

Sistemas Distribuídos

* Definições de S.D:

Consiste de uma coleção de computadores autônomos ligados através de uma rede e equipados com software de sistemas distribuídos” (Coulouris)

Um sistema distribuído é uma coleção de computadores independentes que aparecem para os usuários do sistema como um único computador.” (Tanenbaum)

* **O software de distribuição habilita os computadores a coordenar suas atividades e compartilhar recursos do sistema - hardware, software e dados;**

* **Os usuários de um sistema distribuído devem perceber como uma facilidade de computação única, integrada, mesmo que seja implementada por vários computadores;**

* **Hardware - máquinas autônomas;**

* **Software - os usuários pensam no sistema como um único computador.**

* **O desenvolvimento de sistemas distribuídos iniciou em 1970 seguindo o progresso das redes de computadores de alta velocidade;**

* **Esse crescimento tem sido alcançado devido ao desenvolvimento de software para sistemas distribuídos destinados a suportar o desenvolvimento de aplicações distribuídas.**

* **Fatores que contribuem para a evolução dos Sistemas Distribuídos:** Baixo custo dos microprocessadores; Interconexão e abrangência das redes; Aplicações multimídias em expansão;

-->Aplicações Distribuídas

* **Aplicações de conferência e acesso a informações multimídia;**

* **Aplicações que empreguem representação digital de imagens, áudio, vídeo que suportem usuários realizando ações de ensino a distância, trabalhos ou projetos cooperativos, videoconferências, transações comerciais dentre outras atividades;**

* **Muitas aplicações multimídias executam sobre um sistema distribuído, habilitando muitos usuários a interagirem entre si e compartilhar recursos;**

* **Requisitos especiais a serem considerados em projetos deste tipo:** Vídeo e áudio são mídias contínuas; Aplicações interativas devem apresentar um delay entre o início da conversação em um site e a sua reprodução no outro inferior a 100 ms, pois caso contrário, o atraso torna-se perceptível.

-->Multiprocessadores e Multicomputadores

* **Multiprocessadores -Caracteriza-se pelo uso de dois ou mais processadores no mesmo sistema;**

* **Multicomputadores -Caracteriza-se pela ligação e intercalação de dois ou mais computadores (geralmente dispostos em Cluster/Grid).**

* **Independente de onde esteja sendo implementado o objetivo é realizar processamento paralelo.**

-->Paralelismo-O processamento paralelo pode ser definido como “uma forma eficiente do processamento de informação com ênfase na exploração de eventos concorrentes no processo computacional” (HWANG 1984).

-----> Tipos de Paralelismo

* **Paralelismo Temporal:** É feito por operações sobrepostas, onde os eventos ocorrem em instantes de tempos sobrepostos.

******Pipeline:** Onde a tarefa é subdividida numa sequência de subtarefas cada uma executada por um estágio de hardware específico (E1), que trabalha concorrentemente com outros estágios do Pipeline (E2 ... En), criando o que tecnicamente é chamado de Paralelismo Temporal na execução das subtarefas

* **Paralelismo Espacial Síncrono:** Onde os eventos são simultâneos, pois ocorrem em múltiplos recursos no mesmo instante de tempo. Nesta arquitetura, uma unidade de controle

distribui para vários elementos de processamento, EP, a mesma instrução que será executada em paralelo por estes EPs sobre seus dados no mesmo instante de tempo. O processador matricial, Array Processor, encontrado em arquitetura SIMD, executa tal processo.

*** Paralelismo Espacial Assíncrono: Onde os eventos não são simultâneos. Neste caso, diversos processadores trabalham em paralelo processando suas tarefas concorrentemente de forma assíncrona para certo intervalo de tempo. Esse tipo de concorrência é encontrada em multiprocessadores de arquitetura MIMD.**

Qual a motivação para o estudo de Sistemas Distribuídos?

O uso de sistemas de informação e redes de comunicação torna-se cada vez mais importantes; Com o aumento da complexidade dos sistemas em geral, os sistemas distribuídos passam a ser cruciais nas infra-estruturas de informação de hoje e do futuro;

Vantagens do uso de SDs

Comparando com os sistemas centralizados:

Apresentam economia já que microprocessadores oferecem melhor relação preço/desempenho do que os antigos mainframes; Velocidade à maior poder total da computação do que o mainframe; Distribuição inerente à máquinas espacialmente separadas; Confiabilidade à se uma máquina falha, o sistema como um todo pode ainda sobreviver; Crescimento incremental à poder computacional adicionado em incrementos;

Comparando com os sistemas isolados:

Compartilhamento de dados;
Compartilhamento de dispositivos;
Comunicação à e-mail, por exemplo;
Flexibilidade à distribuição de carga e adaptabilidade;
Desvantagens do uso de SDs
Pouco software disponível;
A rede pode causar problemas de comunicação;
Segurança à vários pontos de acesso.

Arquiteturas não paralelas

RISC - Reduced Instruction Set Computer: Caracteriza-se por executar instruções simplificadas (não exige um poder computacional muito alto) em curto período de tempo.

CISC - Complex Instruction Set Computer: Caracteriza-se por executar instruções complexas (exige alto poder computacional) em um período de tempo, ao menos, 10log-13 maior que o tempo gasto na arquitetura RISC.

Arquiteturas paralelas

SISD - Single Instruction stream Single Data stream: Possui uma unidade de processamento e uma unidade de controle.

SIMD - Single Instruction stream Multiple Data stream: Caracteriza-se por uma arquitetura vetorial onde a mesma operação é executada sobre múltiplos operandos.

MISD - Multiple Instruction stream Single Data stream: Caracteriza-se por executar um Pipeline em processadores cujos dados vão sendo processados e passados para outros processadores.

MIMD - Multiple Instruction stream Multiple Data stream: Várias instruções podem ser executadas ao mesmo tempo em unidades diferentes controladas por unidades de controle independentes.

Tolerância a Falhas em SDs

Modelos de Tempo

Modelos síncrono e Assíncrono podem ser considerados extremos teóricos. Outros modelos misturam características desses extremos.

Sistema Assíncrono

Não existe limite de tempo

Impossível alcançar consenso em sistemas assíncronos

Sistema Síncrono

Existe um limite de tempo finito e conhecido

Falha de componentes pode ser deduzida pela ausência de resposta

Timeout - Detecção de defeitos em nós e perda de mensagens

Exemplos de Falhas

Processador (Crash)

Rede de comunicação (todos os tipos)

Clock (temporização)

Meio de armazenamento (temporização, omissão ou resposta)

Software (resposta)

O modelo de falhas é uma simplificação da realidade

Modelo Lógico

Aplicação distribuída

Conjunto de processos concorrentes

Processos cooperativos

Cada processo é sequencial

Cada processo pode estar em um nó diferente

Progresso Finito

Rede completamente conectada

Topologia não é considerada

Processos

Canais

Middleware

É uma camada de software de distribuição que abstrai a complexidade e a heterogeneidade do ambiente distribuído sobre uma rede com diversas tecnologias, arquiteturas de máquinas, sistemas operacionais e linguagens de programação;

Possui o objetivo de projetar, programar e gerenciar aplicações distribuídas através do provimento de um simples, consistente e integrado ambiente para programação distribuída;

Diferentes plataformas de *middleware* suportam diferentes modelos de programação para objetos distribuídos.

Middleware Baseado em Objetos

Um dos mais populares modelos de *middleware*;

Ambiente de programação

Distribuição + orientação a objetos;

Utilização de IDL's;

Serviços

A comunicação poderá ser síncrona ou assíncrona;

Assume a utilização de um transporte confiável à Protocolo TCP.

Transmissão de Dados

Dados em programas são estruturados enquanto que mensagens carregam informação sequencial:

Necessidade de linearização/restauração de dados;

Heterogeneidade na representação de dados em computadores:

Uso de um formato externo comum;