

CAPITULO4- GESTAO DA MEMÓRIA

1. Existem 3 níveis de estrutura da memória de um SC, assim temos memória cache (pequena, muito rápida, volátil e cara); memória principal (tamanho da ordem dos MB até 1-4GB, volátil, velocidade e preço razoável); memória massa (grande centenas GB, lenta, não volátil e barata).
Esta forma a memória cache e utilizada para conter uma cópia das posições de memória mais frequentemente referenciadas pelo CPU num passado recente. A memória principal, contém a TCP, e toda a informação necessária para a utilização dos processos em execução, contém os processos em execução e que competem pelo CPU.
A memória de massa, tem como função armazenar dados e programas permanentes e contém uma área de swapping que serve de extensão à memória principal para esta não ser um factor limitativo ao nº de processos q podem existir.
2. O princípio da localidade de referência, esta subjacente à organização da memória hierarquicamente.
Este princípio baseia-se no pressuposto de q qto mais afastada do CPU estiver a instrução menos vezes será referenciada, e assim o tempo médio de uma referência aproxima-se do valor mais baixo. Através da constatação heurística do comportamento dos programas em execução estabelece que as referências à memória durante a execução tendem a concentrar-se em fracções bem definidas do seu espaço de endereçamento durante intervalos mais ou menos longos.
3. As principais actividades entre a memória principal e de massa são a manutenção de um registo sobre as partes de memória ocupada e livre, a reserva para os processos, de porções de memória principal, qdo necessário e dp a sua libertação, a transferência para área de swapping de todo ou parte do espaço de endereçamento qd a memória principal é diminuta para todos os processos coexistirem.
4. A imagem binária do espaço de endereçamento de um processo é relocatável pois num ambiente multiprogramado nunca se sabe em q posição de memória o ficheiro vai ser carregado, e com as constantes comutações de processos é necessário que os endereços sejam referenciados no início do modulo ou ficheiro pois o ficheiro pode tb não estar sequencial e desse modo é necessário q o ficheiro seja relocatável.
5. A mais exigente em termos de ocupação de memória e a estática, pois esta carrega para a memória principal todas as funções necessárias da biblioteca, e cada processo tem a sua biblioteca, enquanto que na ligação dinâmica a carga para memória principal faz-se apenas qd é necessário e se processos diversos usarem a mesma biblioteca utilizem uma cópia da mesma reduzindo espaço na memória.
Em termos de processamento a dinâmica é mais exigente pois sempre q encontra uma referência para uma biblioteca tem q determinar o endereço da rotina pretendida e fazer a sua carga para memória principal, qd for precisa novamente esse código já estará residente em memória, na estática não é necessário nada disto pois é carregado tudo logo no início.
6. i) A área partilhada e definida na zona de definição dinâmica na região global, variáveis globais partilhadas pelos vários processos;
ii) não pois o tamanho das zonas de definição estática e de código local/global podem ser diferentes;
iii) ponteiros.
7. O espaço de endereçamento lógico é relocatável e é referenciado ao início do modulo ou do programa, enquanto que a região de memória onde ele é efectivamente carregado em memória principal é o espaço de endereçamento físico. Para ser eficiente e segura entre os 2 espaços lógico e físico é necessário resolver a relocação dinâmica, isto é a capacidade de converter um endereço lógico em físico durante a execução (run-time), de modo a que o endereço físico esteja em q lugar da memória e possa ser movido. A protecção dinâmica, impedimento em run-time de referências a endereços localizados fora do espaço de endereçamento próprio do processo.
8. A organização de memória real exige uma correspondência biunívoca entre espaço endereçamento lógico e físico do processo. As consequências desta organização são: a limitação espaço de endereçamento de um processo, o espaço de um processo não pode ser superior ao tamanho da memória principal disponível; a contiguidade do espaço de endereçamento físico, mais simples e eficiente supõe que o espaço de endereçamento do processo é contíguo; a área de swapping, extensão memória principal, serve para armazenar o espaço de endereçamento dos processos qdo não existe espaço memória principal disponível.
9. A comparação de um endereço lógico com um registo limite, existente na TCP, para saber se o endereço é válido, gerando uma excepção em caso de não ser, soma o endereço lógico ao registo base existente na TCP, que é o endereço físico do início do processo na memória.
10. Uma arquitectura de partições fixas divide a memória principal em partições de tamanho fixo e não necessariamente iguais, cada partição contém o espaço endereçamento físico de um processo, a arquitectura de partições variáveis admite à partida um bloco e cada vez q um processo necessitar de espaço de endereçamento é-lhe atribuído o espaço necessário, sendo depois libertado qdo não for necessário.
Arquitectura partições fixas: (vantagens) – simples implementar: não exige hardware nem estruturas de dados especiais; - eficiente: a selecção é feita rapidamente. (desvantagens) – conduz a uma grande fragmentação interna da mem principal: a parte não utilizada de cada partição é desperdiçada. – de aplicação específica: para minimizar o desperdício adequa-se o tamanho das partições ao tipo de processos.
Arquitecturas partições variáveis: (vantagens) - geral: independente do tipo de processos a ser executados. – implementação pequena complexidade: não exige hardware especial, apenas é necessário 2 listas ligadas. (desvantagens) – grande fragmentação externa da mem principal: a fracção da mem principal desperdiçada face ao tamanho reduzido das regiões em que está dividido. – pouco eficiente: não é possível construir algoritmos que sejam simultaneamente eficientes na reserva e na libertação de espaço.
11. Fragmentação interna, associada às partições fixas, espaço não usado dentro da partição. Fragmentação externa, associado às partições variáveis, espaço resultante das pequenas fracções de memória que não servem para alocar nenhum processo.
12. Next fit: pesquisa até encontrar uma região com tamanho suficiente, fragmentação interna grande. Best fit: pesquisa na totalidade da mem pela região mais pequena com tamanho maior que o necessário, fragmentação interna pequena.
13. Organização de memória virtual existe qdo o espaço de endereçamento lógico e físico estão completamente dissociados.
As consequências destes tipos de organização são: a não limitação do espaço de endereçamento do processo, é possível executar processos cujo espaço de endereçamento é superior ao tamanho da memória disponível; a não contiguidade do espaço de endereçamento físico, os espaços de endereçamento são divididos em blocos e não estão necessariamente contínuos, procurando garantir uma ocupação mais eficiente do espaço disponível; área swapping, extensão da memória principal e serve para manter uma imagem actualizada dos espaços de endereçamento dos processos, nomeadamente a sua parte variável (zonas definição estática, dinâmica e stack).
14. Numa organização de memória virtual, o ee físico e o ee lógico estão totalmente dissociados, sendo que podem estar dispersos na memória principal, ou na área de swapping, ou em ambos. O ee lógico de um processo pode estar segmentado no ee físico, e apenas parcialmente em memória principal com o resto na área de swapping, forma esta que é designada de organização de memória virtual, ou o ee lógico pode estar carregado completamente em memória principal, um bloco único, abordagem designada de organização memória real. Face à organização de memória real, a organização de memória virtual tem como vantagens a possibilidade de obter maior rentabilidade da memória principal, já que o ee físico está segmentado, o facto de permitir que mais processos coexistam em memória, pois apenas alguns blocos de ee lógico estão carregados na memória principal. Como desvantagem a organização de memória virtual causa um maior particionamento e as partições podem estar na área de swapping e não estar em memória principal com consequências ao nível da performance.
15. **Arquitectura paginada**—os blocos (designados de paginas) todos iguais e tamanho correspondente a uma potencia de dois;—mecanismo de divisão de espaço de endereçamento lógico do processo é meramente operacional (bits mais significativos – numero de paginas; bits menos significativos—deslocamento);—memória principal vista como dividida em blocos (frames) de tamanho igual às paginas.
Arquitectura segmentada—constituida por espaços de endereçamentos autónomos (segmentados) de comprimento variável;—segmentos para código, para zona de definição estática (para as variáveis globais e globalmente locais), definição dinâmica local, definição dinâmica global e stack;—pouco interesse pratico—exige a aplicação de técnicas de reserva de espaço para carregamento de um segmento em memória em tudo semelhantes às usadas em organização de memória real de partições variáveis, o que resulta uma enorme fragmentação externa da memória principal. Também os segmentos de dados de crescimento contínuo podem originar problemas com falta de espaço.
Arquitectura segmentada / paginada—divisão do espaço de endereçamento lógico é primariamente segmentada com a atribuição de múltiplos espaços de endereçamentos lineares é dividido em paginas originando um mecanismo de carregamento de blocos em memória principal com todas as características da arquitectura paginada .
Vantagens arquitectura paginada
Geral—o âmbito da sua aplicação é independente do tipo de processo que vão ser executados (numero e tamanho do espaço de endereçamento);
Grande aproveitamento da memória principal— não conduz a fragmentação externa e a fragmentação interna é praticamente desprezável;
Não exige requisitos especiais de hardware – a unidade de gestão de memória existentes nos processadores actuais de uso geral está já preparada para sua implementação;
Desvantagens da arquitectura paginada
-acesso à memória mais longo ----cada acesso à memória transforma-se num duplo acesso por consulta prévia da tabela de paginação ;
- operacionalidade muito exigente--- a sua implementação exige por parte do sistema operativo a existência de um conjunto de operações de apoio que são complexas e que têm de ser cuidadosamente estabelecidas para que não hajam perdas acentuadas de eficiência.
Vantagens da arquitectura Mista
-geral—âmbito da sua aplicação independente do tipo de processos que vão ser executados
- grande aproveitamento da memória principal----não conduz a fragmentação externa e a fragmentação interna é praticamente desprezável
- gestão mais eficiente da memória no que respeita a regiões de crescimento dinâmico
- minimização do numero de paginas que têm que estar residentes em memória principal em cada etapa de execução do processo
Desvantagens da arquitectura Mista
- exige requisitos especiais de hardware – nem todos os processadores estão preparados para a sua implementação.
- acesso à memória mais longo ----cada acesso à memória transforma-se num triplo acesso por consulta prévia das tabelas de segmentação e de paginação (minimizado se a unidade de gestão de memória associativa para armazenamento das entradas das tabelas de paginação recentemente mais referenciadas)
- operacionalidade muito exigente ---a sua implementação por parte do sistema operativo é ainda mais exigente do que a arquitectura paginada.
- 16.
17. Nos computadores pessoais actuais, o seu funcionamento baseia-se numa grande multiplicidade de processos, de complexidade e tamanhos variáveis, pelo que a arquitectura de memória virtual oferece uma maior flexibilidade para lidar com os diferentes tamanhos dos espaços de endereçamento de cada processo, resultando numa optimização dos recursos disponíveis. Como o tamanho dos processos não é fixo e não pode ser previsto, a arquitectura de memória real que se baseia nestes pressupostos não pode ser utilizada.
18. Na arquitectura de memória virtual a área de swapping tem como função enquanto extensão da memória principal servir para manter uma imagem actualizada dos espaços de endereçamentos dos processos que correntemente coexistem, nomeadamente da sua parte variável. Na arquitectura de memória real, a área de swapping tem como função, enquanto extensão da memória principal, servir de arrecadação para armazenamento do espaço de endereçamento dos processos que não podem ser directamente carregados em memória principal por falta de espaço.
19. A tabela de segmentação contém a lista de segmentos dos processos, contendo o modo de acesso a esses segmentos (permissões) e o endereço em memória da tabela de paginação do segmento. A tabela de segmentação é portanto a ligação entre os segmentos e as paginas associadas a cada processo. A tabela de paginação de um segmento é usada para fazer a ligação entre as paginas de um segmento e a frame/bloco em memória/área de swapping. Cada entrada desta tabela tem um campo para indicar se a pagina está residente em memória ou na área de swapping, 3 campos para referenciar o tipo de acesso, 2 campos onde estarão a localização da pagina e localização da pagina na área de swapping, se já lhe tiver sido atribuído espaço.

20. Na arquitectura paginada os blocos são todos iguais. A memória principal também é dividida em blocos de tamanho igual. Na arquitectura segmentada existem espaços de endereçamentos autónomos de tamanho variável (segmento) , existindo segmentos para código, zona da definição estática, zona de definição dinâmica e stack. A arquitectura segmentada pura tem pouco interesse pratico porque, tratando a memória principal como um espaço contínuo, exige a aplicação de técnicas de reserva de espaço para carregamento de um 1 segmento em memória que são em tudo semelhantes às usadas em organizações de memória real de partições variáveis. Isto resulta numa enorme fragmentação externa da memória principal com o consequente desperdício de espaço. Além disso segmentos de dados de crescimento contínuo conduzem a problemas adicionais: são facilmente concebíveis situações em que um acréscimo de tamanho do segmento não poderá ser realizado na sua localização presente, originado a sua transferência na totalidade para outra região da memória ou um caso limite em que não há espaço suficiente, ao bloqueio, ou suspensão do processo, com a remoção do segmento, ou de todo o espaço de endereçamento residente, para a área de swapping.