

PÓS-GRADUAÇÃO

## Integração e fluxo de dados



PÓS-GRADUAÇÃO

# Ferramentas de ETL

## Bloco 1

Thiago Salhab Alves





## ► Ferramentas de ETL

### Objetivos

- Conhecer as ferramentas de ETL.
- Conhecer a ferramenta Microsoft SQL Server Integration Services (SSIS).
- Conhecer sobre arquiteturas e componentes da ferramenta Microsoft SQL Server Integration Services (SSIS).



## ► Ferramentas de ETL

- De acordo com Pall e Khaira (2013), o processo de integração de dados envolve práticas, técnicas arquiteturais e ferramentas para obter acesso consistente e para entregar dados em uma grande área de assuntos e tipos de estrutura em uma empresa.
- A demanda do mercado se torna cada vez mais diversificadas, e os usuários buscam ferramentas que apoiem iniciativa de *Business Intelligence*.



## ► Ferramentas de ETL

- O processo de ETL consiste na extração (leitura dos dados de um ou mais bancos de dados), transformação (conversão dos dados extraídos de sua forma anterior na forma que precisam estar para ser inseridos em um *data warehouse*) e carga (inserção dos dados no *data warehouse*) (TURNBAN et al., 2009).



## ► Ferramentas de ETL

- O processo de carregar dados para um *data warehouse* pode ser realizado por meio de ferramentas de transformação de dados, que oferecem uma GUI (*Graphical User Interface*) para ajudar no desenvolvimento e manutenção das regras de negócio.



## ► Ferramentas de ETL

- De acordo com Pall e Khaira (2013), as ferramentas de ETL são categorizadas em duas categorias:
  - ETL codificado manualmente.
  - ETL baseado em ferramenta.



## ► Ferramentas de ETL

- ETL codificado manualmente são desenvolvidas em linguagens de programação, tais como, Perl, Cobol, C e PL/SQL, sendo programas longos e de difícil documentação.
  - Vantagem: cria metadados que podem ser gerenciados diretamente.
  - Desvantagem: para atender as mudanças nos grandes volumes de dados, os programas precisam ser modificados com frequência, causando impacto no projeto.





## ► Ferramentas de ETL

- ETL baseado em ferramenta, segundo Pall e Khaira (2013), iniciaram suas extrações em *mainframes* para bancos de dados.
- São as ferramentas ETL que hoje fornecem características de transformação, suportam vários arquivos de banco de dados de entrada ou saída, projetos multidimensionais, várias funções de transformação e banco de dados nativo.



## ► Ferramentas de ETL

- De acordo com Pall e Khaira (2013) algumas ferramentas de ETL se destacam:
- IBM *InfoSphere Information Server*: alto nível de satisfação dos clientes, de fácil uso, com grande flexibilidade, porém se torna “pesado” devido aos dados envolvidos, e requer muito poder de processamento.
- Oracle *Data Integrator*: líder do mercado de ETL, com conexão com todos os aplicativos de *data warehousing* da Oracle e integra todas as ferramentas em uma aplicação e um ambiente.



## ► Ferramentas de ETL

- Microsoft *SQL Server Integration Services* (SSIS) apresenta facilidade e velocidade de implementação com integração de dados, recursos em tempo real, custo baixo e excelente modelo de suporte e distribuição.

PÓS-GRADUAÇÃO

## Ferramentas de ETL

Bloco 2

Thiago Salhab Alves






## ► *SQL Server Integration Services (SSIS)*

- De acordo com Cote, Lah e Sarka (2017), o *SQL Server Integration Services (SSIS)* é uma ferramenta que facilita a extração de dados, a consolidação e as opções de carregamento (ETL), apresentando aprimoramentos de codificação do SQL Server e *data warehousing*.
- *SQL Server Integration Services (SSIS)* permite extrair e transformar uma grande variedade de fonte de dados, tais como arquivos simples, arquivos de dados XML, fontes de dados relacionadas e transferir esses dados para um ou mais destinos.




## ► *SQL Server Integration Services (SSIS)*

- De acordo com a Devmedia (2014), a arquitetura do SSIS é formada por alguns componentes:
  - *Solution*: usado para gerenciar um ou mais projetos, reunindo os projetos em um determinada solução de negócios.
  - *Project*: local que se desenvolve os pacotes, podendo possuir vários pacotes.
- 



## ► **SQL Server Integration Services (SSIS)**

- *Package*: local que se desenvolvem os fluxos que podem ser *Control Flow* ou *Data Flow*;
  - *Control Flow*: principal componente do pacote e seus principais itens são containeres e tarefas. *Data Flow* é a tarefa de fluxo de controle para transferência de dados.
- 

**PÓS-GRADUAÇÃO**

## **Teoria em Prática**

**Bloco 3**


Thiago Salhab Alves







## ► Teoria em Prática

- Uma empresa nacional de revenda de cosméticos está enfrentando alguns problemas financeiros, tais como:
  - Dificuldades em acompanhar a demanda por produtos de lançamento, o que muitas vezes acaba por comprometer o capital de giro, investindo em produtos com baixa procura.
  - Produtos que possuem prazo de validade curto, que acabam por vencer e não podem ser trocados por parte do fornecedor.
- 



## ► Teoria em Prática

- Sistema de vendas e controle de estoque, com banco de dados relacional e um processo de marketing pelas redes sociais, porém está tendo dificuldades para a tomada de decisões.
- Para que a empresa possa construir um *data warehouse* e utilizar, posteriormente, um sistema de *Business Intelligence*, como você poderia auxiliar a empresa na obtenção de dados limpos e que possam ser usados para a tomada de decisões?



## ► Teoria em Prática

- R: Realizar um estudo para utilizar uma ferramenta de ETL, como a *SQL Server Integration Services*, para poder extrair, transformar e carregar dados para o *data warehouse*.

## Dica do Professor

### Bloco 4

Thiago Salhab Alves





## ► Dica do Professor

Indicações de leitura:

- Material introdutório **SQL Server Integration Services** (ZHANG, 2019a).
- Material **Instalar o Integration Services** (ZHANG, 2019b).



## ► Referências

COTE, C.; LAH, M.; SARKA, D. **SQL Server 2017 Integration Services Cookbook**: ETL techniques to load and transform data from various sources using SQL Server 2017 Integration Services. Birmingham, 2017.

DEV MEDIA. **Microsoft ETL**: arquitetura SSIS. 2014.  
Disponível em:

<<https://www.devmedia.com.br/microsoft-etl-arquitetura-ssis-sql-server-integration-services/30862>>. Acesso em: 18 mar. 2019.




## ► Referências

KNIGHT, Brian et al. **Professional Microsoft SQL Server 2014 Integration Services**. Indianapolis: John Wiley & Sons, Inc., 2014.

PALL, Amanpartap S.; KHAIRA, Jaiteg S. A comparative Review of Extraction, Transformation and Loading Tools. **Database Systems Journal**, Jalandhar, v. 4, n. 2, p. 42-51, fev./2013.

TURBAN, Efrain et al. **Business Intelligence: Um Enfoque Gerencial para a Inteligência do Negócio**. Porto Alegre: Bookman, 2009.





## ► Referências

ZHANG, J. **SQL Server Integration Services**. Disponível em: <<https://docs.microsoft.com/pt-br/sql/integration-services/sql-server-integration-services?view=sql-server-2017>>. Acesso em: 18 mar. 2019.

\_\_\_\_\_. **Instalar o Integration Services**. Disponível em: <<https://docs.microsoft.com/pt-br/sql/integration-services/install-windows/install-integration-services?view=sql-server-2017>>. Acesso em: 18 mar. 2019.



