Sumário

- Introdução.
- Aspectos.
- Propriedades.
- Dificuldades.

Características dos sistemas distribuídos

- Diversos computadores:

  Com diferentes processadores, memórias, discos, e dispositivos de entrada/saída.

- Interligações:

  As diferentes máquinas podem comunicar entre si.

- Estado partilhado:

  Os computadores cooperam para manter um estado partilhado.

  Existem requisitos de coerência global.

Aspectos dos sistemas distribuídos

- Modos de falha independentes: *um nó não é necessáriamente afectado pela falha de outro nó (infelizmente, isto raramente é verdade).*

  "Um sistema distribuído é aquele que nos impede de trabalhar devido à falha de que nem sequer ouvimos falar." [L. Lamport].

- Comunicação sujeita a falhas: *ruído, partições.*
- Comunicação insegura: *escuta não autorizada.*
- Comunicação de elevado custo: *latência e débito.*

Sistemas distribuídos vs sistemas centralizados

- Segurança:

  Centralizado é fácil de controlar o acesso mas possui pontos únicos de falha.

- Disponibilidade:

  Quanto mais máquinas maior a probabilidade de uma falha.

  É possível aumentar a disponibilidade.

Propriedades dos sistemas distribuídos

- Heterogeneidade: *maquinaria diferente, usando programas diversos.*
- Dispersão geográfica: *máquinas em locais fisicamente dispersos.*
- Interligação por uma rede: *características bastante diferentes.*
- Serviços uniformes: *o mesmo tipo de serviço deve ser acessível em qualquer ponto.*
- Propriedades globais: *propriedades válidas em todos os nós: nomes, acesso, segurança, etc.*

Propriedades globais

- Nomes globais: *nome de máquina, ficheiro, utilizador, é o mesmo em todo o sistema.*
- Acesso global: *as mesmas funcionalidades estão acessíveis em qualquer ponto com desempenho razoável.*
- Segurança global: *o sistema de autenticação funciona em todos os pontos.*
- Gestão global: *a mesma pessoa pode gerir o sistema de qualquer ponto.*
- Disponibilide global: *os mesmos serviços funcionam mesmo após algumas falhas.*

Serviços

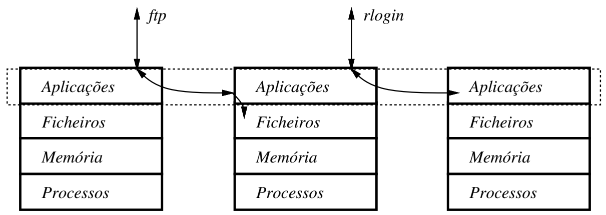
- Serviço de nomes: *emparelhamento entre padrões binários.*
- Chamada a procedimentos remotos: *mecanismo "normalizado" de aceder a serviços remotos.*
- Registo de utilizadores: *capacidade de autenticar e acrescentar novos utilizadores.*
- Tempo: *disponibilidade de uma referência de tempo.*
- Ficheiros: *acesso global (usando replicação se necessário).*
- Gestão: *serviços de gestão global.*

Evolução: técnicas

- Partilha de recursos:

  Sessão remota (rlogin).

  Transferência de ficheiros.

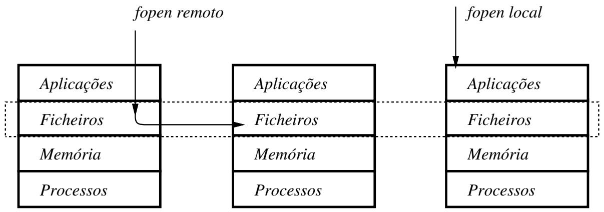


Evolução: técnicas

- Ficheiros Distribuídos:

  Evitar a sessão remota para transferência explicita de ficheiros.

  Partilha eficaz de informação.

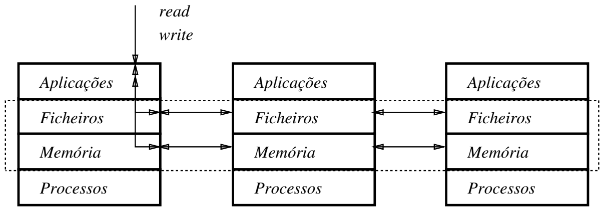


Evolução: técnicas

- Memória distribuída:

  Beneficiar dos recursos de uma rede de máquinas..

  Suportar memória partilhada (distribuída).

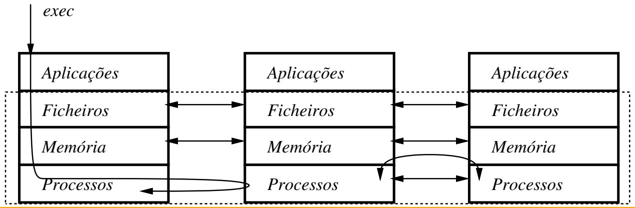


Evolução: técnicas

- Processamento distribuído:

  Actividades que "passam" de máquina para máquina.

  Actividades concorrentes sincronizadas sobre memória distribuída.



Evolução: técnicas

- Aplicações intensamente distribuídas requerem:

  Suporte de sistema adequado.

  Ambientes de desenvolvimento.

Evolução: sistemas

- Arpa Internet: *protocolos, comunicação fiável, serviços distribuídos.*
- Alto/Ethernet: *modelo/cliente; RPC; redes difusão.*
- Locus: *sistema de ficheiros UNIX distribuído.*
- Apollo: *memória virtual distribuída.*
- Tandem: *sistema duplex com disponibilidade elevada.*
- VAXCluster: *time-sharing expansível incrementalmente.*

Evolução: sistemas (continuação)

- Andrew/Athena: *Sistema de ficheiros de grande-escala (AFS).*
- Sprite: *sistema ficheiros distribuído com cache coerente.*
- Coda, Echo: *sistema de ficheiros distribuído com replicação e operação desligada.*
- OSF DCE: interfaces standard cliente/servidor, basedas em RPC.

Projecto ANSA

- Advanced Network System Architecture: modelo de referência Open Disdributed Processing (ODP).
- Modelo distribuído orientado para objectos.

  Acesso a funções do sistema e informação partilhada por invocação remota ou local: cliente-servidor. Clientes e servidores encapsulados por objectos.

  Heterogeneidade e particularidades das máquinas e sistemas operativos escondidadas por abstracção de máquina virtual (a infraestrutura).

  Transparência da distribuição e efeitos associados (ex. replicação, localização, concorrência, etc.

"Transparências" da distribuição

- Transparência de acesso:

  Acesso a objectos remotos (ficheiros, dados) igual ao acesso a objectos locais.

  Comunicação, segurança, fiabilidade e pontualidade.

- Transparência de localização:

  Acesso a objectos, onde quer que estejam, do mesmo modo.

  Nomes, endereços, encaminhamento.

"Transparências" da distribuição

- Transparência de concorrência:

  Acesso concorrente a objectos de modo serializável.

  Sincronização, ordenação.

- Transparência de replicação:

  Acesso a objectos replicados como se fossem um só.

  Algoritmos de gestão da replicação.

"Transparências" da distribuição

- Transparência de faltas:

  Operações correctamente terminadas, na óptica do utilizador, a despeito de faltas (falhas de componentes).

  Processamento de erros, tratamento de faltas.

- Transparência de migração:

  Distribuição dos objectos pelo sistema, e a sua redistribuição durante o funcionamento, transparentemente ao utilizador.

  Configuração (estática) e reconfiguaração (dinâmica).

"Transparências" da distribuição

- Transparência de desempenho:

  Reajuste automático do sistema para compensar variações de carga.

  Redistribuição de carga.

- Transparência de escala:

  Estrutura do sistema, nomeadamente os algoritmos e subsistemas de suporte (comunicação, nomes, etc), é independente da dimensão.

  Independência da dimensão (geográfica, numérica, etc).

"Transparências": vantagens e desvantagens

- Vantagem:

  Manutenção de ambiente integrado de fácil gestão.

- Desvantagens:

  Impossível de conseguir na totalidade.

  Não permite ao programador controlar as relações custo/benefício.

Sincronismo

- Sistemas assíncronos:

  Não existem limites ao tempo de execução.

  Não existem limites ao tempo de trânsito das mensagens.

- Sistemas síncronos:

  Estes limites existem.

Problema dos sistemas assíncronos

- O nó x não responde:

  Estará falhado ou será a ligação que está lenta?

  Impossível distinguir estas duas situações.

- Os sistemas síncronos não padecem deste problema:

  Existe um tempo máximo para a execução de uma acção.

Partições

- Um conjunto de nós pode comunicar entre si mas não consegue comunicar com os restantes nós fora desse conjunto.

  Os restante nós estão falhados ou a rede está falhada?

  Impossível distinguir.

  Em certos casos, não se pode progredir sem correr o risco de tornar o sistema incoerente.

- O que fazer? Parar, continuar numa única partição ou continuar em várias partições?

Filiação

- Detecção de falhas:

  Ausência de comunicação periódica.

  Ausência de resposta a um pedido.

- Acordo acerca das falhas:

  Diferentes nós podem detectar falhas em momentos diferentes.

  Pode resultar em incoerência no sistema.

Problema do acesso a informação

- Informação num único local:

  Quando falha, perde-se a informação.

  Quando se está longe, perde-se muito tempo a obter a informação.

- Informação replicada.

  Quando se actualiza é necessário garantir a coerência das cópias.

  Exemplo: levantamentos/depósitos em conta bancária.

Problema da ordenação de mensagens

- Exemplo:

  Mensagem do nó A: "O nó X falhou".

  Mensagem do nó B: "O nó X recuperou".

  X está falhado ou recuperado?

- Tempos diferentes na propagação de mensagens:

  É necessário ordenar as mensagens.

  Ordem lógica (relativamente umas às outras).

  Ordem temporal.

Problemas induzidos pela distribuição

- Coerência:

  Como o utilizador vê o sistema.

- Ordenação de eventos, acordo:

  A ordem pela qual os eventos são recebidos corresponde à ordem das causas?

  Conseguem todos os processos acordar num valor?

Problemas induzidos pela distribuição

- Confirmação atómica:

  Acordar num valor acerca da confirmação/ cancelamento de uma transacção.

  Assimétrico: um nó pode decidir independentemente, abortar a transacção.

- Exclusão mútua distribuída:

  Só um nó pode aceder a um dado recurso.

- Eleição:

  Escolher um nó para realizar uma acção (ou ser coordenador dessa acção).

Canais exteriores ao sistema

- Através do ambiente.

  Utilizadores que se deslocam.

  Acções que se propagam no ambiente físico.

  Passagem no tempo.

Resumo

- Sistema distribuído:

  Modos de falha independente e estado global.

- Problemas:

  Coerência, acordo, ordenação, eleição, confirmação atómica.