

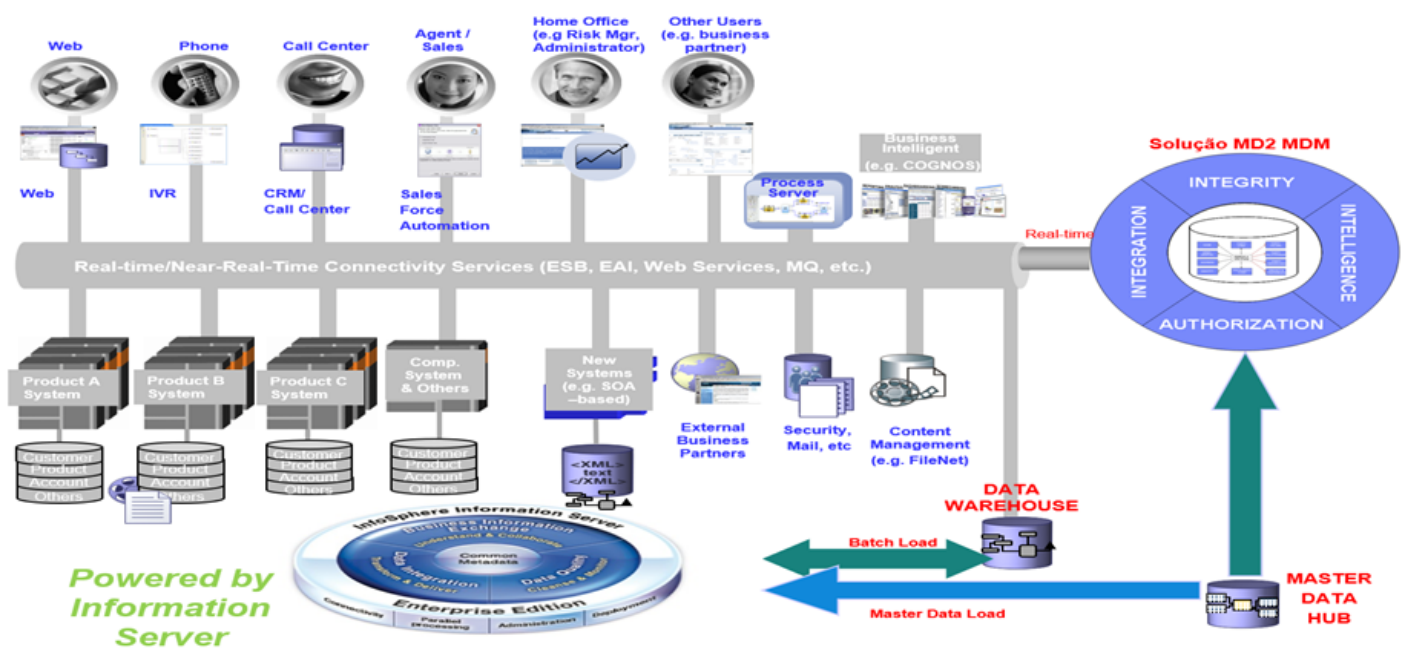
6. Publicação de Dados

- 6.1. Publicação por Serviços Web (SOA)
- 6.2. Publicação por Filas de Mensageria
- 6.3. Publicação de Arquivos
- 6.4. Publicação por Integração de Dados (ETL)

6.1. Publicação por Serviços Web (SOA)

Um Web service é um conjunto de métodos que podem ser invocados por outros programas utilizando tecnologias Web. É utilizado para transferir dados através de protocolos de comunicação padronizados acessíveis diferentes plataformas de forma desacoplada, ou seja, independentemente das linguagens de programação ou tecnologias específicas utilizadas nessas plataformas, integrando mais informação e novas funcionalidades de forma simples e rápida.

O diagrama abaixo ilustra como essa característica da solução pode ser utilizada em um cenário corporativo, através da exposição dos serviços e opcionalmente controlando o acesso via ESB (Enterprise Service Bus).



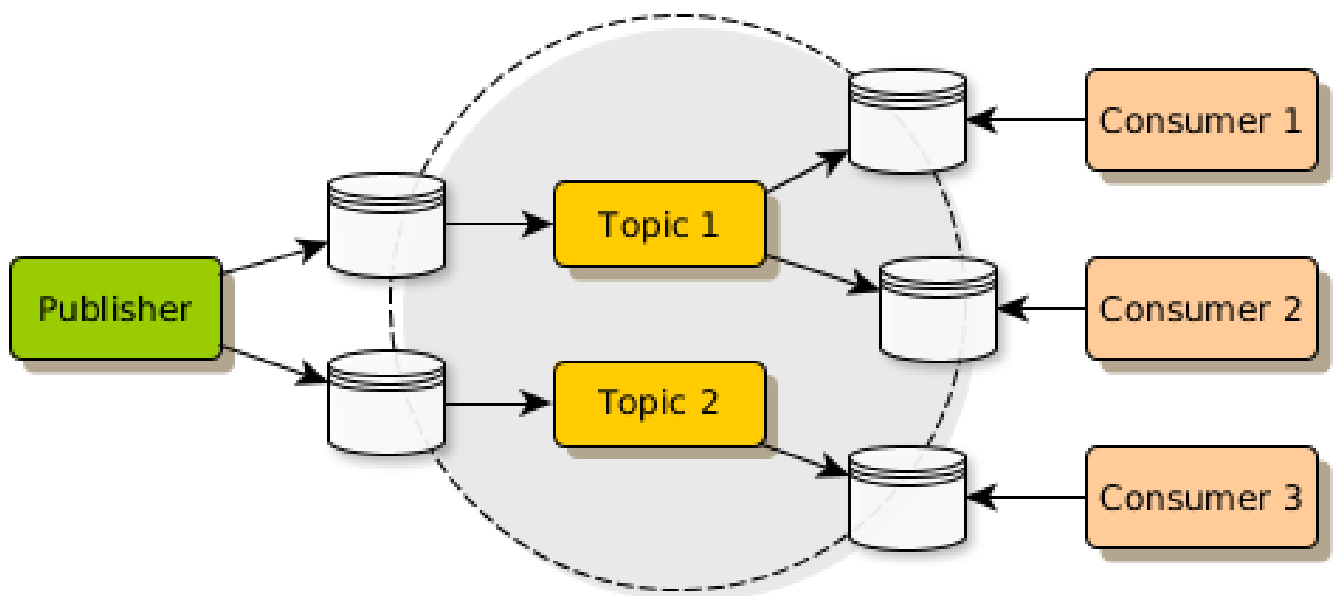
A arquitetura orientada a serviços (SOA) do InfoSphere Information Server garante que a lógica de integração de dados desenvolvida possa ser usada por qualquer processo de negócios. Os melhores dados estão disponíveis o tempo todo, para todas as pessoas e para todos os processos.

A solução MD2 contempla um conjunto de serviços capaz de facilitar o acesso aos dados mestre. Desta forma, as operações de consulta e manipulação de cadastros pode ser feita diretamente, através de um conjunto de serviços disponíveis para este fim.

6.2. Publicação por Filas de Mensageria

A comunicação por troca de mensagens é outra forma de publicação de dados disponível na solução de MDM da MD2. A estratégia de publicar e assinar mensagens é particularmente útil para processos de retroalimentação de dados, ou seja, o envio do Golden Record a partir do HUB MDM para leitura pelos sistemas legado de forma proativa,

O diagrama abaixo ilustra esta estratégia de publicação de dados da solução.



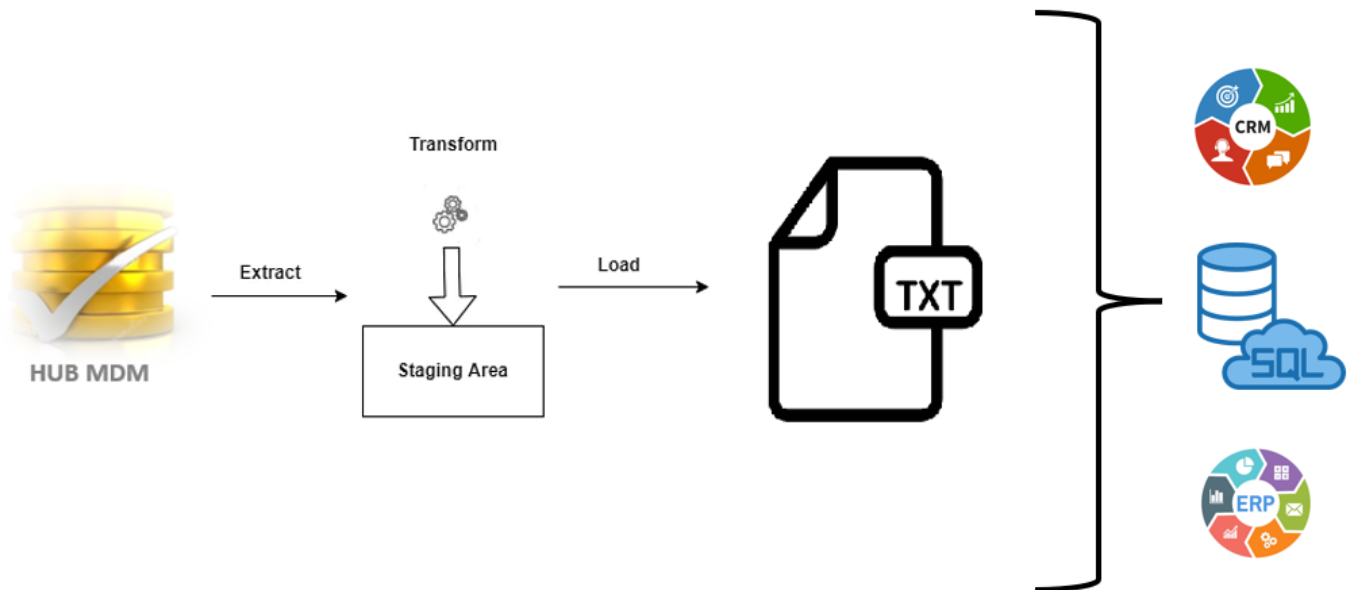
É uma forma de comunicação assíncrona de serviço a serviço usada em arquiteturas de micro serviços. Em um modelo pub/sub, qualquer mensagem publicada para um tópico é imediatamente recebida por todos os assinantes do tópico. As mensagens podem ser usadas para permitir arquiteturas orientadas a eventos, ou para desacoplar aplicações a fim de aumentar o desempenho, confiabilidade e escalabilidade.

Observando o caso de uso específico da solução MDM na perspectiva da estratégia de retroalimentação, essa arquitetura de publicação se mostra extremamente útil e versátil. Toda vez que alguma modificação de dados ocorre no Golden Record, este cadastro masterizado é publicado em uma fila de mensagem, normalmente utilizando uma estrutura genérica representada por um modelo canônico de dados. Dessa forma, todos os assinantes dessa fila, ou seja, outros sistemas legado que também possuem interligação lógica com este Golden Record, recebem esta atualização de forma assíncrona e após realizarem as respectivas adequações

estruturais e as traduções de domínios de dados, podem atualizar sua base de dados com a versão mais rica possível dos dados deste indivíduo que foram centralizadas no HUB MDM.

6.3. Publicação de Arquivos

Em contextos onde os dados a serem publicados são de natureza tabular (dados organizados em tabela), é comum que estejam armazenados numa base de dados e sejam facilmente transportados para algum arquivo adequado ao armazenamento de dados tabulares, como arquivos CSV ou TXT.

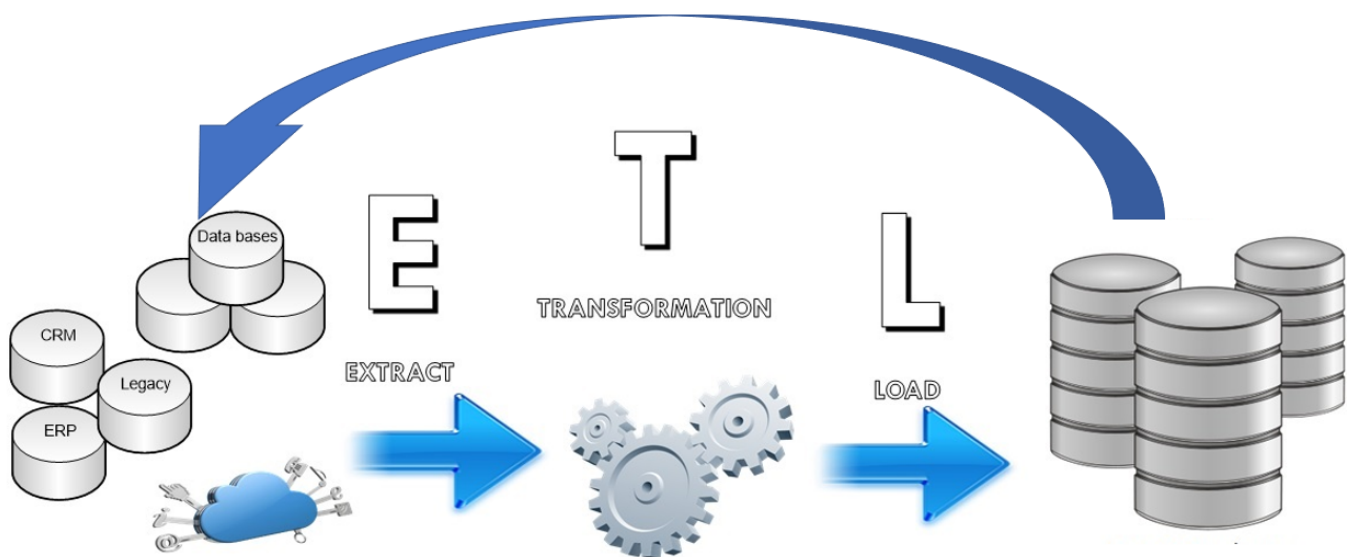


Esta estratégia de publicação de dados se faz extremamente útil quando as aplicações de destino já possuem algum tipo de funcionalidade para a importação de dados a partir de arquivos. Desta forma, a geração dos dados neste arquivo pode utilizar as robustas rotinas ETL para gerar arquivos de dados na estrutura correta adequada à aquela aplicação em particular.

6.4. Publicação por Integração de Dados (ETL)

ETL é um acrônimo para as palavras em inglês Extract, Transform e Load. Refere-se aos procedimentos de extração, transformação e carregamento de dados para um ambiente integrado. Os processos ETL são altamente eficientes no quesito integração de dados, estabelecendo regras de otimização e manipulação padronizada dos dados a fim de facilitar sua inserção em ferramentas ou ambientes integrados.

O diagrama abaixo ilustra estratégia de publicação de dados por ETL.



A publicação de dados mestre através de ETL permite alta flexibilidade na formatação de dados e na aplicação de regras mais robustas de integração ou qualidade de dados, porém cria maior acoplamento entre aplicações uma vez que grava os dados mestre diretamente nas bases de dados das aplicações interessadas, criando uma forma de publicação mais invasiva na ótica dos administradores responsáveis por estes outros sistemas.