O Papel dos Bancos de Dados SQL e NoSQL na Engenharia de Dados

Atualmente, é imprescindível a presença de um banco de dados. Seja para uma corporação, um hospital ou análises meteorológicas, um banco de dados é um recurso essencial para gerenciamento, manutenção e previsão de informações. Posto isto, neste artigo descrevemos em linhas gerais os papéis dos bancos de dados para engenharia de dados.

**O que é um banco de dados SQL?**

Antes de tudo, precisamos entender o que é um banco relacional (SQL). SQL vem de **Structed Querry Language,** ou **Linguagem de Consulta Estruturada** uma linguagem para bancos relacionais. Bancos de Dados Relacionais são voltados para relacionamentos entre os dados, onde temos situações onde podemos abstrair as informações em um esquema de relacionamento de entidades. Posto isso, este tipo de banco de dados é utilizado quando queremos uma normalização dos dados sem criar muitas duplicidades. Além disso, o modelo relacional é bastante utilizado quando queremos um esquema de dados rígido, isto é quando conhecemos de antemão como nossos dados serão armazenados e como eles irão se relacionar com outros dados.

**O que é um banco de dados NoSQL?**

Em contrapartida, os bancos não relacionais (Not Only Structed Querry Language) são bancos de dados que não seguem uma estrutura relacional. Este modelo vem crescendo como uma alternativa para dados que não necessitam de modelo de dados rígidos.

Com essas definições em mente, é notório que o mundo vem produzindo uma grande quantidade de dados atualmente. Com a internet das coisas, temos uma rede de dispositivos capazes de coletar, transmitir e processar dados e com essa grande quantidade de dados, o armazenamento e processamento destes é o novo desafio. Os bancos de dados relacionais vêm sendo utilizados por grandes corporações, entretanto, a produção elevada de dados imputam, aos bancos relacionais,dificuldades de escalabilidade e performance, considerando o respeito às propriedades ACID e as formas normais.

Com o surgimento da IoT (Internet das coisas) o volume gerado e a dificuldade de manipulação desses dados crescem a cada dia. Além do grande volume de dados a serem armazenados, temos os desafios de extrair destes dados informações úteis. A forma de armazenamento implica diretamente em como manipular e na facilidade de cumprir com essa função.

Essa nova demanda de flexibilidade e escalabilidade das aplicações por conta do volume de dados, velocidade de processamento e/ou falta de estrutura fixa dos dados, em muitos casos emerge o uso de sistemas gerenciados de banco de dados não relacionais, também conhecidos como SGBDs NoSQL, ressalta a importância da utilização deste para diferentes tipos de SGBD, promovendo a ferramenta ao armazenamento de dados, de acordo com a necessidade.

Portanto, com o crescimento contínuo na geração de dados, esta elevada quantidade impacta diretamente na escolha de um SGBD a ser utilizado no desenvolvimento de um software. Neste momento os engenheiros de dados devem considerar os prós e contras de cada tipo de SGBD, tendo em vista que os banco de dados são essenciais para a sociedade moderna. Todo banco de dados cria alguma abstração sobre os dados, sendo essa característica importante para definição de qual banco de dados será utilizado em uma solução de software ou sistema.