

## \*.\* Distribuição de informação – Replicação \*.\*

### Resumo

- Objetivo simples: uma cópia dos mesmos dados em várias máquinas
- A replicação pode servir vários propósitos
  - Alta disponibilidade
  - Operação desconectada
  - Latência
  - Escalabilidade
- A replicação é um problema complicado
  - exige pensar cuidadosamente sobre a concorrência e sobre todas as coisas que podem dar errado
  - lidar com as consequências dessas falhas
- Três abordagens principais para a replicação:
  - Single-leader replication
  - Multi-leader replication
  - Leaderless replication
- Cada abordagem possui vantagens e desvantagens
  - A replicação Single-leader é popular porque é bastante fácil de entender e não há resolução de conflitos para se preocupar
  - A replicação Multi-leader e sem liderança pode ser mais robusta na presença de nós defeituosos, interrupções de rede e espiras de latência - com o custo de ser mais difíceis de argumentar e fornecendo apenas garantias de consistência muito fracas
- A replicação pode ser síncrona ou assíncrona
  - opção com efeito profundo sobre o comportamento do sistema quando falha
  - a replicação assíncrona pode ser rápida quando o sistema está funcionando sem problemas
  - Se um líder falhar, dados recentemente committed podem ser perdidos em atualização assíncrona de seguidores
- Discutido alguns modelos de consistência que são úteis para decidir como uma aplicação se deve comportar em atraso de replicação:
  - consistência de leitura e gravação: um utilizador deve sempre ver os dados que eles se submeteram
  - Leituras monotônicas: depois que um utilizador viu os dados em um ponto no tempo, eles não deveriam ver os dados de um ponto anterior
  - O prefixo consistente de leitura: os utilizadores devem ver os dados em um estado que faz sentido causal, por exemplo, ver uma pergunta e sua resposta na ordem correta
- Discutido os problemas de concorrência que são inerentes às abordagens de replicação multi-líder e sem liderança