Operating Systems

Lecture # 9

Department of Computer

4th Class

Process Schedulers

By

Dr. Ahmed Khudhair Abbas

Computer and Internet Center

جدولة العملية Process scheduling هي نشاط مدير العملية الذي يتولى إزالة العملية الجارية من وحدة المعالجة المركزية واختيار عملية أخرى على أساس إستراتيجية معينة.

وتعد جدولة العمليات جزءًا أساسيًا من أنظمة تشغيل البرمجة المتعددة Multiprogramming operating systems وتسمح أنظمة التشغيل هذه بتحميل أكثر من عملية واحدة في الذاكرة القابلة للتنفيذ في وقت واحد ، وتشارك العملية المحملة وحدة المعالجة المركزية باستخدام مضاعفة الوقت.

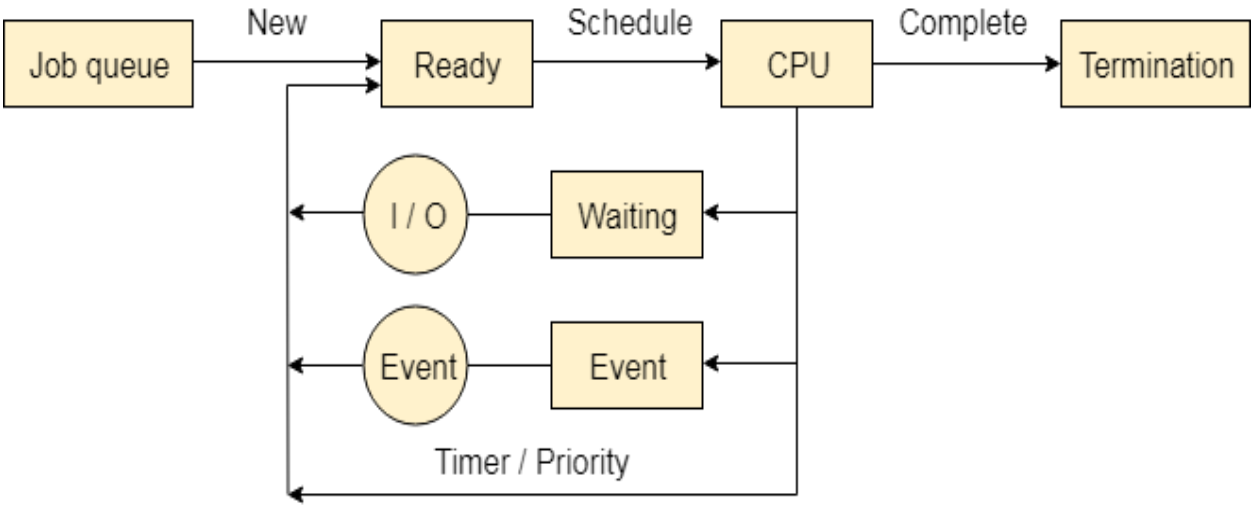
فئات الجدولة Categories of Scheduling

هناك فئتان من الجدولة:

1. غير استباقي **Non-preemptive** : هنا لا يمكن أخذ المورد resource من العملية حتى تكتمل العملية التنفيذ. ويحدث تبديل الموارد عندما تنتهي العملية الجارية وتنتقل إلى حالة الانتظار Waiting
2. استباقية **Preemptive :** هنا يخصص نظام التشغيل الموارد لعملية لفترة محددة من الوقت. أثناء تخصيص الموارد ، تنتقل العملية من حالة التشغيل Running إلى حالة الاستعداد Ready أو من حالة الانتظارWaiting إلى حالة الاستعداد Ready. يحدث هذا التبديل لأن وحدة المعالجة المركزية قد تعطي الأولوية للعمليات الأخرى وتستبدل العملية بأولوية أعلى بالعملية الجارية.

قوائم انتظار المعالجة Process Queues

يدير نظام التشغيل أنواعًا مختلفة من قوائم الانتظار لكل حالة من حالات العملية. يتم أيضًا تخزين PCB المتعلق بالعملية في قائمة الانتظار من نفس الحالة. إذا تم نقل العملية من حالة إلى حالة أخرى، فسيتم أيضًا إلغاء ربط PCB الخاص بها من قائمة الانتظار المقابلة وإضافتها إلى قائمة انتظار الحالة الأخرى التي يتم فيها الانتقال.



وهذه هي قوائم الانتظار التي يحتفظ بها نظام التشغيل.

1. طابور المهام Job Queue

في البداية، يتم تخزين جميع العمليات في قائمة انتظار المهام ويتم الاحتفاظ بها في الذاكرة الثانوية ثم يقوم برنامج الجدولة طويلة المدى (برنامج جدولة المهام) باختيار بعض المهام ووضعها في الذاكرة الأساسية.

2. طابور الجاهزية Ready Queue

يتم الاحتفاظ بقائمة الانتظار الجاهزة في الذاكرة الأساسية. يقوم برنامج الجدولة قصير المدى باختيار المهمة من قائمة الانتظار الجاهزة وإرسالها إلى وحدة المعالجة المركزية للتنفيذ.

3. طابور الانتظار Waiting Queue

عندما تحتاج العملية إلى بعض عمليات الإدخال / الإخراج لإكمال تنفيذها، يقوم نظام التشغيل بتغيير حالة العملية من التشغيل إلى الانتظار. ثم يتم تخزين ال PCB المرتبط بالعملية في قائمة الانتظار التي سيستخدمها المعالج عندما تنتهي العملية من الإدخال / الإخراج.

S.N.	State & Description
1	**Running:** When a new process is created, it enters into the system as in the running state.
2	**Not Running:** Processes that are not running are kept in queue, waiting for their turn to execute. Each entry in the queue is a pointer to a particular process. Queue is implemented by using linked list. Use of dispatcher is as follows. When a process is interrupted, that process is transferred in the waiting queue. If the process has completed or aborted, the process is discarded. In either case, the dispatcher then selects a process from the queue to execute. يتم الاحتفاظ بالعمليات التي لا تعمل في قائمة الانتظار، في انتظار دورها للتنفيذ. كل إدخال في قائمة الانتظار هو مؤشر لعملية معينة. يتم تطبيق قائمة الانتظار باستخدام القائمة المرتبطة. عند مقاطعة عملية ما ، يتم نقل هذه العملية في قائمة الانتظار. إذا اكتملت العملية أو تم إحباطها، يتم تجاهل العملية. في كلتا الحالتين، يقوم المرسل بعد ذلك بتحديد عملية من قائمة الانتظار ليتم تنفيذها.

المجدولات Schedulers

المجدولات هي برنامج نظام خاص يتعامل مع جدولة العمليات بطرق مختلفة. مهمتهم الرئيسية هي تحديد الوظائف التي سيتم تقديمها في النظام وتحديد العملية التي سيتم تشغيلها. المنظمون من ثلاثة أنواع-

1. جدولة طويلة المدى Long-Term Scheduler
2. جدولة قصيرة المدى Short-Term Scheduler
3. جدولة متوسطة المدى Medium-Term Scheduler

جدولة طويلة الامد Long-Term Scheduler

ويسمى أيضًا بجدولة المهام ويقوم المجدول طويل الأجل بتحديد البرامج التي يتم قبولها في النظام للمعالجة. وكذلك يقوم بتحديد العمليات من قائمة الانتظار وتحميلها في الذاكرة للتنفيذ. ثم يتم تحميل العملية في الذاكرة لجدولة وحدة المعالجة المركزيةCPU.

الهدف الأساسي لجدولة الوظائف job scheduler هو توفير مزيج متوازن من المهام، مثل ربط الإدخال / الإخراج والمعالج. كما يتحكم في درجة البرمجة المتعددة multiprogramming إذا كانت درجة البرمجة المتعددة مستقرة، فيجب أن يكون متوسط معدل إنشاء العملية مساويًا لمتوسط معدل المغادرة للعمليات التي تغادر النظام.

في بعض الأنظمة، قد لا يكون المجدول طويل المدى متاحًا أو ضئيلًا كما لا تحتوي أنظمة تشغيل مشاركة الوقت على برنامج جدولة طويل المدى وعندما تتغير حالة العملية من جديد إلى جاهز، يكون هناك استخدام جدولة طويلة المدى.

جدولة قصيرة المدى Short-Term Scheduler

ويسمى أيضًا باسم جدولة وحدة المعالجة المركزية CPU scheduler هدفها الرئيسي هو زيادة أداء النظام وفقًا لمجموعة المعايير المختارة. إنه تغيير من حالة الاستعداد ready state إلى حالة تشغيل العملية running state يحدد برنامج جدولة وحدة المعالجة المركزية عملية من بين العمليات الجاهزة للتنفيذ ويخصص وحدة المعالجة المركزية لإحدى هذه العمليات.

يقوم المجدول على المدى القصير باتخاذ قرار بشأن العملية التي سيتم تنفيذها بعد ذلك. يعتبر المجدول على المدى القصير أسرع من المجدول على المدى الطويل.

جدولة متوسطة المدى Medium-Term Scheduler

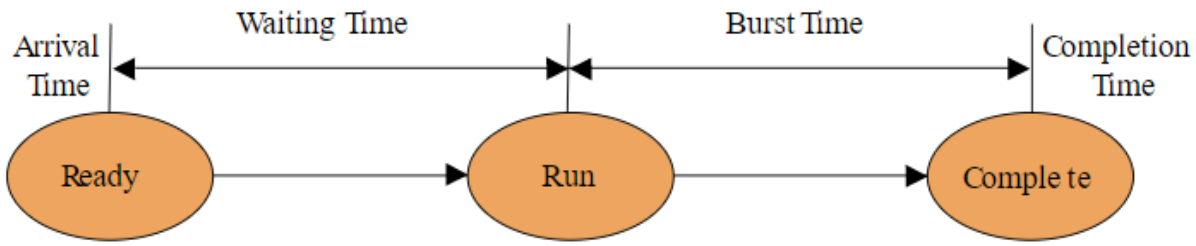
الجدولة متوسطة المدى هي جزء من المبادلة swapping يقوم بإزالة العمليات من الذاكرة ويقلل من درجة البرمجة المتعددة multiprogramming . ان المجدول على المدى المتوسط هو المسؤول عن التعامل مع العمليات الخارجية المبادلة.

قد يتم تعليق العملية الجارية suspended إذا قدمت طلب إدخال / إخراج ولا يمكن للعمليات المعلقة إحراز أي تقدم نحو الانتهاء. في هذه الحالة، لإزالة العملية من الذاكرة وإفساح المجال space للعمليات الأخرى، يتم نقل العملية المعلقة إلى التخزين الثانوي secondary storage تسمى هذه العملية بالمبادلة swapping ويقال إن العملية يتم تبديلها أو طرحها swapped out or rolled out وقد يكون التبادل ضروريًا لتحسين مزيج العملية.

Comparison among Scheduler مقارنة بين المجدولات

S.N.	Long-Term Scheduler	Short-Term Scheduler	Medium-Term Scheduler
1	It is a job scheduler	It is a CPU scheduler	It is a process swapping scheduler.
2	Speed is lesser than short term scheduler	Speed is fastest among other two	Speed is in between both short- and long-term scheduler.
3	It controls the degree of multiprogramming	It provides lesser control over degree of multiprogramming	It reduces the degree of multiprogramming.
4	It is almost absent or minimal in time sharing system	It is also minimal in time sharing system	It is a part of Time-sharing systems.
5	It selects processes from pool and loads them into memory for execution	It selects those processes which are ready to execute	It can re-introduce the process into memory and execution can be continued.

Times related to the Process أوقات العملية



$$CT - AT = WT + BT$$

$$TAT = CT - AT$$

$$\text{Waiting Time} = TAT - BT$$

TAT ⟶ Turn around time

BT ⟶ Burst time

AT ⟶ Arrival time

1 .وقت الوصول Arrival Time

يُطلق على الوقت الذي تدخل فيه العملية إلى قائمة الانتظار الجاهزة ready queue بوقت الوصول.

2 .وقت الاندفاع Burst Time

يُطلق على إجمالي الوقت الذي تتطلبه وحدة المعالجة المركزية لتنفيذ العملية بأكملها اسم .Burst Time هذا لا يشمل وقت الانتظار waiting time

3 .وقت الإنجاز Completion Time

يسمى الوقت الذي تدخل فيه العملية في حالة الإكمال completion state أو الوقت الذي تكمل فيه العملية تنفيذها بوقت الانتهاء.

4 .الوقت المستغرق Turnaround time

يُطلق على إجمالي الوقت الذي تقضيه العملية من وصولها إلى اكتمالها بوقت التحول.

5 .وقت الانتظار Waiting Time

يُطلق على إجمالي الوقت الذي تنتظره العملية حتى يتم تعيين وحدة المعالجة المركزية (CPU) لها بوقت الانتظار.

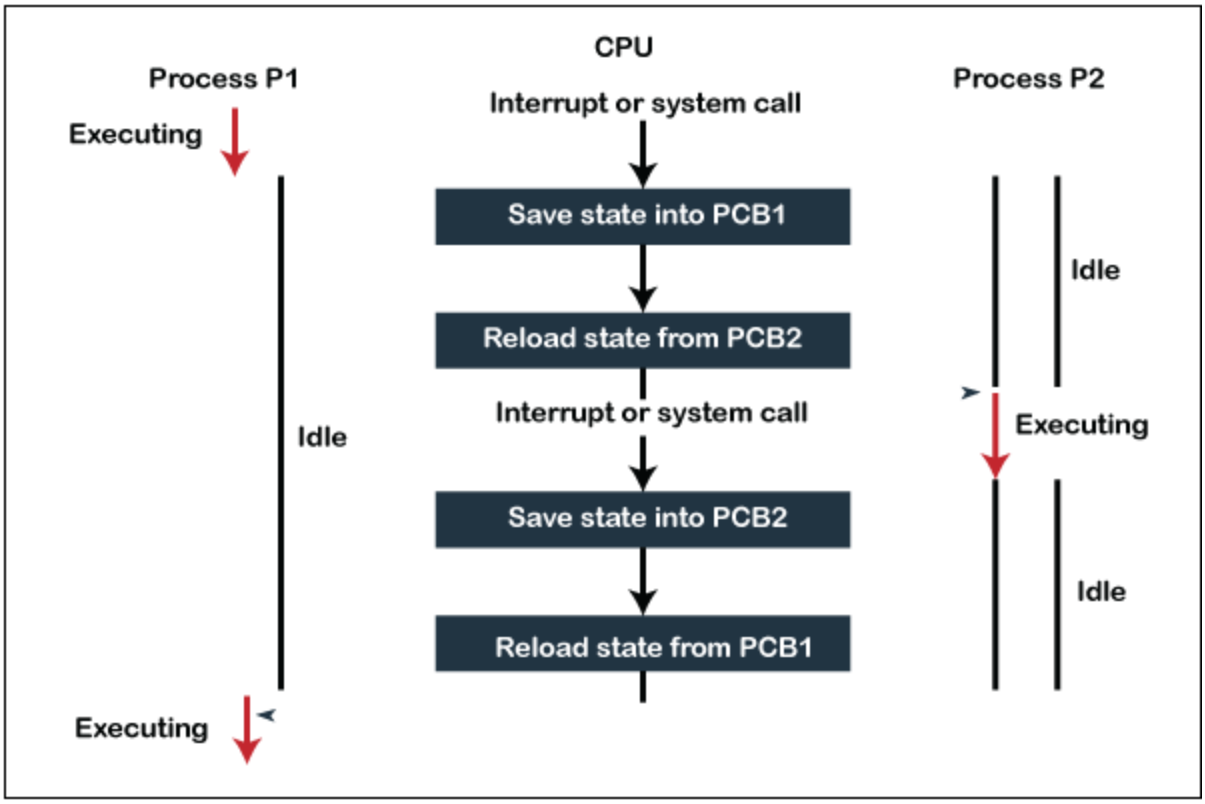
6 .زمن الاستجابة Response Time

يسمى الفرق بين وقت الوصول arrival time والوقت الذي تحصل فيه العملية على وحدة المعالجة المركزية لأول مرة بوقت الاستجابة.

تبديل السياق Context Switch

تبديل السياق هو آلية لتخزين واستعادة حالة أو سياق وحدة المعالجة المركزية في كتلة التحكم في العملية PCB بحيث يمكن استئناف تنفيذ العملية من نفس النقطة في وقت لاحق. باستخدام هذه التقنية، يتيح مبدل السياق عمليات متعددة لمشاركة وحدة معالجة مركزية واحدة ويعد تبديل السياق جزءًا أساسيًا من ميزات نظام التشغيل متعدد المهام multitasking operating system.

عندما يقوم المجدول بتحويل وحدة المعالجة المركزية من تنفيذ عملية إلى أخرى، يقوم مبدل السياق بحفظ محتوى جميع سجلات المعالج processor registers للعملية التي تتم إزالتها من وحدة المعالجة المركزية، في واصف العملية الخاص بها ويتم تمثيل سياق العملية في كتلة التحكم في العملية PCB ويمكن أن يؤثر تبديل السياق بشكل كبير على الأداء نظرًا لأن أجهزة الكمبيوتر الحديثة بها الكثير من السجلات العامة وسجلات الحالة التي يتعين حفظها.



الشكل أعلاه يمثل تنفيذ عمليتين (Process 1 and Process 2) داخل وحدة المعالجة المركزية CPU والية التبديل بينهما وكذلك يبين الشكل عملية حفظ وتحميل ال PCB الخاصة بكل عملية

معايير الجدولة Scheduling Criteria

عملية الجدولة من أهم الخصائص في تشغيل العمليات حيث ينظم دخول العمليات المراد تنفيذها إلى وحده المعالجة المركزية (CPU) وتعتمد هذه العملية على العديد من المعايير التي تحدد من هي العملية التي يجب تنفيذها و من أهمها:

1. **استغلال وحده المعالجة المركزية CPU utilization :**
   استغلال كل وقت وحده المعالجة المركزية(CPU) في تنفيذ العمليات. اي ان تكون وحده المعالجة المركزية (CPU) مشغولة بقدر الامكان ليتم استغلالها الاستغلال الأمثل

2. ***كمية البيانات المتدفقة في الثانية الواحدة*Throughput:**
   تبيّن معدل العمليات التي يمكن انجازها في وقت معين . **مثال توضيحي /** استطيع انجاز بمعدل 4 عمليات في 3 ثواني.

3. **الوقت المستغرق Turnaround time (TAT) :**
   الوقت اللازم لتنفيذ عملية ما (الوقت المستغرق من بداية تنفيذ العملية إلى نهايتها) بعدالة. (العملية ممكن ان تكون single CPU burst انبثاق واحد لوحده التحكم او قد تكون جزء thread)
   مثال توضيحي/ابتدأ العملية في الثانية الثالثة وانتهت في الثانية السابعة الوقت اللازم لتنفيذ هذه العملية هو 4 ثواني

4. **وقت الانتظار Waiting time :**
هو الوقت الذي تستغرقه العملية في الانتظار داخل مصفوفة الانتظار(ready queue) قبل دخولها إلى وحده المعالجة المركزية(CPU)

5. **زمن الاستجابة Response time :** هو الوقت الذي يحتاجه البرنامج لبدأ فعليا. **مثال توضيحي /**منذ الضغط على البرنامج ضغطاً مزدوج إلى أن يعمل البرنامج فعليا.

هذه المفاهيم تمكنا من اختيار أفضل طريقة للجدولة ولكن من الصعب تحقيقها جميعا

حتى تصل وحده المعالجة المركزية إلى الكمال لابد من توفر التالي:

- ***Minimize latency:*** وهو تقليل تأخر زمن الاستجابة واكتمال تنفيذ البرنامج.
- **Maximize throughput:** وهو زيادة الوقت الذي يتم تنفيذ العمليات خلاله
- **Maximize utilization:** هو زيادة الاستفادة من وحدة المعالجة المركزية وإبقائها مشغولة طوال الوقت .
- **Fairness:** وهو الإنصاف بين البرامج حيث يسمح لجميع البرامج من استخدام وحدة المعالجة المركزية ولا تكون الصلاحيات جميعها لبرنامج واحد.