Mestrado Integrado em Engenharia Informática

Sistemas de Armazenamento Configuráveis e Seguros

Tânia da Conceição Araújo Esteves
tania.c.araujo@inesctec.pt

João Tiago Paulo e José Orlando Pereira

2 de novembro de 2018

Contextualização

- Crescimento exponencial de informação digital
- Dados são armazenados em serviços na nuvem
- **Desafios:**
  - Assegurar confidencialidade e integridade dos dados
  - Aplicar funcionalidades de armazenamento orientadas ao conteúdo

Problema e Objetivos

- Dados cifrados com esquemas probabilísticos
- Incompatibilidade com técnicas orientadas ao conteúdo
- **Objetivo:**
  - Integrar soluções de *hardware* confiável em sistemas programáveis e empilháveis

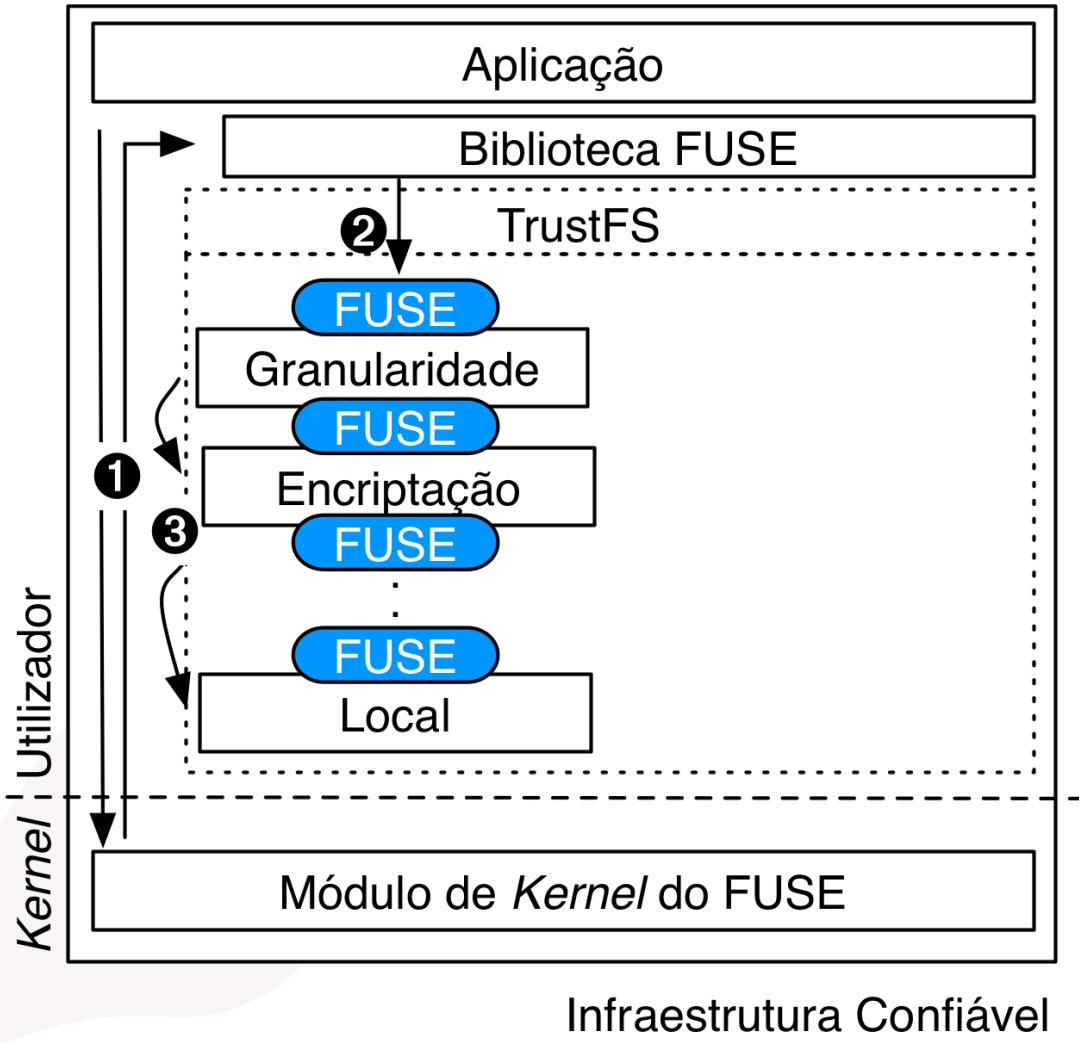
Contribuições

- Estudo e comparação das tecnologias de *Hardware* Confiável
- Plataforma TrustFS
- Deduplicação de dados segura por épocas
- Avaliação Experimental

Estado da Arte

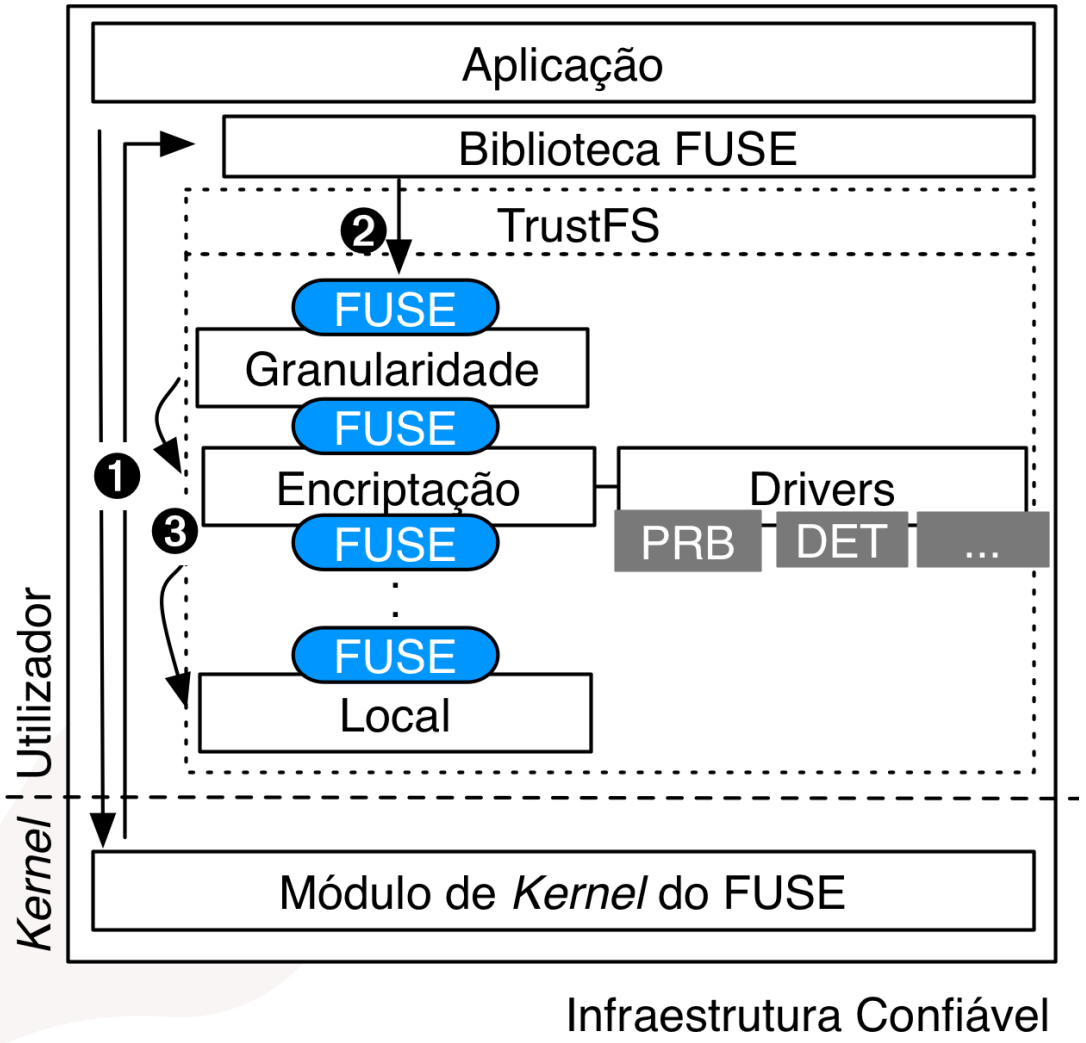
- Intel SGX utilizado em diversas áreas:
  - Base de dados, sistemas de armazenamento, *machine learning*, etc.
- Deduplicação Segura:
  - Criptografia Convergente
    - Uma chave por bloco/ficheiro
    - Ataques de confirmação do ficheiro ou aprendizagem da restante informação
  - Apenas uma solução com SGX
    - Com perdas a nível do espaço poupado

Arquitetura



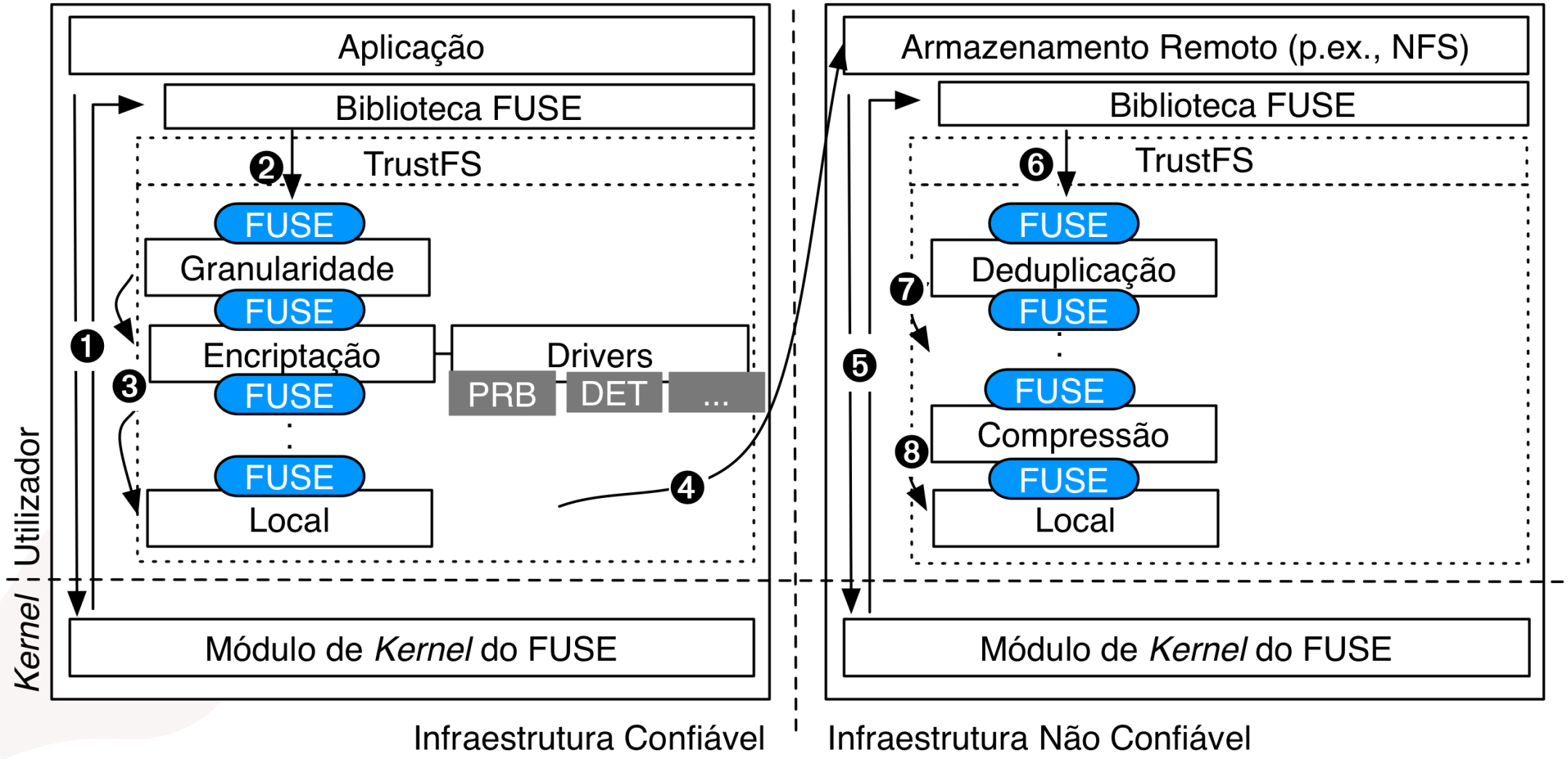
Infraestrutura Confiável

Arquitetura

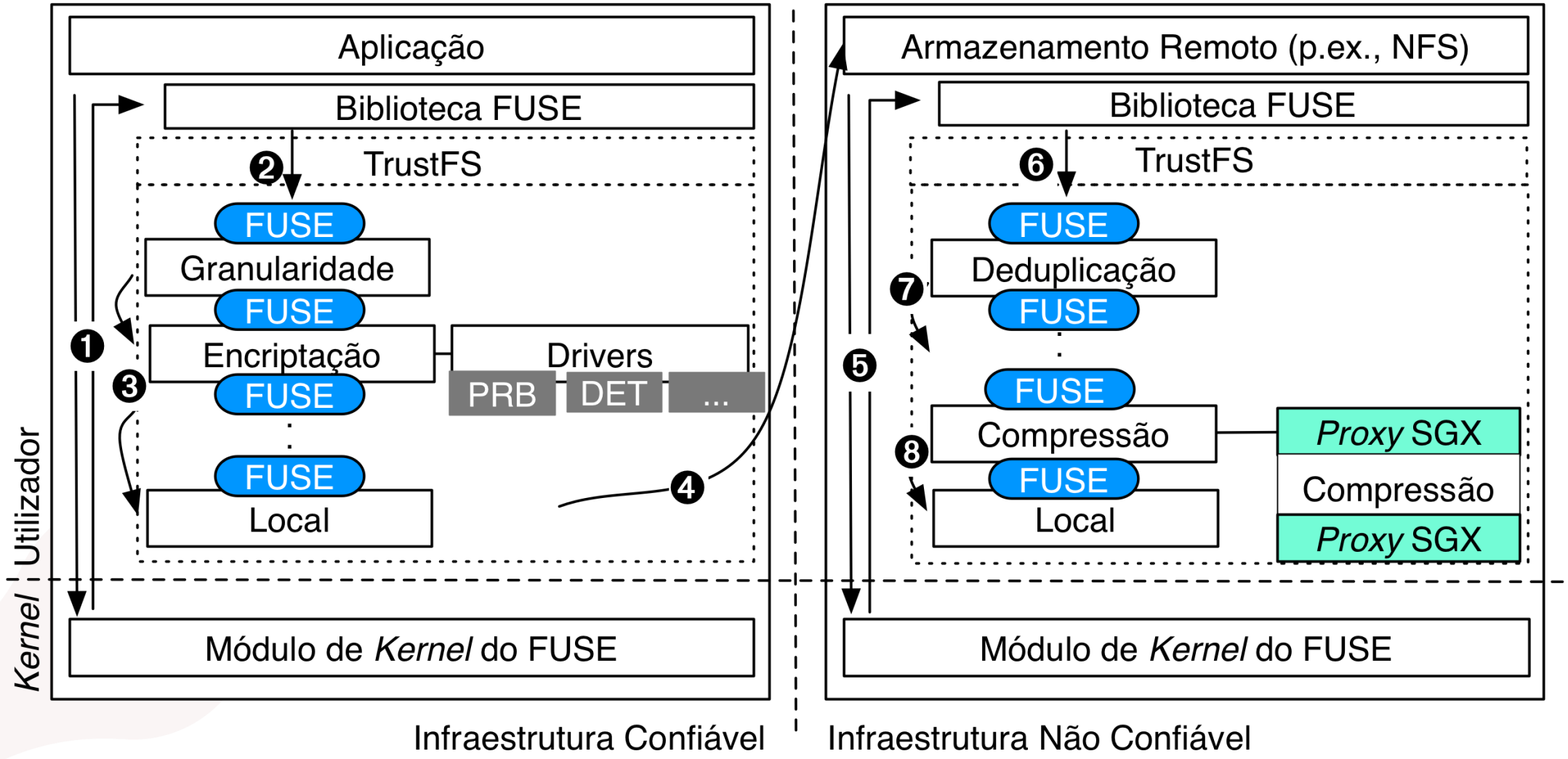


Infraestrutura Confiável

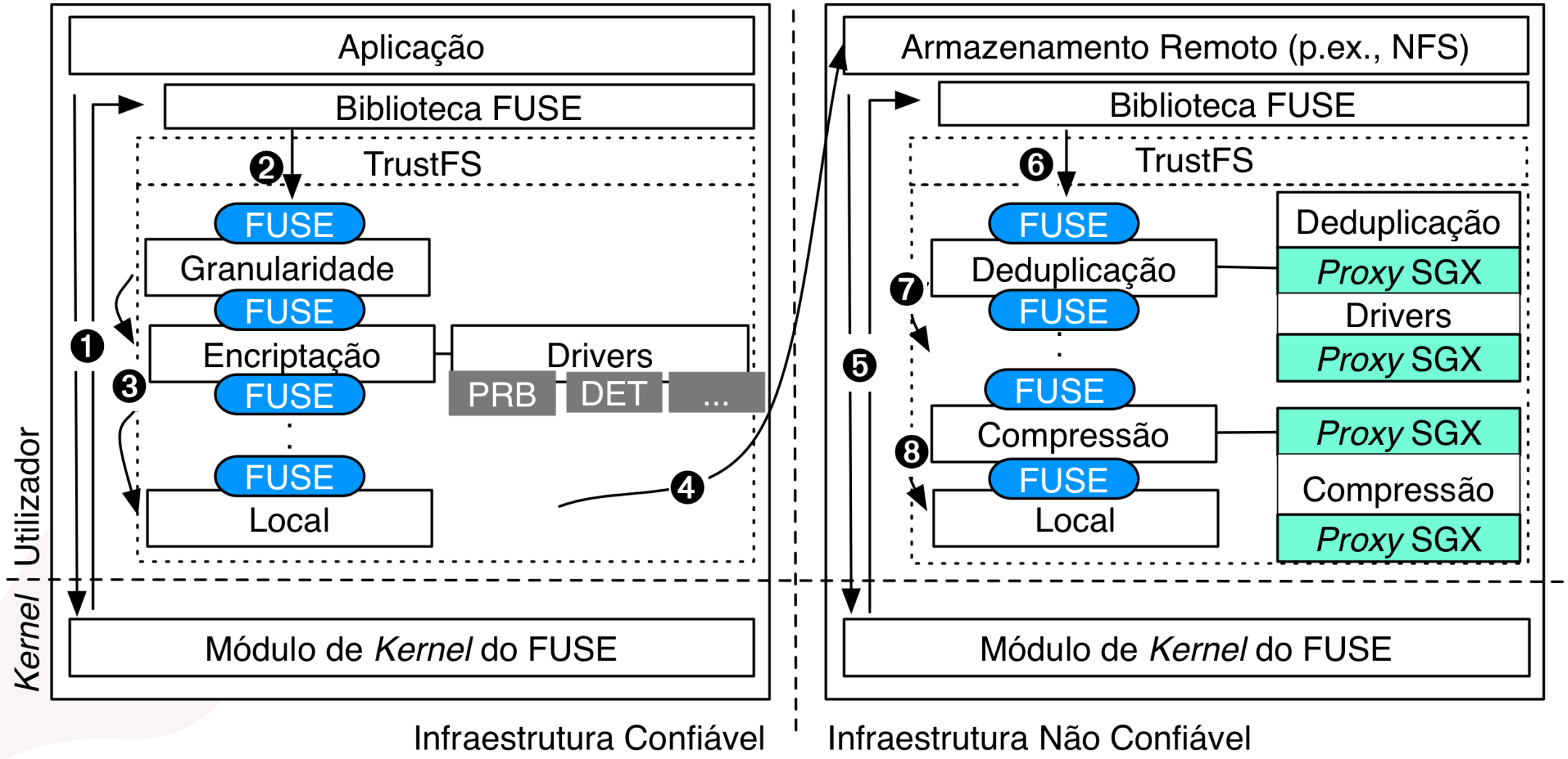
Arquitetura



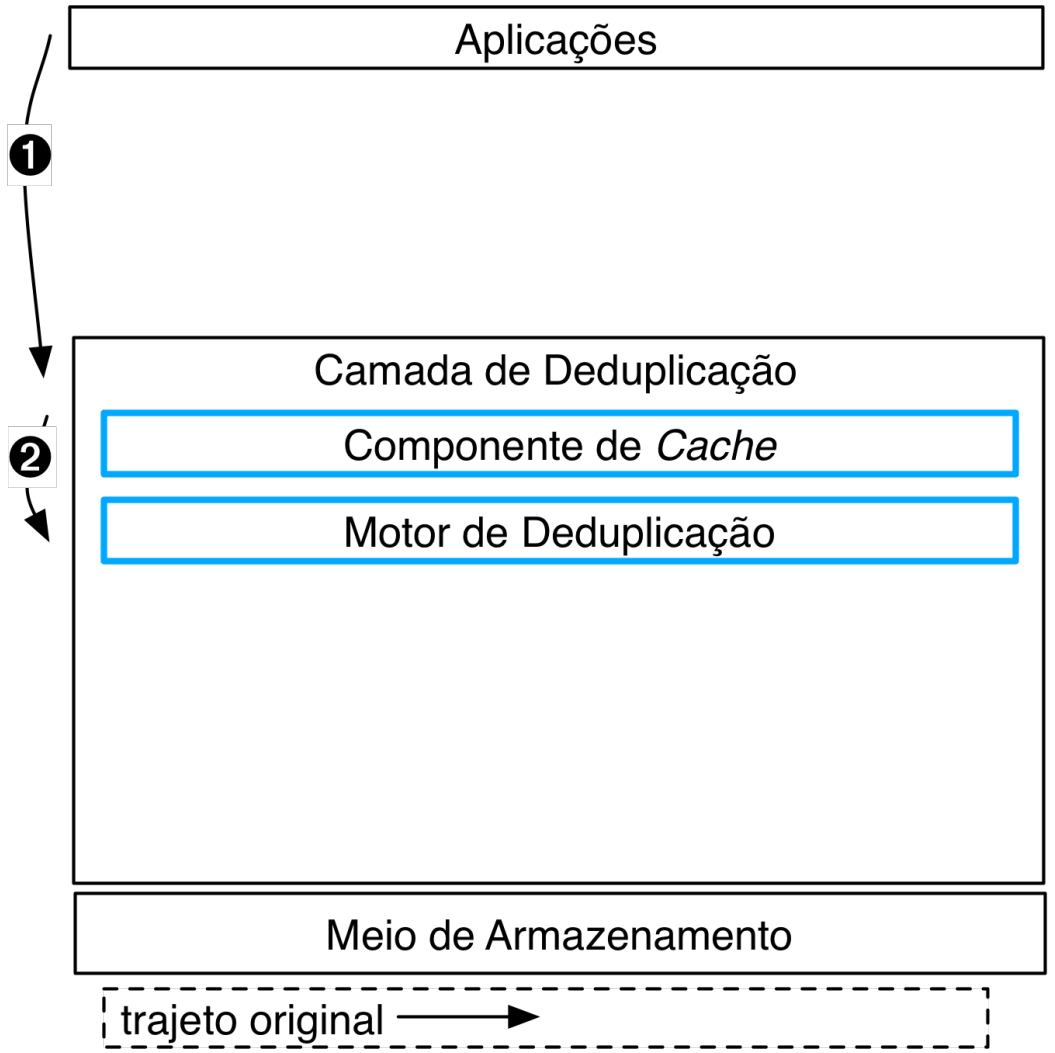
Arquitetura



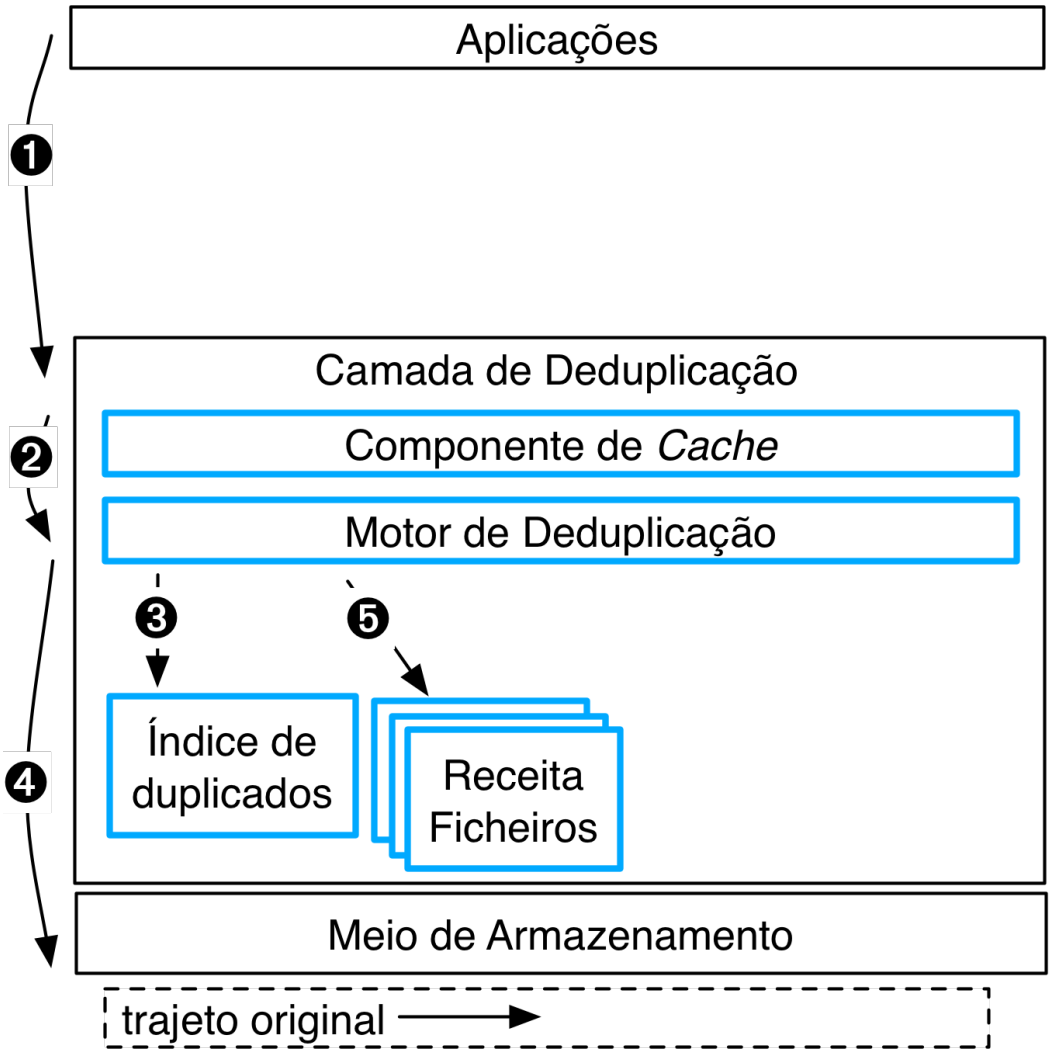
Arquitetura



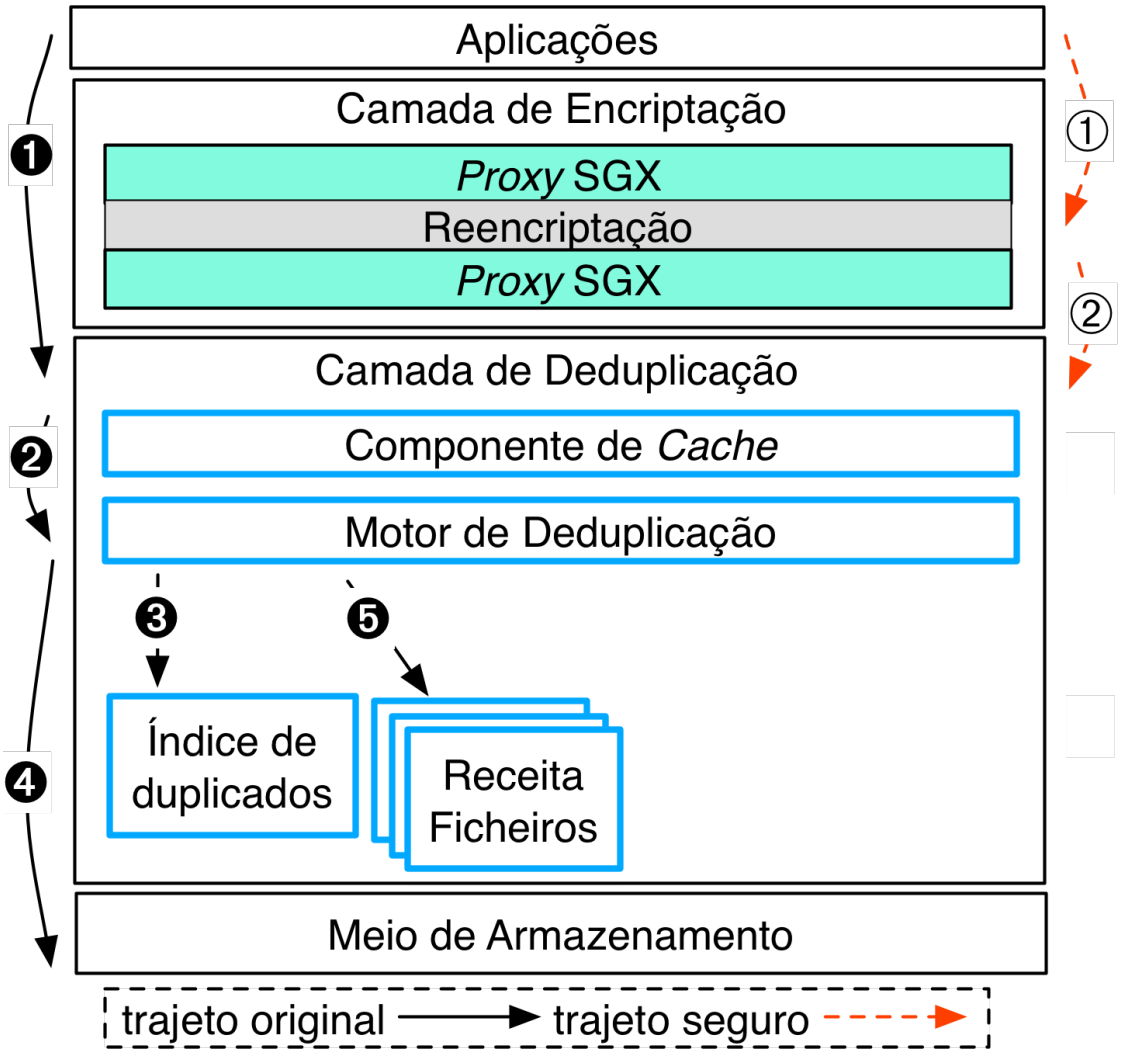
Protótipo | Camada de deduplicação



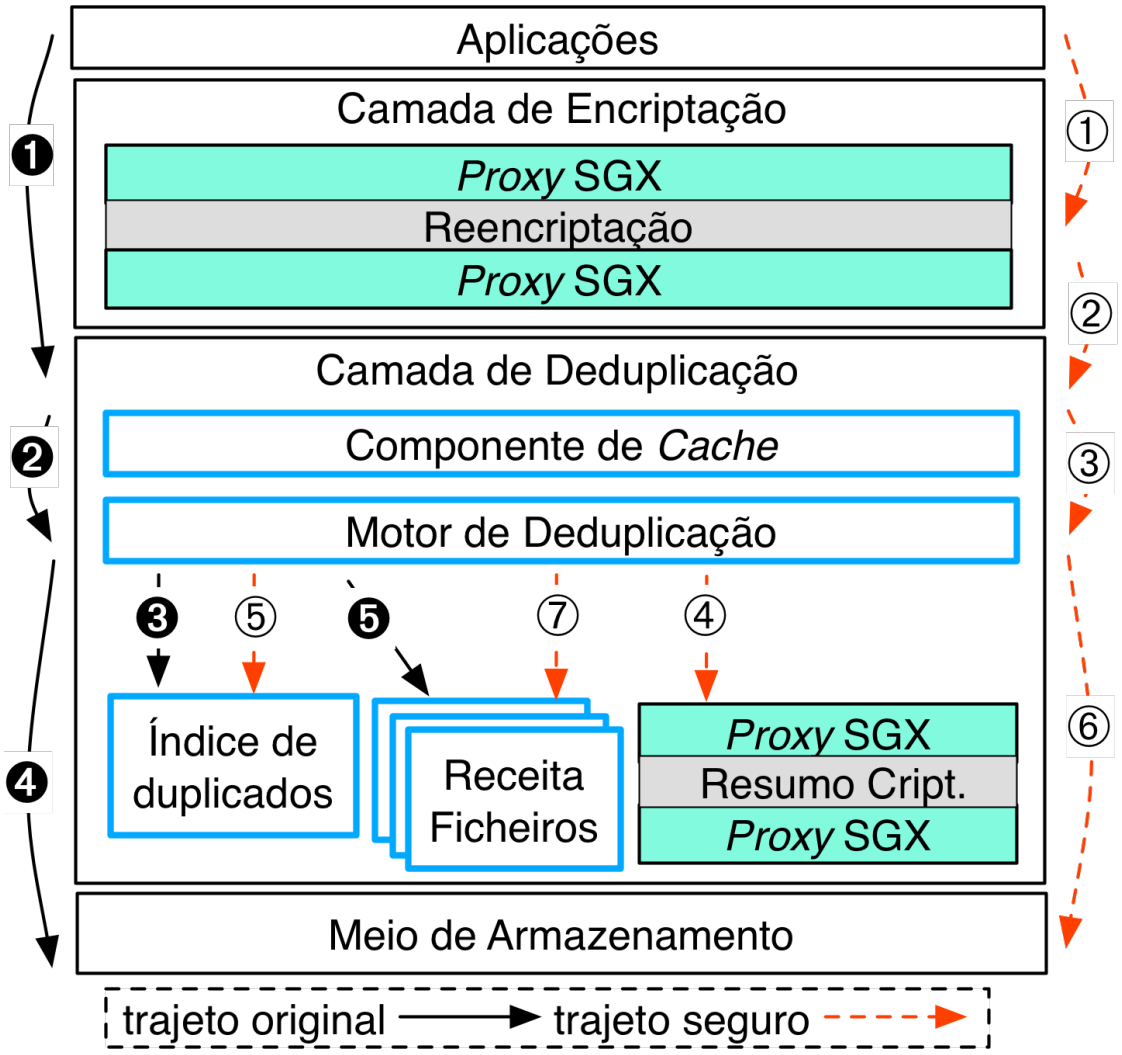
Protótipo | Camada de deduplicação



Protótipo | Camada de deduplicação



Protótipo | Camada de deduplicação



Protótipo | Deduplicação por épocas

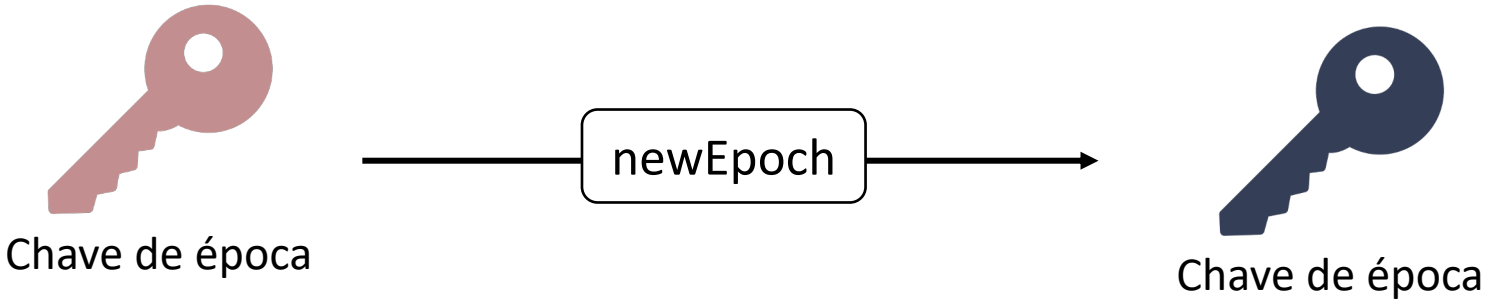
Esta solução:

- Protege os dados contra um atacante que esteja à escuta da rede
- Mas continua a ser alvo do ataque de confirmação de ficheiros

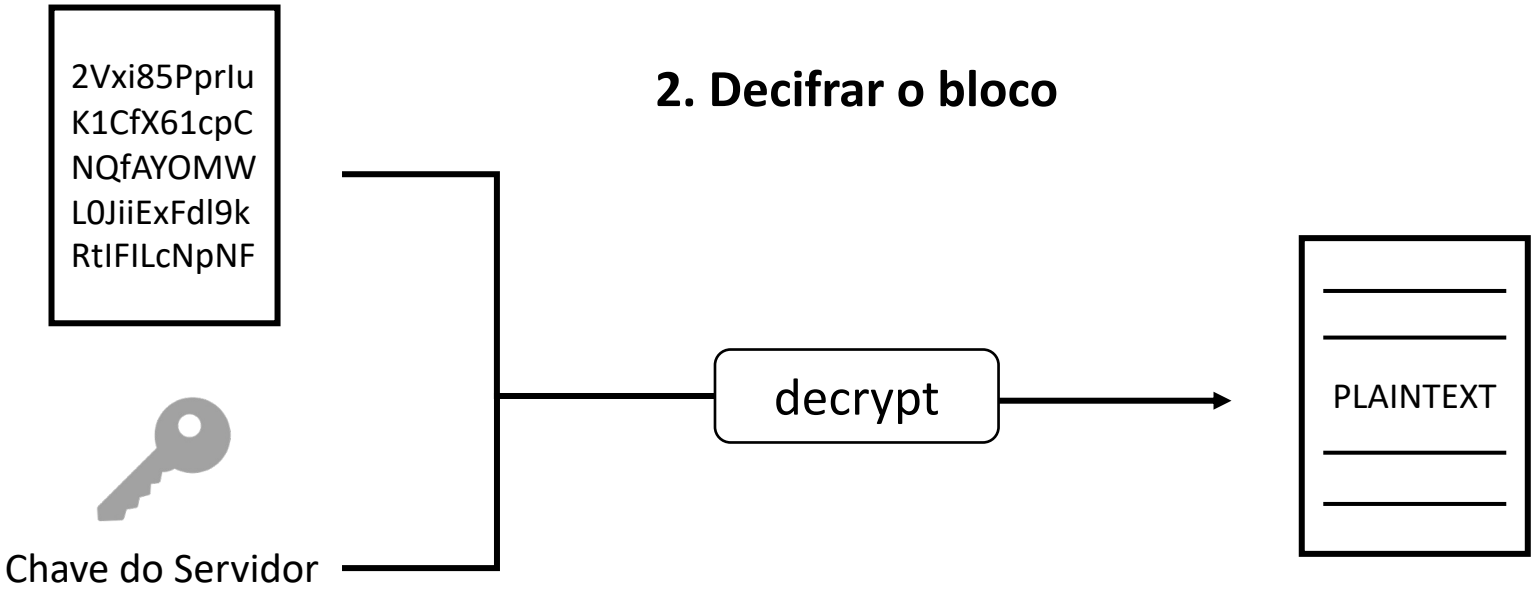
Para colmatar este problema introduzimos um novo esquema de deduplicação por épocas

Protótipo | Deduplicação por épocas

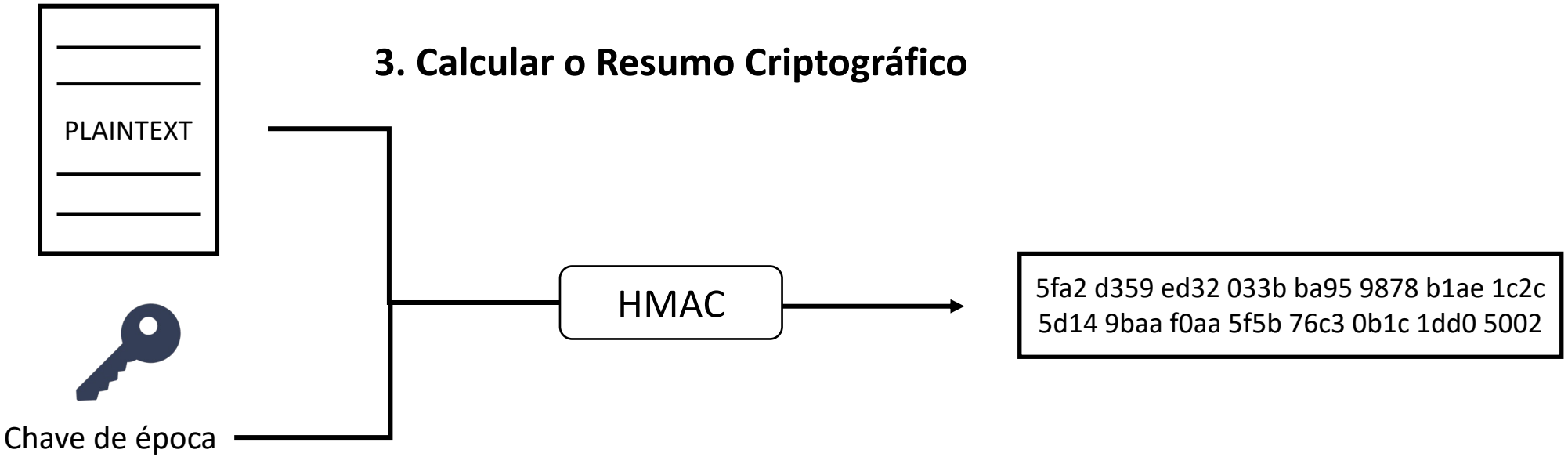
1. Mudar de época



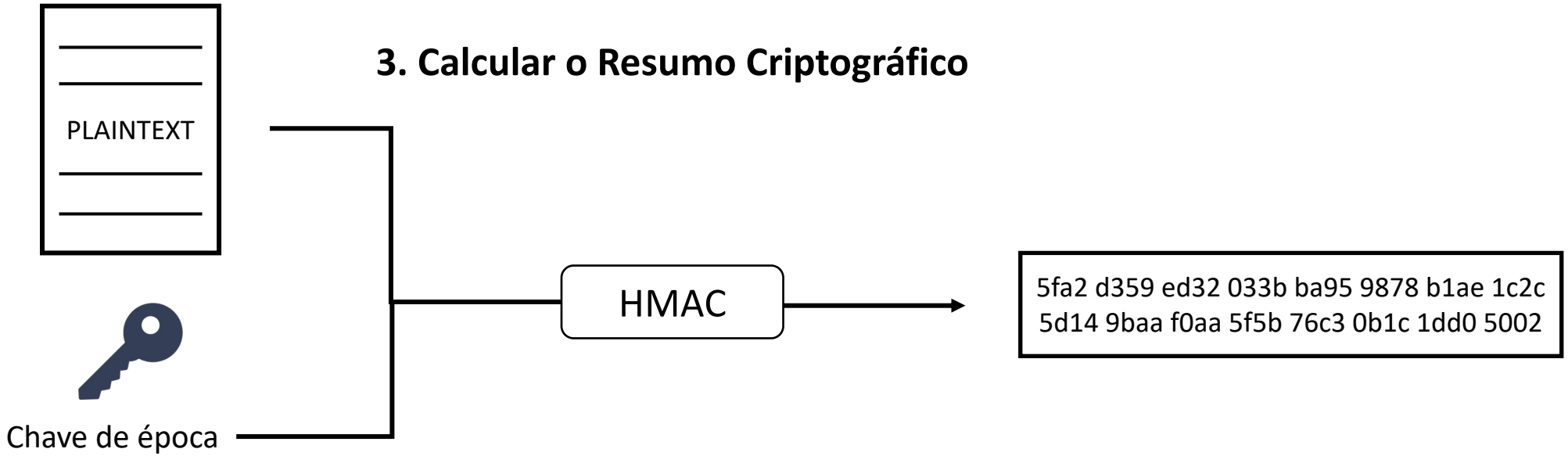
Protótipo | Deduplicação por épocas



Protótipo | Deduplicação por épocas

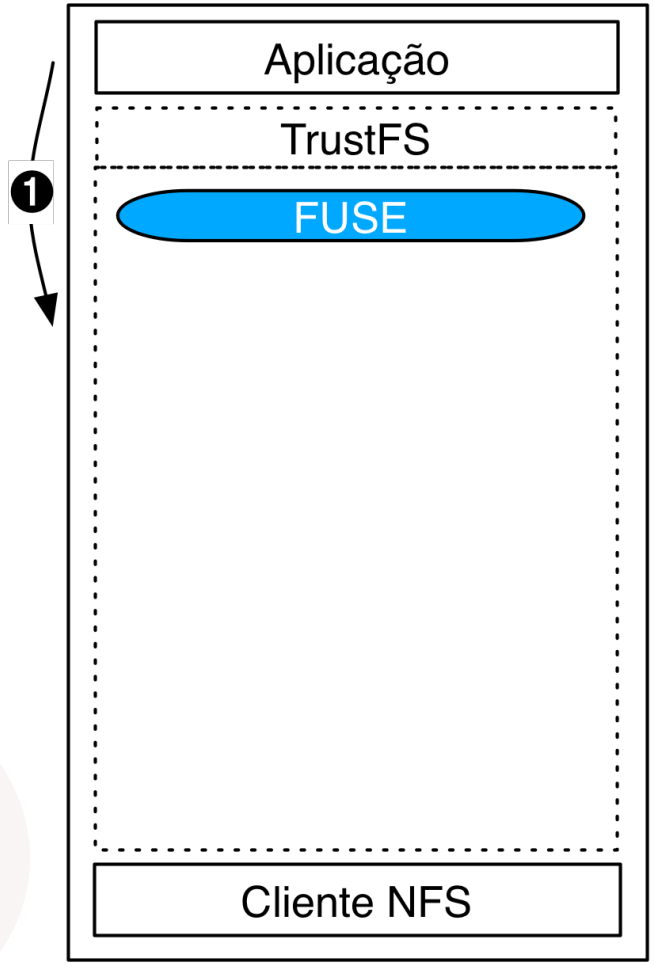


Protótipo | Deduplicação por épocas



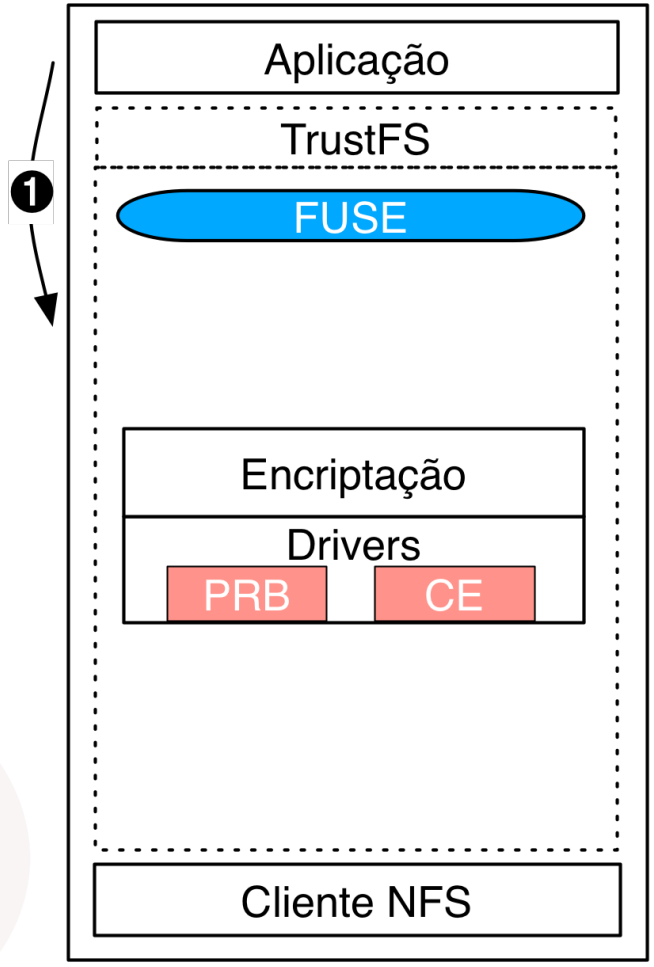
- Deduplicação apenas entre cópias armazenadas na mesma época
- Compromisso entre segurança e economia de espaço

Protótipo | Configurações das camadas



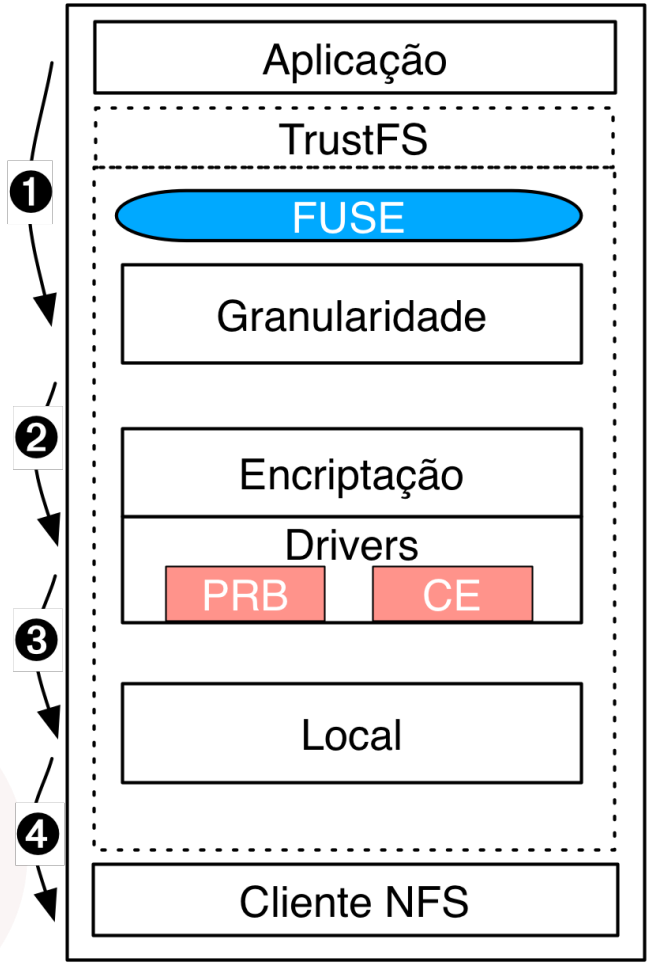
Máquina Cliente

Protótipo | Configurações das camadas



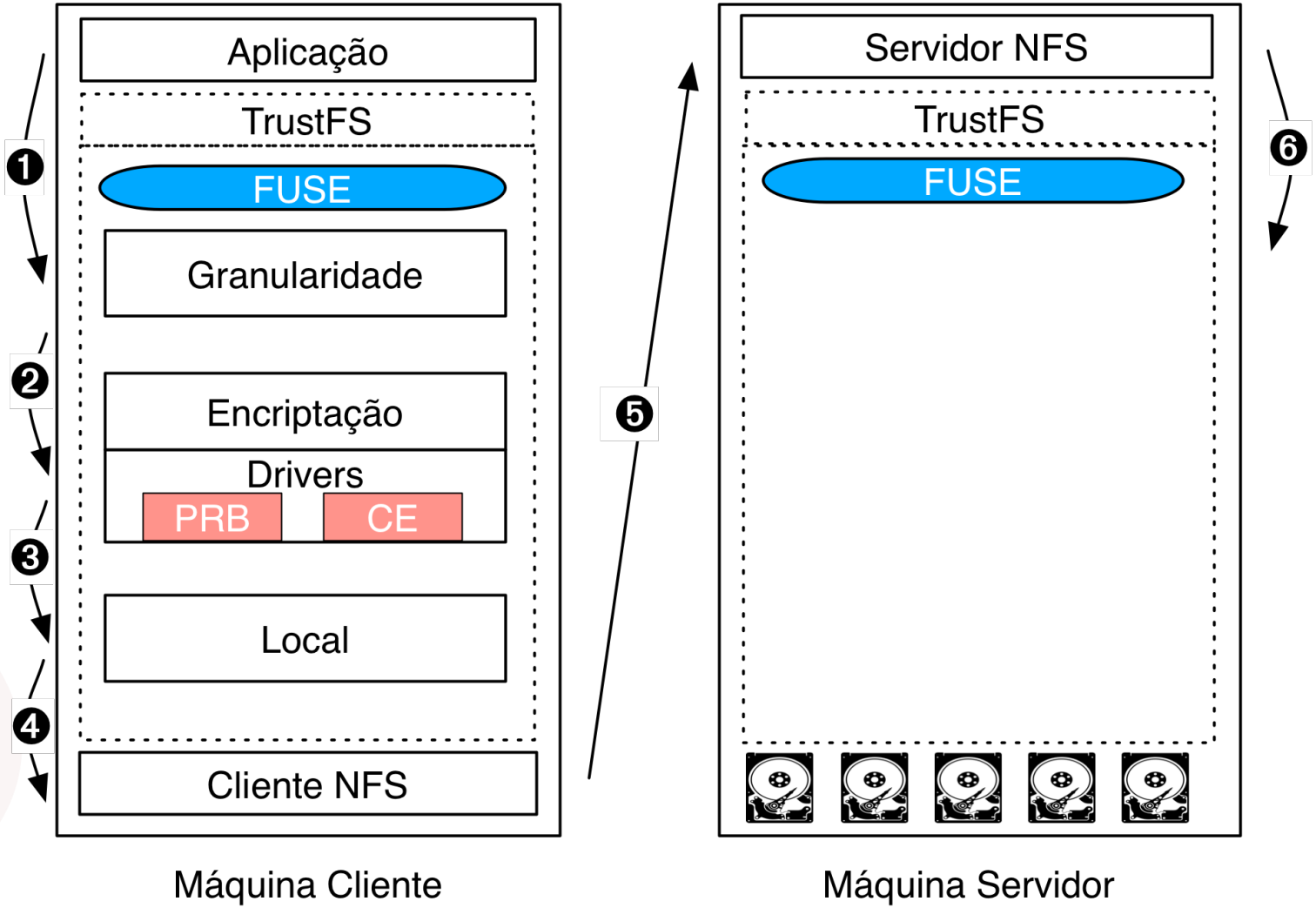
Máquina Cliente

Protótipo | Configurações das camadas

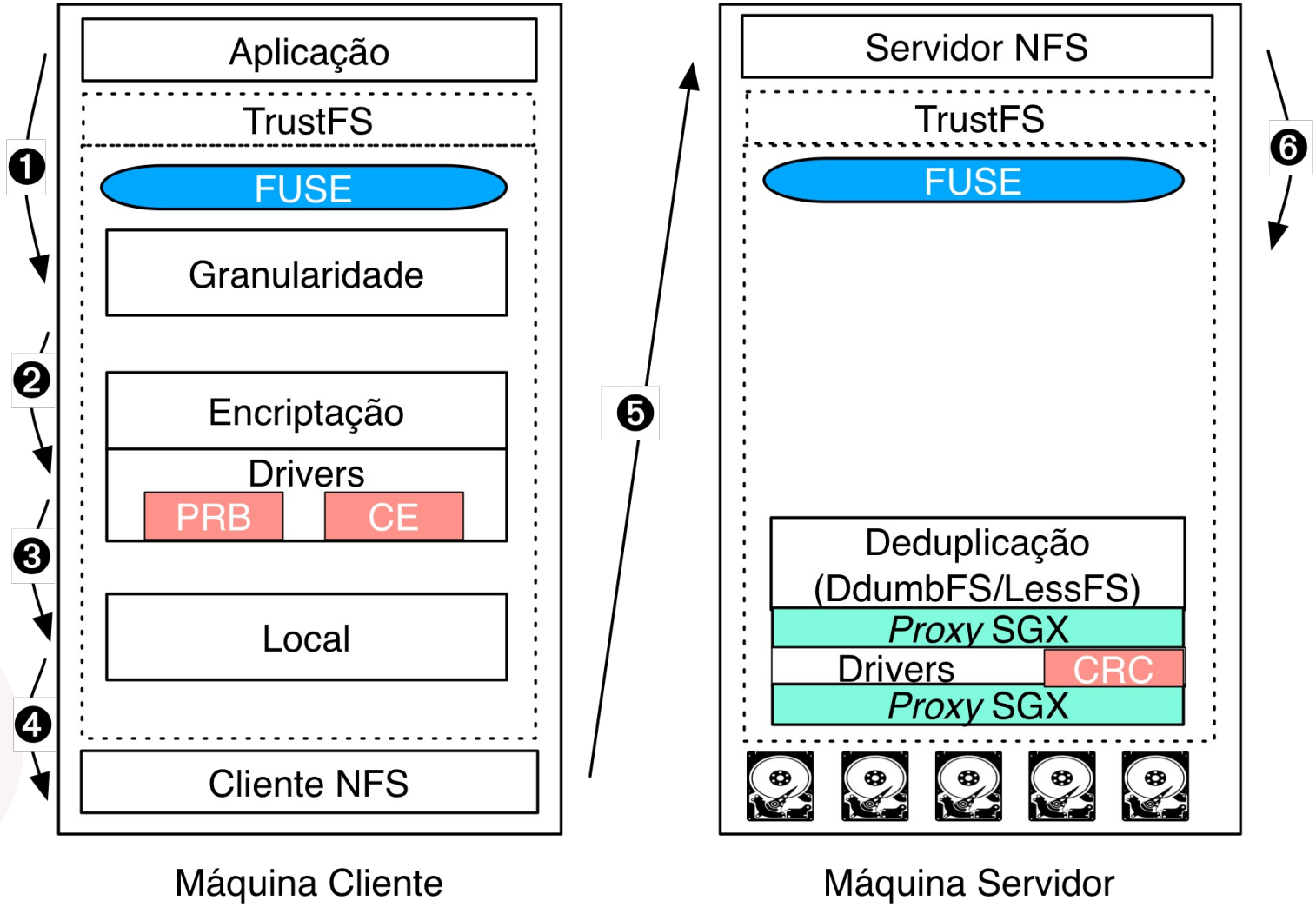


Máquina Cliente

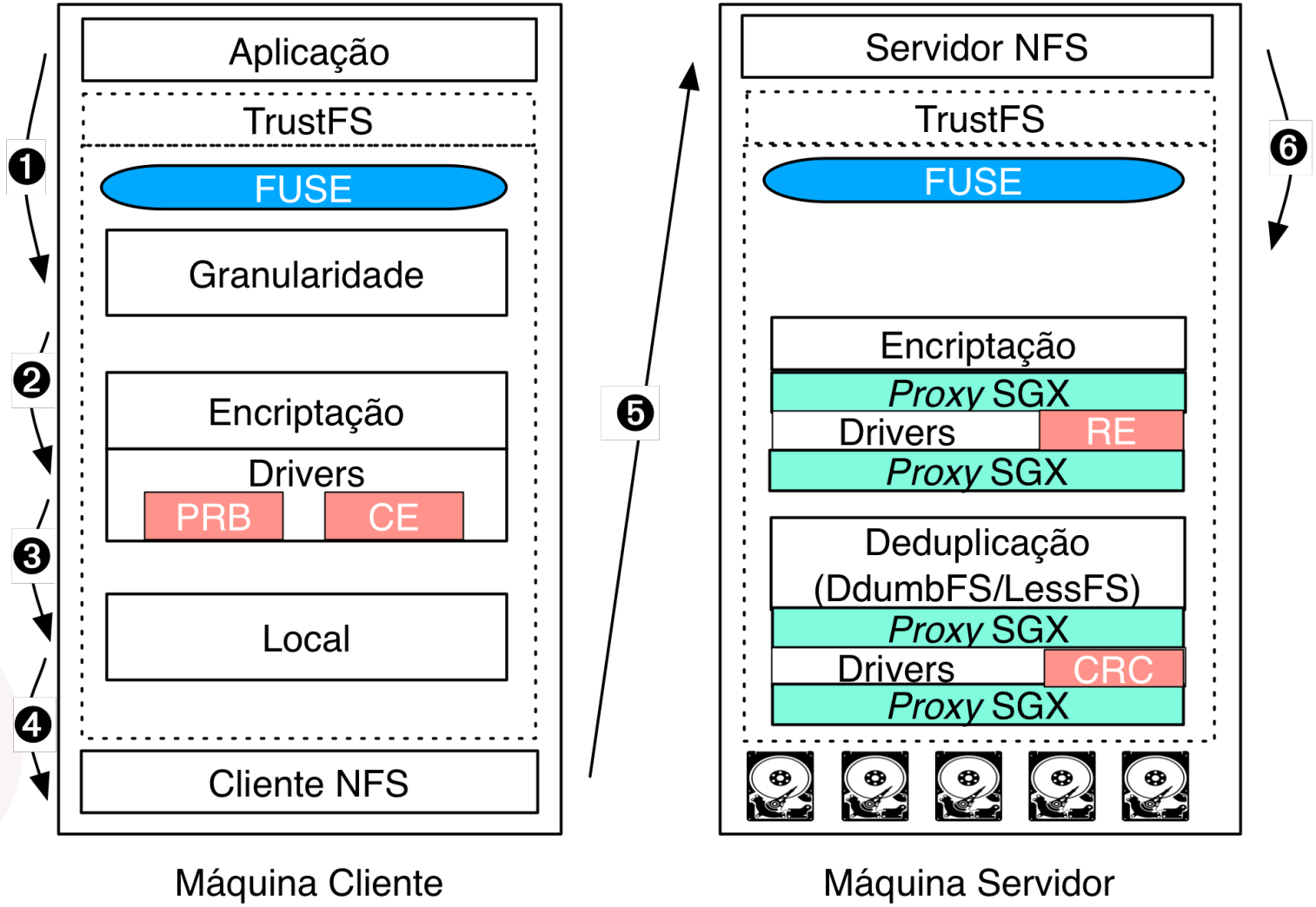
Protótipo | Configurações das camadas



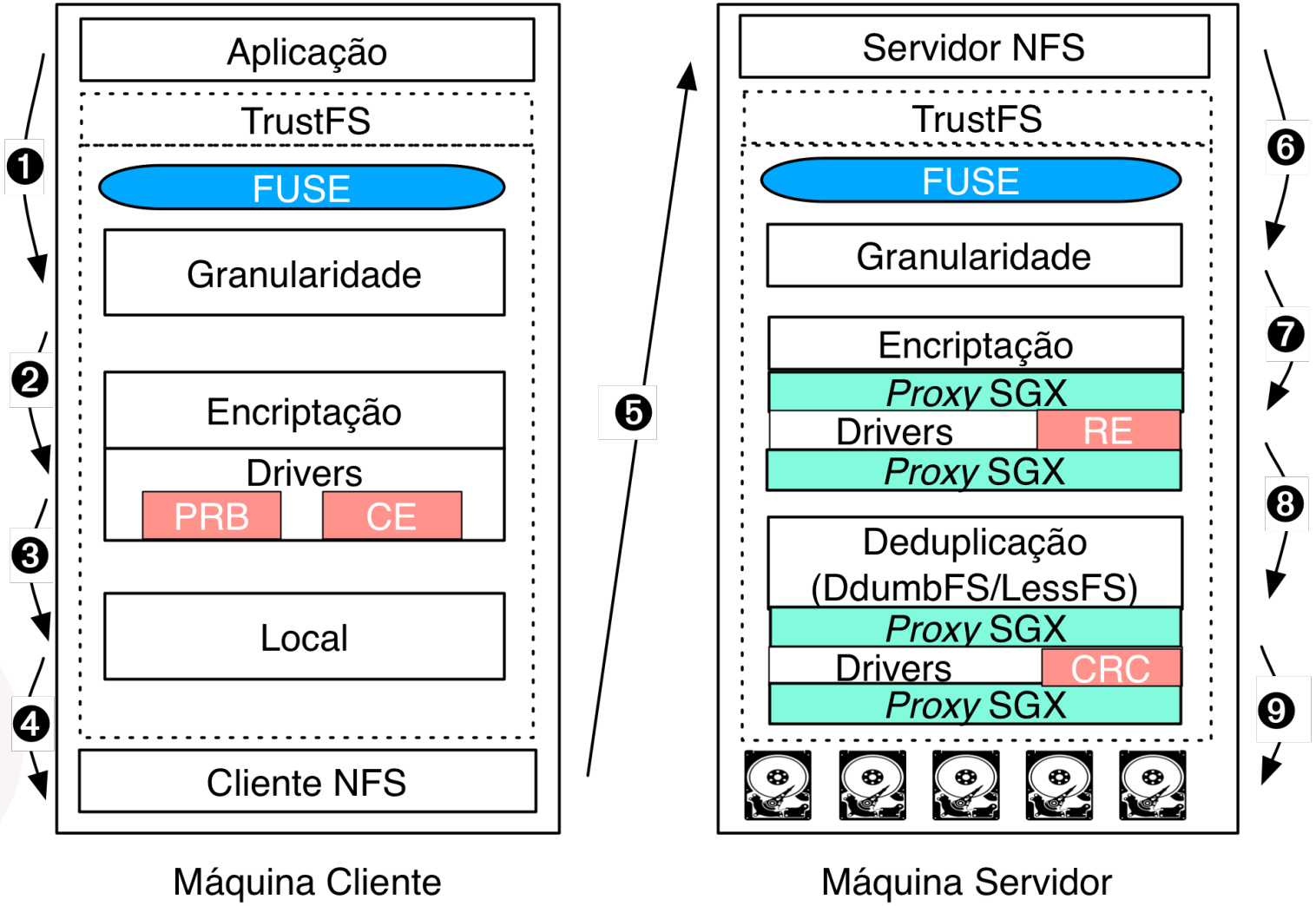
Protótipo | Configurações das camadas



Protótipo | Configurações das camadas



Protótipo | Configurações das camadas



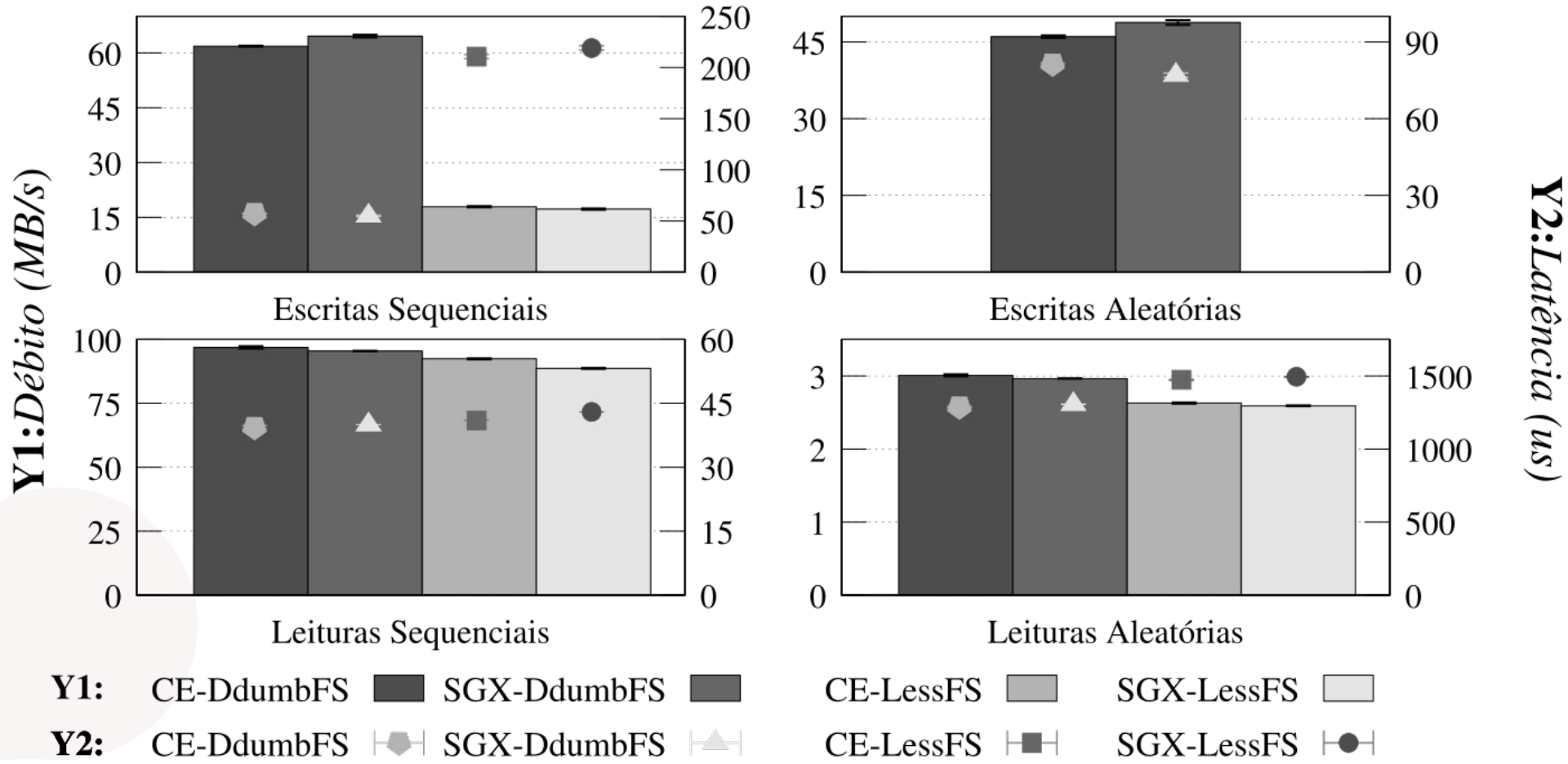
Metodologia

- **Micro Testes** **[5x · 10mins]**
  - OPENSSL vs SGX_SDK vs SGX_SSL
  - Tamanho I/O: 4KB, 16KB, 32KB, 64KB, 128KB

- **Testes Remotos** **[3x · (32GB | 20mins)]**
  - Testes de Impacto de Integração
  - CE vs SGX

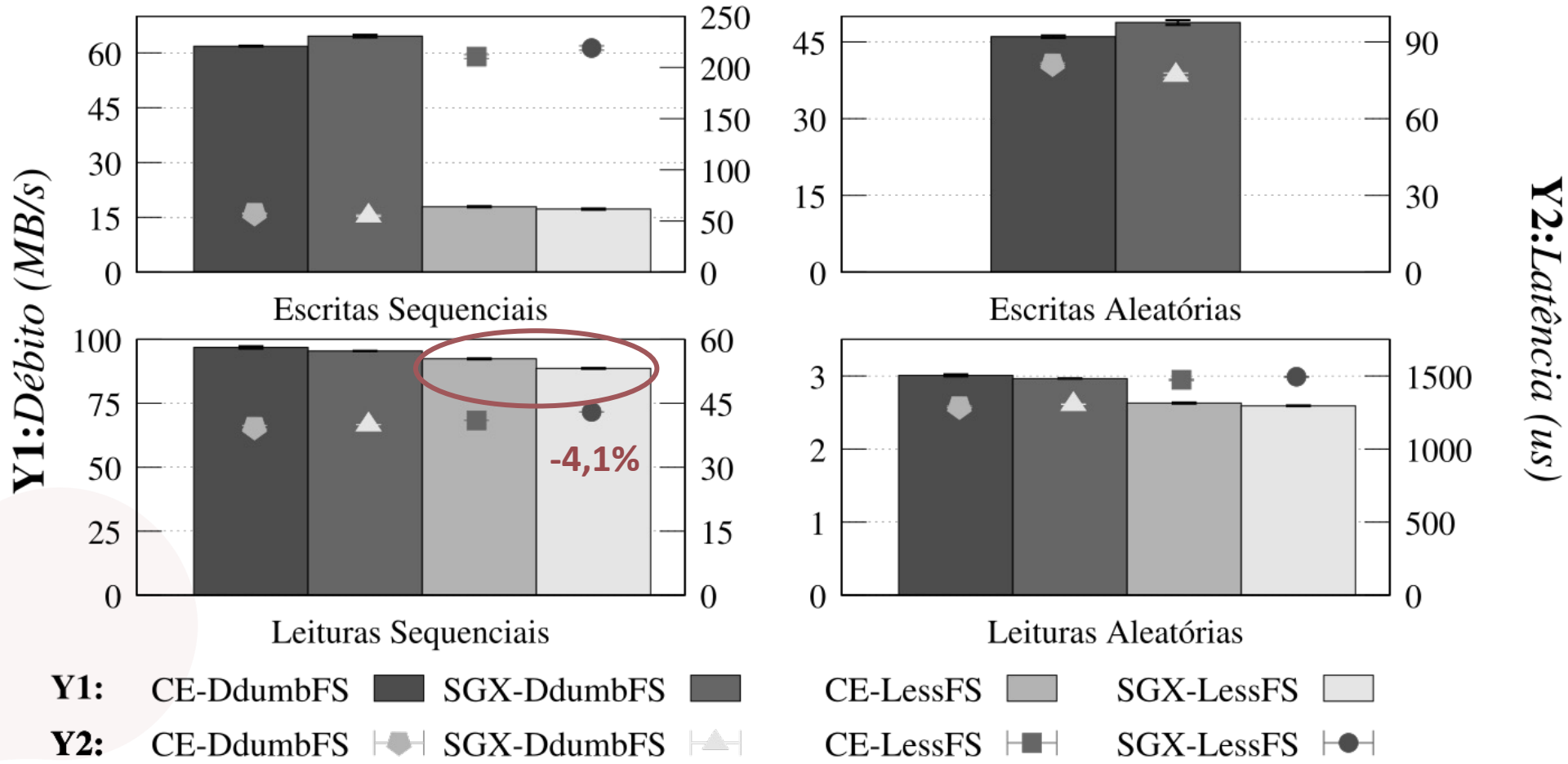
Avaliação Experimental | Testes Remotos

Desempenho da Solução Segura de Deduplicação com SGX



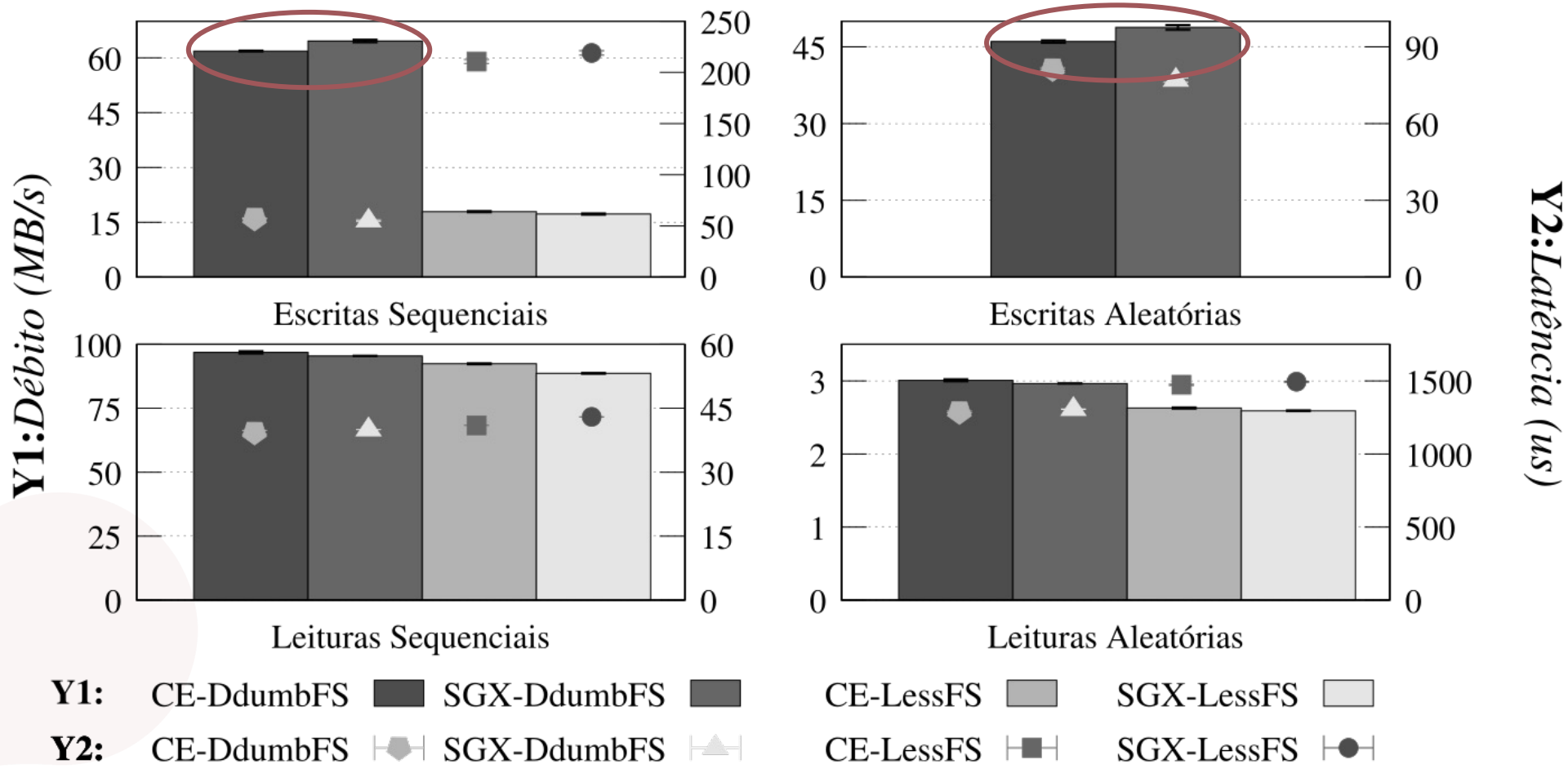
Avaliação Experimental | Testes Remotos

Desempenho da Solução Segura de Deduplicação com SGX



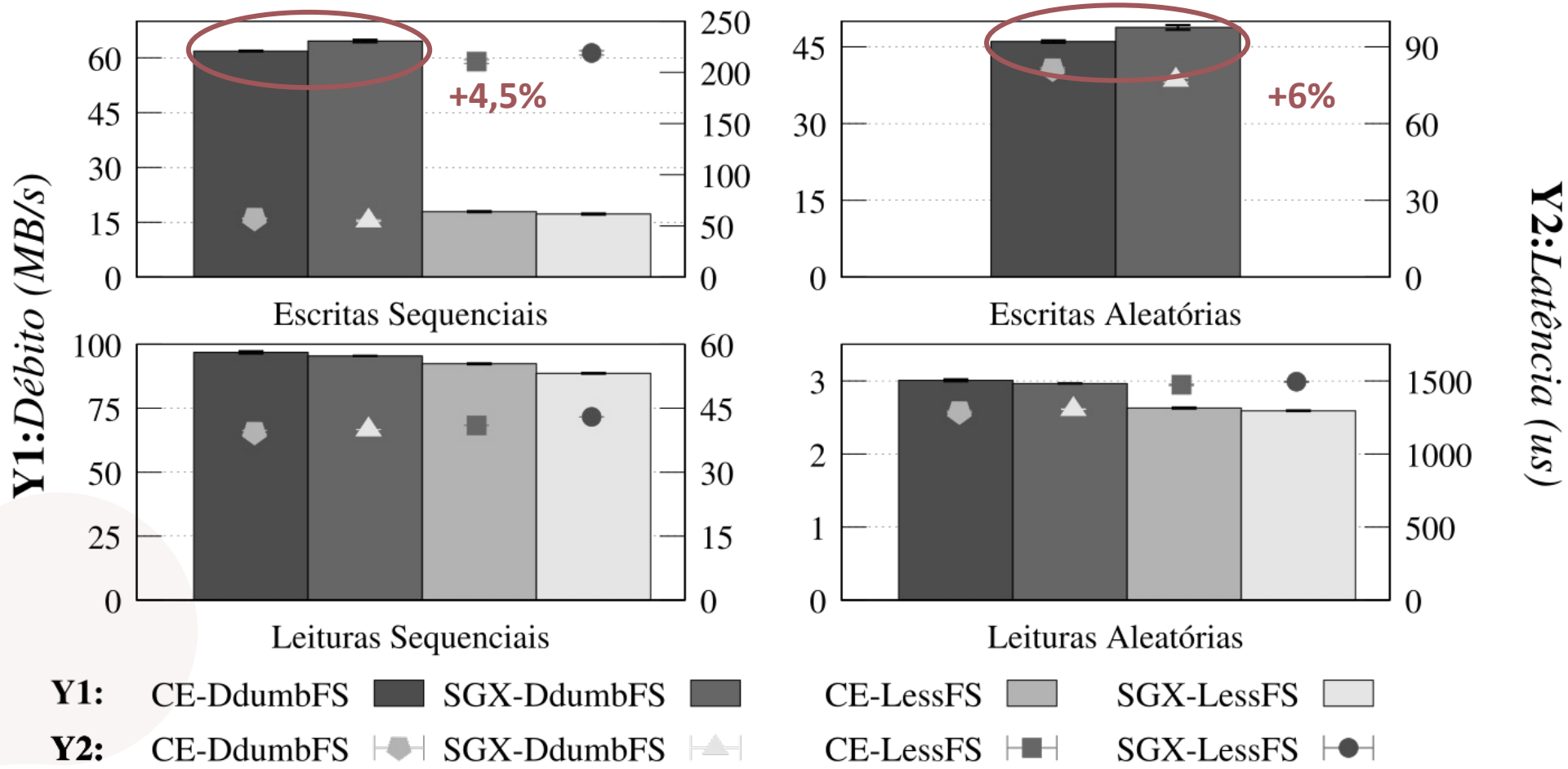
Avaliação Experimental | Testes Remotos

Desempenho da Solução Segura de Deduplicação com SGX



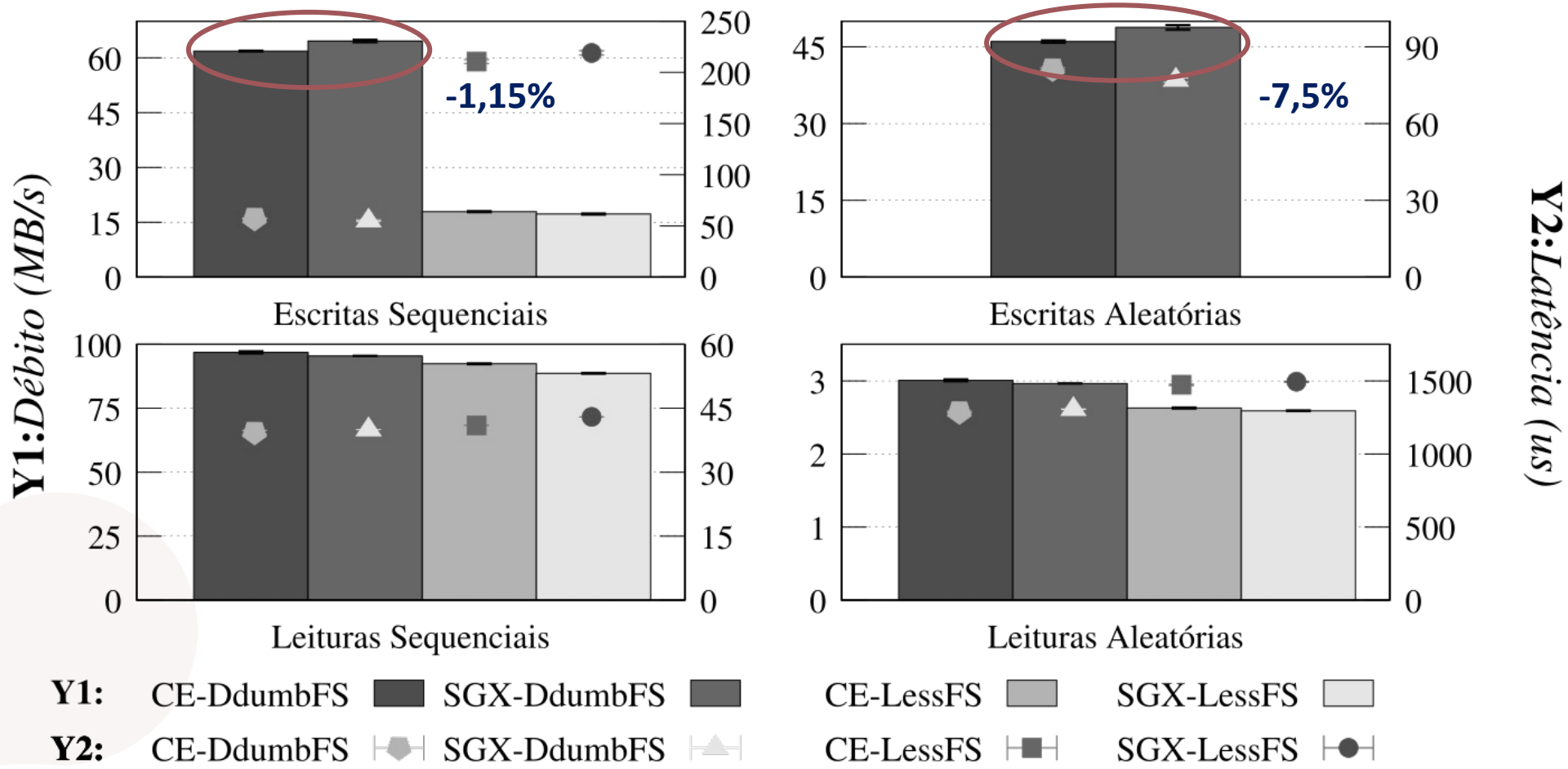
Avaliação Experimental | Testes Remotos

Desempenho da Solução Segura de Deduplicação com SGX



Avaliação Experimental | Testes Remotos

Desempenho da Solução Segura de Deduplicação com SGX



Avaliação Experimental | Espaço poupado

Redução de espaço de armazenamento alcançada

Distribuição	1 GB	5 GB	10 GB	Sem Épocas
dist_highperformance	**5.38 GB**	5.42 GB	5.46 GB	5.60 GB
dist_personalfiles	**0.85 GB**	1.51 GB	2.08 GB	3.72 GB

- ***dist_highperformance***: redundância provém de blocos com **grande** quantidade de duplicados
- ***dist_personalfiles***: redundância provém de blocos com **pequena** quantidade de duplicados

Avaliação Experimental | Espaço poupado

Redução de espaço de armazenamento alcançada

Distribuição		1 GB	5 GB	10 GB	Sem Épocas
dist_highperformance	-4%	**5.38 GB**	5.42 GB	5.46 GB	5.60 GB
dist_personalfiles		**0.85 GB**	1.51 GB	2.08 GB	3.72 GB

- ***dist_highperformance***: redundância provém de blocos com **grande** quantidade de duplicados
- ***dist_personalfiles***: redundância provém de blocos com **pequena** quantidade de duplicados

Avaliação Experimental | Espaço poupado

Redução de espaço de armazenamento alcançada

Distribuição		1 GB	5 GB	10 GB	Sem Épocas
dist_highperformance	-4%	**5.38 GB**	5.42 GB	5.46 GB	5.60 GB
dist_personalfiles	-77%	**0.85 GB**	1.51 GB	2.08 GB	3.72 GB

- ***dist_highperformance***: redundância provém de blocos com **grande** quantidade de duplicados
- ***dist_personalfiles***: redundância provém de blocos com **pequena** quantidade de duplicados

Conclusões

- Plataforma para sistemas de armazenamento dotada de SGX
  - Impacto reduzido no desempenho (2% a 8%)
- Esquema seguro de deduplicação por épocas
  - Compromisso entre segurança e espaço de armazenamento poupado
  - Dependente das características temporais dos dados

Trabalho Futuro

- Nova camada de deduplicação que tire o maior partido do SGX 2.0
- Integrar protocolos de estabelecimento de canais seguros
- Adicionar novas funcionalidades de armazenamento ao TrustFS

Publicações

Submissão de um artigo para a conferência USENIX FAST'19 (26 de setembro de 2018):

- Tânia Esteves, **Ricardo Macedo, Bernardo Portela, João Paulo, Danny Harnik e José Pereira. "** *TrustFS: A Secure SGX-enabled Stackable File System Framework*".