheduling

(Xem lại slide chương 3)

cuu duong than cong . com

cuu duong than cong . com





Các bộ định thời (tt)

■ Short-term scheduling

- Xác định process nào trong ready queue sẽ được chiếm CPU để thực thi kế tiếp (còn được gọi là định thời CPU, CPU scheduling)
- Short term scheduler còn được gọi với tên khác là dispatcher
- Bộ định thời short-term có thể được gọi khi một process:
 - (1) Chuyển từ trạng thái running tới waiting
 - (2) Chuyển từ trạng thái running tới ready
 - (3) Chuyển từ waiting tới ready
 - (4) Kết thúc





Dispatcher

- Dispatcher sẽ chuyển quyền điều khiển CPU về cho process được chọn bởi bộ định thời ngắn hạn
- Bao gồm:
 - Chuyển ngữ cảnh (sử dụng thông tin ngữ cảnh trong PCB)
 - Chuyển chế độ người dùng
 - Nhảy đến vị trí thích hợp trong chương trình ứng dụng để khởi động lại chương trình (chính là program counter trong PCB)
- Công việc này gây ra phí tổn
 - Dispatch latency: thời gian mà dispatcher dừng một process và khởi động một process khác





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- **Các tiêu chuẩn định thời CPU**
- Các giải thuật định thời
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling





Các tiêu chuẩn định thời CPU

UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

■ Hướng người dùng (User-oriented)

- **Thời gian đáp ứng** (Response time): khoảng thời gian process nhận yêu cầu đến khi yêu cầu đầu tiên được đáp ứng (time-sharing, interactive system) → cực tiểu
- **Thời gian quay vòng** (hoàn thành) (Turnaround time): khoảng thời gian từ lúc một process được nạp vào hệ thống đến khi process đó kết thúc → cực tiểu
- **Thời gian chờ** (Waiting time): tổng thời gian một process đợi trong ready queue → cực tiểu





Các tiêu chuẩn định thời CPU (tt)

■ Hướng hệ thống (System-oriented)

- **Sử dụng CPU** (processor utilization): định thời sao cho CPU càng bận càng tốt → cực đại
- **Công bằng** (fairness): tất cả process phải được đối xử như nhau [duong than cong . com](http://duongthancong.com)
- **Thông lượng** (throughput): số process hoàn tất công việc trong một đơn vị thời gian → cực đại

[cuu duong than cong . com](http://cuuduongthancong.com)





Hai yếu tố của giải thuật định thời

- Hàm chọn lựa (selection function): dùng để chọn process nào trong ready queue được thực thi (thường dựa trên độ ưu tiên, yêu cầu về tài nguyên, đặc điểm thực thi của process,...)
- Quy ước:
 - w = tổng thời gian đợi trong hệ thống
 - e = thời gian đã được phục vụ
 - s = tổng thời gian thực thi của process (bao gồm cả “e”)





Hai yếu tố của giải thuật định thời (tt)

- Chế độ quyết định (decision mode): chọn thời điểm thực hiện hàm chọn lựa để định thời
- Có hai chế độ:
 - Không trưng dụng (Non-preemptive)
 - ▶ Khi ở trạng thái running, process sẽ thực thi cho đến khi kết thúc hoặc bị blocked do yêu cầu I/O
 - Trưng dụng (Preemptive)
 - ▶ Process đang thực thi (trạng thái running) có thể bị ngắt nửa chừng và chuyển về trạng thái ready
 - ▶ Chi phí cao hơn non-preemptive nhưng đánh đổi lại bằng thời gian đáp ứng tốt hơn vì không có trường hợp một process độc chiếm CPU quá lâu





Hai yếu tố của giải thuật định thời (tt)

- Định thời CPU được thực hiện khi xảy ra một trong 4 trường hợp sau:

(1) Khi một process chuyển từ trạng thái running sang waiting

(2) Khi một process chuyển từ trạng thái running sang ready

(3) Khi một process chuyển từ trạng thái waiting sang ready

(4) Khi một process kết thúc

- Trường hợp (1) và (4) là định thời nonpreemptive
- Trường hợp (2) và (3) có thể có lựa chọn, do đó là định thời preemptive





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- **Các giải thuật định thời**
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling



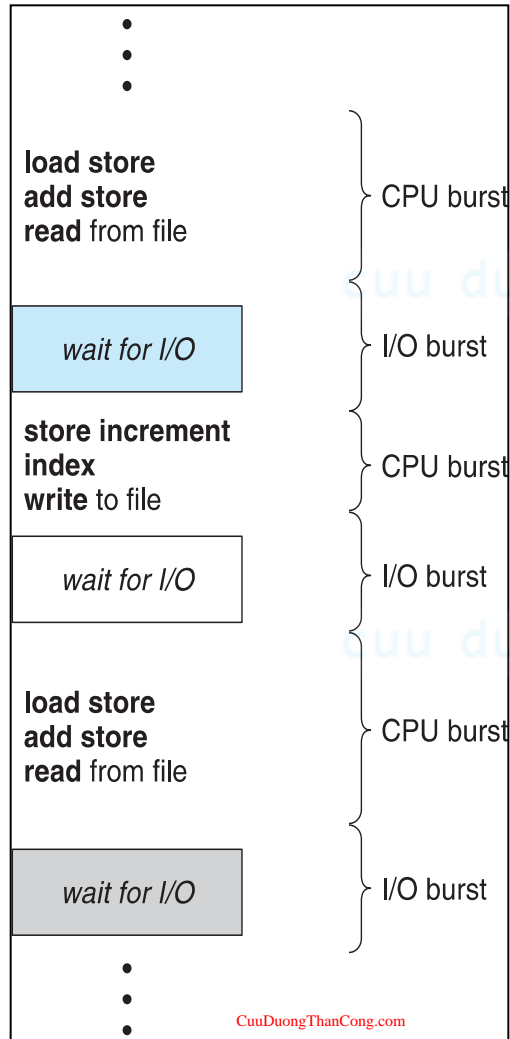


Khảo sát giải thuật định thời

UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

Service time = thời gian process cần CPU trong một chu kỳ CPU-I/O

- Process có service time lớn là các **CPU-bound process**





Các giải thuật định thời

- First-Come, First-Served (FCFS)
- Shortest Job First (SJF)
- Shortest Remaining Time First (SRTF)
- Priority Scheduling
- Round-Robin (RR)
- Highest Response Ratio Next (HRRN)
- Multilevel Queue
- Multilevel Feedback Queue





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- **Các giải thuật định thời**
 - **First-Come, First-Served (FCFS)**
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling





First-Come, First-Served (FCFS)

■ Hàm lựa chọn:

- Tiến trình nào yêu cầu CPU trước sẽ được cấp phát CPU trước
- Process sẽ thực thi đến khi kết thúc hoặc bị blocked do I/O

■ Chế độ quyết định: non-preemptive algorithm

■ Hiện thực: sử dụng hàng đợi FIFO (FIFO queues)

- Tiến trình đi vào được thêm vào cuối hàng đợi
- Tiến trình được lựa chọn để xử lý được lấy từ đầu của queues



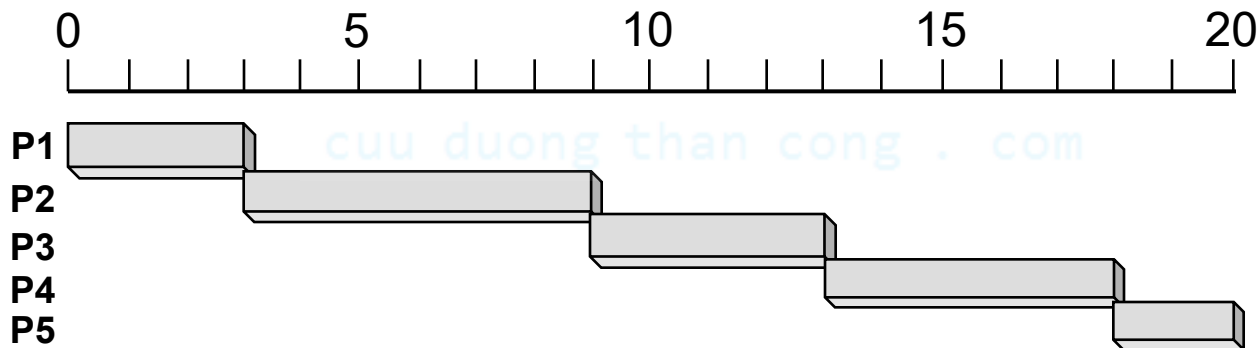


First-Come, First-Served (FCFS)

UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

Ví dụ FCFS:

Process	Arrival Time	Service Time
1	0	3
2	2	6
3	4	4
4	6	5
5	8	2





First-Come, First-Served (FCFS) (tt)

■ Ví dụ :

Process	Burst Time
P1	24
P2	3
P3	3

Gantt Chart



■ Giả sử thứ tự vào của các tiến trình là

- P1, P2, P3

■ Thời gian chờ

- P1 = 0
- P2 = 24
- P3 = 27

■ Thời gian chờ trung bình?

- $(0+24+27)/3 = 17$





First-Come, First-Served (FCFS) (tt)

■ Ví dụ :

Process	Burst Time
P1	24
P2	3
P3	3

Gantt Chart



■ Giả sử thứ tự vào của các tiến trình là

- P2, P3, P1

■ Thời gian chờ

- P1 = 6

- P2 = 0

- P3 = 3

■ Thời gian chờ trung bình?

- $(6+0+3)/3 = 3$

➔ tốt hơn





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- **Các giải thuật định thời**
 - First-Come, First-Served (FCFS)
 - **Shortest Job First (SJF)**
 - **Shortest Remaining Time First (SRTF)**
 - Priority Scheduling





Shortest-Job-First (SJF)

UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

- Định thời biểu công việc ngắn nhất trước
- Khi CPU được tự do, nó sẽ cấp phát cho tiến trình yêu cầu ít thời gian nhất để kết thúc (tiến trình ngắn nhất)
- Liên quan đến chiều dài thời gian sử dụng CPU cho lần tiếp theo của mỗi tiến trình
- Sử dụng những chiều dài này để lập lịch cho tiến trình với thời gian ngắn nhất





Shortest-Job-First (SJF) (tt)

■ Cách 1: Non-preemptive

- Khi CPU được trao cho quá trình nó không nhường cho đến khi nó kết thúc chu kỳ xử lý của nó

■ Cách 2: Preemptive

- Nếu một tiến trình mới được đưa vào danh sách với chiều dài sử dụng CPU cho lần tiếp theo nhỏ hơn thời gian còn lại của tiến trình đang xử lý, nó sẽ dừng hoạt động tiến trình hiện hành → Shortest-Remaining-Time-First (SRTF)

- SJF là tối ưu – cho thời gian chờ đợi trung bình tối thiểu với một tập tiến trình cho trước



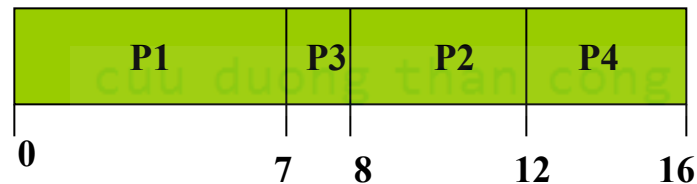


Non-Preemptive SJF

■ Ví dụ:

Process	Arrival Time	Burst Time
P1	0	7
P2	2	4
P3	4	1
P4	5	4

Gantt Chart



$$\text{Average waiting time} = (0 + 6 + 3 + 7)/4 = 4$$





Preemptive SJF (SRTF)

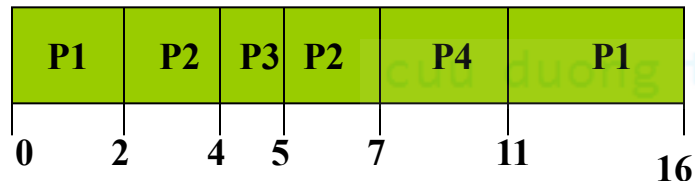
UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

■ Ví dụ 1:

Process	Arrival Time	Burst Time
P1	0	7
P2	2	4
P3	4	1
P4	5	4

■ Ví dụ 2:

Gantt Chart



$$\text{Average waiting time} = (9 + 1 + 0 + 2)/4 = 3$$

Process	Arrival Time	Burst Time
P1	0	8
P2	1	4
P3	2	9
P4	3	5





Nhận xét về giải thuật SJF

- Có thể xảy ra tình trạng “đói” (starvation) đối với các process có CPU-burst lớn khi có nhiều process với CPU-burst nhỏ đến hệ thống.
- Cơ chế non-preemptive không phù hợp cho hệ thống time sharing (interactive)
- Giải thuật SJF ngầm định ra độ ưu tiên theo burst time
- Các CPU-bound process có độ ưu tiên thấp hơn so với I/O-bound process, nhưng khi một process không thực hiện I/O được thực thi thì nó độc chiếm CPU cho đến khi kết thúc





Nhận xét về giải thuật SJF (tt)

- Tương ứng với mỗi process cần có độ dài của CPU burst tiếp theo
- Hàm lựa chọn: chọn process có độ dài CPU burst nhỏ nhất
- Chứng minh được: SJF tối ưu trong việc giảm thời gian đợi trung bình
- Nhược điểm: Cần phải ước lượng thời gian cần CPU tiếp theo của process
- Giải pháp cho vấn đề này?





Nhận xét về giải thuật SJF (tt)

UIT
TRƯỜNG ĐẠI HỌC
CÔNG NGHỆ THÔNG TIN

- (Thời gian sử dụng CPU chính là độ dài của CPU burst)

Như thế nào để ước lượng CPU burst tiếp theo?

- Trung bình tất cả các CPU burst đo được trong quá khứ
 - Nhưng thông thường những CPU burst càng mới càng phản ánh đúng hành vi của process trong tương lai
- Một kỹ thuật thường dùng là sử dụng trung bình hàm mũ (exponential averaging)

$$\tau_{n+1} = \alpha t_n + (1 - \alpha)\tau_n \quad 0 \leq \alpha \leq 1$$

Với τ_{n+1} : là giá trị ước lượng của độ dài của CPU burst thứ $n + 1$

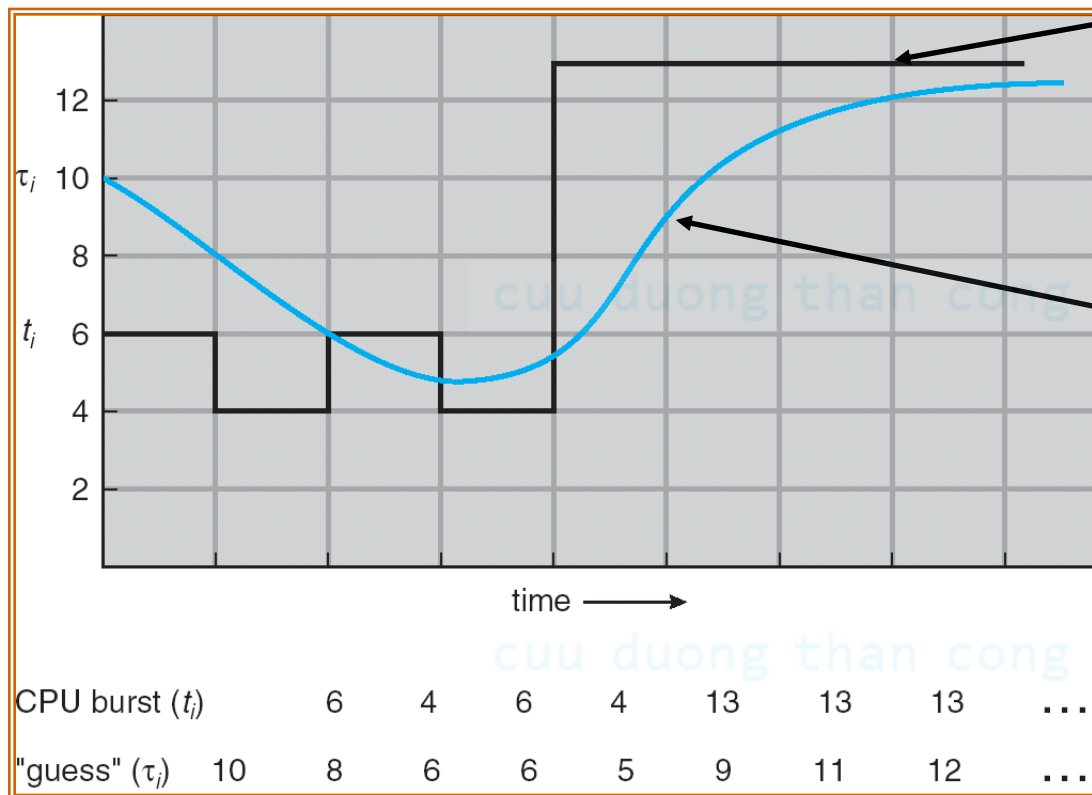
t_n : là giá trị thật sự của độ dài của CPU burst thứ n

Nếu chọn $\alpha = 1/2$ thì có nghĩa là trị đo được t_n và trị dự đoán τ_n được xem quan trọng như nhau.





Dự đoán thời gian sử dụng CPU



Độ dài CPU burst
đo được

Độ dài CPU burst
dự đoán, với
 $a = \frac{1}{2}$ và $\tau_0 = 10$





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- **Các giải thuật định thời**
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - **Priority Scheduling**





Priority Scheduling

- Mỗi process sẽ được gán một độ ưu tiên
- CPU sẽ được cấp cho process có độ ưu tiên cao nhất
- Định thời sử dụng độ ưu tiên có thể:
 - Preemptive hoặc
 - Nonpreemptive





Priority Scheduling (tt)

- SJF là một giải thuật định thời sử dụng độ ưu tiên với độ ưu tiên là thời-gian-sử-dụng-CPU-dự-đoán
- Gán độ ưu tiên còn dựa vào:
 - Yêu cầu về bộ nhớ
 - Số lượng file được mở
 - Tỷ lệ thời gian dùng cho I/O trên thời gian sử dụng CPU
 - Các yêu cầu bên ngoài ví dụ như: số tiền người dùng trả khi thực thi công việc





Priority Scheduling (tt)

- Vấn đề: trì hoãn vô hạn định – process có độ ưu tiên thấp có thể không bao giờ được thực thi
- Giải pháp: làm mới (aging) – độ ưu tiên của process sẽ tăng theo thời gian

cuu duong than cong . com





Priority Scheduling (tt)

Ví dụ:

<u>Process</u>	<u>Burst Time</u>	<u>Priority</u>
P_1	10	3
P_2	1	1
P_3	2	4
P_4	1	5
P_5	5	2

Gantt chart



$$\text{Average waiting time} = (6 + 0 + 16 + 18 + 1)/5 = 8.2$$





Nội dung

- Các khái niệm cơ bản về định thời
- Các bộ định thời
- Các tiêu chuẩn định thời CPU
- Các giải thuật định thời
 - First-Come, First-Served (FCFS)
 - Shortest Job First (SJF)
 - Shortest Remaining Time First (SRTF)
 - Priority Scheduling





Câu hỏi ôn tập

- Có bao nhiêu bộ định thời? Kể tên?
- Nêu các tiêu chí định thời?
- Nêu ưu điểm và nhược điểm của các giải thuật định thời FCFS, SJF, SRTF, Priority?

cuu duong than cong . com





Bài tập

- Sử dụng các giải thuật FCFS, SJF, SRTF, Priority để tính các giá trị thời gian đợi, thời gian đáp ứng và thời gian hoàn thành trung bình

<u>Process</u>	<u>Burst Time</u>	<u>Priority</u>
P_1	2	2
P_2	1	1
P_3	8	4
P_4	4	2
P_5	5	3

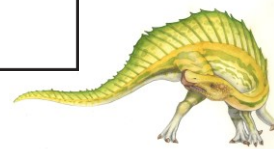




Bài tập (tt)

- Sử dụng các giải thuật FCFS, SJF, SRTF, Priority để tính các giá trị thời gian đợi, thời gian đáp ứng và thời gian hoàn thành trung bình

Thread	Priority	Burst	Arrival
P_1	40	20	0
P_2	30	25	25
P_3	30	25	30
P_4	35	15	60
P_5	5	10	100
P_6	10	10	105



Kết thúc chương 4-1

