

70 VEZES MAIS RÁPIDO!

Mário Mendes Barbosa Junior

Licardino Siqueira Pires

Antes de explicar o título da matéria, abordaremos de forma simples sobre o que é um Big Data.

Imagine que você tenha perdido um relógio nas dependências do Tribunal de Contas do Estado de Goiás. Suponha que não exista um serviço de achados e perdidos e que você tenha que procurá-lo sozinho. Eventualmente, seus primeiros passos seriam vasculhar os ambientes por onde percorreu, do mais para o menos provável. Em todo esse processo, alguns minutos, talvez horas, e quem sabe dias, seriam gastos para concluir a busca, podendo ainda contar com o insucesso.

De uma forma simplificada, a busca de informações nos bancos de dados tradicionais poderia ser comparada ao exemplo acima. Alguns índices (lugares mais prováveis) podem ser criados e associados a uma tabela para facilitar a busca por um registro, por um dado. Todavia, ainda é uma busca solitária, muitas vezes sequencial, ou seja, indo de um ponto a outro até vasculhar tudo.

Agora imagine que você consiga mobilizar mais pessoas para te ajudar a encontrar o relógio; cada uma responsável por um ou mais ambientes. As chances de se encontrar o relógio não só podem aumentar como, certamente, diminuirá drasticamente o tempo de busca. E nesse caso, quanto mais pessoas ajudando, melhor!

Mas o que o tal do Big Data tem a ver com isso? Lembra daquelas pessoas mobilizadas para ajudar na busca pelo relógio? Pois é, o Big Data é um sistema assim. Ele é distribuído, ou seja, ele conta com vários “nós” de processamento, de forma que, toda informação nele armazenada é distribuída para todos esses “nós” que garantem não apenas uma excelente performance na busca, mas também confiabilidade! Todas as informações são replicadas em vários “nós”, de forma que se algum deles falha ou está sobrecarregado, os outros são acionados.

Outra vantagem é a seguinte, os bancos de dados tradicionais guardam as informações apenas de forma estruturada, no formato de tabelas. Já o Big Data não, ele armazena de tudo, documentos de texto, imagens, vídeos, áudios e também tabelas. Você deve ter notado o poder de armazenamento e de busca de uma plataforma de Big Data, uma vez que ele

armazena e busca tanto os dados estruturados de bancos de dados e muito mais, com uma performance ainda melhor.

Então vamos para um caso prático.

A Gerência de TI deste Tribunal implantou em seu parque tecnológico a plataforma de Big Data a fim de receber, com segurança e performance, as bases de dados integrais dos jurisdicionados. Em uma delas, consta a informação dos Restos a Pagar de todas as unidades do Estado em sua posição diária (se já foi pago, se foi cancelado ou se permanece como está). Essa base de dados tem mais de 130 milhões de registros e por isso ela foi escolhida para a realização de testes de performance.

A consulta a esses dados, que são estruturados, demorou 77 segundos no banco de dados tradicional utilizado pelo TCE-GO. Já em uma primeira tentativa de busca no Big Data, a mesma consulta levou 56 segundos. Nada de impressionante até aqui. O próximo passo foi aumentar a quantidade de “nós” para ver se o desempenho seria melhor. Olha que interessante: o tempo reduziu para 29 segundos, ou seja, 38% mais rápida!

Como a equipe da Gerência de TI é composta por bons cientistas, e conta com uma consultoria especializada em implantação de Big Data, encabeçada pelo professor e doutor pela Universidade Federal de Goiás, Sávio Teles, foram alterados alguns parâmetros de configuração da ferramenta que faz a busca no Big Data. O resultado obtido foi impressionante! A busca de mais de 130 milhões de registros levou menos de 1 segundo! Sim, retornou os dados num piscar de olhos!

Naturalmente, como toda boa ciência requer, não podemos generalizar esse ganho alcançado para todos os casos daqui para frente, até porque teremos vários tipos de informações guardadas na plataforma de Big Data, e muitas delas nem referência existirá para comparação se a busca foi realmente muito melhor. Todavia, esse teste demonstra o avanço para um novo patamar tecnológico que esta Corte está adentrando, não só para ganho de desempenho, mas mudando o paradigma de obtenção e disponibilização de informação aos membros e auditores de controle externo.

Vale lembrar que existe um serviço de achados e perdidos aqui no TCE-GO.