

Sistemas Distribuídos - Introdução

Prof. DSc Marcelo
Lisboa Rocha



O que é um Sistema Distribuído?

- A definição varia entre autores, muitas vezes é mais uma “idéia” do que uma definição
- Redes de Sistemas Computacionais:
 - Infra-estrutura servindo um conjunto de sistemas computacionais (SC) interconectados através de links de comunicação de diversas topologias e mídias.
- Sistema Distribuído:
 - Composto de vários sistemas computacionais (SC) se comunicando através de uma rede, onde os SC abrigam conjuntos de processos que se comunicam através de um conjunto de protocolos distribuídos para assistir a execução coerente das atividades distribuídas.
- Internet: rede ou sistema distribuído?

O que é um Sistema Distribuído?

- Segundo Andrew Tanenbaum: “Um sistema distribuído é uma coleção de computadores independentes que parecem ao usuário ser um computador único”.
- Dois aspectos na definição do Tanenbaum: hardware e software.
- Hardware: computadores independentes.
- Software: ilusão de um único computador.

O que é um Sistema Distribuído?

- Resumo: uma coleção de sistemas computadorizados (computadores, PDA's, celulares) autônomos com as seguintes características:
 - Ligados por uma rede
 - Utilizam software para produzir uma capacidade de computação integrada.

O que é um Sistema Distribuído?

- Considerando a existência de vários processadores, temos:
 - Multiprocessador: compartilhamento de recursos físicos, tal como memória principal, disco e relógio (fortemente acoplado).
 - Multicomputador: computadores com seu hardware básico independente, interconectados por rede de comunicação de alta velocidade (acoplamento por proximidade)
 - Sistemas Distribuídos: atrasos de comunicação (rede) significativos e incertos (fracamente acoplado)

O que é um Sistema Distribuído?

- Características:
 - Probabilidade independente de falha
 - Comunicação é não-confiável, com atraso variáveis (solução QoS)
 - Velocidade e banda moderadas (aumentando)
 - Dificuldade em determinar a ordem de ocorrência de eventos e computar estado global
 - Um nodo se comunica com outro apenas através de troca de mensagens
 - Pode não ser possível determinar o estado global do sistema

Qual o tamanho de um Sistema Distribuído?

- De acordo com a topologia de rede
 - Local Area Network (LAN/Rede Local)
 - Metropolitan Area Network (MAN/Rede Metropolitana)
 - Wide Area Networks (WAN/Rede de Longa Distância) —→ Internet

Exemplos de Aplicações Distribuídas Atuais

- UNIX (SO) Distribuído
- Aplicações WAN (internet)
 - E-mail.
 - BBS/Chatt
 - WWW
 - Teleconferência
 - Acesso a informações multimídia

Características Chave de Sistemas Distribuídos

1. Compartilhamento de Recursos
2. Gerenciamento de Recursos
3. Seguir Padrões Abertos
4. Escalabilidade
5. Tolerância a Falhas
6. Transparência

Compartilhamento de Recursos

- Recursos
 - hardware: discos e impressoras
 - software: arquivos, objetos e/ou banco de dados
- Compartilhando Hardware por:
 - conveniência
 - redução de custo
- Compartilhando Dados por:
 - consistência – compiladores e bibliotecas
 - troca de informações – banco de dados
 - Trabalho cooperativo - groupware

Gerenciamento de Recursos

- Módulos de software que gerenciam um conjunto de recursos.
- Cada recurso exige sua própria política e método de gerenciamento.
- Modelo cliente-servidor: processos do servidor agem como gerentes de recursos para um conjunto de recursos e um conjunto de clientes.
- Modelos baseados em objetos (recursos são objetos que podem ser movidos)

Seguir Padrões Abertos

- A forma como pode ser estendido.
- Aberto ou fechado com respeito a: hardware ou software.
- Aberto:
 - Especificações e interfaces públicas.
 - Interfaces padronizadas.
- Unix e Linux são SO's relativamente abertos:
 - A linguagem C está disponível
 - Chamadas de sistema documentadas
 - Novos drivers de hardware são fáceis de adicionar
 - Aplicações são independentes de hardware.

Concorrência

- Multi-programação
- Multi-processamento
- Execução em paralelo em sistemas distribuídos
 - Muitos usuários usando o mesmo recurso
 - Interações das aplicações
 - Muitos servidores respondendo a pedidos dos clientes

Escalabilidade

- Categorização
 - Sistemas pequenos: poucos computadores e um servidor de arquivos numa única rede
 - Sistemas grandes: internet atual
- Características
 - O software não deve mudar para suportar escalabilidade (aumento do número de participantes)
 - Área de pesquisa – para redes grandes e de alto-desempenho.

Escalabilidade

- Cuidados
 - Evite centralização para suportar escalabilidade
 - Escolha o seu esquema de nomeação e/ou numeração cuidadosamente
 - Manipule problemas de tempo com cacheamento (caching) e replicação de dados.

Tolerância a Falhas

- Computadores falhas, por isto necessitamos:
 - Redundância de hardware
 - Recuperação de/por software
- Para que?
 - Aumentar a disponibilidade dos recursos/serviços
 - A rede normalmente não é redundante

Transparência

- Um SD necessita de transparência de:
 - Acesso
 - Localização
 - Concorrência
 - Replicação
 - Falha
 - Migração
 - Desempenho
 - Escalabilidade

Sistemas Distribuídos X Processamento Paralelo

- **Sistemas Distribuídos:**
 - Sistemas e aplicações em geral cujas peças executam em múltiplos computadores autônomos interligados por uma rede de comunicação.
- **Processamento Paralelo:**
 - Um problema que exige alto poder de processamento é “quebrado” em partes que são computadas ao mesmo tempo (ou seja, em paralelo).
 - Processamento de alto desempenho é área de computação científica.

Quando Distribuir?

- Se não precisar, não distribua.
- Problemas de natureza descentralizada:
 - Rede de manufatura com atividades concorrentes de engenharia em locais remotos.
 - Teleconferência ou CSCW (suporte computadorizado a trabalho cooperativo).
 - Automação industrial.

Quando Distribuir?

- Quando técnicas de distribuição são artefatos úteis da solução:
 - Um banco com base de dados transacional.
 - Distribuição para aumentar o desempenho.
 - Replicação de dados para aumentar disponibilidade.
- Adaptação e evolução na atividade/localização das organizações:
 - Adaptabilidade a um cenário de negócios extremamente dinâmico (modularidade, expansibilidade e escalabilidade).

Prós e Contrás de SD's

- Melhor relação custo/desempenho de conjunto de CPUs.
- Escalabilidade de processamento: sem limite na capacidade computacional de uma única CPU.
- Menor tempo de resposta, menor custo de comunicação e maior disponibilidade em aplicações com distribuição inerente.
- Aplicações como CSCW, jogos e treinamento.

Prós e Contrás de SD's

- Aplicações de “alta confiabilidade” e “alta disponibilidade”, obtidos via replicação de dados e processamento.
- Flexibilidade, crescimento incremental de um sistema.
- Compartilhamento de dados e recursos de hardware.

Prós e Contras de SD's

- Falta de experiência em projetar, implementar e usar software distribuído.
- Falta de um controle central.
- Depende do funcionamento correto da rede.
- Segurança: oferece uma oportunidade para violação.

Para Próxima Aula

- Cite as principais diferenças entre sistemas centralizados e distribuídos.
- Por que atualmente os meios de interligação (links) entre os sistemas computacionais são considerados “gargalos” para os SD’s?
- Como os SD’s podem melhorar a relação custo/desempenho de um conjunto de recursos computacionais?
- Em qual(is) situação(ões) uma empresa se beneficiaria em utilizar SD’s?
- Em alguma disciplina do curso você já viu algum sistema/método/técnica/ algoritmo que seja distribuído? Cite qual e como o mesmo trabalha.
- Se não precisar, não distribua. Qual a sua opinião sobre este conceito? Seja explícito e se possível, utilize-se de recursos do material fornecido ou outros para embasar a sua resposta.