

A HISTÓRIA DOS SISTEMAS OPERACIONAIS

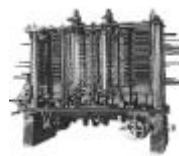
Os sistemas operacionais (SO) estiveram intimamente associados à arquitetura dos computadores nos quais eles rodam. Por esta razão vamos examinar as sucessivas gerações de computadores para ver como eram seus SOs.

O primeiro computador digital verdadeiro foi projetado pelo matemático inglês Charles Babbage (1792-1871). Ele gastou a maior parte de sua fortuna e de sua vida na tentativa de construir uma máquina analítica. Era um projeto puramente mecânico composto de engrenagens e discos perfurados. Babbage foi o primeiro a perceber que uma máquina computadora deveria consistir em um dispositivo de entrada, uma memória, uma unidade central de processamento, e um dispositivo de saída. Ele usava uma "impressora" como dispositivo de saída e, como dispositivo de entrada, usava um leitor de cartões que era inspirado nos cartões perfurados de Jacquard. Para saber mais sobre Babbage visite o link <http://www.geocities.com/CapeCanaveral/Hall/3608/babbage1.htm>



C. Babbage

Embora tenha investido a fortuna e tempo, Babbage nunca conseguiu fazê-la funcionar adequadamente. Babbage sabia que precisaria de software para a sua máquina analítica. Assim ele contratou uma jovem mulher, Ada Lovelace, filha de um famoso poeta britânico (Lord Byron), como programadora. Ela é considerada a primeira programadora do mundo. Não houve também outro programador antes dela. O nome da linguagem de programação Ada[®] é uma homenagem a ela.



Máquina Analítica

A Primeira Geração de computadores (1945-55) marca a época das válvulas e painéis de conectores. Note que temos uma ausência de progresso na área, entre Babbage e a II Guerra Mundial. Em meados da década de 40, Harvard Aiken, de Harvard, John von Neuman, do Instituto de Estudos Avançados de Princeton, J. Presper Eckert e William Mauchly, da Universidade da Pensilvânia, e Konrad Zuse, na Alemanha, dentre outros tiveram sucesso na construção de máquinas de cálculo utilizando válvulas. Essas máquinas eram enormes, ocupando salas inteiras com dezenas de milhares de válvulas, e eram muito mais lentas que as mais simples calculadoras de hoje. Naquele tempo um único grupo de pessoas projetava, construía, programava, operava e mantinha cada máquina. Toda a programação era feita em linguagem de máquina, frequentemente ligando com fios painéis de conectores para controlar as funções básicas da máquina. As linguagens de programação eram desconhecidas e ninguém nunca tinha ouvido falar de SOs. O modo normal de operação era o programador reservar um período de tempo na folha de reserva na parede, depois descer até o lugar da máquina, inserir suas conexões no computador e gastar algumas horas esperando que nenhuma das aproximadamente 20.000 válvulas queimasse durante a execução. Praticamente todos os problemas eram simples cálculos numéricos, tais como gerar tabelas de senos e de co-senos.



Ada Lovelace

No início da década de 50 a rotina melhorou um pouco com a introdução dos cartões perfurados, tornando-se possível gravar programas em cartões e lê-los, em vez de usar cabos conectores.

A Segunda Geração (1955-63) marca a fase dos transistores e sistemas de lote. Com a introdução dos transistores os computadores tornaram-se mais confiáveis, o bastante para serem fabricados e vendidos para clientes com expectativa de que continuariam a funcionar por tempo suficiente para realizar algum trabalho útil. Pela primeira vez estabeleceu-se uma distinção clara entre projetistas, construtores, operadores, programadores e o pessoal de manutenção. As máquinas eram guardadas em salas especiais com ar condicionado e equipes de operadores profissionais para mantê-las funcionando. Somente as grandes corporações ou importantes órgãos do governo ou universidades podiam ter recursos para arcar com seu preço, na casa dos milhões de dólares. Para executar um programa ou conjunto de programas (job), um programador primeiro escrevia o programa em papel (em FORTRAN ou Assembly), e então o transformava em cartões perfurados. Os cartões eram então levados para a sala de entrada onde eram recebidos por um dos operadores. Quando o computador terminava o job, um operador ia até a impressora, removia a saída e a levava para a sala de saída, onde ficava a disposição do programador. Em seguida o operador pegava um dos conjuntos de cartões que haviam sido trazidos para a sala de entrada e os inseria na máquina para leitura. Muito tempo de máquina se perdia enquanto os operadores se deslocavam nas salas para implementar as operações em computador. A busca de economia de tempo levou ao sistema de processamento em lotes (batch system). Coletava-se uma bandeja completa de jobs na sala de entrada e depois os mesmos eram lidos sobre uma fita

magnética, utilizando-se um computador pequeno e mais barato (para os computadores da época). Nesta fase SO típicos era o FMS (FORTRAN Monitor System) e o IBSYS da IBM.

Terceira Geração (1965-80). A maioria dos fabricantes de computadores tinha duas linhas de produto distintas e totalmente incompatíveis. De um lado tinha-se os computadores científicos de grande escala, baseados em palavras, que eram utilizados para cálculos numéricos em ciência e engenharia. Por outro lado havia os computadores comerciais, baseados em caracteres, amplamente usados para classificar e para imprimir fitas para bancos e para companhias de seguro. Desenvolver duas linhas distintas de produtos era uma proposta cara para os fabricantes. A IBM tentou resolver estes problemas em uma única tacada introduzindo o System/360. O 360 era uma série de máquinas compatíveis ao nível de software que variavam desde a capacidade de 1401 até um muito mais do que o 7094, desenvolvido durante a segunda geração. As máquinas diferiam só no preço e no desempenho (memória máxima, velocidade de processamento, número de dispositivos de E/S permitidos, etc.). Como todas as máquinas tinham a mesma arquitetura e conjunto de instruções, programas escritos para uma máquina podiam executar em todas as outras. Assim uma única família de máquinas podia satisfazer as necessidades de todos os clientes. O 360 era a primeira linha importante de computadores usando Circuitos Integrados, CIs, de pequena escala. Isto proporcionava vantagens importantes de preço/desempenho sobre máquinas de segunda geração, que eram baseadas em transistores como componentes. O 360 foi um êxito imediato e a idéia de uma família de computadores compatíveis logo foi adotada por todos os outros fabricantes importantes.

A maior força da idéia de "uma família" era ao mesmo tempo sua maior fraqueza. A intenção era que todo software, incluindo o SO, tinha que trabalhar em todos os modelos. Ele tinha que ser bom em sistemas com poucos e com muitos periféricos. Ele tinha que funcionar bem em ambientes comerciais e em ambientes científicos. Tinha de ser eficiente para para todo uso. Não havia como a IBM (ou qualquer outra pessoa) conseguir escrever um software para atender a todos estes requisitos contraditórios. O resultado era um SO extraordinariamente complexo e grande. Consistia de milhões de linhas de linguagem assembler, escritas por milhares de programadores e com milhares de "bugs". Apesar disto o SO/360 e os sistemas operacionais semelhantes de terceira geração produzidos por outros fabricantes de computador realmente satisfizeram a maioria dos seus clientes. Eles também popularizaram várias técnicas-chaves em SO de segunda geração. A mais importante delas era a multiprogramação. Para evitar a ociosidade da CPU quando se fazia uma pausa para troca de fitas ou outra operação, dividia-se a memória em várias partições, com um job diferente em cada uma delas. Assim quando se pausava a execução de um job outro entrava em operação, ocupando o tempo ocioso da CPU. Veja a figura ao lado, ilustrando a partição da memória:

Outro importante recurso da geração foi a capacidade de ler Jobs de cartões para disco assim que chegavam na sala do computador. Desta forma quando uma execução terminava já se podia iniciar outra. Esta técnica é chamada spooling (sigla de simultaneous peripheral operation on line - operação periférica simultânea on line). A necessidade de tempo de resposta rápido abriu caminho para uma variante da técnica de multiprogramação chamada compartilhamento de tempo. Num sistema de tempo compartilhado, se 20 usuários faziam login e 17 deles estavam pensando ou tomando café, a CPU podia estar alocada para os outros 3 jobs.

Outro importante desenvolvimento durante a terceira geração foi o crescimento fenomenal dos minicomputadores dando origem a toda uma nova indústria. O preço destas máquinas giravam em torno de 120 mil dólares e venderam como pãozinho quente da padaria. Um dos cientistas de computação da Bell Labs, Ken Thompson, descobriu um pequeno microcomputador que ninguém estava usando e começou a escrever uma versão simplificada monousuária que evoluiu para o SO UNIX®, que se tornou popular no mundo acadêmico, entre órgãos do governo e entre muitas empresas.

Quarta Geração (1980 até hoje). O desenvolvimento dos circuitos LSI (Large Scale Integration - Integração em Larga Escala), chips contendo milhares de transistores por cm² de silício foi o alvorecer da era do computador pessoal. Assim como os microcomputadores da terceira geração possibilitou que cada departamento em uma empresa ou uma universidade tivesse seu computador, o microprocessador possibilitou que as pessoas tivessem seu computador pessoal. Os primeiros foram chamados estações de trabalho. Na verdade estas estações nada mais são do que computadores pessoais de maior porte, normalmente conectados por uma rede. A ampla disponibilidade do poder da computação abriu caminho para o crescimento de uma importante indústria de software para computadores pessoais. Grande parte desse software tinha uma interface amigável, ou seja, era projetada para usuários que não

Job 1
Job 2
Job 3
Sistemas Operacionais

Partição de Memória

só ignoravam tudo sobre computadores como também não tinham nenhuma intenção de aprender. Dois sistemas operacionais inicialmente dominaram a cena do computador pessoal e da estação de trabalho: o MS-DOS da Microsoft e o UNIX. O MS-DOS era amplamente usado na IBM PC e em outras máquinas que empregavam CPU 8088 e seus sucessores 80286, 80386 e 80486. Mais tarde o Pentium e o Pentium PRO. Depois veio o sucessor do MS-DOS, o Windows, originalmente rodando sobre o DOS. Neste sentido os Windows 3.1, 3.11 era mais como uma shell do que um SO. Mas em 1995 foi lançada uma versão independente, o Windows95[®], de modo que o MS-DOS deixou de ser utilizado com suporte do Windows.

O UNIX é outro importante concorrente que dominante em estações de trabalho e em outros computadores topo de linha, como os servidores de rede.

Um desenvolvimento importante que começou durante os meados da década de 80 é o crescimento de redes de computadores pessoais, executando sistemas operacionais de rede e sistemas operacionais distribuídos. Em um sistema operacional de rede, os usuários estão cientes da existência de múltiplos computadores e podem se comunicar e podem se conectar com máquinas e copiar arquivos de uma máquina para outra. Cada máquina executa localmente seu SO e tem seu próprio usuário (ou usuários).

Os SO de rede não são fundamentalmente diferentes dos SO monoprocessadores. É claro que eles necessitam de uma interface de rede e de algum software de baixo nível para controlá-la, bem como programas para permitir conexões remotas e acesso a arquivos remotos. Entretanto tal condição não muda essencialmente a estrutura de um SO.

Verdadeiros sistemas operacionais distribuídos requerem mais do que adição de pequenas porções de código a um SO monoprocessador, porque os sistemas operacionais distribuídos diferem significativamente. Por exemplo eles com frequência permitem que os aplicativos executem vários processadores simultaneamente. Por isso requerem um algoritmo de agendamento de processador mais complexo a fim de otimizar a quantidade de paralelismo.