

Desempenho em SMR: paralelismo, comunicação e gerência de estados

Da prática na indústria brasileira à pesquisa acadêmica na Suíça

Quem sou eu

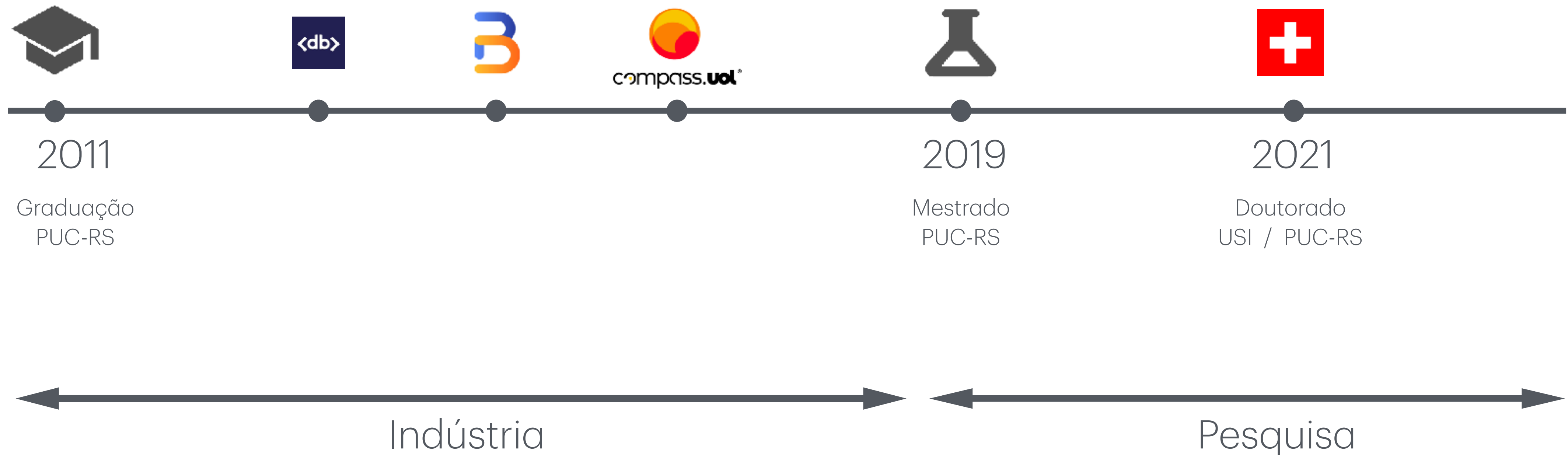
Eliã Batista

- **Doutorando, pesquisador, professor assistente, casado, pai da Eloise**
- Natural de São Leopoldo, gaúcho de Panambi
- Pesquisador em Sistemas Distribuídos à 6 anos
- 10 + anos de experiência na indústria
- 4 + anos morando em Lugano, Suíça 🇨🇭



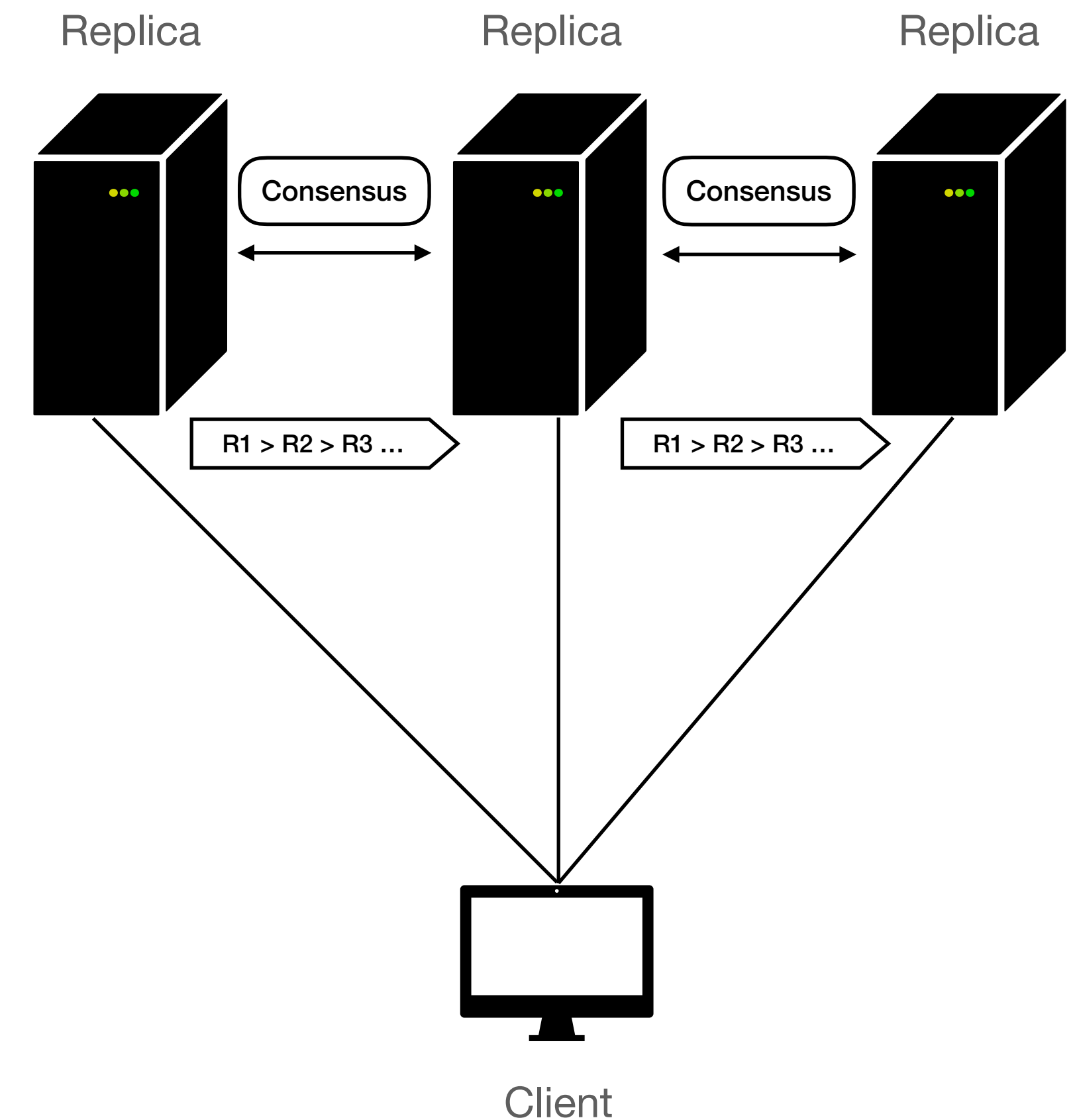
Minha trajetória

Linha do tempo



Pesquisa

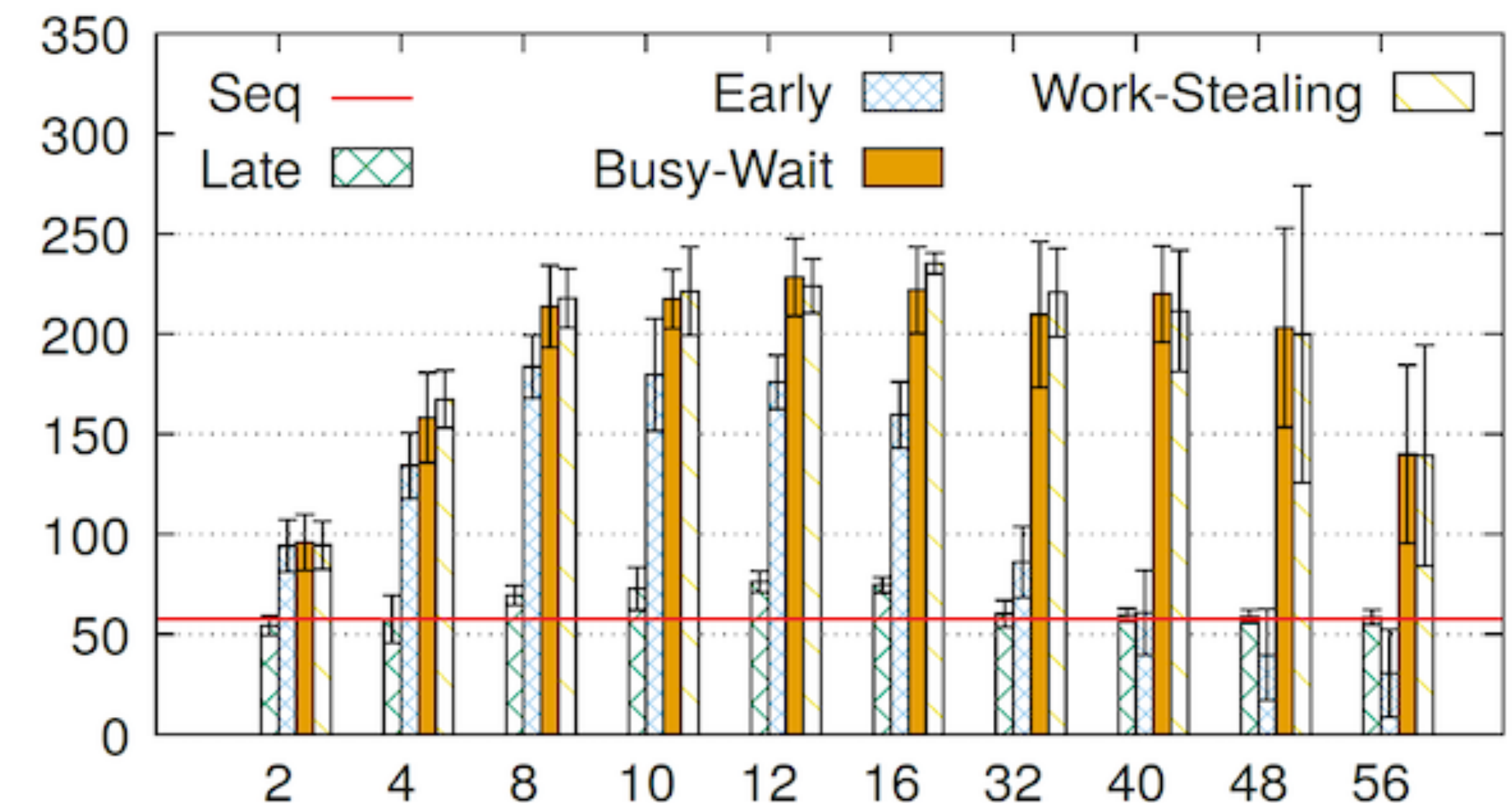
- Replicação Máquina de Estados (SMR)
- Tolerância a falhas em Sistemas Distribuídos
- Consenso entre diferentes replicas mantém um estado único
- Principais aspectos na performance:
 - Execução, Comunicação e Gerência de Estados



Mestrado

Escalonamento em SMR Paralelas - Execução

- **Late Scheduling:** ordena antes de escalonar; alto overhead.
- **Early Scheduling:** Usa classes e mapeamento de threads para evitar overhead. Simples e eficiente, mas restritivo.
- Melhorias:
 - **Busy wait:** sincronização ativa em vez de barreiras.
 - **Work stealing:** threads ociosas roubam tarefas.



Brasil 🇧🇷 → Suíça 🇨🇭

Um salto para a colaboração internacional

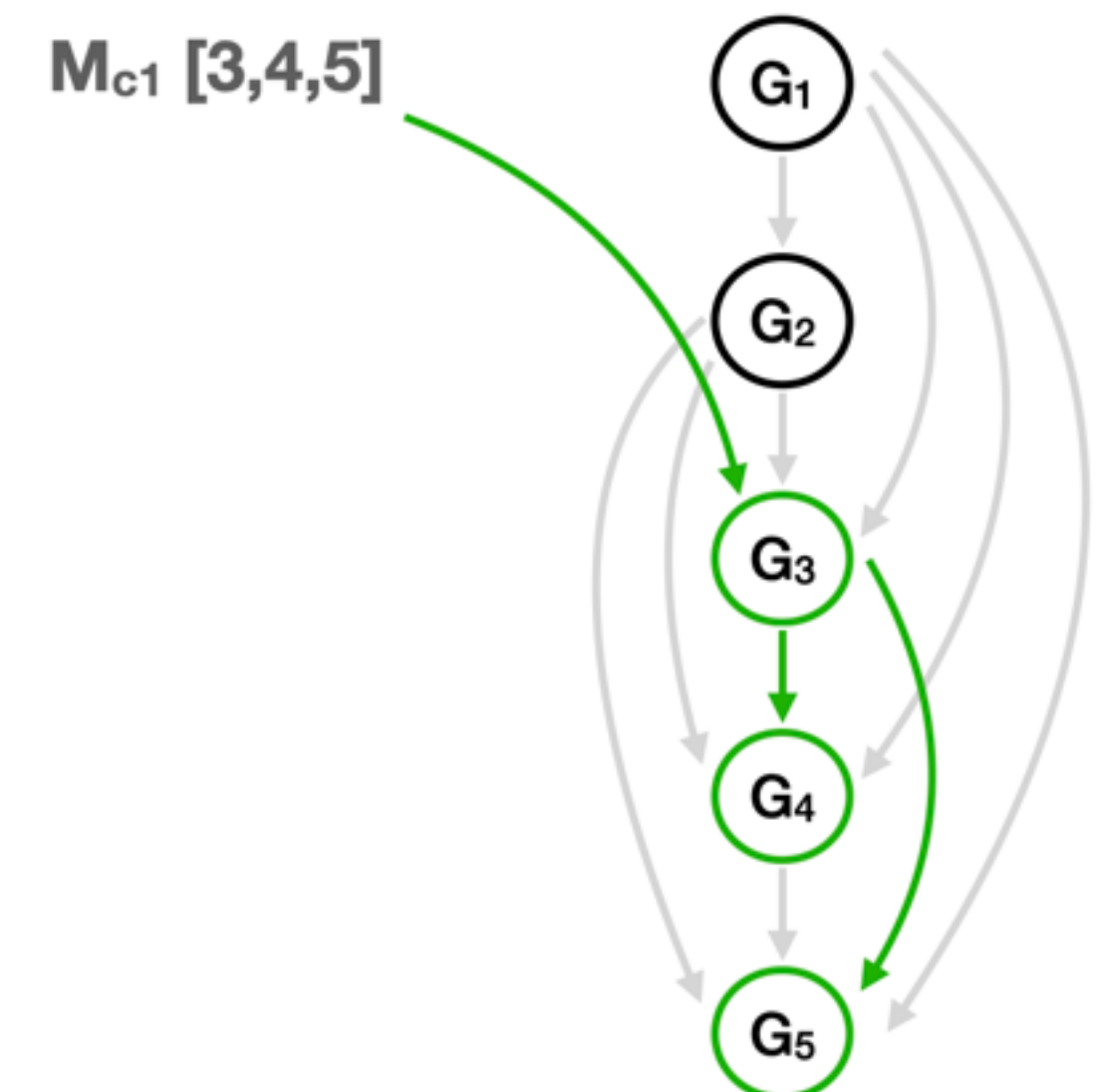
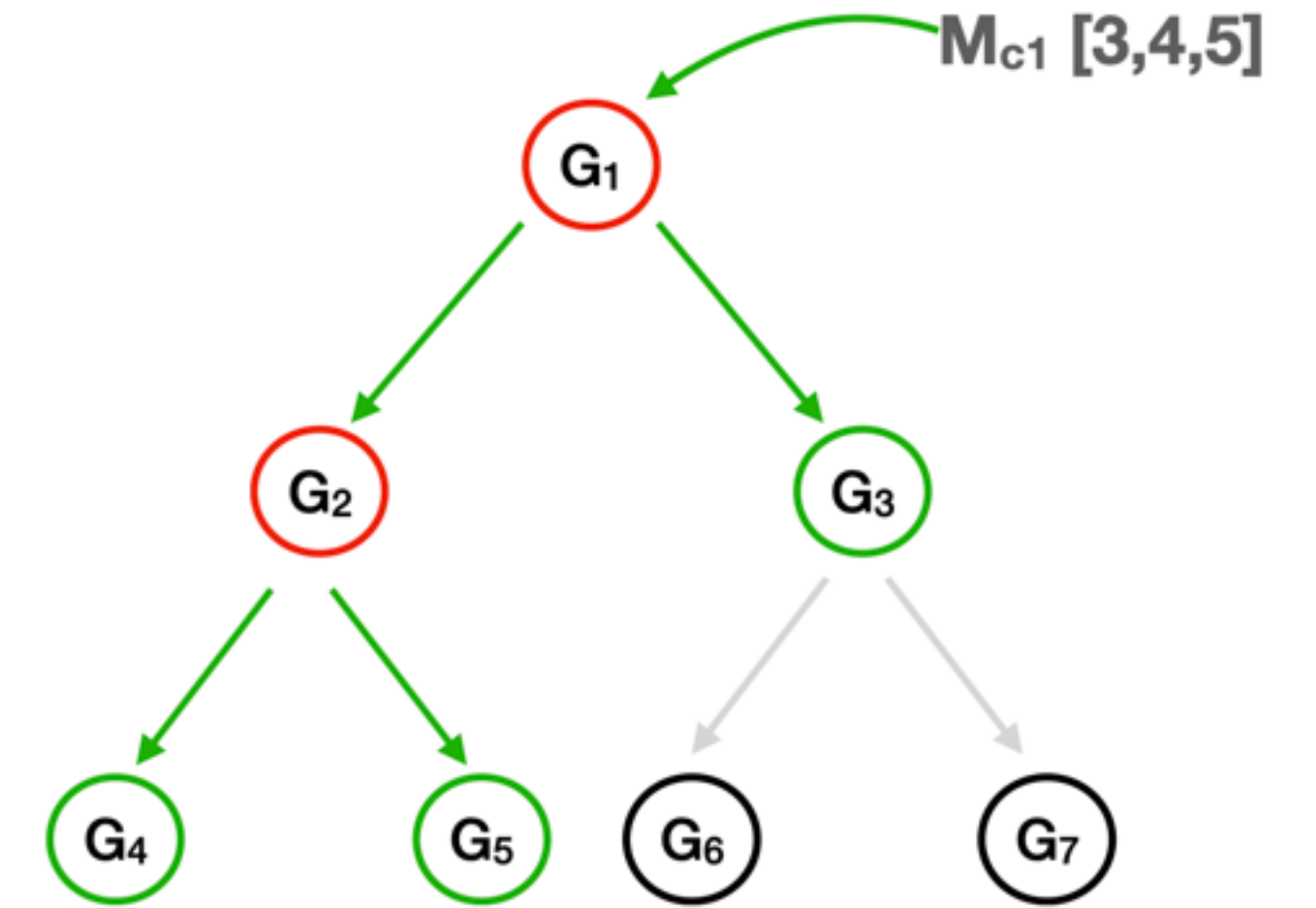
- Intercâmbio USI - 2019
- Doutorado USI / PUC-RS
- Mudança para Suíça - 2021
- Desafios: Nova cultura, língua, distância, vida acadêmica no exterior



Doutorado

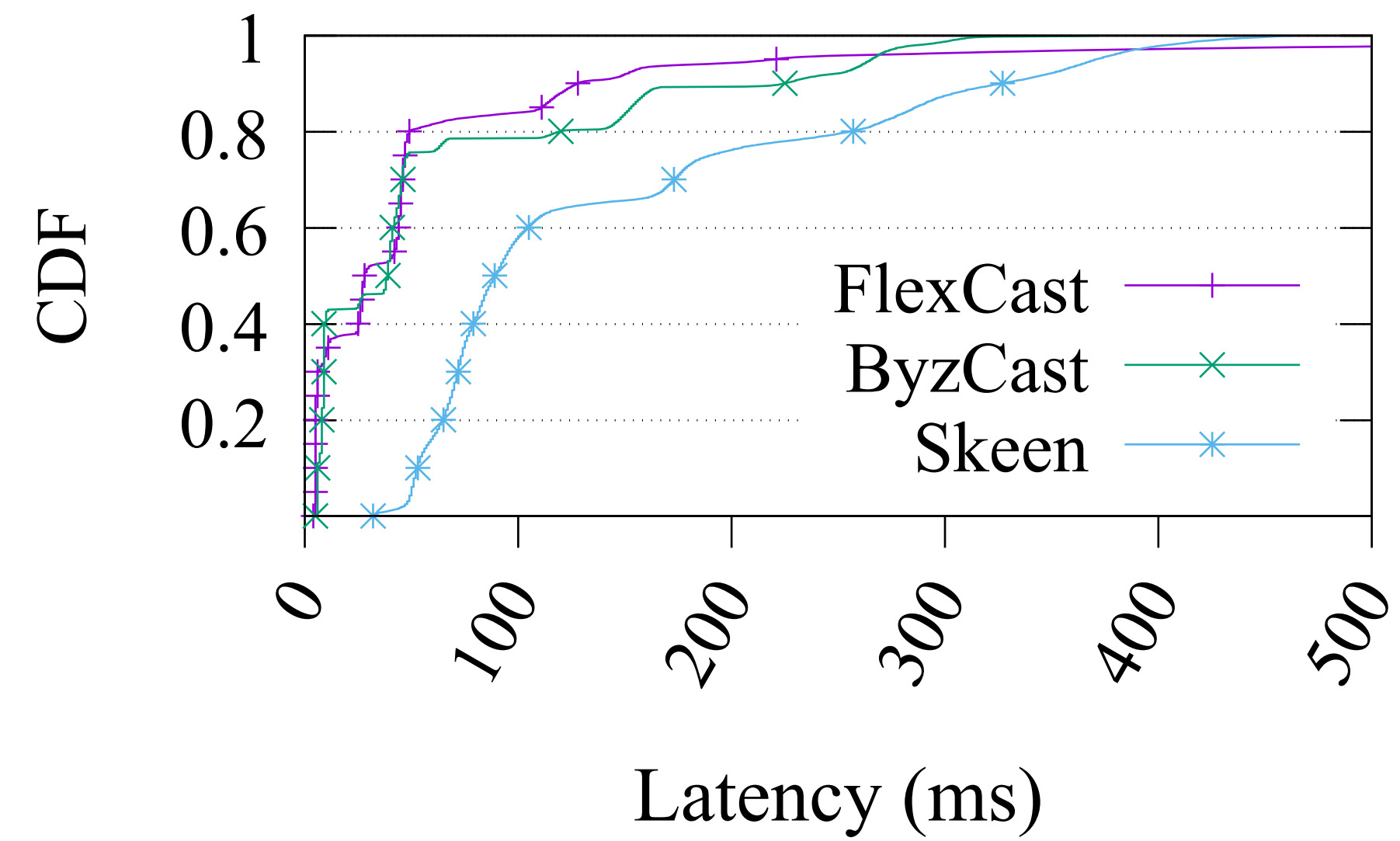
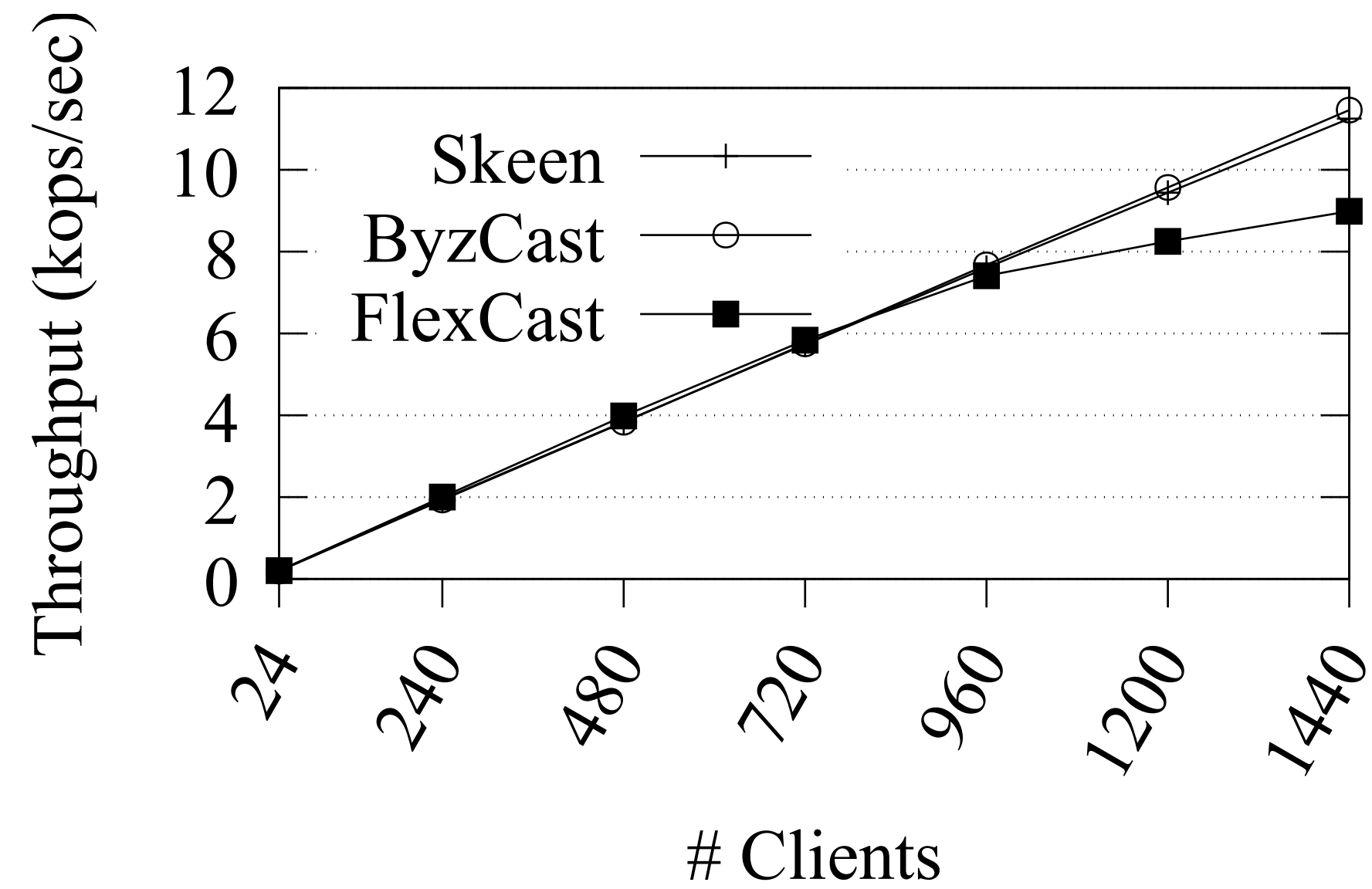
Atomic Multicast em SMR: Comunicação

- SMRs particionadas usam **multicast** atômico para enviar mensagens a subgrupos de replicas.
- **FlexCast**: primeiro protocolo genuíno de multicast atômico baseado em overlays de comunicação.
- Usa um C-DAG para permitir comunicação **genuína**.
- Garante ordem total acíclica através de **históricos e mensagens auxiliares**.
- Processos filhos seguem a ordem definida por processos **ancestrais** na hierarquia.



FlexCast

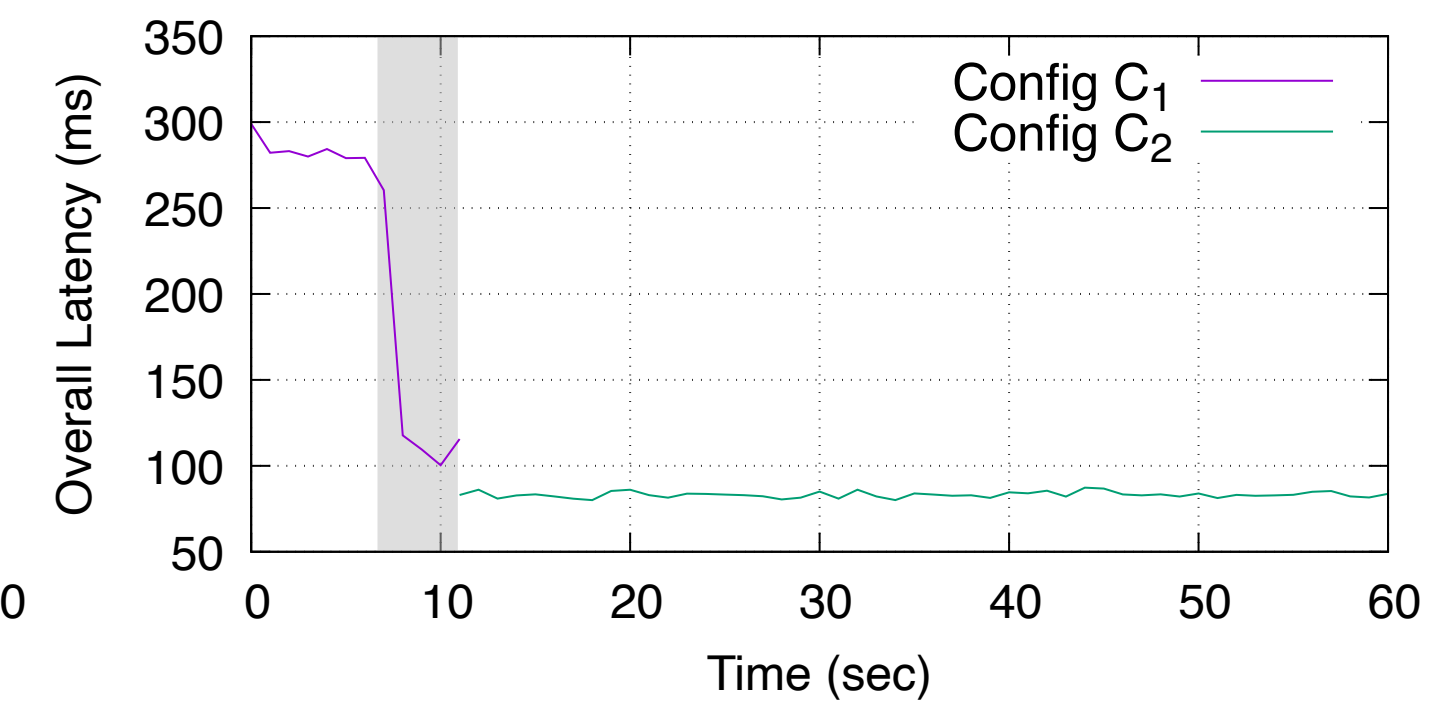
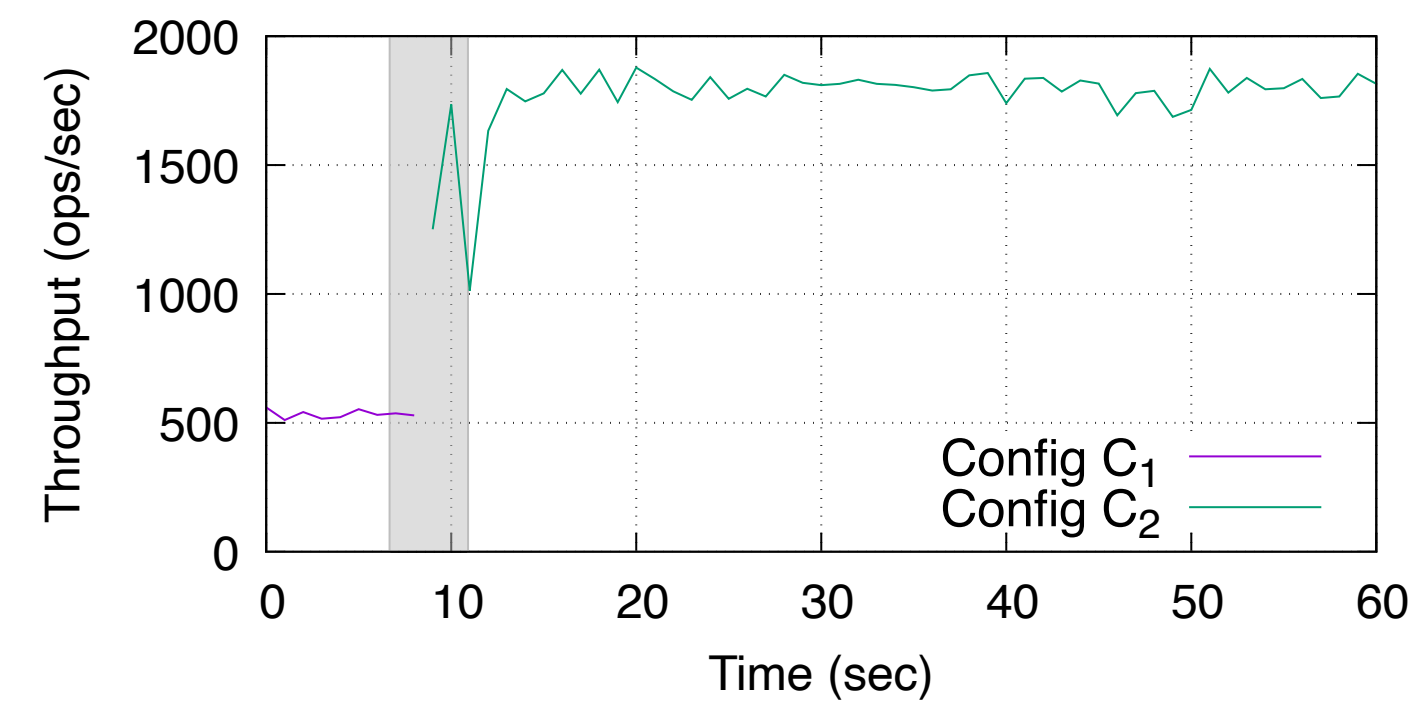
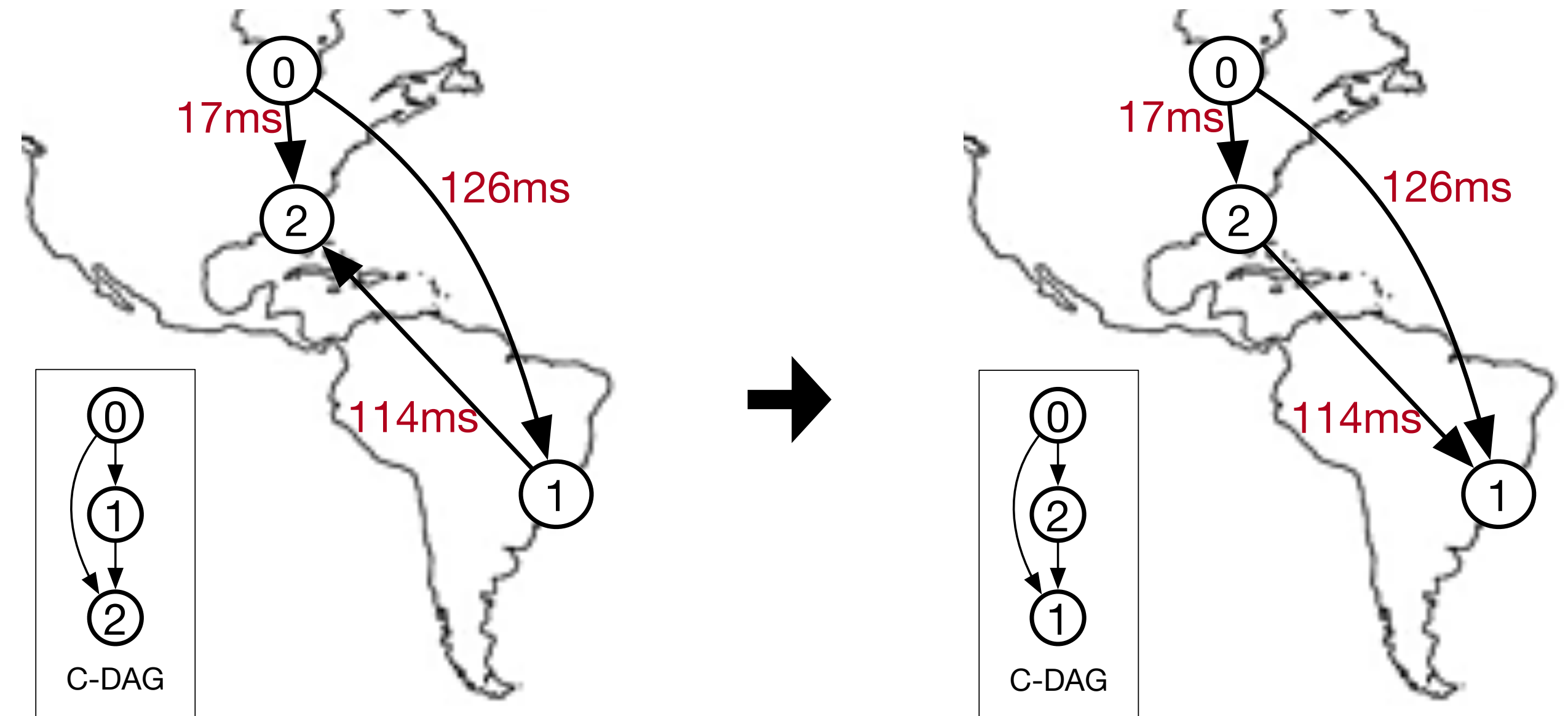
Resultados



FlexCast

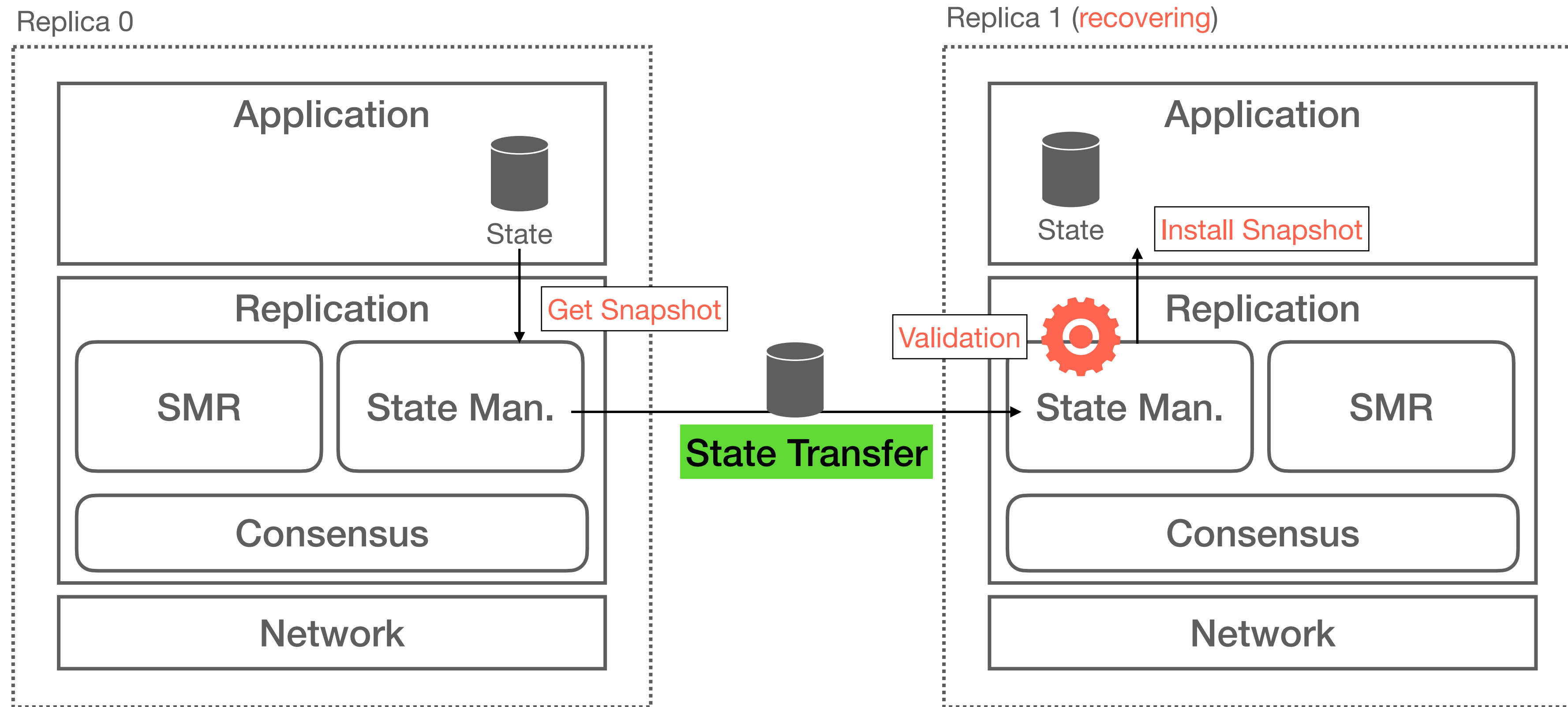
Reconfigurando

- Reconfiguração ajusta o sistema para reduzir a latência, conforme a carga de trabalho atual.
- Uma função de otimização calcula um novo overlay, adaptado à situação do sistema.
- Essa nova configuração é disseminada e instalada dinamicamente (on-the-fly), sem interromper o serviço.



Gerência de Estados em SMR

Transferência de estado



Gerência de Estados em SMR

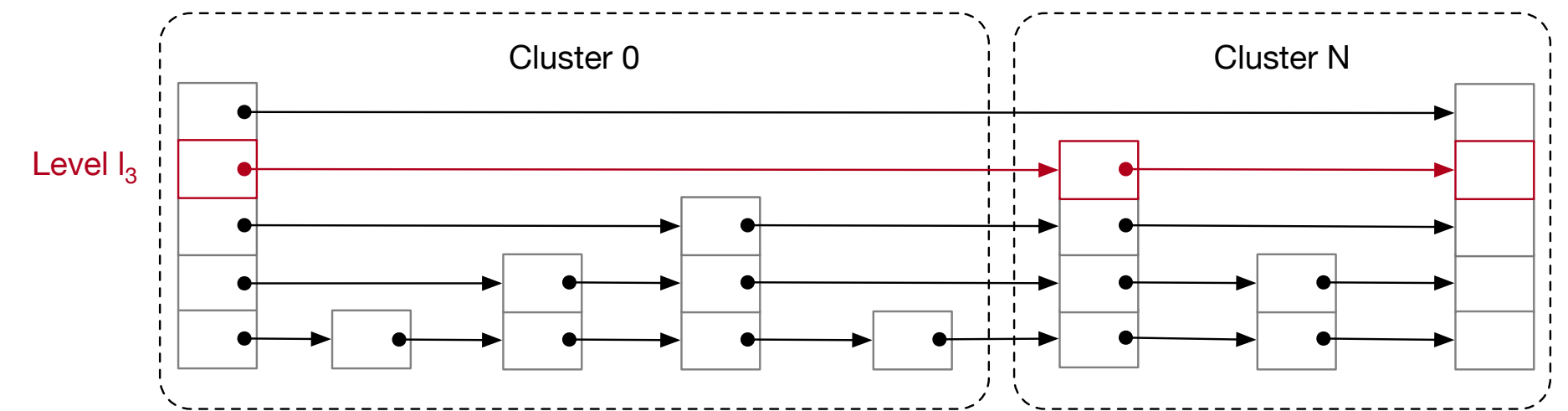
Transferência de estado

- Gerência de estado é crucial para o desempenho em SMR, especialmente na recuperação.
- Garante consistência entre réplicas mesmo com falhas (especialmente bizantinas).
- Problemas comuns: Transferência lenta; Validação ineficiente

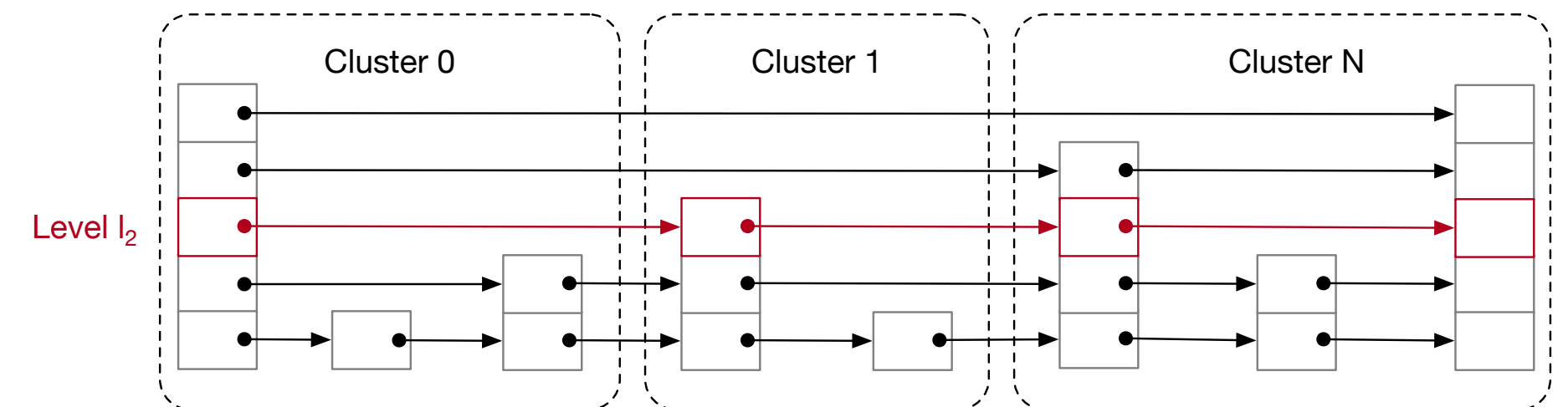
Gerência de Estados em SMR

SkipLists particionadas e auto-verificáveis

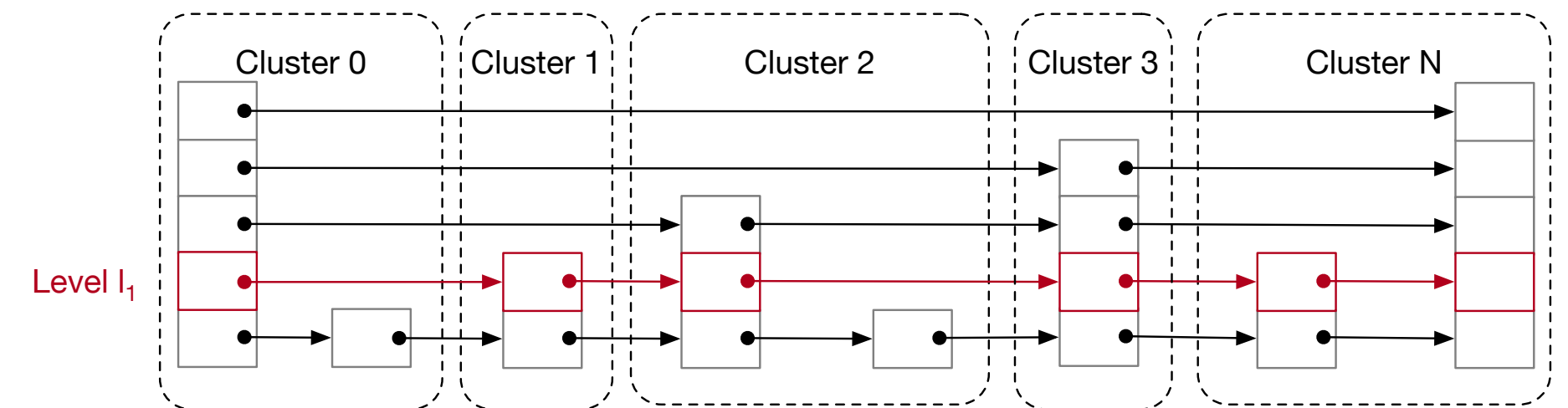
- Estudo de SkipLists auto-verificáveis inspiradas em Merkle-trees.
- Introdução de mecanismo de particionamento de SkipLists.
- Partições são verificáveis separadamente e podem ser transmitidas em paralelo durante recuperação.



(a)



(b)

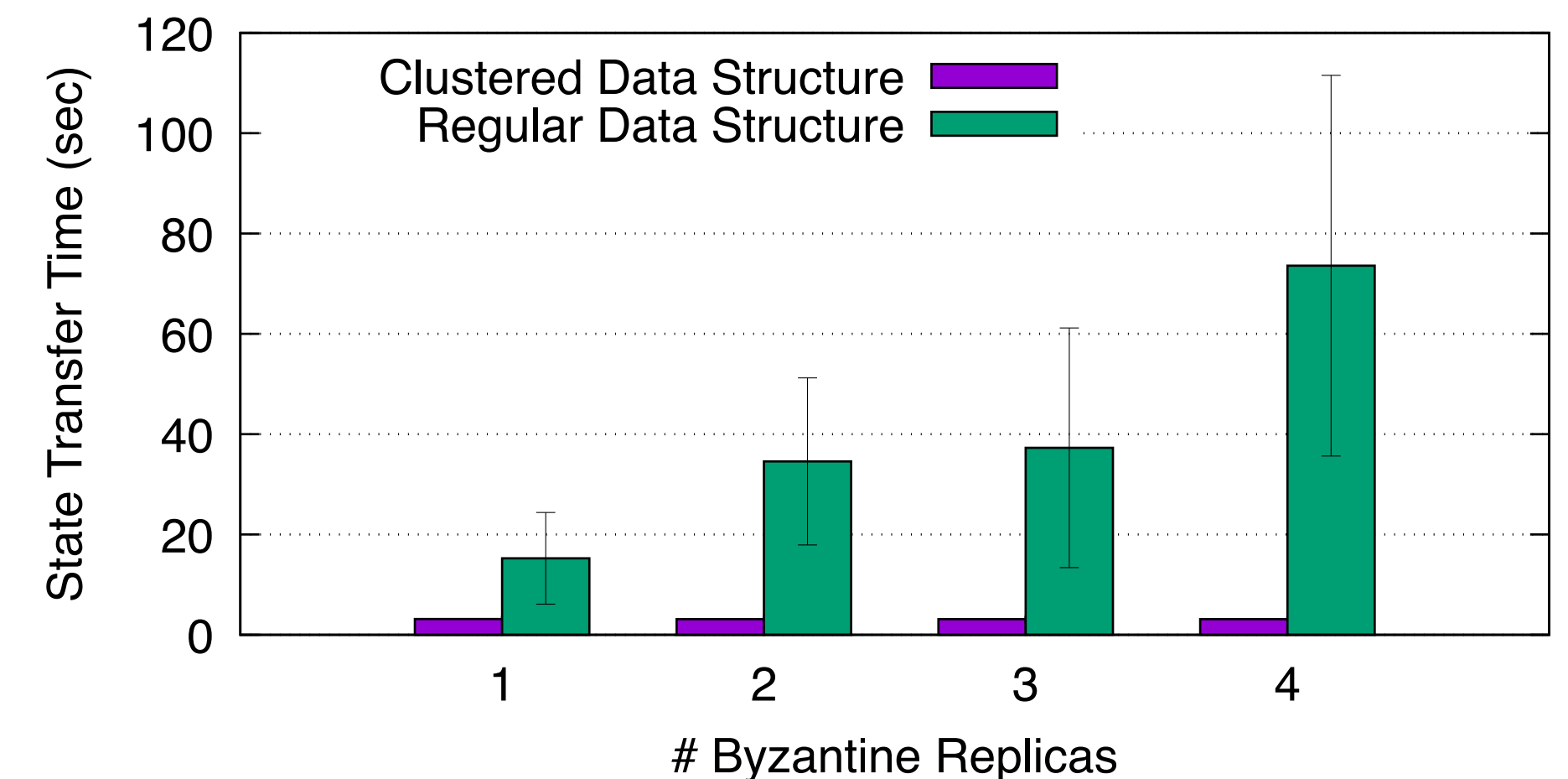


(c)

Gerência de Estados em SMR

SkipLists particionadas e auto-verificáveis

- Em cenários com falhas bizantinas, estruturas auto-verificáveis e particionadas permitem detectar corrupção com baixo custo e recuperar apenas partes afetadas.
- Estruturas tradicionais só detectam erros após baixar todo o estado, causando mais atraso e tráfego.
- **O que vem pela frente**
 - Integração com framework SMR
 - Validação de desempenho da transferência de estados



Lições Aprendidas & Conclusão

- Simplicidade tem preço: modelos simples como *early scheduling* são eficientes, mas exigem cuidados com restrições.
- Mapa sob medida: overlay do FlexCast desenha caminhos otimizados para cada carga e topologia.
- Adaptar é essencial: reconfiguração dinâmica do FlexCast responde à carga dinâmicas.
- Especialização importa: estruturas otimizadas melhoram desempenho e recuperação.
- Confiabilidade e particionamento: validação parcial é chave contra falhas bizantinas.

Obrigado!

Contato



delime@usi.ch



[linkedin.com/in/e-batista](https://www.linkedin.com/in/e-batista)

