Integração e Governança de Dados

A necessidade de gerenciar a complexidade, juntamente com os seus custos decorrentes, são os motivos para arquitetar a integração de dados em uma perspectiva corporativa. Um projeto corporativo de integração de dados é comprovadamente mais eficiente e econômico do que soluções distribuídas ou ponto a ponto, visto que o desenvolvimento de soluções ponto a ponto entre sistemas pode acabar sobrecarregando até mesmo os recursos de suporte de TI mais eficientes (DAMA, 2017).

Data Hubs, como *Data Warehouses* e soluções de Gestão de Dados Mestres auxiliam na minimização de tais problemas por meio da consolidação e provimento de dados consistentes para usuários e outros sistemas. Da mesma forma, a complexidade do gerenciamento de dados operacionais e transacionais que precisam ser compartilhados em toda a organização pode ser bastante simplificada por meio da utilização de técnicas de integração de dados corporativos, como integração *hub-and-spoke* e modelos de mensagens canônicas (DAMA, 2017).

Sabe-se que Integração e interoperabilidade de dados é algo crítico para *Data Warehousing* e *Business Intelligence*, bem como para a Gestão de Dados Mestres e de Referência, visto que esses tópicos concentram-se na transformação e integração de dados de sistemas de origem para *Data Hubs* (*hubs* de dados) consolidados, e destes para sistemas de destino, de onde os dados podem ser entregues tanto para outros sistemas quanto diretamente para para humanos (DAMA, 2017).

O modelo *hub-and-spoke* de *Data Hub* mostra-se como uma alternativa às integrações do tipo ponto-a-ponto. Este modelo consolida dados compartilhados (fisicamente ou virtualmente) em um *hub* de dados central em que muitos aplicativos podem utilizar. Deste modo, todos os sistemas que desejam trocar dados o fazem por meio de um sistema central comum de controle de dados, em vez de diretamente entre si (ponto a ponto).

Os *Enterprise Service Buses* (ESB) constituem um tipo de solução de integração para compartilhamento de dados quase em tempo real entre vários sistemas, onde o *hub* é um conceito virtual do formato padrão ou do modelo canônico para compartilhamento de dados na organização.

Uma arquitetura *hub-and-spoke* nem sempre pode ser a melhor solução. A latência ou o desempenho do modelo pode às vezes não ser adequada para a necessidade. O próprio *hub* cria sobrecarga em uma arquitetura *hub-and-spoke*. Uma solução ponto a ponto não exigiria o *hub*, no entanto, os benefícios do mesmo superam as desvantagens da sobrecarga assim que três ou mais sistemas estão envolvidos no compartilhamento de dados. O uso do padrão de design *hub-and-spoke* para o intercâmbio de dados pode reduzir drasticamente a proliferação de soluções de integração e transformação de dados e, assim, simplificar massivamente o suporte organizacional necessário (DAMA, 2017).

Revisão #1
Criado 2022-03-21 12:42:37 UTC por FLAVIO LOPES DE MORAIS
Atualizado: 2022-12-14 09:02:06 UTC por FLAVIO LOPES DE MORAIS