

Replicação de objetos genéricos

Qual a estrutura de dados replicada?

- Registo
 - Vimos atrás (algoritmo ABD)
- Espaço de tuplos
 - Vimos atrás (algoritmo Xu-Liskov)
- Qualquer estrutura de dados, oferecendo uma qualquer interface remota
 - Exige algoritmos de replicação genéricos
 - Não são tão otimizados como os anteriores, pois não podem recorrer a otimizações específicas de uma estrutura de dados
 - Tema desta aula

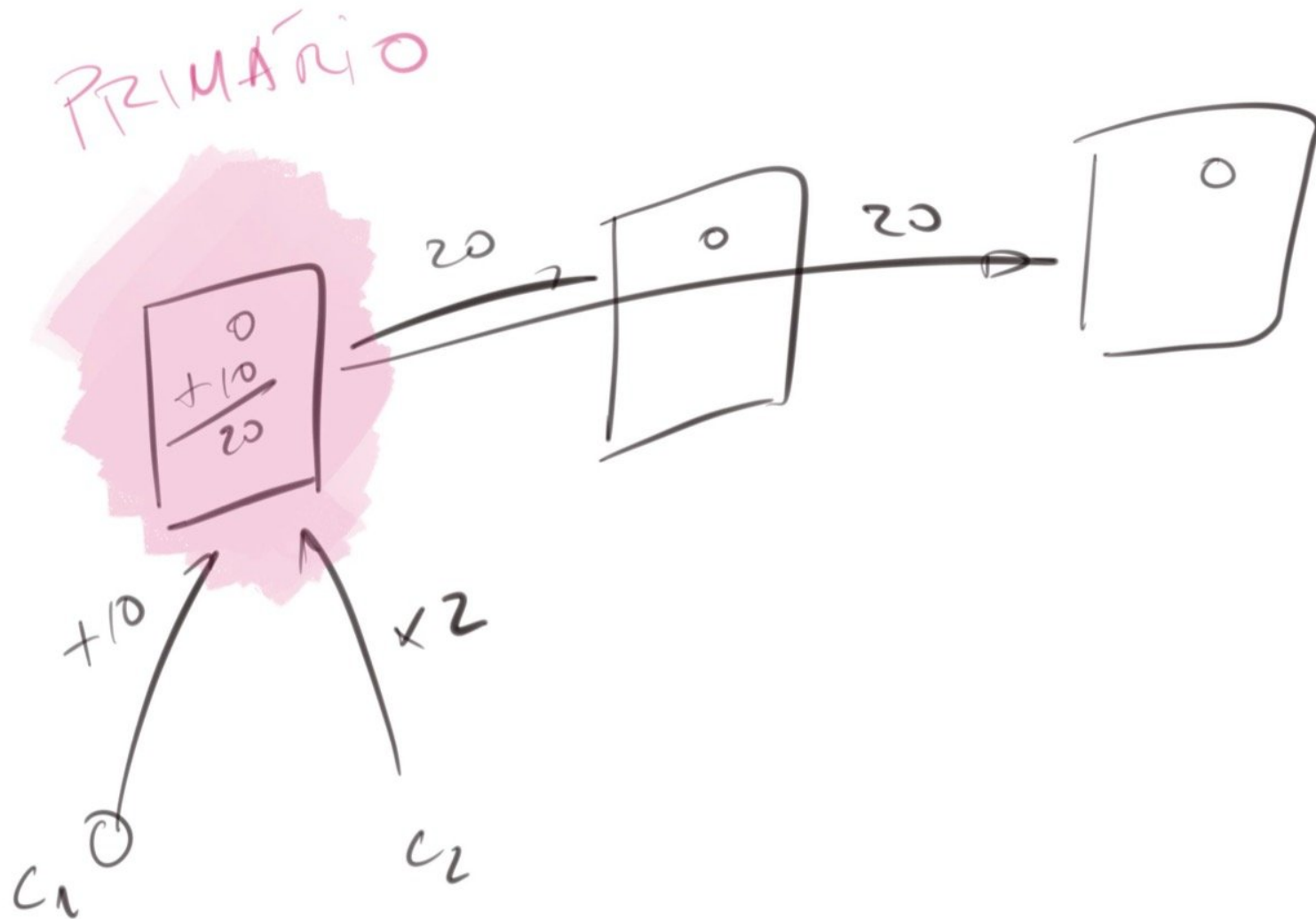
Tentemos construir algoritmos de replicação genéricos

- Duas variantes principais
 - Primário-secundário
 - Replicação de máquina de estados
- Objetivo: assegurar **linearizabilidade** (coerência forte)

Primário-secundário

- Um processo é eleito como primário
- Os clientes enviam pedidos ao primário
- Para cada pedido recebido, primário:
 - Executa o pedido
 - Propaga o estado para os secundários e aguarda confirmação de todos
 - Responde ao cliente
 - (E processa o próximo pedido que esteja na fila de pedidos)

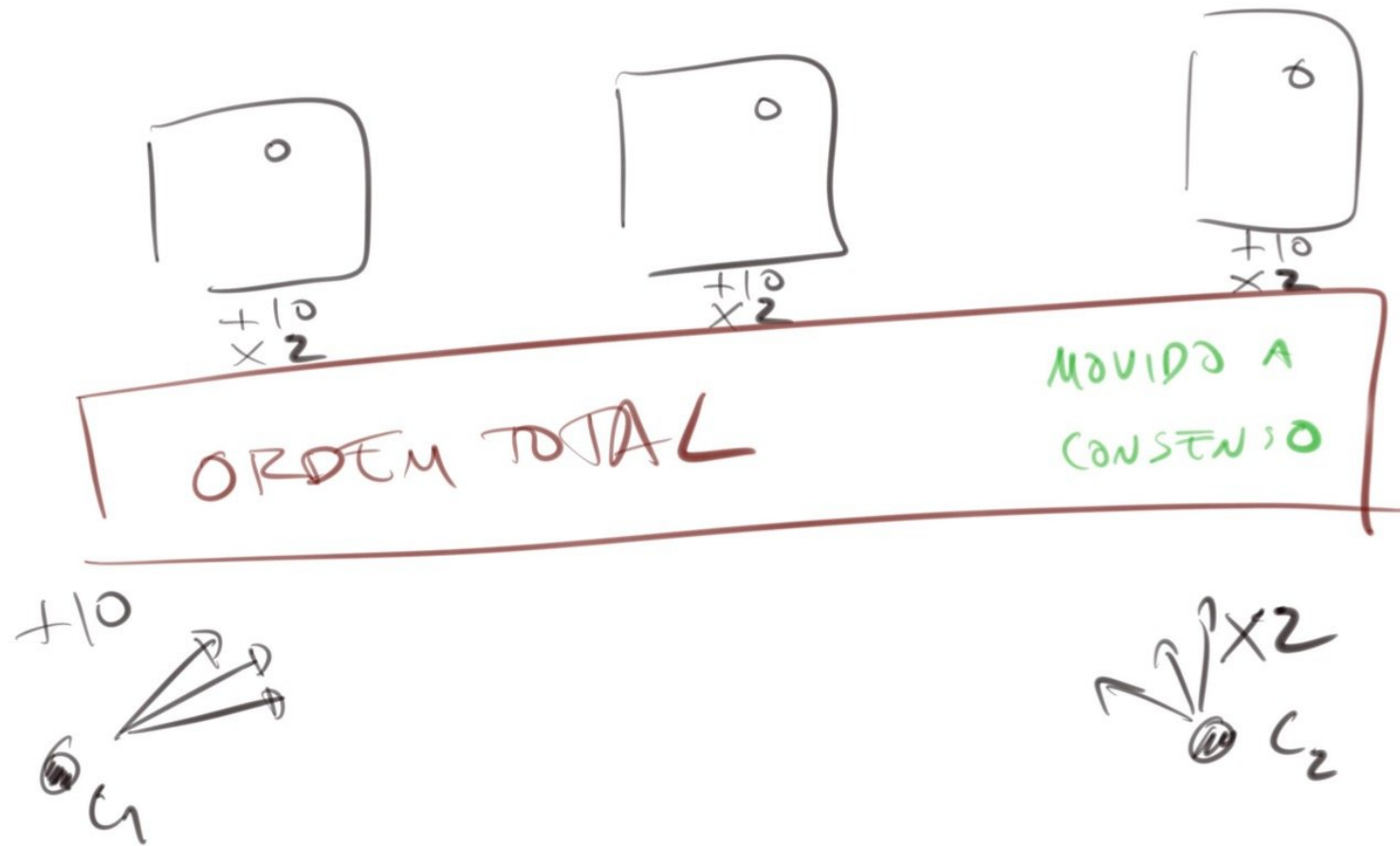
Primário-secundário



Replicação de máquina de estados

- Os clientes enviam os pedidos para todas as réplicas
- Todos os pedidos são ordenados por ordem total
- Todas as replicas processam os mesmos pedidos, pela mesma ordem
 - Assume-se que operações são determinísticas => replicas ficarão idênticas

Replicação de máquina de estados (um esboço)



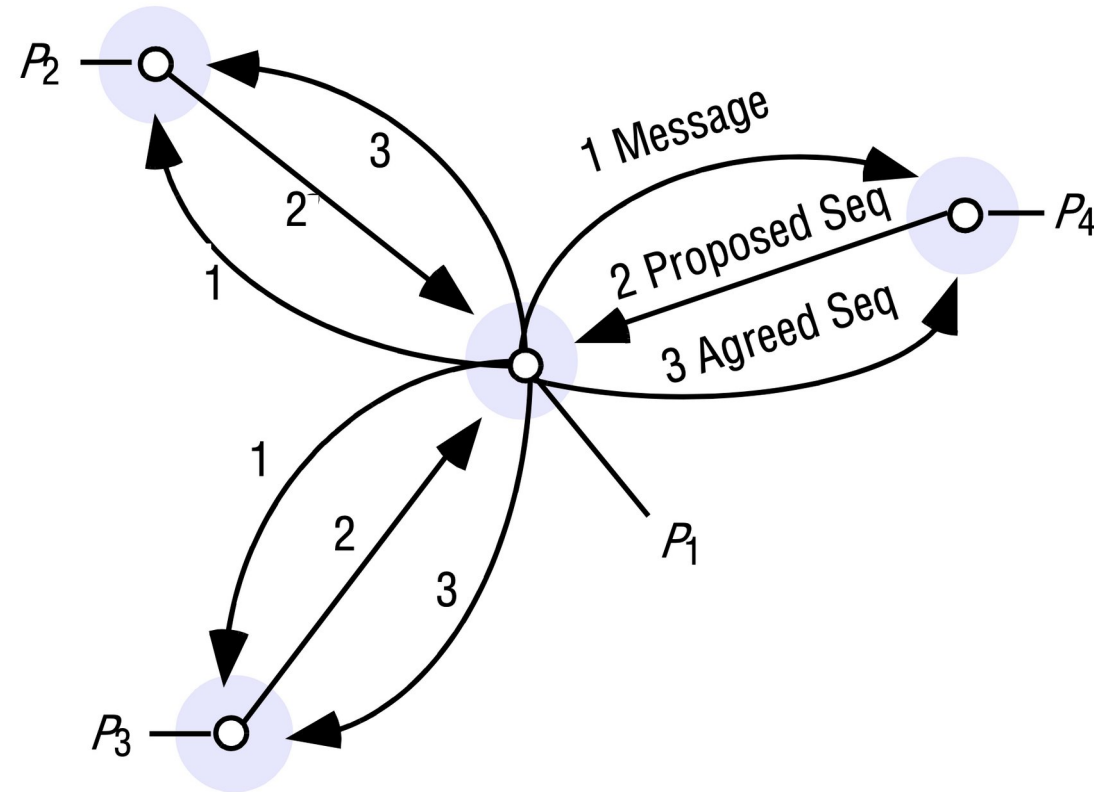
Dois algoritmos para ordem total

- Ordem total baseada em sequenciador
- Ordem total baseada em acordo colectivo

Ordem total baseada em sequenciador

- As mensagens são enviadas para todas as réplicas usando um algoritmo de difusão fiável
- Uma das réplica é eleita líder, decide qual a ordem pela qual as mensagens devem ser processadas, e envia esta informação para as réplicas restantes

Acordo coletivo



Algoritmo explicado nas páginas 671-672 do livro

Sequenciador vs. acordo coletivo

- Cenário I: um objeto replicado por um dado grupo de réplicas
- Qual a solução mais escalável?

Sequenciador vs. acordo coletivo (II)

- Cenário II: múltiplos objetos, potencialmente replicados em grupos de nós diferentes
 - Atenção: um nó pode replicar mais que um objeto, logo pertencer a mais que um grupo!
- Usando cada algoritmo, como assegurar que cada nó entrega as mensagens (dos objetos que replica) de acordo com ordem total global?
 - Com sequenciador: precisamos de um sequenciador global, que processa todas as mensagens.
 - Acordo coletivo: este algoritmo é correto mesmo que cada mensagem seja enviada para um conjunto diferente de nós.
- Qual a solução mais escalável neste cenário?

Primário secundário vs RME

- Primário-secundário
 - Suporta operações não deterministas (o líder decide o resultado)
 - Se o líder produzir um valor errado, este valor é propagado para as réplicas
- Replicação máquina de estados
 - Se uma réplica produzir um valor errado não afecta as outras réplicas
 - As operações necessitam de ser deterministas