

Consistência e Replicação

CAPÍTULO V

Consistência e Replicação

Os dados são replicados para melhorar a confiabilidade ou melhorar o desempenho.

Mas como se manter as réplicas consistentes?

Consistência e Replicação

Razões para a replicação

- Confiabilidade;
- Desempenho.

Consistência e Replicação

Confiabilidade

“... dados são replicados para aumentar a confiabilidade de um sistema”

- Um sistema replicado pode, em caso de falha, funcionar a partir de uma réplica;
- É mais difícil perder dados.

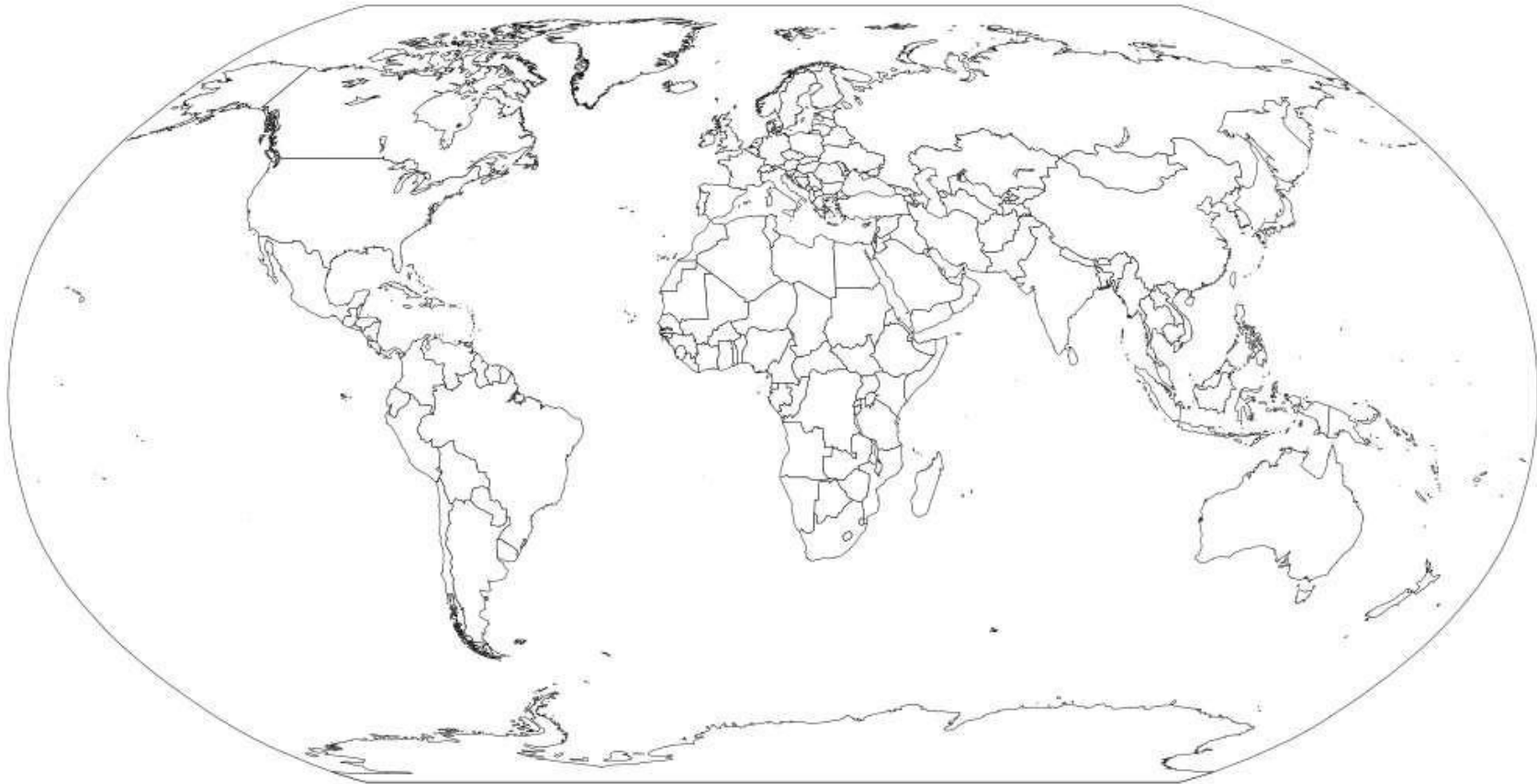
Consistência e Replicação

Desempenho

“Replicação para melhoria de desempenho é importante quando um sistemas distribuído precisa ser aumentado em quantidade e área geográfica”

- **Quantidade:**
 - Acontece quando um número maior de processos precisa aceder a dados que são geridos por um único servidor;
 - Aumento do desempenho por replicação do servidor.
- **Geográfica:**
 - Colocação de uma réplica mais próximo ao local de acesso ao processo em execução;

Consistência e Replicação



Consistência e Replicação

Desempenho

- Problemas com a replicação:
- Manter várias cópias exige um esforço para manter a consistência;
- A modificação de um cópia a torna-a diferente das restantes;
- O preço da replicação é determinado “quando e como é que essas modificações precisam de ser executadas”;

Consistência e Replicação

Desempenho

- Exemplo - acesso a uma página web:
 - Os Browsers Web podem armazenar localmente uma cópia de uma página acedida previamente (cache);
 - O utilizador se pedir novamente a mesma página, o browser utiliza a cópia local;
 - O tempo de acesso é excelente;
 - Mas, a página pode ter sido atualizada e ainda não está propagar na cache local;

Consistência e Replicação

Replicação como técnica de crescimento

- Replicação e cache para melhoria de desempenho são muito utilizadas:
 - Usar cópias locais aumenta a percepção de maior largura de banda.
- No entanto, é necessário uma análise prévia:
 - Justifica colocar um réplica tão próxima?
 - A estratégia de atualização de réplica é eficiente?

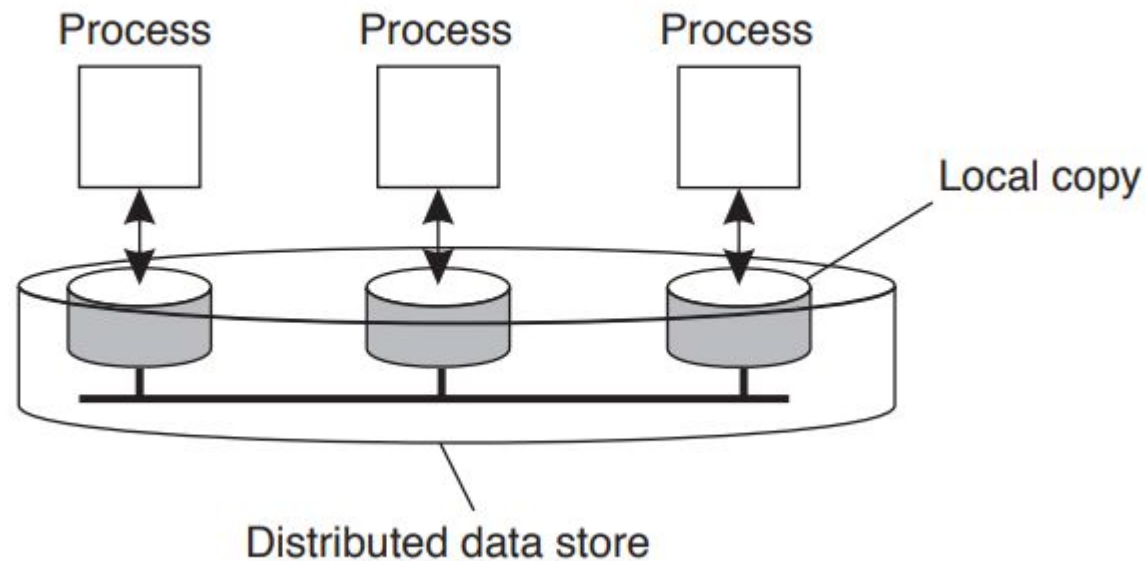
Consistência e Replicação

Consistência ou Escalabilidade?

- Consistência com replicação síncrona:
 - Atualizar todas as cópias com uma única transação.
 - Como?
 - Necessário definir quando ...
 - Pode trazer um mau desempenho ...
 - Carece de sincronização global ...
 - Solução:
 - Permitir que nem todas as caches estejam sincronizadas => consistência de dados

Consistência e Replicação

Consistência de dados



Consistência e Replicação

Consistência de dados

- Uma operação de dados é classificada como uma operação de escrita quando altera os dados, caso contrário é classificada como uma operação de leitura”.
- “Um modelo de consistência é um contrato entre os processos e o data store. Ele diz que, se os processos concordarem em obedecer a certas regras, o depósito promete funcionar de maneira correta.
- Como definir com precisão qual operação é a última operação de escrita a ser realizada?
- Consistência sequencial e consistência causal

Consistência e Replicação

Consistência Sequencial

- O resultado de qualquer execução é o mesmo que seria se as operações (de leitura e escrita) realizadas por todos os processos no data store fossem executados na mesma ordem sequencial que individualmente no seu programa.

Consistência Casual

- Idêntico ao anterior, mas diferencia eventos conflitantes de não conflitantes.

Consistência e Replicação

Operações em grupo

- Não é o relevante o resultado de cada operação singular. Apenas é relevante o resultado final de um conjunto de operações.

Consistência e Replicação

Modelo de consistências centrados no cliente

- Mostrar como evitar uma consistência global, focando no que clientes específicos precisam, em vez do que deveria ser mantido pelos servidores
- Inconsistência aceitável:
 - Utilização de cache. O utilizador ter acesso a páginas Web desatualizadas.
 - Mas seria aceitável dados de um monitor cardíaco desatualizados?
 - O que seria considerado não aceitável numa informação sobre tropas inimigas?

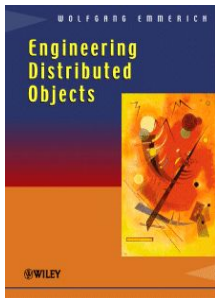
Consistência e Replicação

Gestão de réplicas

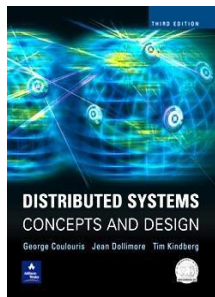
Uma questão fundamental para qualquer sistema distribuído que suporte replicação é decidir onde, quando e por quem é que as réplicas devem ser posicionadas, e quais os mecanismos para manter as réplicas consistentes.

- Posicionar servidores de réplicas:
 - Encontrar as melhores localizações para colocar um servidor que pode ser um data store (ou parte dele).
- Posicionar conteúdo:
 - Quais os melhores servidores para colocar o conteúdo.

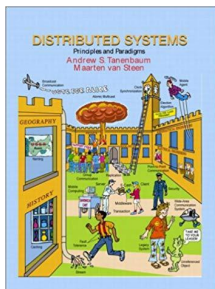
Bibliografia



From: Wolfgang Emmerich
Engineering Distributed Objects
John Wiley & Sons, Ltd 2000



From: Coulouris, Dollimore and Kindberg
Distributed Systems: Concepts and Design
Edition 4 © Addison-Wesley 2005



From: Andrew S., Tanenbaum and Van Steen, Maarten
Distributed Systems: Principles and Paradigms
Edition 2 © Pearson 2013

Questões?