

CHAPITRE 2 :

# GESTION DES PROCESSUS

## Objectifs spécifiques

- Connaître la notion de processus,
- Connaître les caractéristiques d'un processus ainsi que son contexte
- Connaître la notion d'interruptions et de ressources
- Connaître les étapes du cycle de vie d'un processus.

## Eléments de contenu

- I. Définition d'un processus
- II. Caractéristique d'un processus
- III. Notion d'interruptions
- IV. Ressources et allocation de ressources
- V. Cycle de vie d'un processus

## Volume Horaire :

**Cours** : 1 heure 30

**TD** : 0 heure

## 2.1 Définition d'un processus

Un processus est une entité dynamique correspondant à l'exécution d'une suite d'instructions : un programme qui s'exécute, ainsi que ses données, sa pile, son compteur ordinal, son pointeur de pile et les autres contenus de registres nécessaires à son exécution.

**Attention** : ne pas confondre un processus (aspect dynamique, exécution qui peut être suspendue, puis reprise), avec le texte d'un programme exécutable (aspect statique).

## 2.2 Caractéristiques

- Un processus possède un identifiant unique qui est généralement un entier incrémental (le premier processus 1, le second 2, ... etc.) et qui désigne de façon unique le processus dans le système (PID : Process Identifier)
- Les instructions à exécuter sont stockées dans une pile de données contenant les valeurs des variables du programme.

- Un contexte d'exécution : contenant entre autre le compteur ordinal qui indique l'adresse de l'instruction qui va être exécuté
- Les adresses des ressources utilisées
- Identifiant du processus parent

Les appels système relatifs aux processus permettent généralement d'effectuer au moins les actions suivantes :

- création d'un processus (fils) par un processus actif (d'où la structure d'arbre de processus gérés par un SE)
- destruction d'un processus
- mise en attente, réveil d'un processus
- suspension et reprise d'un processus, grâce à l'Ordonnanceur de processus (scheduler)
- demande de mémoire supplémentaire ou restitution de mémoire inutilisée
- attendre la fin d'un processus fils
- remplacer son propre code par celui d'un programme différent
- échanges de messages avec d'autres processus
- spécification d'actions à entreprendre en fonction d'événements extérieurs asynchrones
- modifier la priorité d'un processus

Dans une entité logique unique, généralement un mot, le SE regroupe des informations-clés sur le fonctionnement du processeur : c'est le mot d'état du processeur (Processor Status Word, PSW). Il comporte généralement :

- la valeur du compteur ordinal
- des informations sur les interruptions (masquées ou non)
- le privilège du processeur (mode maître ou esclave)
- etc.... (format spécifique à un processeur)

A chaque instant, un processus est caractérisé par son état courant ou contexte : c'est l'ensemble des informations nécessaires à la poursuite de son exécution (valeur du compteur ordinal, contenu des différents registres, informations sur l'utilisation des ressources). A cet effet, à tout processus, on associe un bloc de contrôle de processus (BCP). Il comprend généralement :

- une copie du PSW au moment de la dernière interruption du processus
- l'état du processus : prêt à être exécuté, en attente, suspendu, ...
- des informations sur les ressources utilisées
- mémoire principale
- temps d'exécution
- périphériques d'E/S en attente
- files d'attente dans lesquelles le processus est inclus, etc...
- et toutes les informations nécessaires pour assurer la reprise du processus en cas d'interruption

Les BCP sont rangés dans une table en mémoire centrale à cause de leur manipulation fréquente.

## 2.3 Les interruptions

Une interruption est une commutation de l'état (contexte) d'un processus provoquée par un signal généré par le matériel. Ce signal est la conséquence d'un événement interne à un processus, résultant de son exécution, ou bien extérieur et indépendant de son exécution. Le signal va modifier la valeur d'un indicateur qui est consulté par le SE. Celui-ci est ainsi informé de l'arrivée de l'interruption et de son origine. A chaque cause d'interruption est associé un niveau d'interruption. On distingue au moins 3 niveaux d'interruption :

- les interruptions externes : panne, intervention de l'opérateur, ...
- les déroutements qui proviennent d'une situation exceptionnelle ou d'une erreur liée à l'instruction en cours d'exécution (division par 0, débordement de mémoire, ...)
- les appels système

Le chargement d'un nouveau mot d'état provoque l'exécution d'un autre processus, appelé le traitant de l'interruption. Le traitant réalise la sauvegarde du contexte du processus interrompu (compteur ordinal, registres, indicateurs,...). Puis le traitant accomplit les opérations liées à l'interruption concernée et restaure le contexte et donne un nouveau contenu au mot d'état : c'est l'acquiescement de l'interruption.

Généralement un numéro de priorité est affecté à un niveau d'interruption pour déterminer l'ordre de traitement lorsque plusieurs interruptions sont positionnées.

## 2.4 Les ressources

On appelle ressource tout ce qui est nécessaire à l'avancement d'un processus (continuation ou progression de l'exécution) : processeur, mémoire, périphérique, bus, réseau, compilateur, fichier, message d'un autre processus, etc... Un défaut de ressource peut provoquer la mise en attente d'un processus.

Un processus demande au SE l'accès à une ressource. Certaines demandes sont implicites ou permanentes (la ressource processeur). Le SE alloue une ressource à un processus. Une fois une ressource allouée, le processus a le droit de l'utiliser jusqu'à ce qu'il libère la ressource ou jusqu'à ce que le SE reprenne la ressource (on parle en ce cas de ressource préemptible, de préemption).

On dit qu'une ressource est en mode d'accès exclusif si elle ne peut être allouée à plus d'un processus à la fois. Sinon, on parle de mode d'accès partagé. Un processus possédant une ressource peut dans certains cas en modifier le mode d'accès.

**Exemple** : un disque est une ressource à accès exclusif (un seul accès simultané), une zone mémoire peut être à accès partagé.

Le mode d'accès à une ressource dépend largement de ses caractéristiques technologiques. Deux ressources sont dites équivalentes si elles assurent les mêmes fonctions vis à vis du processus

demandeur. Les ressources équivalentes sont groupées en classes afin d'en faciliter la gestion par l'Ordonnanceur.

## 2.5 Cycle de vie d'un processus

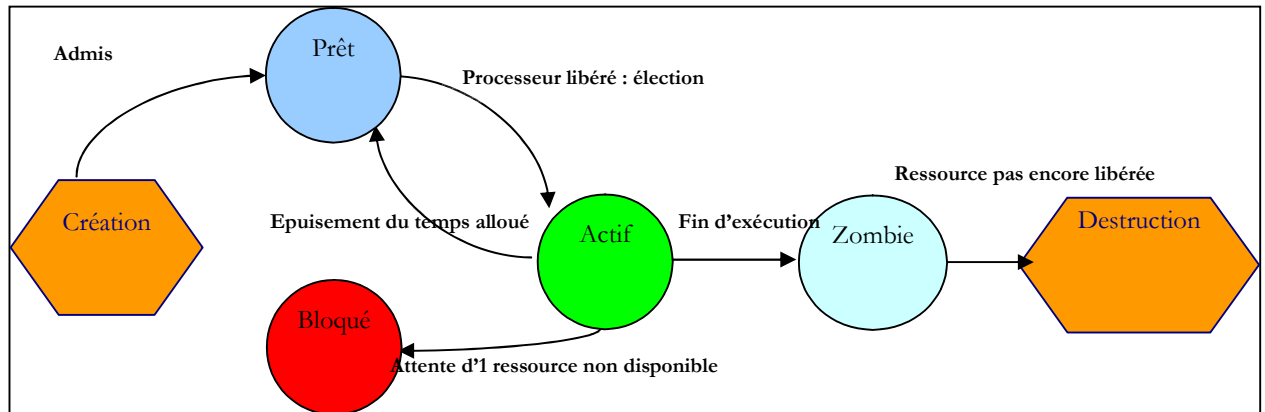


Figure 2 : Cycle de vie d'un processus

- **Création** : chargement des instructions, allocation de mémoires et des ressources (statiquement), il passe directement à l'état prêt.

Il existe 4 événements pour créer un processus :

- ✿ L'initialisation du système : au chargement du système il y a création automatique du processus racine père de tous les processus utilisateurs (id=0)
- ✿ Un processus peut lancer un autre processus, il en devient le parent, l'autre dernier sera désigné comme processus fils. (Un processus père ne se termine que lorsque tous ses fils sont terminés. On a donc une structure arborescente de processus).
- ✿ Une requête de l'utilisateur
- ✿ Initiation d'un travail en traitement par lot

- **La destruction d'un processus** : Lors de la destruction le processus libère les ressources allouées. Il y a quatre causes possibles de la destruction d'un processus :

- ✿ Arrêt normal : volontaire, lorsque le processus termine sa tâche.
- ✿ Arrêt pour erreur : volontaire suite à une erreur pour une instruction illégale
- ✿ Arrêt pour erreur fatale : involontaire tel que les mauvais paramètres de l'exécution du processus
- ✿ Arrêt volontaire par un autre processus

- **L'état prêt** : Le processus est prêt à être exécuté. Il est mis en attente jusqu'à ce qu'on lui libère le processeur (dispatch de l'Ordonnanceur), il passera alors à l'état Actif

- **L'état actif ou élu** :

- ✿ Le processus est en cours d'exécution par le processeur.
- ✿ Si le processus épuise le temps qui lui est alloué par le SE, il est remis en file d'attente des Prêts.

- Si il a besoin d'une ressource non disponible (opérations sur les périphériques), il est mis en attente prolongée (Interruption : état bloqué) jusqu'à la libération de la ressource nécessaire.
  - Si le processus atteint son terme (se termine) il passe à l'état Zombie
- **L'état bloqué** : Le processus est en attente d'une ressource pour terminer. Dès sa libération il repasse à l'état Prêt
- **L'état zombie** : Le processus a terminé son exécution et il ne peut plus évoluer mais les ressources qu'il a allouées ne sont pas encore libérées

### 3 L'ordonnancement

On appelle ordonnancement la stratégie d'attribution des ressources aux processus qui en font la demande. Différents critères peuvent être pris en compte :

- temps moyen d'exécution minimal
- temps de réponse borné pour les systèmes interactifs
- taux d'utilisation élevé de l'UC
- respect de la date d'exécution au plus tard, pour le temps réel, etc...