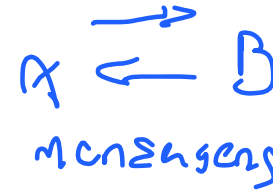


Conceitos de Sistemas Distribuídos

Wednesday, March 2, 2016 13:43

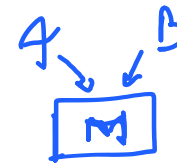
Qual é, para esta disciplina, conceito de sistema distribuídos?

- Coleção de processos, executando em processadores autônomos, que se comunicam usando trocas de mensagens sobre uma rede, cooperando para realizar tarefas específicas.
 - o Uma alternativa para a comunicação entre os processadores é a memória compartilhada.



Quais tarefas são executadas pelos sistemas distribuídos?

- Variam imensamente!
 - o Desde aplicações específicas
 - Por exemplo: sistema de reserva de passagem aérea, sistemas bancários
 - o Até plataformas de computação distribuída de propósito geral
 - Por exemplo: CORBA
 - Jgroups: ferramenta para construir um paradigma de grupos em sistemas distribuídos
 - Ensemble
 - Zookeeper: consenso distribuído



É possível criar sistemas distribuídos utilizando sockets, mas dá um grande trabalho. Plataformas como as acima fornecem abstrações para simplificar a criação de um sistema distribuído. Antigamente haveria a "utopia" de que um paradigma seria suficiente, mas aos poucos surgiram outros mais específicos.

Existem versões distribuídas de

- Sistemas operacionais
- Bancos de dados
- Etc.

- O foco desta disciplina está em algoritmos clássicos, que são usados em todos os contextos de sistemas distribuídos.
- Por exemplo, algoritmos de diagnóstico de falha. Necessário para qualquer sistema distribuído.
 - Os processos executando esses algoritmos estão na camada de aplicação do modelo TCP/IP. Pode ser necessário mudar isso dependendo do contexto, mas isto simplifica o entendimento.
 - o Ou ainda acima da camada de aplicação, sobre um *middleware*.

Características dos Sistemas Distribuídos

1. Compartilhamento de Recursos

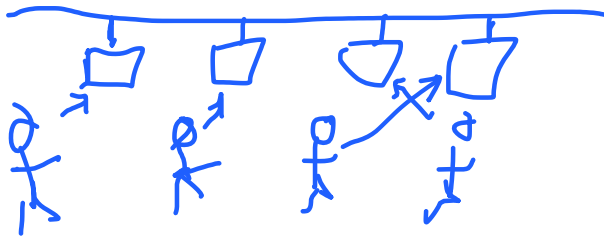
Recursos variam imensamente, desde *hardware* até dados.

- É preciso um sistema de nomes (identificadores)
- Acesso concorrente

2. Escalabilidade

O sistema escala?

- Aumentando o sistema ou o tamanho do problema, ainda é possível resolver o problema?
 - Em alguns casos, não é possível. Um problema pode se tornar tão grande que a quantidade de recursos necessária é impossível.
- Qual é o gargalo do sistema?
 - O que está impedindo a escalabilidade do sistema? CPU, memória, disco, rede?
 - Uma solução para resolver os gargalos é o uso de réplicas.
 - A replicação tem como requisito básico a garantia de consistência.



- Como garantir a integridade dos dados em réplicas recebendo requisições simultâneas?

3. Transparência

O sistema é visto pelo usuário como se fosse um sistema centralizado, ou é um conjunto de partes independentes?

Modelos de sistemas distribuídos

Modelo: representação de um sistema real que elimina todos os detalhes que não são relevantes para o objetivo do estudo.

Objetivos diferentes -> modelos diferentes.

Um modelo é sempre uma **simplificação** do sistema real.

Em sistemas distribuídos, os processos executam um algoritmo distribuído (protocolo) que faz uso de um conjunto de mensagens; as ações executadas por um processo dependem da sequência de mensagens que ele recebe.

- Os modelos de falhas são dos mais importantes, além dos **modelos temporais**.

Sistemas distribuídos **síncronos**: Existe um limite máximo de tempo **conhecido** para a transmissão de uma mensagem entre processos e existe um limite de tempo **conhecido** para um processo executar uma ação. Praticamente impossível numa rede convencional.

Praticamente impossível numa rede convencional.

Sistemas distribuídos **assíncronos**: Não existem quaisquer premissas sobre os limites de tempo, nem de transmissão de mensagens, nem de execução de ações.

Sistemas distribuídos **parcialmente síncronos**:



São muitos modelos de sistemas parcialmente síncronos. Por exemplo, saber o limite para a troca de mensagens e não o da execução de ações, ou vice-versa.

Problema para a próxima aula:

Um problema importante:

A coordenação sobre um canal de comunicação não-confiável.

Considere um sistema que consiste de dois processos A e B que nunca falham, mas se comunicam sobre um canal que, eventualmente, perde mensagens.

Pense uma solução distribuída para que ambos os processos:

(A) Ou executam a mesma ação dentre duas ações possíveis (α, β)

(B) Ou executam nenhuma ação.

Programa de simulação: SMPL

