

# NOVOS ALGORITMOS DE JUNÇÃO PARA BANCOS DE DADOS IN-MEMORY

XXXVII Encontro de Iniciação Científica

Thiago de Sousa Garcia, Jose Maria da Silva Monteiro Filho

O surgimento das plataformas de computação em nuvem possibilitou o desenvolvimento de aplicações de abrangência global, como as redes sociais e os sites de vendas online, por exemplo. Tais aplicações envolvem grandes volumes de dados e possuem requisitos rígidos quanto ao tempo de processamento das transações. Neste contexto, surgiram os sistemas de bancos de dados em memória principal (In-Memory Database Systems ou IMDSs), os quais buscam explorar a crescente quantidade de memória principal disponível, além dos enormes caches de dados sub-processados ou descartados. Um IMDS é um sistema de gerenciamento de banco de dados que armazena os dados completamente na memória principal. Trabalhar com dados na memória é muito mais rápido do que escrever e ler a partir de um sistema de arquivos. Porém, os sistemas de banco de dados (SBDs) atuais foram projetados e desenvolvidos sob a premissa de que os dados estariam armazenados em discos rígidos (HDs). Neste sentido, os módulos que compõem um sistema de gerenciamento de bancos de dados (SGBD) necessitam ser reescritos ou adaptados com o objetivo de se adequar ao fato dos dados estarem armazenados completamente em memória principal. A operação de junção é, sem dúvida, uma das operações mais importantes em SGBDs relacionais, uma vez que é executada frequentemente e apresenta um custo bastante elevado. Por estes motivos, novos algoritmos de junção, que possam tirar proveito do fato dos dados estarem armazenados em memória principal tornam-se necessários. Esses algoritmos poderiam explorar diferentes estratégias, tais como a execução em pipeline, paralelismo e adaptabilidade, visando diminuir a geração de resultados intermediários e, conseqüentemente, reduzindo a utilização da memória principal e o número de operações de escrita em disco, que são necessárias quando os dados temporários não cabem na memória principal.

Palavras-chave: Bancos de Dados. Memória Principal. Operação de junção. Algoritmo de Junção.