

Caracterização de Compartilhamento de Dados para as Aplicações do *Benchmark NPB*

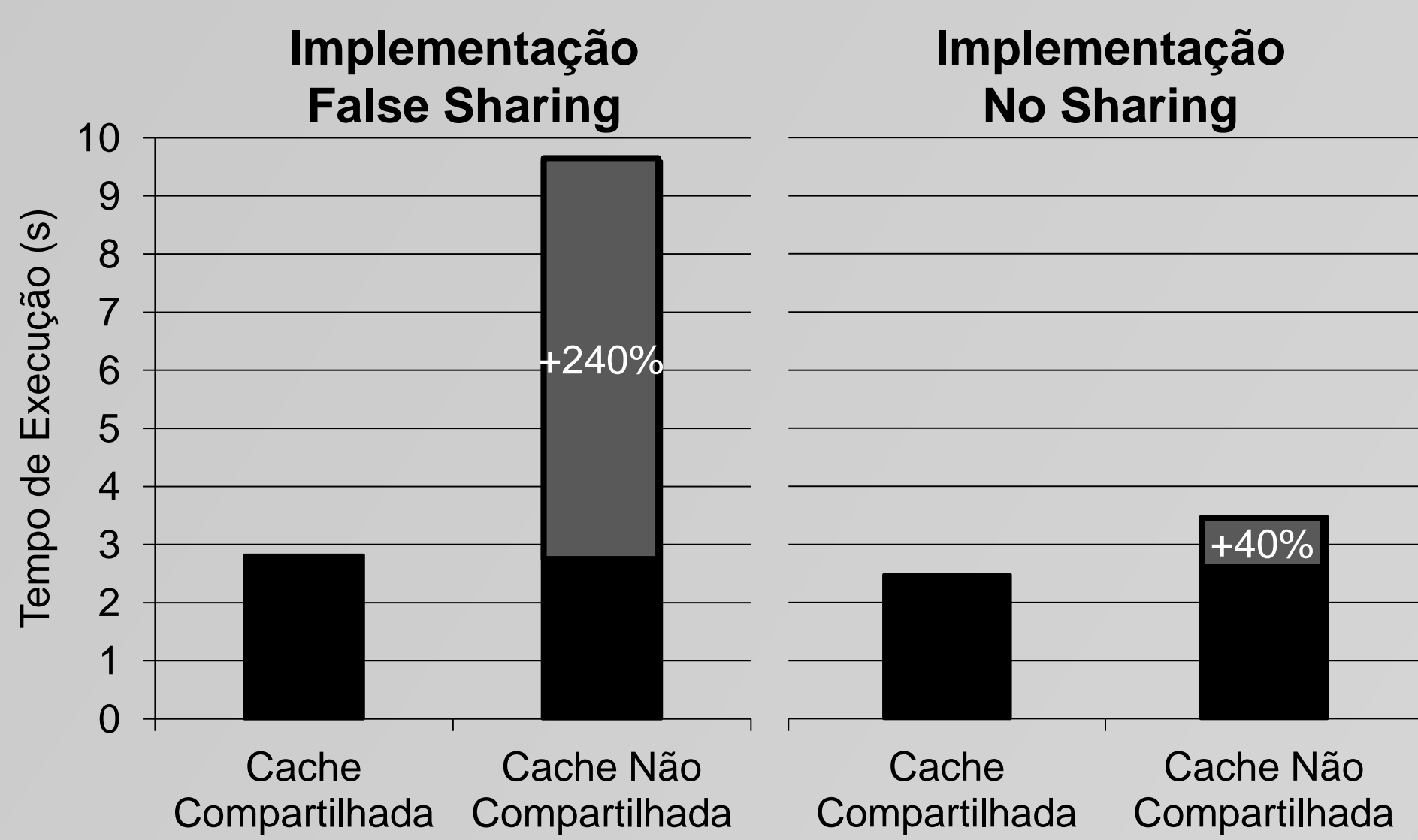
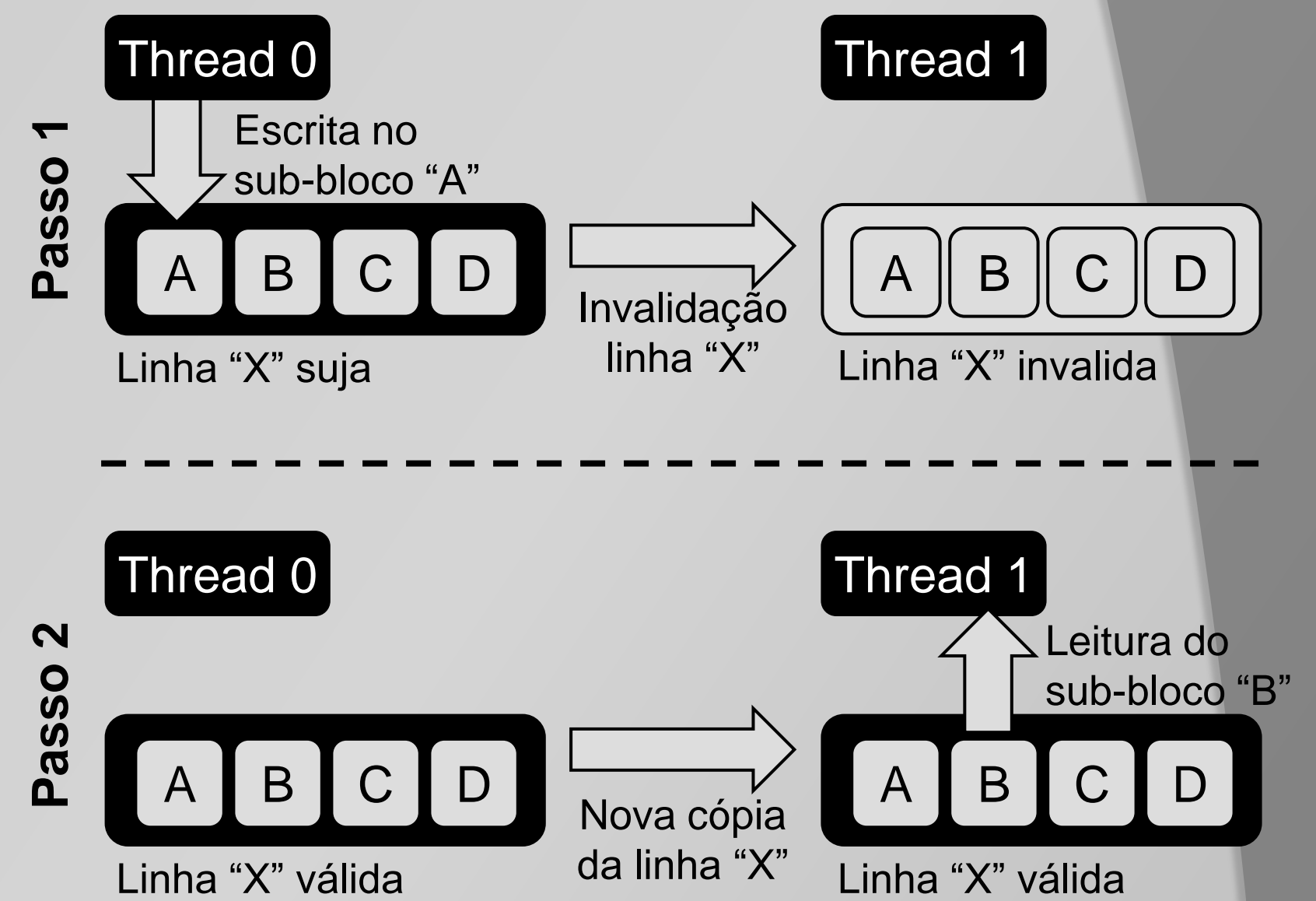
Lucas T. Cardozo, Marco A. Z. Alves, Matthias Diener, Philippe O. A. Navaux
 Grupo de Processamento Paralelo e Distribuído, Instituto de Informática
 Universidade Federal do Rio Grande do Sul, Porto Alegre, Brasil
 {lcardozo, mazalves, mdiener, navaux}@inf.ufrgs.br

Introdução

O falso compartilhamento ocasiona uma redução considerável do desempenho e aumenta o consumo de energia e a contenção das interconexões.

Este problema pode ser exemplificado por duas threads em execução em diferentes núcleos com caches privadas. A **Thread 0** grava alguns sub-blocos dentro de certa linha de cache (**X**) várias vezes, enquanto outra **Thread 1** lê diferentes sub-blocos da mesma linha de cache (**X**).

Dessa forma, se os sub-blocos estivessem em duas linhas de cache diferentes (**X** e **Y**), ambas as threads poderiam realizar as suas operações ao mesmo tempo. Como os sub-blocos pertencem à mesma linha de cache (**X**), o protocolo de coerência de cache irá gerar invalidações na linha (**X**) toda vez que a **Thread 0** gravar dados na linha (**X**). Além disso, após cada invalidação, se a **Thread 1** ler a linha (**X**) será necessário pedir uma nova cópia da linha de cache (**X**) prejudicando ainda mais o desempenho.



Motivação

O impacto do falso compartilhamento sobre o desempenho pode ser visto usando um benchmark produtor-consumidor sintético, que consiste em duas threads, que escrevem em diferentes locais de um mesmo vetor.

Na implementação **“False Sharing”**, as threads escrevem em posições distintas da mesma linha da cache, gerando muitas mensagens de invalidação, o que resulta na diminuição do desempenho.

No código **“No Sharing”**, obtemos a mesma funcionalidade sem o problema do falso compartilhamento.

Podemos ver que o falso compartilhamento tem maior impacto no tempo de execução em memórias cache não compartilhadas (240%), devido ao aumento de invalidações e faltas de dados na cache durante a execução.

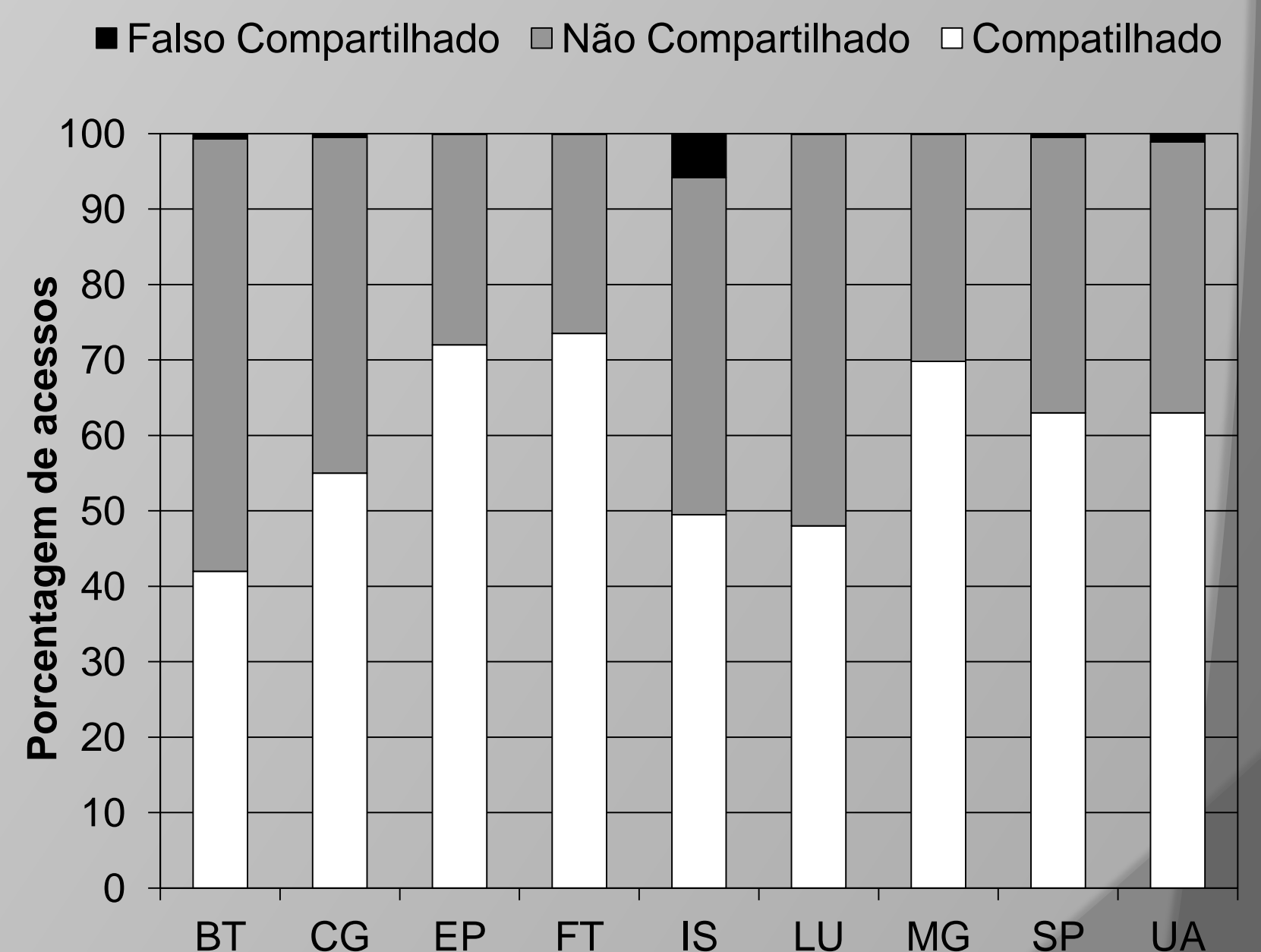
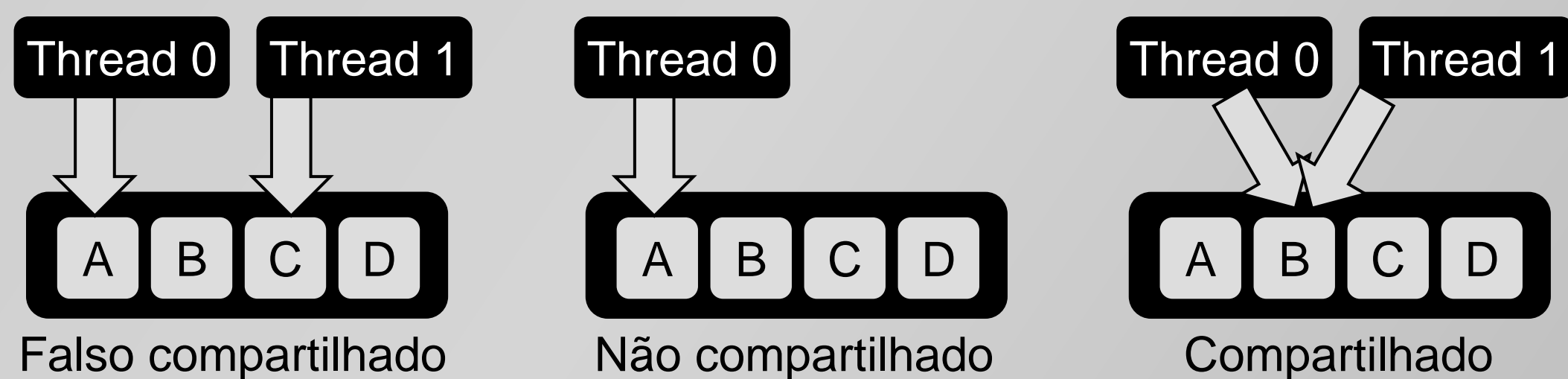
Metodologia e Resultados

Para medir a quantidade de falso compartilhamento, foi desenvolvido um simulador de memória cache baseado na ferramenta Pin da Intel.

O simulador, utiliza informações sobre qual thread acessou cada sub-bloco na linha da cache para classificar os acessos em: **Compartilhado**, **Não Compartilhado** ou **Falso Compartilhado**.

Como carga de trabalho, utilizamos a implementação OpenMP dos benchmarks paralelos NAS (NPB) com o tamanho de entrada *W*, executando com 8 threads.

Embora as aplicações do benchmark NAS serem otimizadas para evitar falso compartilhamento, notamos que esse problema ainda afeta as aplicações IS e UA.



Sumário e Conclusões

Experimentos variando a **associatividade**, **tamanho da linha** e **tamanho da cache** mostraram maior quantidade de falso compartilhamento para, IS(3,2%) e UA(2,6%) quando aumentamos o tamanho da linha da cache.

Além disso, nossas experiências com o benchmark sintético demonstram que o falso compartilhamento pode ser um sério problema de desempenho para arquiteturas multi-core modernas. Portanto, programadores de aplicações HPC precisam levar em conta este problema e programar de uma forma que minimize falso compartilhamento utilizando melhores algoritmos ou técnicas como padding.

Atributo	Falso Compartilhamento	Influência
Maior Associatividade	Menos	Média
Maior Tamanho da Linha	Mais	Alta
Maior Tamanho da Cache	Mais	Baixa