Série d'exercices 6 : Gestion de la mémoire

6.1 Pour chacune des adresses suivantes 0x16EB, 0x3183, 0x4B53 et 0x27A29, dites quel est le numéro de page et le décalage, pour des pages de taille 4Kio?

Il faut séparer le champ numéro de page (bits les plus significatifs) et le champ décalage (log2(taille de page) bits les moins significatifs, soit 12 bits pour 4Kio). Ceci peut se faire en convertissant en binaire et en extrayant les champs voulus. On peut aussi prendre la partie entière de la division par la taille de page (numéro de page) ainsi que le reste de cette division (décalage) en utilisant les valeurs en décimal. En hexadécimal, on a 0x16EB/0x1000 = 1 et 0x16EB % 0x1000 = 0x6EB. En décimal, on a 5867/4096 donne 1 comme numéro de page et 5867 % 4096 donne 1771 comme décalage. Nous avons de la même manière (3, 387), (4, 2899) et (39, 2601) pour les trois autres adresses. L'important est d'indiquer la base dans laquelle on travaille.

6.2 Un ordinateur possède des adresses virtuelles de 32 bits et des pages de 8Kio. La table de page est d'un seul niveau et chaque entrée de la table de pages fait 32 bits. Quelle est la taille maximale de la table de page?

La taille maximale de mémoire adressable avec 32 bit est de 4Gio. La table de pages contient donc au maximum 4Gio / 8Kio = 524288 entrées. Si chaque entrée fait 4 octets, la taille est donc de 2Mio.

6.3 Un ordinateur utilise une table de pages à trois niveau. Les adresses virtuelles sont donc décomposées en 4 champs (a, b, c, d), a contenant les bits les plus significatifs. Que détermine la taille de chacun de ces champs? Il y a souvent une relation précise entre la taille relative de ces champs sur la plupart des systèmes, expliquez?

Le champs d détermine la taille des pages (2^d). Les champs a, b et c déterminent le nombre d'entrées dans les sections de tables de niveau 1, 2 et 3 respectivement. Usuellement, une page est utilisée pour chaque section de table, quel que soit le niveau. Le nombre d'entrée dans la section est donc la taille d'une page divisée par la taille d'une entrée. De ce fait, a, b et c sont généralement égaux. La taille d'une entrée est usuellement la taille de l'adresse virtuelle (32 bits pour un système 32 bits et 64 bits pour un système 64 bits); ceci permet d'y stocker le numéro de page physique ainsi que quelques bits pour les permissions, le statut...

6.4 Un ordinateur possède des adresses virtuelles de 32 bits et des pages de 4Kio. Un programme n'occupe que la première page (pour son code et ses variables) et la dernière page de la table (pour sa pile). Quel est l'espace requis pour la table de pages si le système utilise une table à un seul niveau? A deux niveaux?

Une table à un niveau requiert 4Gio / 4Kio = 1Mi entrées soit 4Mio. Dans une table à 2 niveaux, une page est utilisée au premier niveau avec deux entrées utilisées et deux pages au second niveau avec une entrée utilisée chacune. Il faut donc 3 pages pour la table de pages soit 12Kio.

6.5 Un ordinateur utilise une table de pages avec 1024 entrées pour chaque processus. La lecture dans la table de pages prend 5ns. Une cache de pré-traduction d'adresse (TLB) contient 32 entrées et peut être accédée en 1ns (au lieu de 5ns). Quel est le taux de succès requis afin d'avoir un temps d'accès moyen de 2ns?

Le temps moyen sera pour un taux de succès s : s * 1ns + (1 – s) * 5ns = 2ns. On a s + 5 – 5s = 2, en isolant s on obtient s = ¾ = 75%.

6.6 Un étudiant propose de construire un compilateur capable de produire une liste des pages qui seront accédées en séquence par un programme, afin d'implémenter l'algorithme optimal de remplacement de pages. Est-ce possible? Que peut-on faire pour aider l'algorithme de remplacement de pages?

Puisqu'on ne peut pas prédire même seulement si un programme peut se terminer ou non, il est encore moins possible pour le cas général de prédire à la compilation le chemin d'exécution complet dans le programme. Par contre, il est possible d'utiliser certaines informations à la compilation (e.g. arbre d'appel) pour optimiser la suite des opérations. En effet, les fonctions qui s'appellent l'une l'autre peuvent être placées sur une même page. On peut aussi donner des indices sur les pages qui risquent d'être requises peu de temps après une certaine page, par exemple dans le cas d'accès séquentiels. Il est aussi intéressant de récolter de l'information à l'exécution. Par exemple, la séquence d'amorçage de l'ordinateur est toujours la même. Certains systèmes notent les pages lues une première fois et utilisent ensuite cette information pour demander une pré-lecture de ces pages en parallèle avec le début du démarrage.

6.7 Un petit ordinateur contient 4 pages physiques que se disputent 8 pages en mémoire virtuelle. Pendant l'exécution, la séquence d'accès des pages virtuelles est la suivante : [0, 1, 7, 2, 3, 2, 7, 1 ,0, 3]. Combien de fautes de pages surviennent avec la stratégie FIFO? LRU?

En premier, les pages 0172 sont chargées en mémoire. Ensuite, 3 remplace 0, puis 2, 7 et 1 sont présentes, il faut recharger 0 qui remplace 1, et 3 est là. Le total est 6 fautes. Avec LRU, les pages 0172 sont chargées en mémoire. Ensuite, 3 remplace 0, puis 2, 7 et 1 sont là, il faut recharger 0 qui remplace 3 (inutilisé depuis plus longtemps de 2, 7 et 1). Il faut finalement à nouveau lire 3 qui remplace 2. Le total est 7 fautes.

6.8 Un ordinateur fournit à chaque processus un espace de 65536 octets divisé en pages de 4Kio. Un programme utilise un section text de 32768 octets, data 16386

octets et pile 15870 octets. Est-ce que ce programme peut entrer dans l'espace disponible?

Il ne faut pas oublier que chaque région a ses propres pages. En effet, la section text vient d'un fichier exécutable et a une protection lecture/exécution, alors que la pile croit vers le bas et vient avec lecture et écriture, par exemple. Dans le cas qui nous intéresse, l'espace contient 64Kio/4Kio = 16 pages, 8 prises par le text, 5 requises pour les données (16386 > 4 x 4096) et 4 requises pour la pile, ce qui est plus que l'espace disponible. L'exécution de ce programme pourrait fonctionner, mais sa performance sera réduit en raison des fautes de page pour la pagination.

6.9 Une page peut-elle faire partie de l'espace de travail de deux processus en même temps?

Bien sûr, il est tout à fait possible pour deux processus d'utiliser le même code exécutable (programme ou librairie partagée), le même fichier attaché en mémoire ou la même zone de mémoire partagée.

6.10 Est-il possible qu'une page soit en lecture seulement pour un processus et en lecture et écriture pour un autre?

Oui, c'est tout à fait possible, l'information sur les accès permis est dans la table de pages de chaque processus. On peut imaginer un cas ou un processus maître s'occupe des mises à jour et des processus esclaves ne font que lire cette information.

6.11 Une instruction assembleur charge une valeur de 32 bits dans un registre sur une architecture n'ayant pas de contrainte d'alignement. Combien de défauts de pages pourraient subvenir à cause de cette instruction?

Cette architecture n'ayant pas de contrainte d'alignement, il est possible que l'instruction et la valeur à charger soit à la fin d'une page et s'étende sur la suivante. Si cette situation survient, il se produit un total de 4 fautes de pages pour cette seule instruction.