

## Dica técnica 13 – Alternativas para Melhor Desempenho

Ago 2021

O crescimento massivo dos dados disponíveis tem gerado necessidades de gerenciamento de bases de dados cada vez maiores. Esse aumento quase exponencial muitas vezes dificulta a análise direta dos dados, demandando ferramentas de BI que transformam dados isolados em informação valiosa para os negócios.

Para tornar mais racional e ágil o tratamento desses dados, há algumas alternativas que podem ajudar nesse processo, mesmo em casos de volume de dados transacionais. Propõe-se aqui 3 grupos de abordagens que podem ser utilizados para melhorar o desempenho do **SuperBI**.

### Grupo 1 - Boas práticas de transformação dos dados

Este grupo é o primordial e orientado ao consultor, analista e/ou desenvolvedor que define os dados a serem disponibilizados à ferramenta de BI. Orienta-se a:

- a) Mapear criteriosamente os dados a serem transferidos para o *datamart* (ou “cubo”) de BI. Cabe ponderar quais são realmente necessários, i.e., que agregam valor à tomada de decisão, evitando inchar o cubo com dados operacionais que podem ser consultados diretamente no aplicativo de negócios ou ERP. Este exercício é benéfico independente da limitação de volume, resultando em cubos mais otimizados e alinhados à proposta do BI;
- b) Procurar usar colunas com descrições sucintas ou abreviadas, ao invés de usar colunas de descrições com texto longo, que tem custo maior de processamento. Exemplos:
  - a. Nome de fantasia (do cliente, fornecedor etc.) ao invés de razão social.
  - b. Nas decomposições de data em trimestre, como o título da coluna (“Trimestre”) já explica o conteúdo, neste basta inserir “1”, “2”, “3” e “4”, ao invés de “1º trimestre”, “2º trimestre” etc;
- c) Agrupar os dados, sumarizando-os por tempo, tipo e/ou outra dimensão usando a cláusula GROUP BY no SQL script;
- d) Diminuir o número de linhas do cubo, limitando o período (intervalo de data) de transações a serem transferidas da base de dados transacional para o cubo;
- e) Havendo coluna(s) calculada(s) no cubo, cujo tratamento tende a ter custo de processamento maior que as colunas calculadas no programa extrator, verificar a real necessidade de mantê-las;
- f) Ao invés de criar um único cubo contendo grande quantidade de colunas, tentar distribuir as colunas em mais de um cubo;
- g) Estabelecer processos intermediários de tratamento dos dados transacionais/operacionais, de modo a sumarizá-los/sintetizá-los por meio de um repositório denominado *data warehouse*.

### Grupo 2 - Escalabilidade da Arquitetura de Hardware

A arquitetura de hardware pode definitivamente impactar no desempenho do **SuperBI**. Quando o volume de dados começa a refletir significativamente na performance de entrega dos objetos de BI, **tendo já sido adotadas as boas práticas indicadas no Grupo 1**, é possível revisar os recursos de hardware do servidor onde se encontra o **SuperBI** para melhorar seu desempenho.

A observação de um especialista pode indicar exatamente quais os recursos que devem ser incrementados.

De modo geral, as experiências indicaram que podem ser incrementados os seguintes itens no servidor:

- a) Incrementar o número/velocidade de CPUs do servidor, para pelo menos 2 CPUs;
- b) Incrementar os recursos de armazenamento em disco, nas seguintes configurações:
  - a. Adotar SSDs espelhados para o Sistema Operacional do Servidor e do **SuperBI**. Esta é a melhor configuração;
  - b. Adotar RAID, que são subsistemas processadores redundantes de acesso a disco, que melhoram especialmente a velocidade de leitura de dados. Sugere-se optar pelas seguintes configurações:
    - i. RAID 0+1, que apresenta o melhor desempenho/segurança;
    - ii. RAID 5 ou RAID 6, que apresentam ótimo desempenho de leitura, porém com gravação mais lenta que o RAID 0+1;
- c) Avaliar a possibilidade de incremento da capacidade de RAM para pelo menos 8 GB. O **SuperBI** não é exigente em demasia de recursos de RAM, considerando-se que o que mais pode afetar este item é o número de usuários simultâneos do cliente.

### Grupo 3 – SGBD MariaDB com ColumnStore Engine

A adoção das orientações do **Grupo I** e **II** acima podem contribuir significativamente no desempenho do **SuperBI**. Porém, nas situações em que o volume de dados armazenado no repositório de cubos atinge patamares muito elevados, pode-se adotar a técnica de segregação do repositório de cubos. Por padrão, este repositório é instalado em um SGBD MariaDB no mesmo servidor IIS.

Para adotar esta funcionalidade, é necessário que o cliente tenha o licenciamento do **SuperBI** na modalidade **Advanced**. Esta modalidade permite segregar o SGBD MariaDB em um servidor Linux independente, que deverá residir na mesma *farm* de servidores (*datacenter*) em que se encontra o servidor IIS, com um barramento de integração em rede de altíssima performance entre estes dois servidores.

Nesta configuração de servidores, o SGBD MariaDB deverá ser instalado com a opção de *Engine ColumnStore*, que utiliza técnicas de recuperação de informações próprias para grandes volumes de dados, considerados extremamente grandes.