

Les Processus

1. Notion de processus

La notion de processus constitue un modèle simple et efficace pour représenter l'exécution concurrente des tâches au sein d'un système d'exploitation multitâche.

Tout le travail effectué par l'ordinateur est réalisé par des processus séquentiels qui représentent chacun une instance d'un programme séquentiel entrain de s'exécuter. Le processeur physique commute entre les processus sous la direction d'un ordonnancement.

Dans le cas du système à temps partagé, tous les processus progressent dans le temps mais un seul s'exécute à la fois.

La vitesse d'exécution d'un processus donné ne serait pas nécessairement la même si l'ensemble des processus est exécuté à nouveau, car l'état du système au lancement (position des têtes du disque, occupation de la mémoire,...) ainsi que le traitement d'événement synchrone (transmission réseau,...) peut conduire à des ordonnancement très différents.

2. Relation entre processus

Un processus ne pouvant pas être conceptuellement créé qu'au moyen d'un appel système c.à.d par un autre processus ayant réalisé cet appel. On voit apparaître naturellement la notion d'arborescence des processus, à partir d'un ancêtre commun à tous créé au démarrage du système (processus **init** sous Unix).

La notion de groupe de processus définie comme des sous-arborescences est également utilisée par certains systèmes pour limiter la portée des commandes de contrôle des processus (par exemple la commande kill)

3. Etats d'un processus

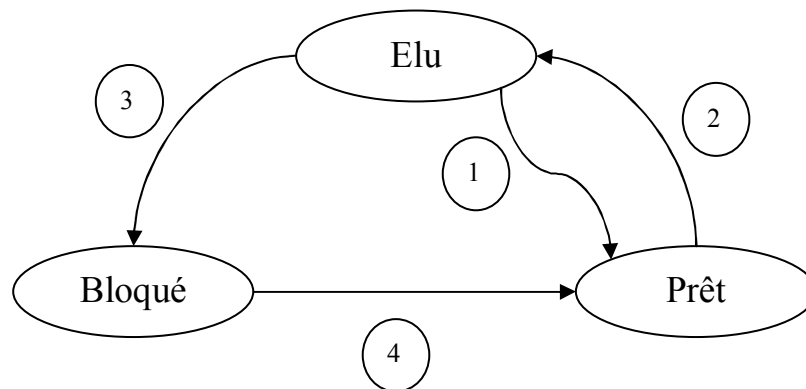
Les processus sont définis par des états discrets différents.

On dit qu'un processus est :

- Elu s'il est en cours de l'exécution sur le processeur. Dans le cas d'une machine multiprocesseur, plusieurs processus peuvent être élus en même temps, mais il y a toujours en plus autant de processus que de processeurs. (On peut l'avoir en tapant la commande **ps**).
- Prêt s'il est suspendu en faveur d'un autre. Un processus est prêt s'il ne lui manque que le processeur pour s'exécuter.

- Bloqué s'il est en attente d'un événement externe (entrées/sorties,...).

Le passage d'un état à un autre nécessite des conditions légales différentes.



Les significations des 4 transitions sont les suivantes :

- 1) Le processus a épuisé son quantum de temps qui lui a été attribué. L'ordonnanceur appelé de façon asynchrone par interruption lorsque le temps est écoulé et élit un autre processus parmi les processus prêts.
- 2) L'ordonnanceur élit ce processus parmi les processus prêts.
- 3) Le processus s'endort en attente d'un événement externe (décompte d'horloge, attente de donnée,...). L'ordonnanceur appelé de façon explicite par le processus courant élit un autre processus parmi les processus prêts.
- 4) L'événement attendu par le processus se produit. C'est le système d'exploitation qui gère son traitement de façon asynchrone, en interrompant temporairement le déroulement du processus actuellement élu pour traiter les données reçues et faire passer le processus de l'état bloqué à l'état prêt.

Avec ce modèle un processus peut progresser de deux façons :

- De façon synchrone, lorsque le processus est élu et dispose de la ressource processeur.
- De façon asynchrone, lorsqu'un autre processus est élu mais que celui-ci est interrompu par le système. (Le processus ne progresse pas lui-même puisque son compteur ordinal reste forcé où le processus a été mis dans l'état bloqué, mais du travail est réalisé dans son intérêt).

4. Gestion des processus

La gestion des processus et des interruptions présente toujours la couche la plus basse d'un système d'exploitation, car elle doit préexister pour concevoir le reste du système comme un ensemble de traitement séquentiel et concurrent.

Pour mettre en œuvre pratiquement le modèle des processus, le système dispose d'une table appelée table des processus, jamais swappée (toujours en mémoire) dont chaque entrée correspond à un processus particulier et contenant des informations telles que :

- Les valeurs de compteur ordinal, son pointeur de pile et des autres registres du processeur.
- Son numéro de processus, son état, sa priorité, son vecteur d'interruption et les autres informations nécessaires à la gestion du processus par le système.
- Son occupation mémoire.
- La liste des fichiers ouverts.

5. Gestion des interruptions

Le traitement des interruptions par le système s'effectue en appelant une routine de traitement associée à chaque type d'interruption, dont l'adresse est stockée dans les cases d'un tableau indexé par le type d'interruption appelé vecteur d'interruption. Par exemple sous linux on distingue les types d'interruptions suivants :

N° d'interruption	Gestionnaire
0	Horloge
1	Disque
2	Terminaux
3	Autres périphériques
4	Logiciels