



# Integration Services SSIS

## Guia de boas práticas

Liliam Leme  
Sergio Fonseca  
Renata Festa

# AGENDA

Otimizando  
Inserts



Componentes

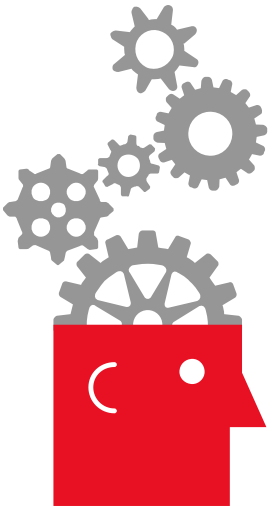


SSIS e o Azure

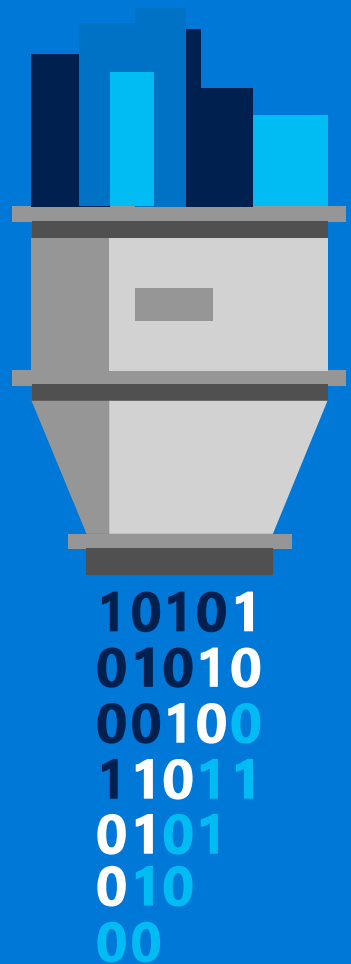


# SSIS

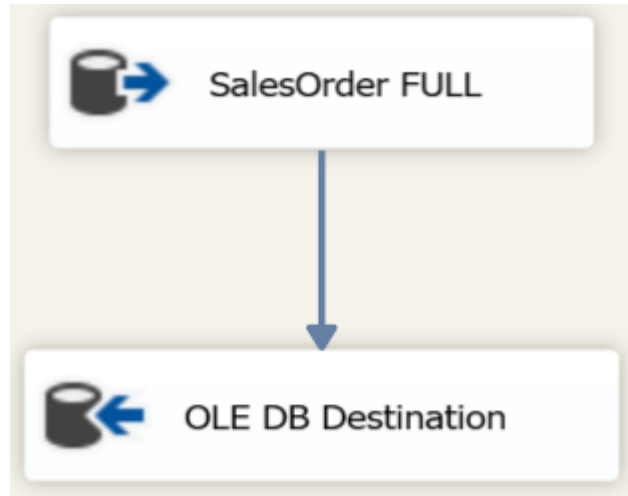
O SQL Server Integration Services (**SSIS**) é uma ferramenta de **ETL**: Extract Transformation Load e pode ser usado para resolver problemas complexos de negócio, permitindo a **transferência de dados, copiar arquivos, enviar email** em resposta a eventos, **tratar a informação** e **administrar objetos** no SQL Server



# Otimizando Inserts



# OLEDB Destination



## Connection Manager

Mappings

Error Output

Specify an OLE DB connection manager, a data source, or a data source view, and select the data access mode. If using the SQL command access mode, specify the SQL command either by typing the query or by using Query Builder. For fast-load data access, set the table update options.

OLE DB connection manager:

localhost.SANDBOX

New...

Data access mode:

Table or view fast load

Name of the table or the view:

[NONBLOCKING]

New...

☐ Keep identity

☒ Table lock

☐ Keep nulls

☒ Check constraints

Rows per batch:

1000000

Maximum insert commit size:

1000000

# Não esqueça...



Use *fast load* em *data access mode* quando possível

Use Tablelock se o cenário for aplicável

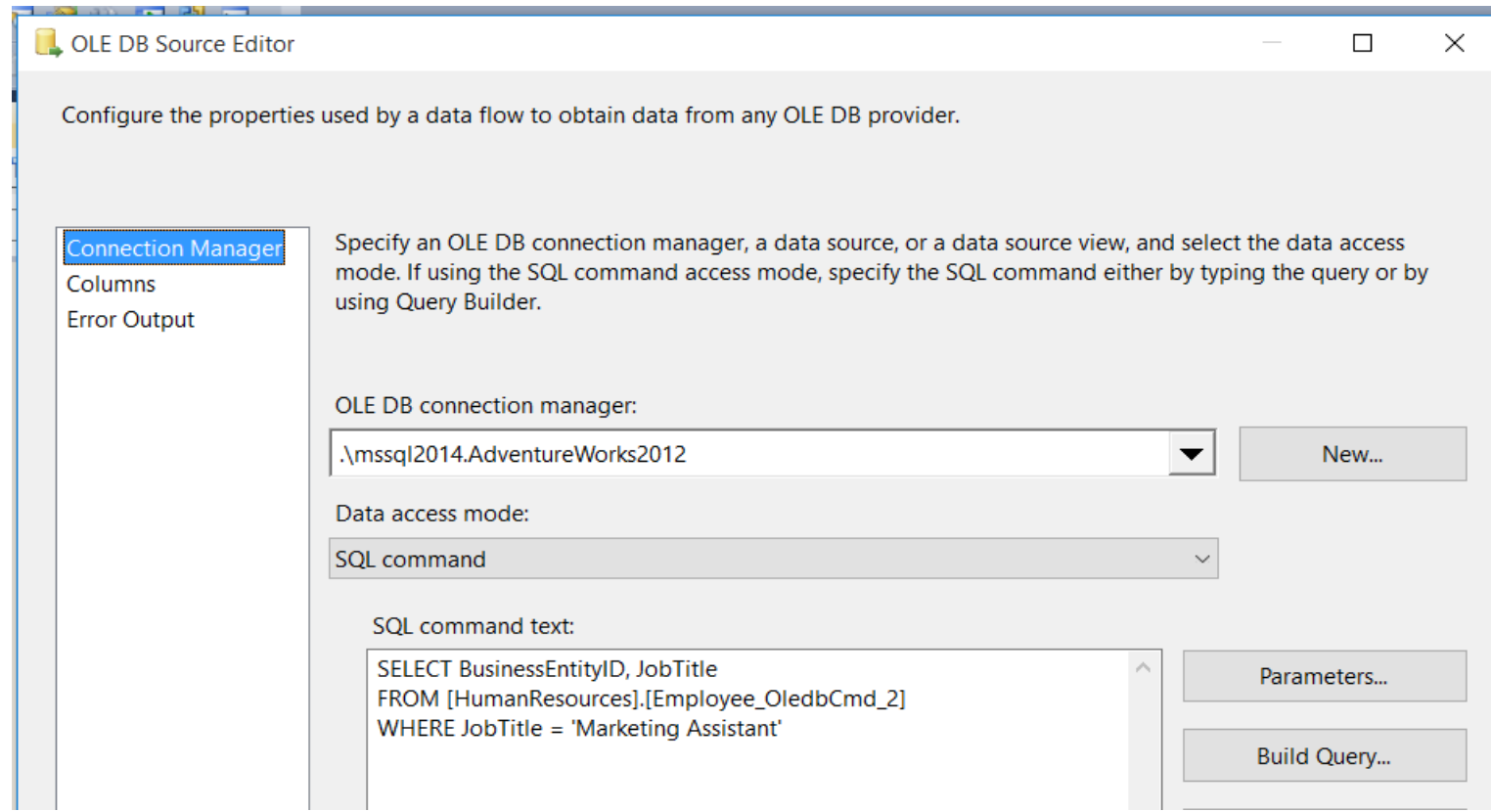
Avalie commit size e rows por batch

# Minimize operações logadas

- Use *Bulk mode* ao invés do *insert* linha a linha.
- Ao invés de *delete*, use *truncate*. O *delete* cria uma entrada para cada linha deletada no log, já o *truncate* remove todos os dados da tabela e apenas informa isso no log.
- Use *Switch partition* para tabelas particionadas

# Otimizar sua transferência de dados

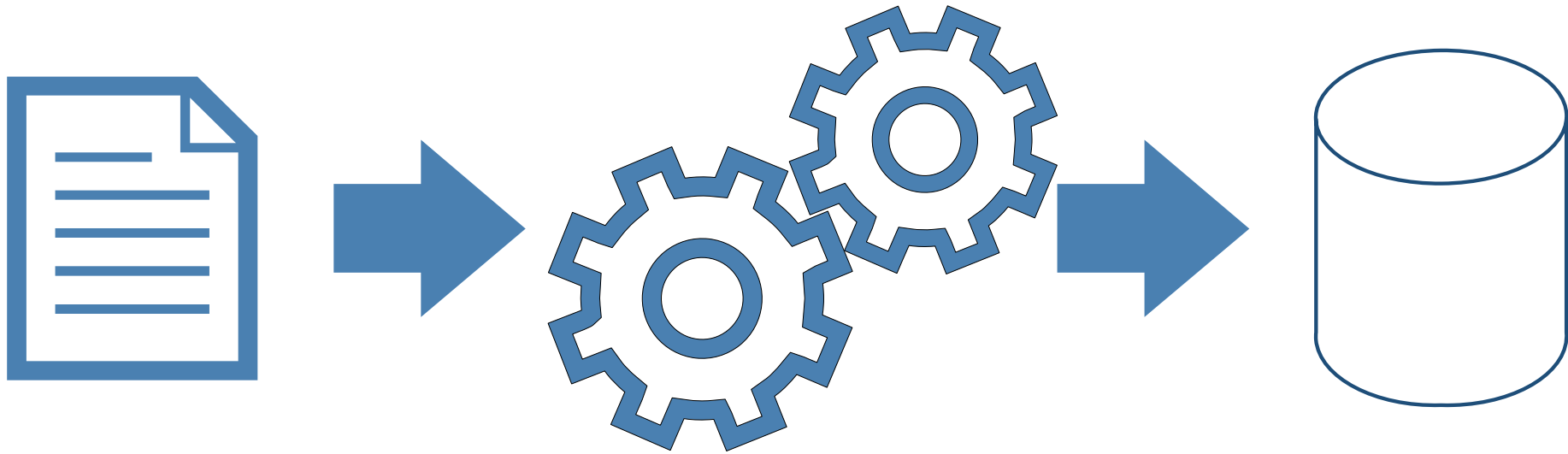
- Selecione apenas as colunas necessárias.





# Carga para Arquivos

- Configure adequadamente os tipos de dados do arquivo.
- Configure o *Fast Parse*



# Carga para Arquivos

- Tipos de dados

The screenshot shows the 'Flat File Connection Manager Editor' window. At the top, there are fields for 'Connection manager name' (Flat File Connection Manager) and 'Description'. Below these is a sidebar with tabs: 'General', 'Columns', 'Advanced', and 'Preview'. The 'Columns' tab is active, showing a list of columns with 'Codigo' selected. The main area is titled 'Configure the properties of each column.' and displays a table of properties for the selected column.

Name	Codigo
ColumnDelimiter	Vertical Bar { }
ColumnType	Delimited
InputColumnWidth	0
DataPrecision	0
DataScale	0
DataType	eight-byte signed integer [DT_I8]
OutputColumnWidth	0
TextQualified	True

Below the table, there is a field labeled 'Name'.

# Carga para Arquivos

- Int e datetime

Advanced Editor for Flat File Source

The advanced editor provides access to the low-level properties of data flow components. Additionally, the advanced editor can be used to configure components that do not have a custom user interface.

Connection Managers | Component Properties | Column Mappings | **Input and Output Properties**

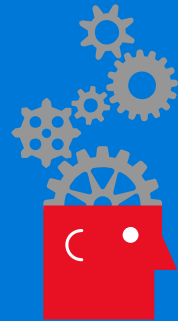
Specify properties for the inputs and outputs of the data flow component.

Inputs and outputs:

- Flat File Source Output
  - External Columns
    - Codigo
    - Descricao
    - Codigo Ident
  - Output Columns
    - Codigo
    - Descricao
    - Codigo Ident
- Flat File Source Error Output

ErrorOrTruncationOperatic	Conversion
ErrorRowDisposition	RD_FailComponent
ExternalMetadataColumnI	9
ID	10
IdentificationString	Flat File Source.Outputs[Flat F
LineageID	10
MappedColumnID	0
Name	Codigo
SortKeyPosition	0
SpecialFlags	0
TruncationRowDisposition	RD_FailComponent
<b>Custom Properties</b>	
FastParse	False
UseBinaryFormat	True
<b>Data Type Properties</b>	
CodePage	0
DataType	eight-byte signed integer [DT
Length	0
Precision	0

# Não esqueça...



Dê preferência a operações minimamente logadas.

Não traga mais colunas ou mais dados que o necessário

Configure adequadamente os tipos de dados do arquivo.

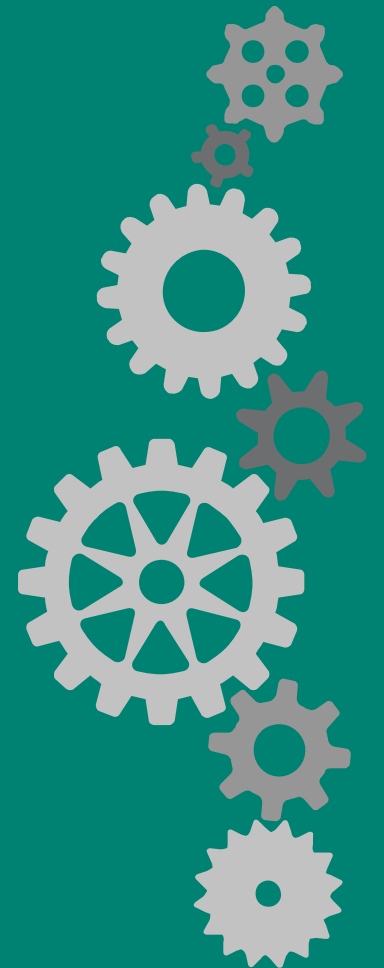
Use fast parse para flat file

# DEMO

Inserts / Fast Parse



# Componentes



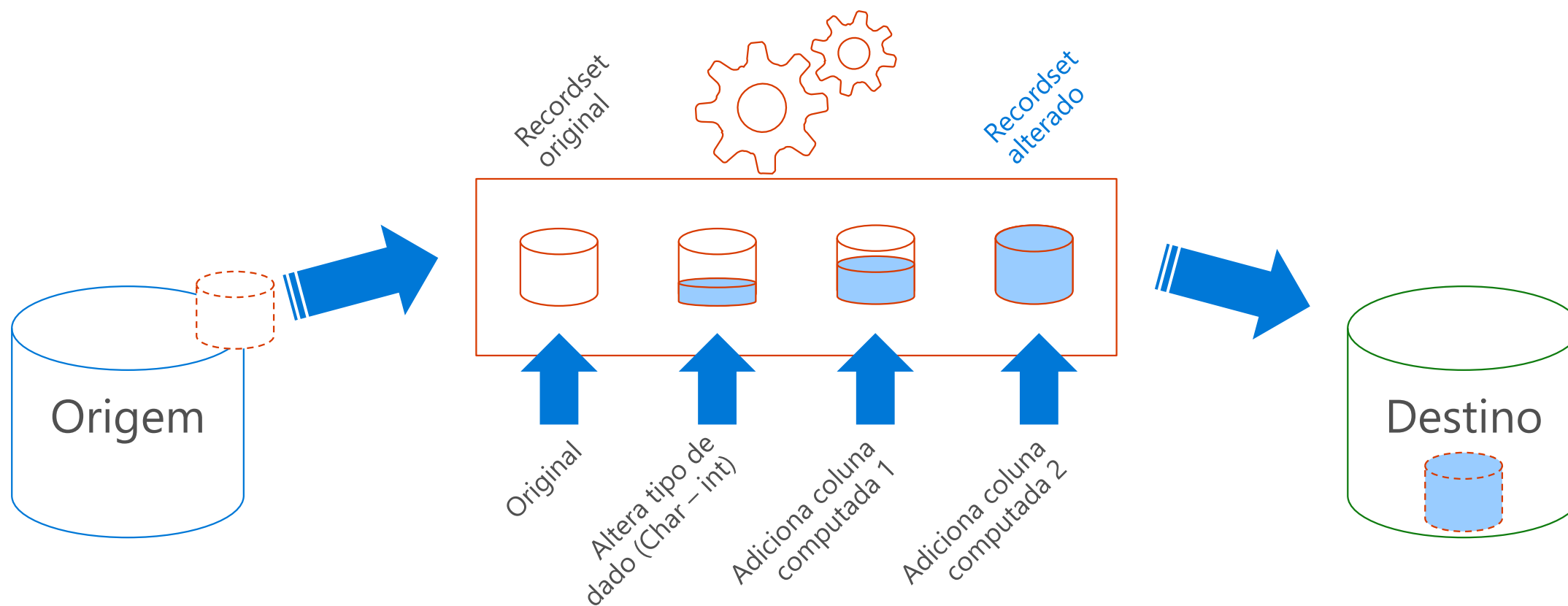
# Full, Partially or Non Blocking operations

Nonblocking  
or row  
transformations

Partially  
blocking  
transformation

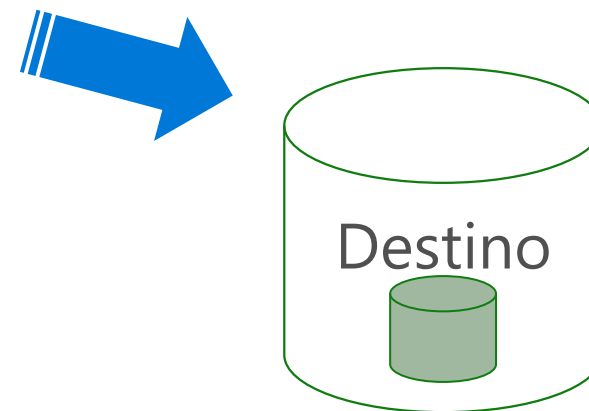
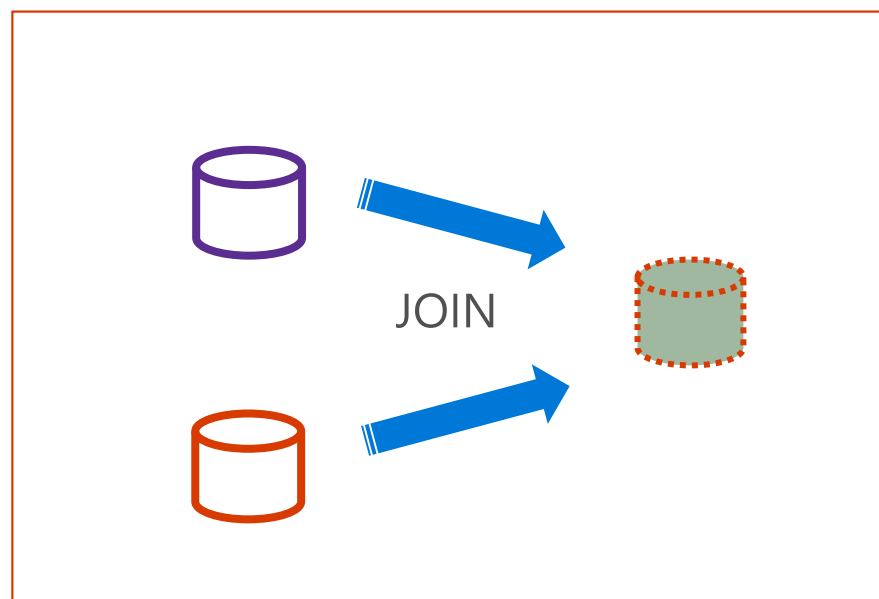
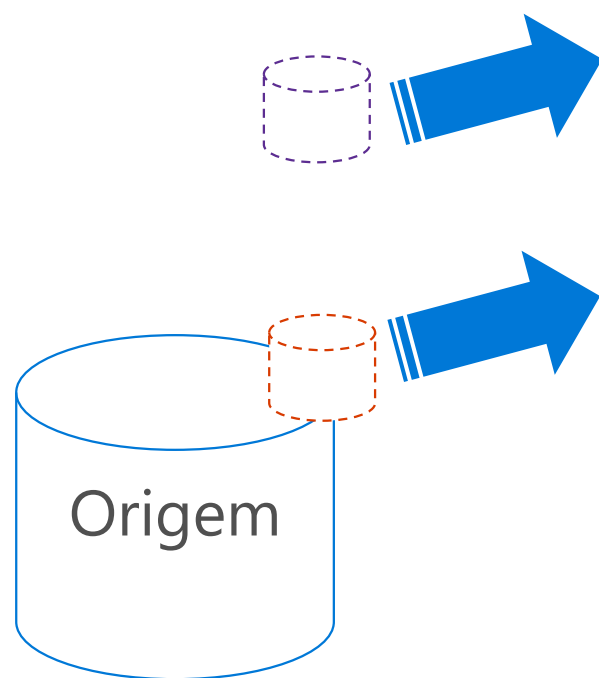
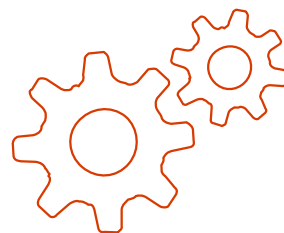
Full  
blocking  
transformations

# Nonblocking ou row transformations

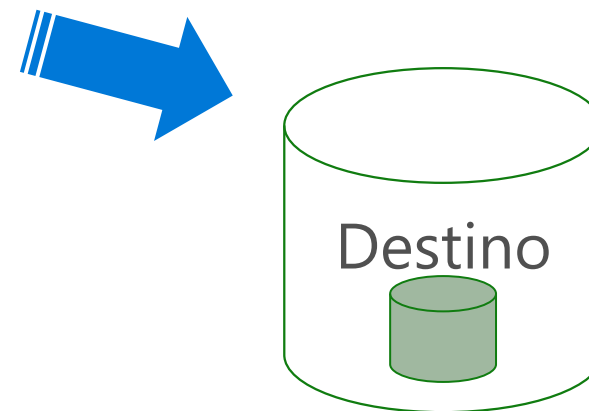
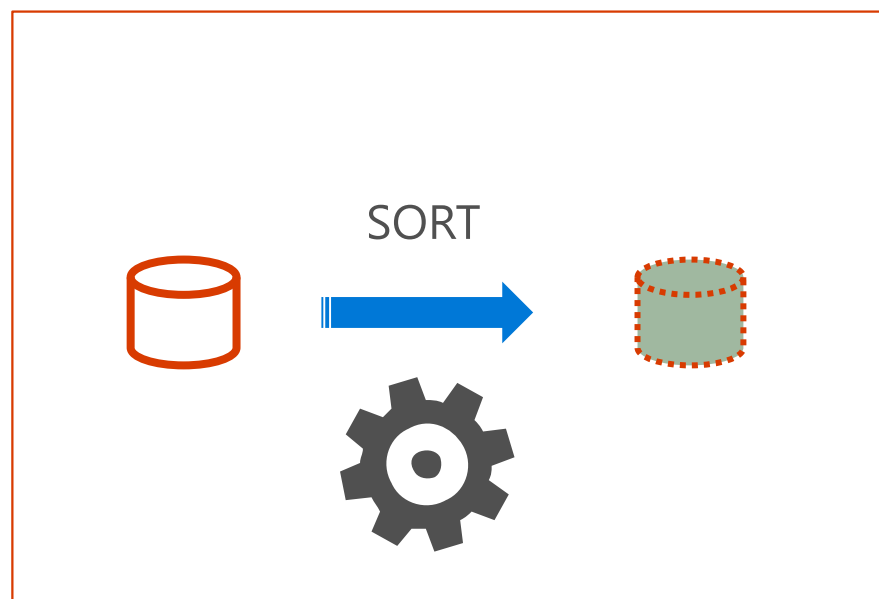
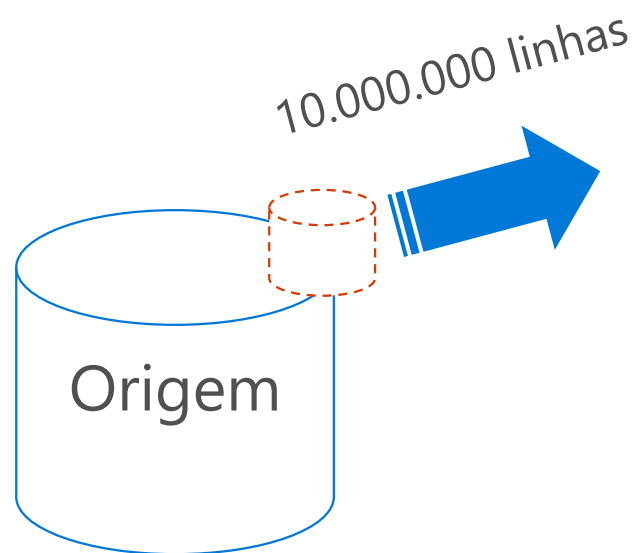
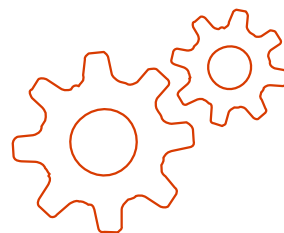




Partially  
blocking  
transformation

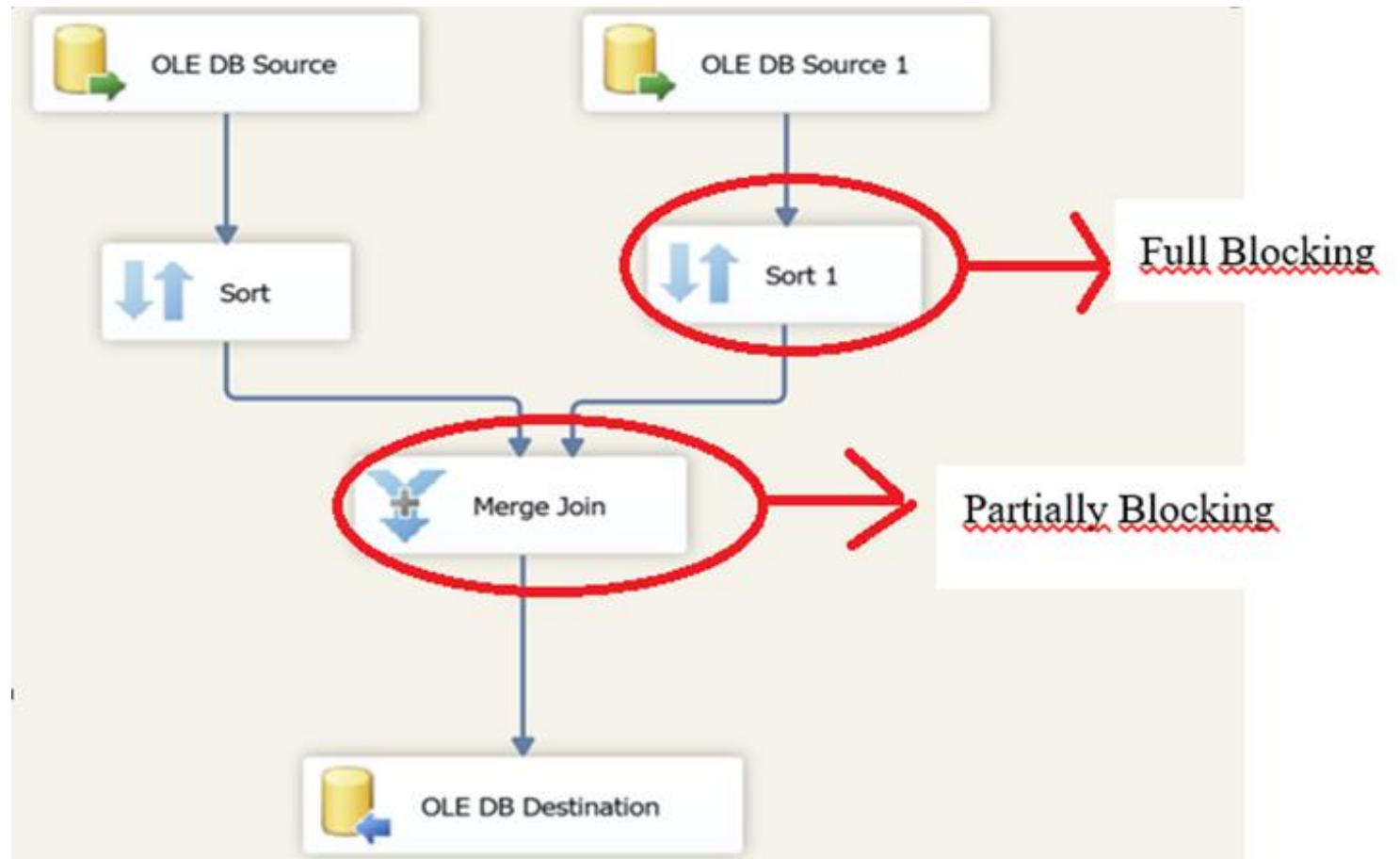


Full  
blocking  
transformations



# Full, Partially ou Non Blocking operations

- ✓ Esses dados que foram ordenados pelo Sort, poderiam ser ordenados na origem?
  - SELECT com ORDER BY
- ✓ É necessário fazer o join durante o processo de carga do SSIS?
  - SELECT com JOIN
  - 2 DB Diferentes
- ✓ Quantas linhas são trafegadas nessa transformação?



# SELECT com ORDER BY

The image shows a screenshot of Microsoft Visual Studio with an SSIS (SQL Server Integration Services) package open. The package is named "Demo - Microsoft Visual Studio". The "Data Flow Task" is selected, and the "Semi blocking Left Join - Without Sort" is chosen. The Data Flow Task is configured with two OLE DB Sources and one OLE DB Destination. The first OLE DB Source is connected to the Merge Join task, and the second OLE DB Source is also connected to the Merge Join task. The Merge Join task is connected to the OLE DB Destination.

The OLE DB Source Editor dialog is open, showing the configuration for the OLE DB Source. The dialog is titled "OLE DB Source Editor" and contains the following fields:

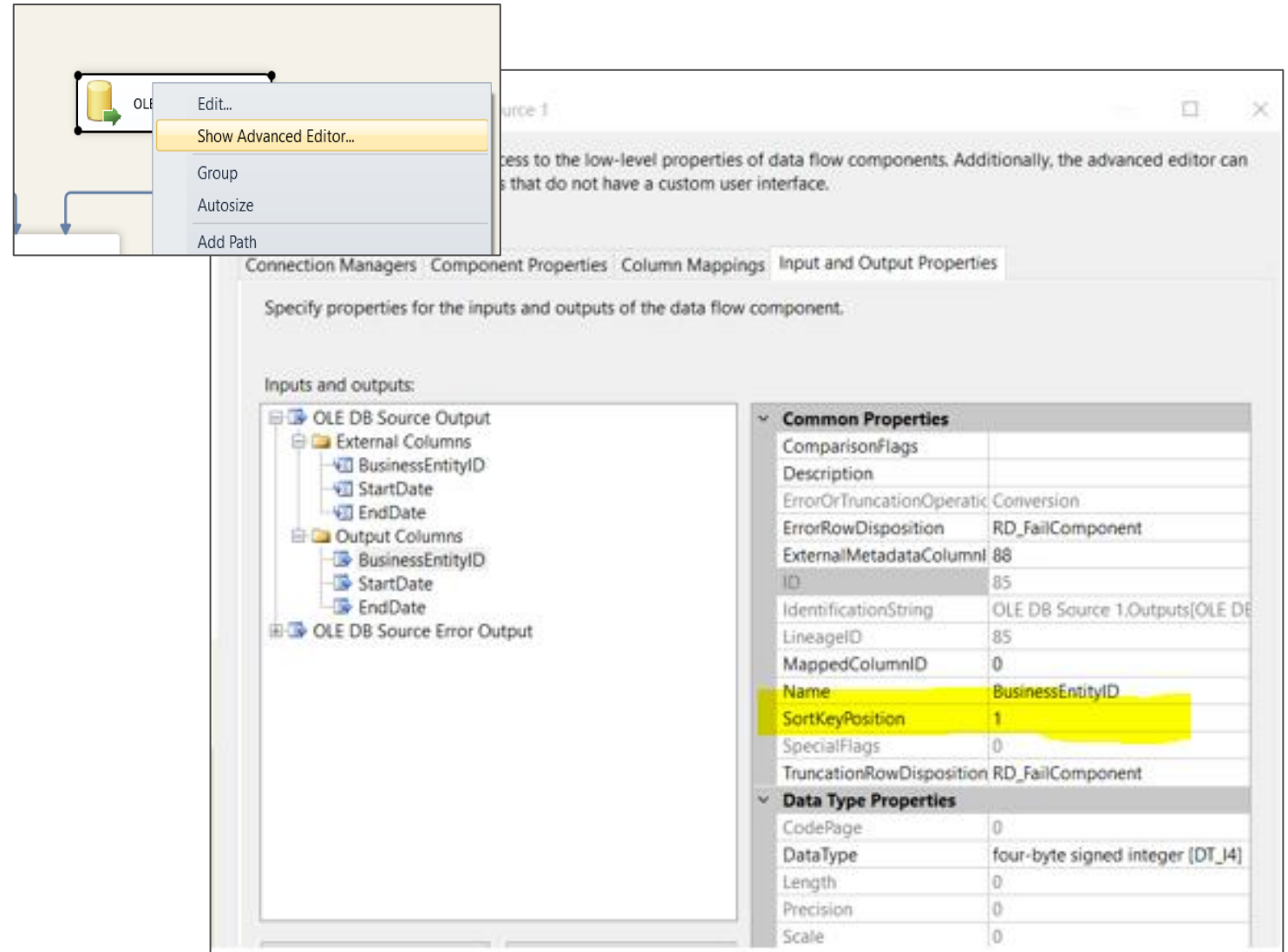
- Connection Manager: .\mssql2014.AdventureWorks2012
- Data access mode: SQL command
- SQL command text:

```
SELECT LoginID, JobTitle, BusinessEntityID
FROM [HumanResources].[Employee]
ORDER BY BusinessEntityID
```

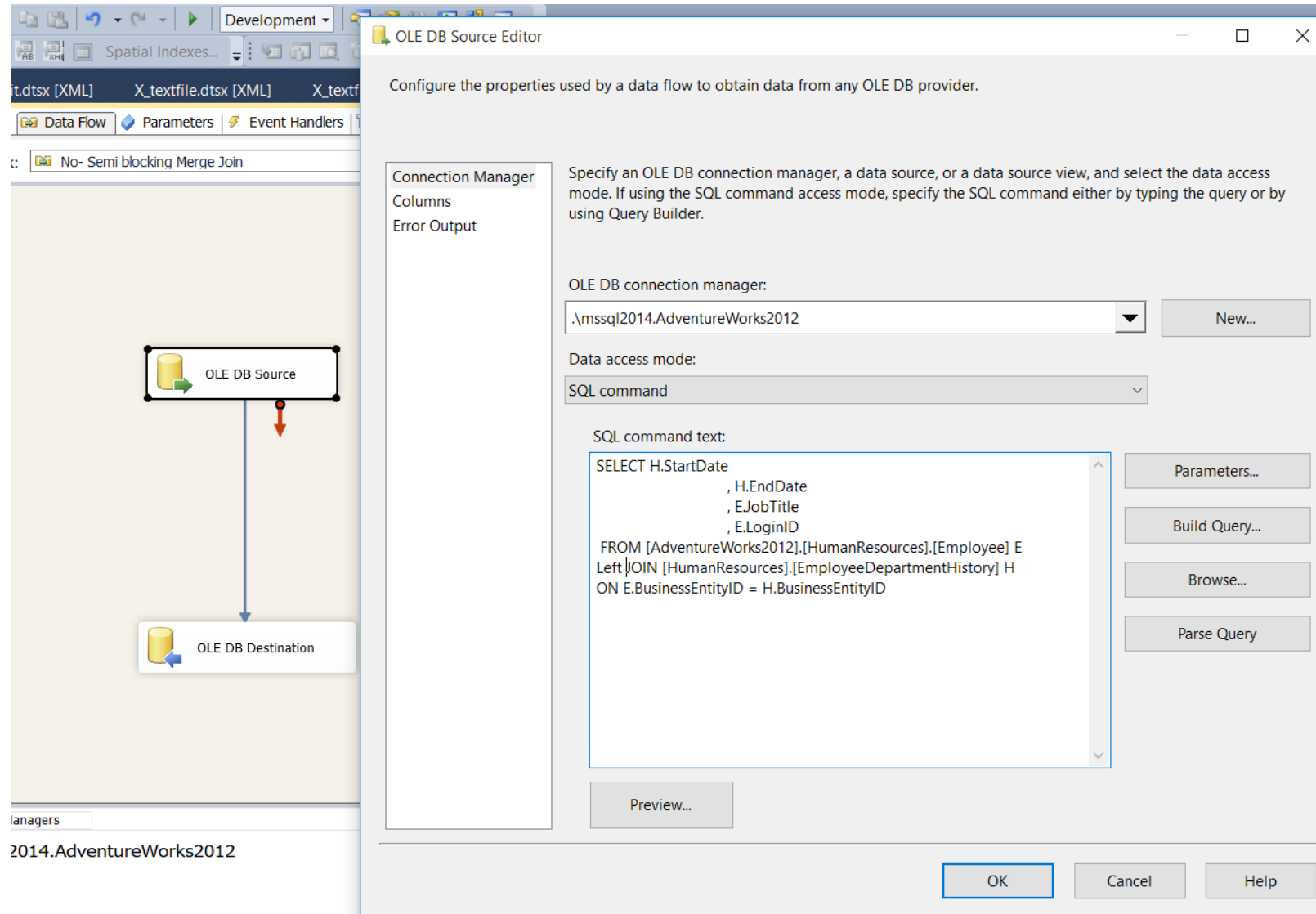
The dialog also includes buttons for "New...", "Parameters...", "Build Query...", "Browse...", "Parse Query", and "Preview...". The "OK" button is highlighted.

# SELECT com ORDER BY

1. Clica-se com o botão direito do mouse em cima do no componente **oledb source**,
2. Seleciona-se a opção **show advanced editor**.
3. Seleciona-se a aba **input and output properties**,
4. Expande-se **Oledb Source output**, e depois expande-se **Output columns**.
5. Selecione a coluna pela qual os dados foram ordenados (exemplo BussineesEntityId), e em common properties (do lado direito da tela) coloque em **sort key position** o valor **1**.



# É necessário fazer o join durante o processo de carga do SSIS?



# Full, Partially ou Non Blocking operations

	Non-blocking	Semi-blocking	Fully-blocking
Synchronous/asynchronous	Síncrono	Assíncrono	Assíncrono
Número de linhas de saída igual a entrada	Sim	Geralmente não	Geralmente não
Precisa ler todos os dados antes de processar	Não	Não	Sim
Novo buffer é criado?	Não	Sim	Sim
Nova thread se faz necessária?	Não	Geralmente Sim	Sim

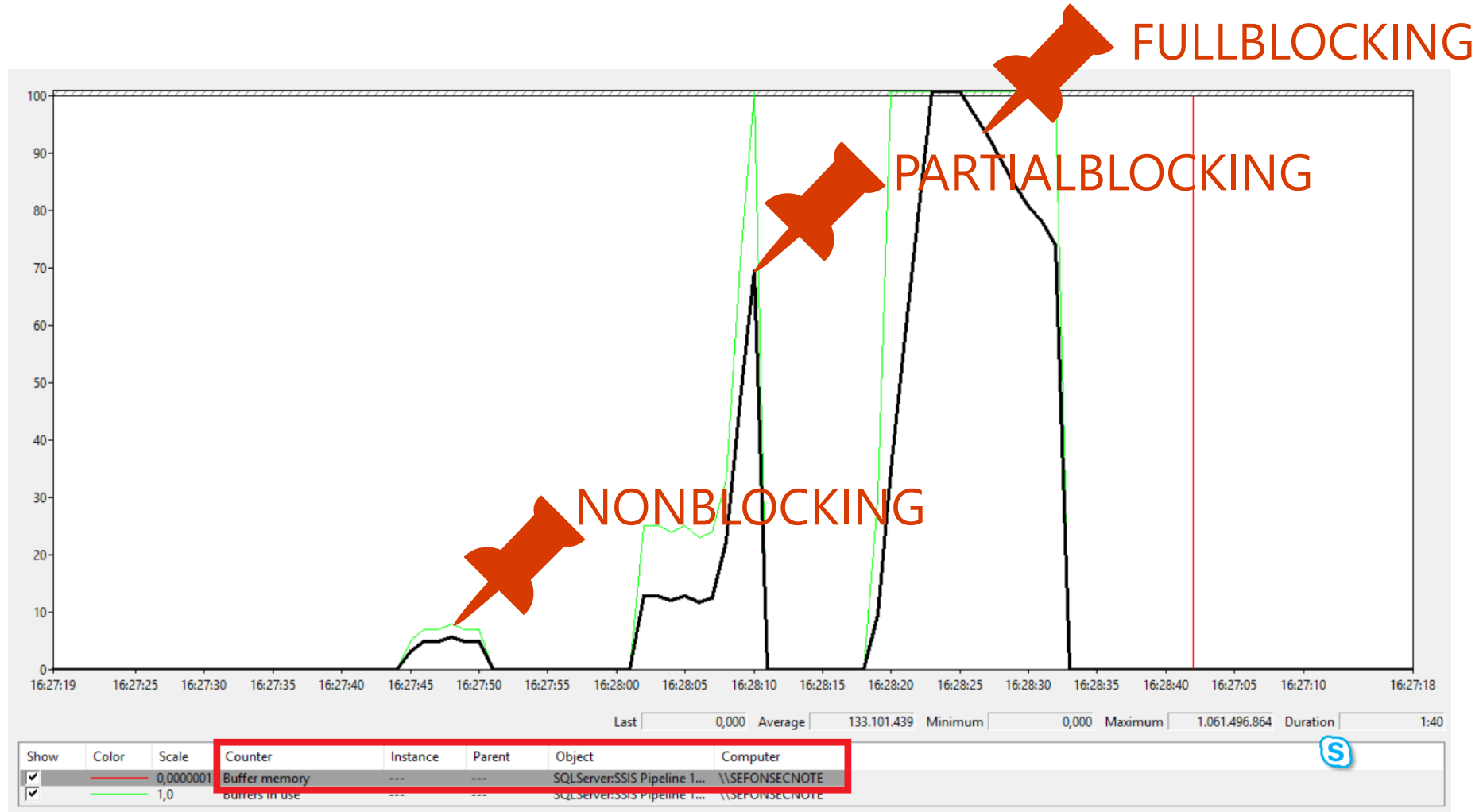
# Full, Partially or Non Blocking operations

Non-Blocking transformations	Semi-blocking transformations	Blocking transformations
Audit	Data Mining Query	Aggregate
Character Map	Merge	Fuzzy Grouping
Conditional Split	Merge Join	Fuzzy Lookup
Copy Column	Pivot	Row Sampling
Data Conversion	Unpivot	Sort
Derived Column	Term Lookup	Term Extraction
Lookup	Union All	
Multicast		
Percent Sampling		
Row Count		
Script Component		
Export Column		
Import Column		
Slowly Changing Dimension		
OLE DB Command		



# Full, Partially ou Non Blocking operations

- Perf Counter: SSIS Pipeline: Buffer Memory

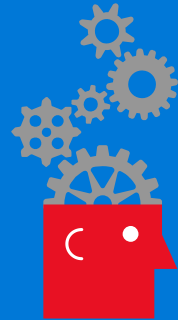


# DEMO

Full, Partially ou Non Blocking operations



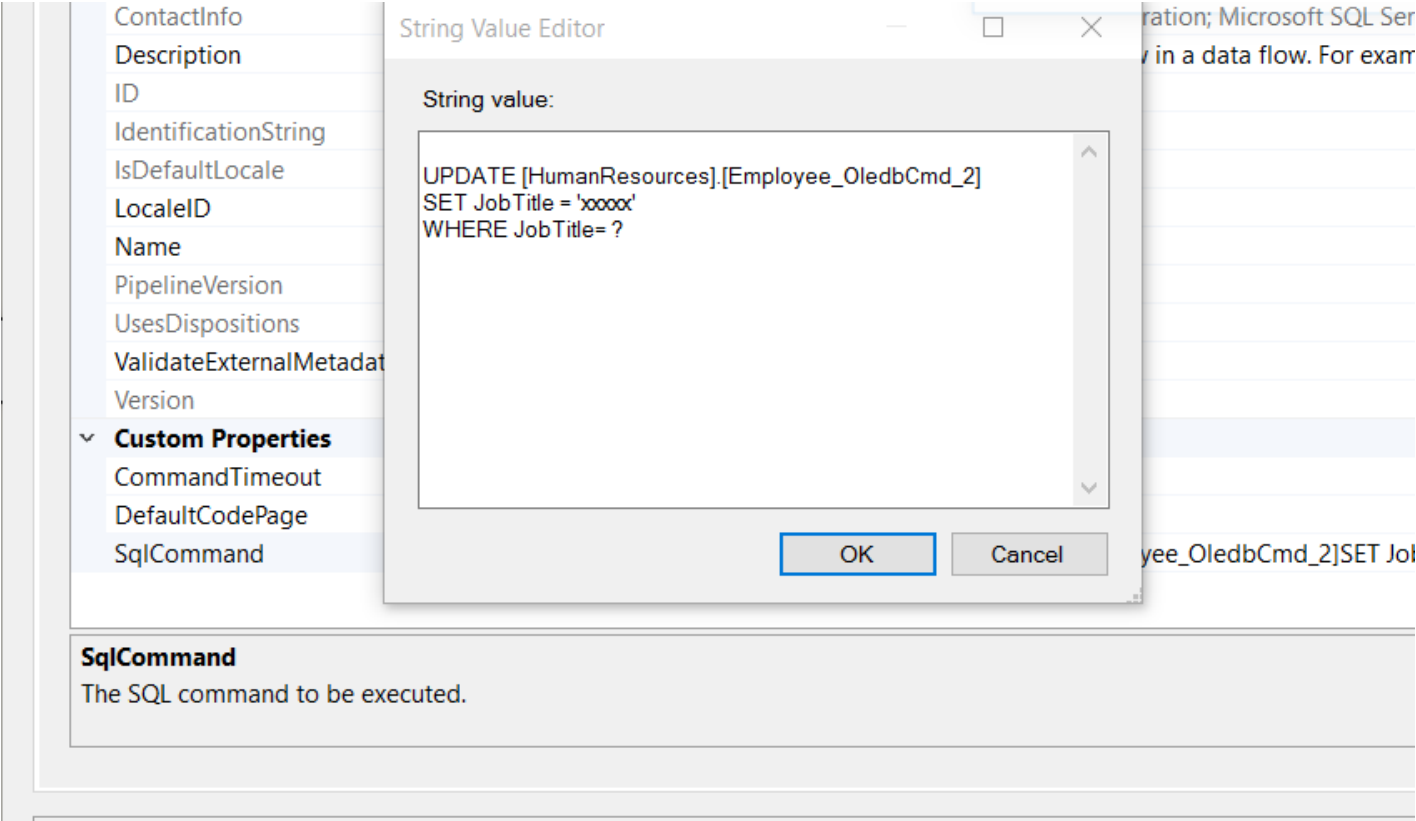
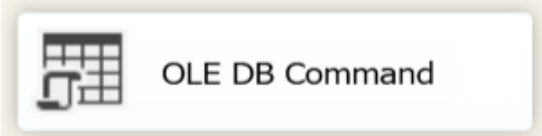
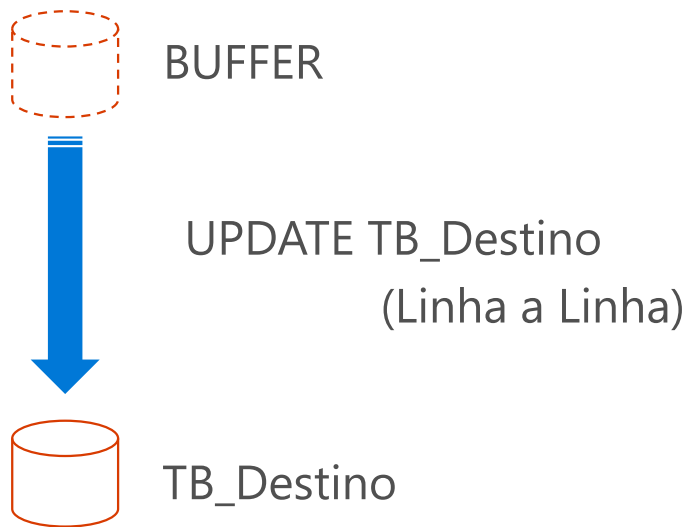
# Não esqueça...



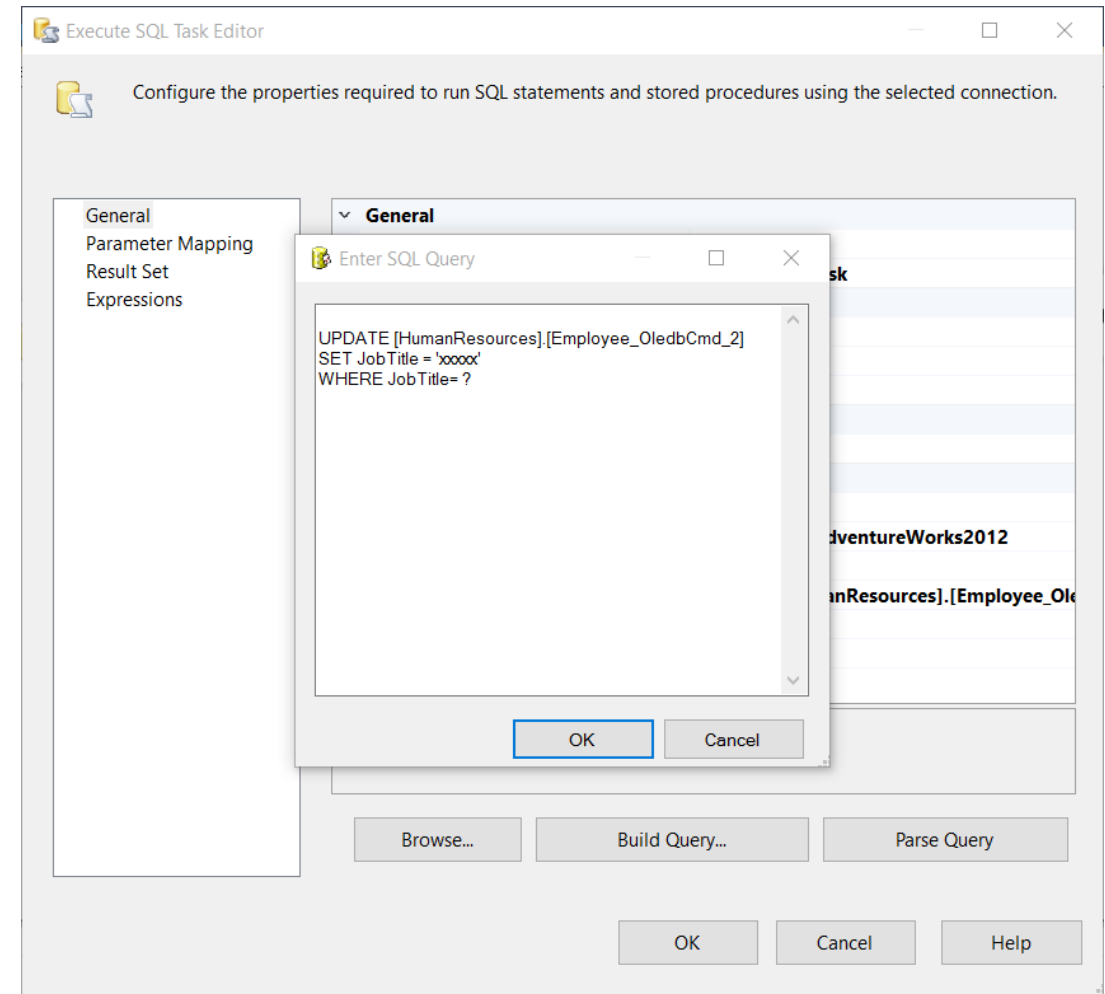
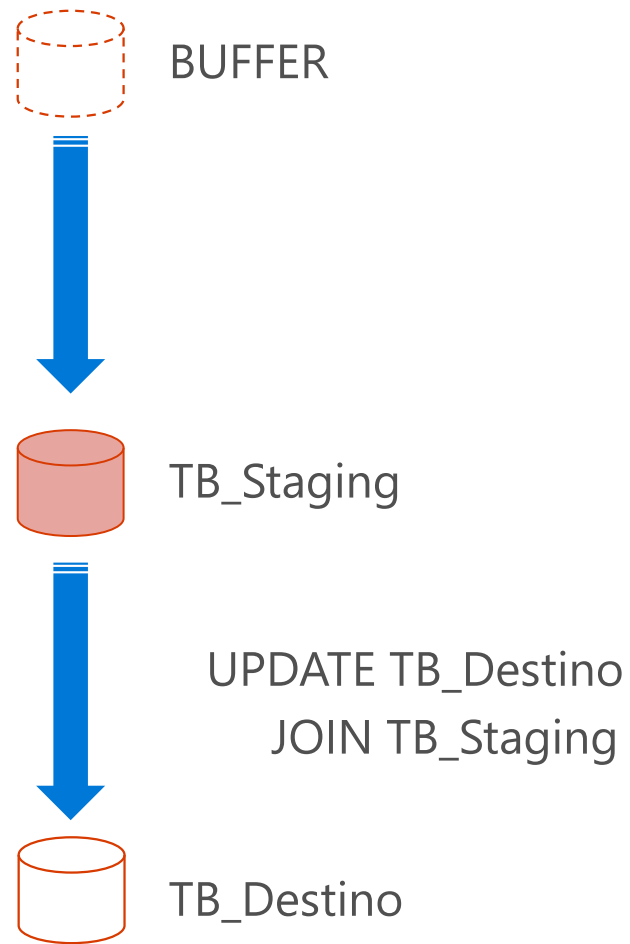
Prefira componentes  
non blocking

Evite, se possível,  
transformações full  
blocking ou partially  
blocking

# OLEDB Command



# OleDb Command -> Execute SQL Task

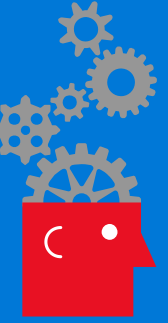


# DEMO

## OleDb Command



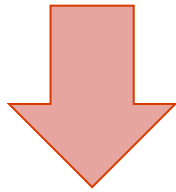
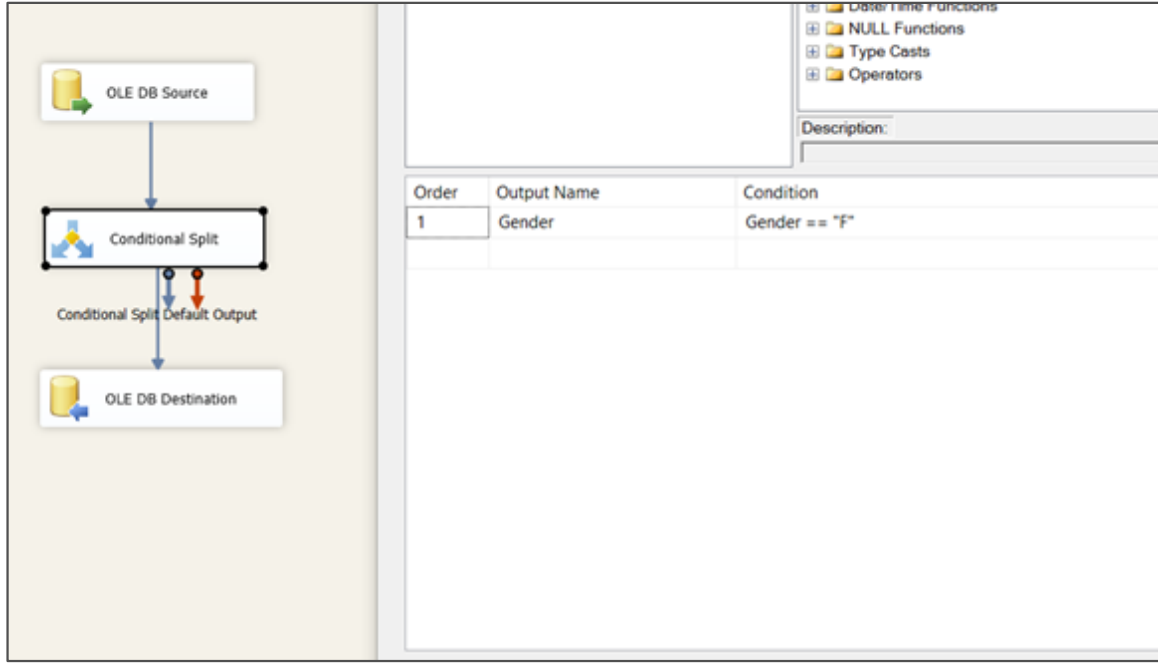
# Não esqueça...



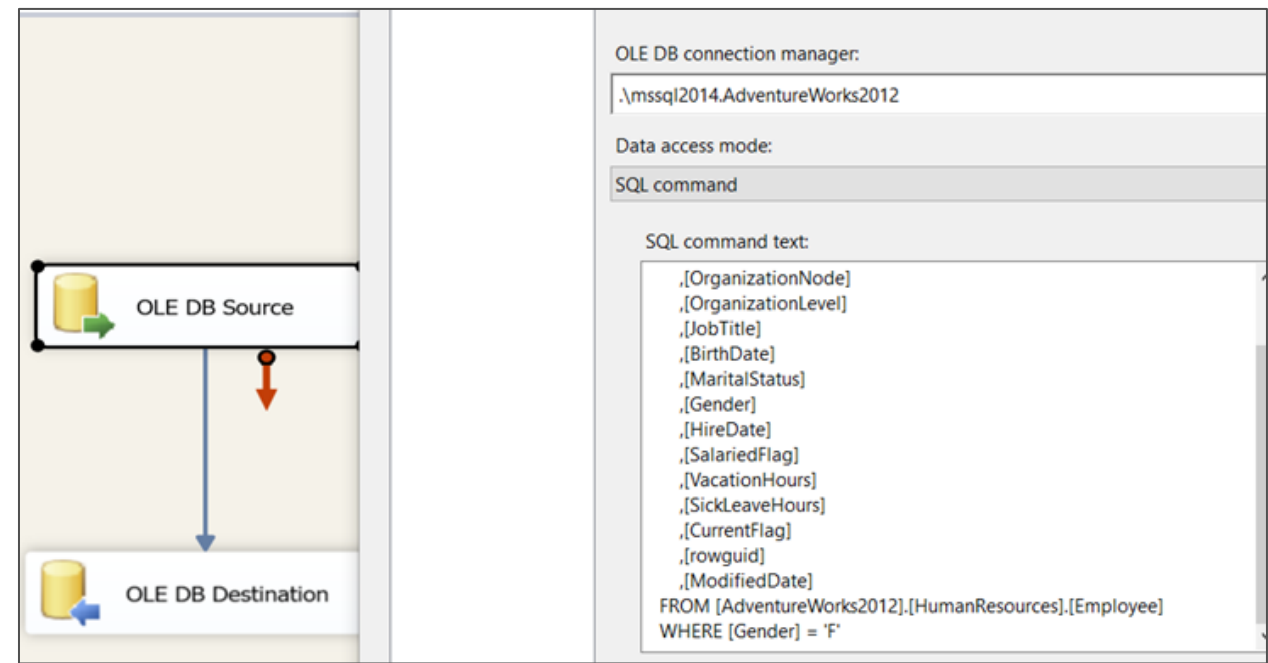
Evite transformações  
**linha a linha**

Procure fazer o  
**processamento dos  
dados em blocos**,  
principalmente se for  
para um grande  
volume de dados

# Conditional Split vs Where



+LENTO  
+DADOS



+RAPIDO  
-DADOS

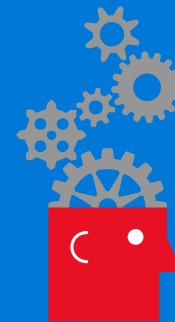


# DEMO

## Conditional Split vs Where



# Não esqueça...



Evite, filtrar os dados  
com conditional Split.

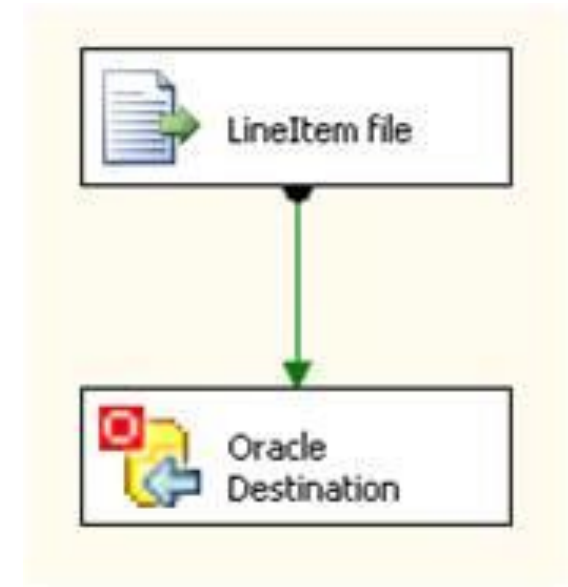
Se for possível filtre  
na origem dos dados

# Oracle para SQL lento

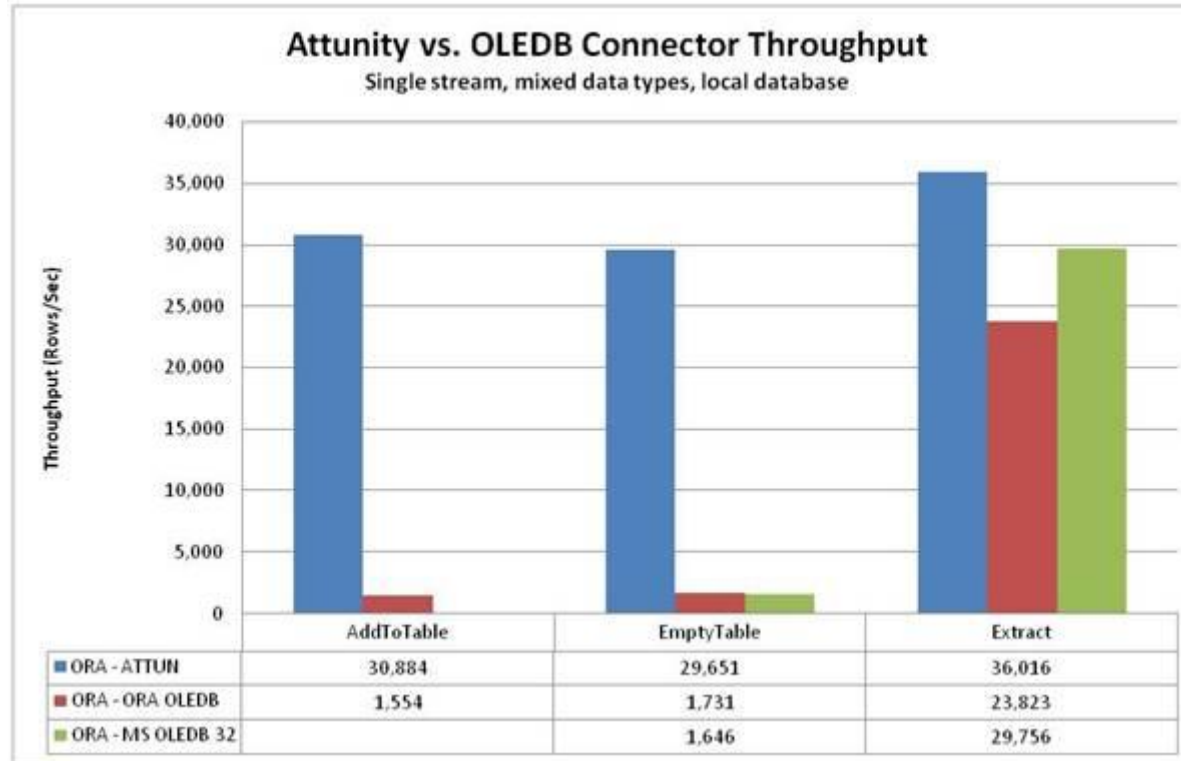
Usar Microsoft Connector for Oracle by [Attunity](#) (API Oracle)

## Versões suportadas:

- **Oracle Attunity 1.0**
  - SSIS 2008 ( Enterprise ou Development version)
  - Oracle Client version 9.2.0.4, 10.x ou 11.x
- **Oracle Attunity 2.0**
  - SSIS 2012 ( Enterprise ou Development version)
  - Oracle Client version 9.2.0.4, 10.x ou 11.x
- **Oracle Attunity 3.0**
  - SSIS 2014 ( Enterprise ou Development version)
  - Oracle Client version 10.x , 11.x ou 12c



# Performance

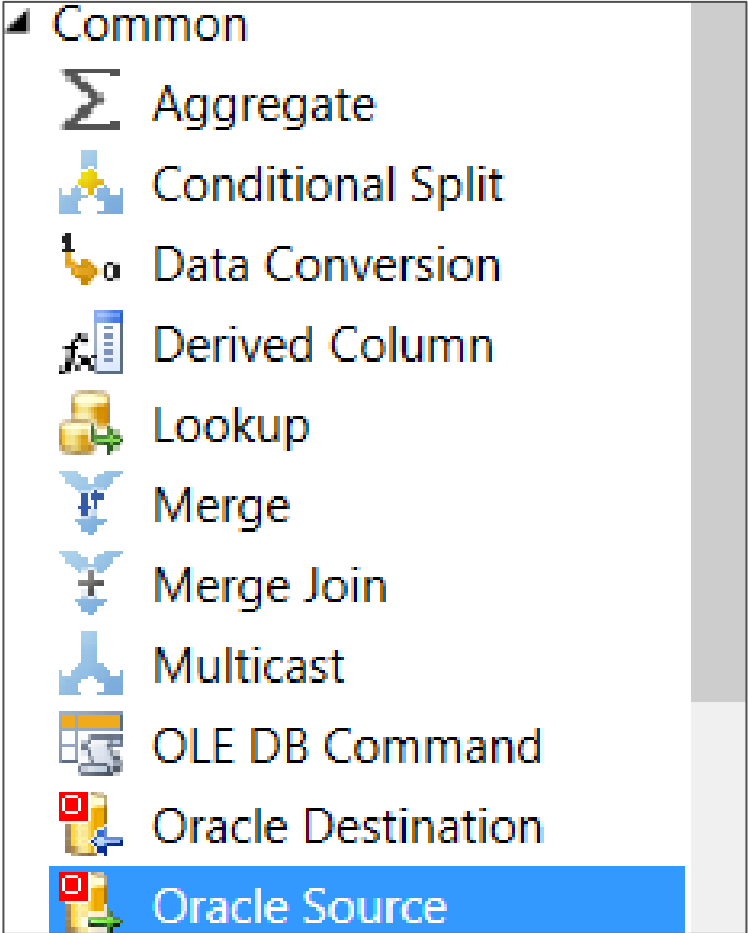
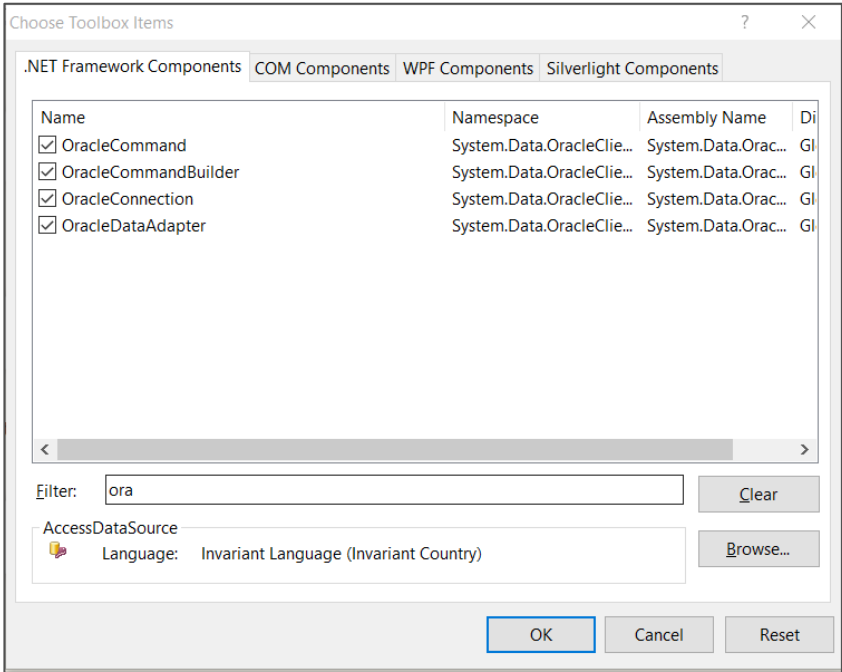
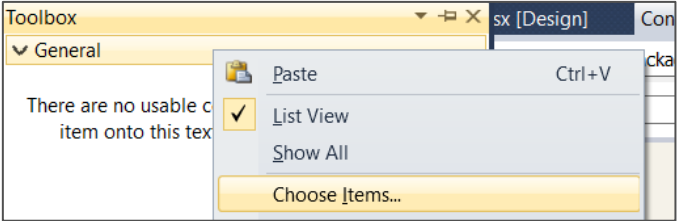
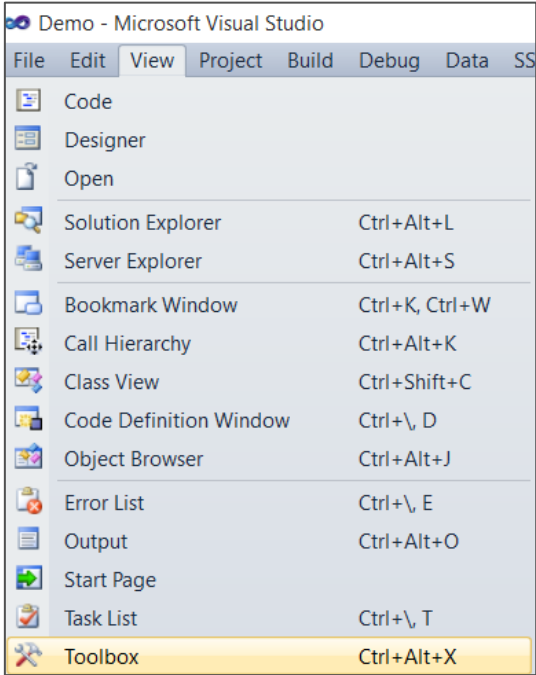


Ref: [https://technet.microsoft.com/en-us/library/ee470675\(v=sql.100\).aspx](https://technet.microsoft.com/en-us/library/ee470675(v=sql.100).aspx)

# Oracle Attunity - Passos da instalação

- Instale o Oracle Attunity 64 bit connector
- Instale o client do Oracle 32 bits
  - É necessário para funcionar no Visual Studio que roda em 32 bits
- Instale o client do Oracle 64 bits
  - O Serviço do SSIS roda em modo 64 bits
- Reinicie o computador

# Oracle Attunity - Configuração

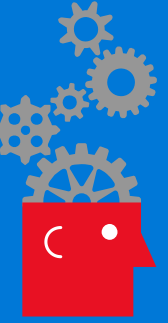


# DEMO

Oracle Attunity



# Não esqueça...



Para extrair ou inserir dados no Oracle, use Microsoft Connector for Oracle by Attunity



# T-SQL vs Coluna Derivada

The screenshot displays the SQL Server Data Tools (SSDT) interface. On the left, a Data Flow Task is shown in Design view, featuring an OLE DB Source connected to a Derived Column task. The Derived Column task is represented by a rectangle with a formula icon. On the right, the Derived Column Transformation Editor is open, showing a list of functions and a table for defining derived columns.

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

- Variables and Parameters
- Columns
- Mathematical Functions
- String Functions
- Date/Time Functions
- NULL Functions
- Type Casts
- Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Le
Derived Column 1	<add as new column>	(DT_I1)CORE_ID_TIPO	single-byte signed inte...	

# T-SQL vs Script Tasks

```
public override void AcquireConnections(object Transaction)
{
    IDTSConnectionManager100 connMgr = this.Connections.con0;
    con = (OracleConnection)connMgr.AcquireConnection(null);
    IDTSConnectionManager100 connMgrSipro = this.Connections.con1;
    Conexao = (SqlConnection)connMgrSipro.AcquireConnection(null);
}

public override void PreExecute()
{
    base.PreExecute();
    try
    {
        obterDados();
    }
    catch (Exception ex)
    {
        preExecuteResult = ex.Message;
    }
}

private void obterDados()
{
    string dataBaseName = Variables.oracleDatabaseName;
    OracleCommand oracleCmd = new OracleCommand(dataBaseName + ".pck_teste_teste", con);
    oracleCmd.CommandType = CommandType.StoredProcedure;
    OracleParameter pi_numfunc = oracleCmd.Parameters.Add("codigo ", OracleType.Number);
    pi_numfunc.Direction = ParameterDirection.Input;
    pi_numfunc.Value = DBNull.Value;
}
```

# DEMO

T-SQL vs (Col Derivada / Script Task)



# Não esqueça...



Use mais T-SQL

Nem tudo fica mais rápido no SSIS.

# Try \ Catch no .Net

```
public void Main()
{
    bool fireAgain = true;
    Dts.Events.FireInformation(0, string.Empty, "Começando", string.Empty, 0, ref fireAgain);

    try
    {
        Dts.Events.FireWarning(0, string.Empty, "Irá gerar um erro, só espera", string.Empty, 0);
        throw new ArgumentNullException();
    }

    catch (Exception e)
    {
        // Raise event so it can be logged by SSIS
        Dts.Events.FireError(0, string.Empty, "Sabia! Um erro ocorreu !!!!!!!!!!!!!", "", 0);
        Dts.TaskResult = (int)ScriptResults.Failure;
    }
}
```

# Release connection no .Net

Não esqueça de liberar a conexão para reuso pelo connection pool, uma vez que esta conexão tenha sido adquirida pelo acquire connection através do Script Task ou Script Componente

```
SqlConnection myADONETConnection = new SqlConnection();  
  
myADONETConnection = (SqlConnection)(Dts.Connections["localhost.AdventureWorks2012_ADONET"].AcquireConnection(Dts.Transaction) as SqlConnection);  
  
Dts.Connections["localhost.AdventureWorks2012_ADONET"].ReleaseConnection(Dts.Transaction);
```

# DEMO

Try \ Catch e Release Connection



# Não esqueça...



- Se estiver usando Script Componente ou Script Task:

Libere as conexões através do comando `release connection` ao adquiri-las no pacote por meio do `acquire connection`.

Use Try catch para tratar erros



# SSIS LOG



Name	Description	Configuration
<input type="checkbox"/> SSIS log provider for Text files	Writes log entries for events to a CSV file	
<input checked="" type="checkbox"/> SSIS log provider for Windows Event Log	Writes log entries for events to the Event Log	
<input type="checkbox"/> SSIS log provider for SQL Server Profiler	Generates SQL traces that may be captured in SQL Server Profiler	
<input type="checkbox"/> SSIS log provider for SQL Server	Writes log entries for events to a SQL Server database	
<input type="checkbox"/> SSIS log provider for XML files	Writes log entries for events to an XML file	

Logging...

Digital Signing...

Variables

Connections

Work Offline

Log Events

☒ Debug Progress Reporting

New Connection...

SSIS Toolbox

Getting Started

Add Annotation

Edit Breakpoints...

Zoom

Cut

Ctrl+X

Copy

Ctrl+C

Paste

Ctrl+V

Delete

Del

Select All

Ctrl+A

Parameterize...

Properties

Alt+Enter

Configure SSIS Logs: Package

Create and configure a new log to capture log-enabled events that occur at run time.

Containers:

☒ Package

☒ Data Flow Task

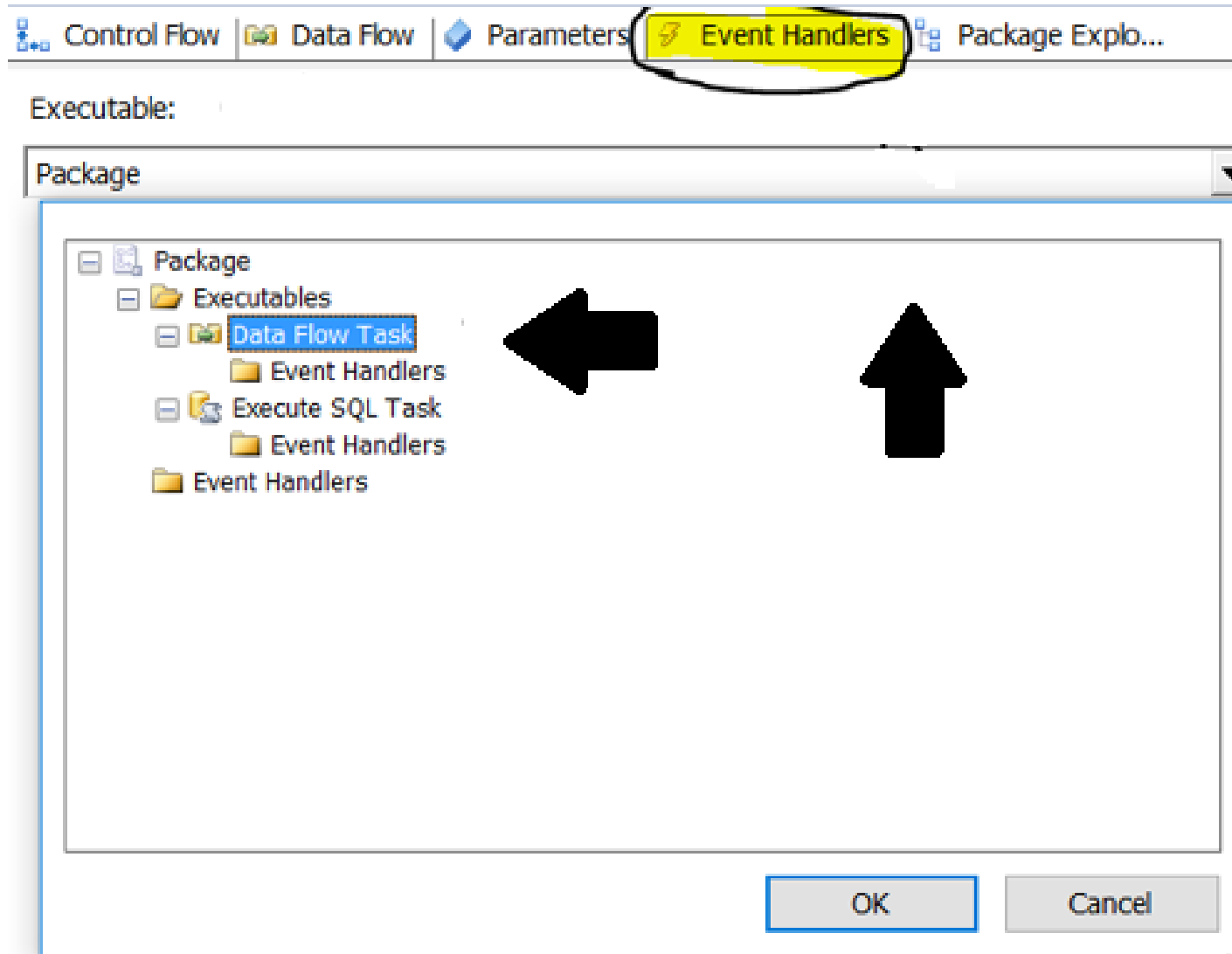
Providers and Logs

Details

Select the events to be logged for the container:

Events	Description
<input checked="" type="checkbox"/> OnError	Handles error events. Use to define ac...
<input type="checkbox"/> OnExecStatusChanged	Handles changes of execution status. ...
<input type="checkbox"/> OnInformation	Handles information events. The mea...
<input type="checkbox"/> OnPipelinePostComp...	Data flow engine has finished a call t...
<input type="checkbox"/> OnPipelinePostEndOf...	A component has finished processing...
<input type="checkbox"/> OnPipelinePostPrime...	A component has returned from its Pr...
<input type="checkbox"/> OnPipelinePreComp...	Data flow engine will call a compone...
<input type="checkbox"/> OnPipelinePreEndOf...	A component will be given the end of...
<input type="checkbox"/> OnPipelinePrePrime...	PrimeOutput will be called on a com...
<input type="checkbox"/> OnPipelineRowsSent	Rows were provided to a data flow c...
<input type="checkbox"/> OnPostExecute	Handles post-execution events. Use t...
<input type="checkbox"/> OnPostValidate	Handles post-validation events. Use t...
<input type="checkbox"/> OnPreExecute	Handles pre-execution events. Use to ...
<input type="checkbox"/> OnPreValidate	Handles pre-validation events. Use to...
<input type="checkbox"/> OnProgress	Handles progress notifications. Use to...
<input type="checkbox"/> OnQueryCancel	Handles cancel events. Called periodi...
<input type="checkbox"/> OnTaskFailed	Handles task failures. Use to define ac...
<input type="checkbox"/> OnVariableValueCha...	Handles value changes in variables w...
<input checked="" type="checkbox"/> OnWarning	Handles warning events. Use to defin...
<input type="checkbox"/> Diagnostic	Logs package diagnostics informatio...
<input type="checkbox"/> DiagnosticEx	Logs more information on package e...

# Event Handlers

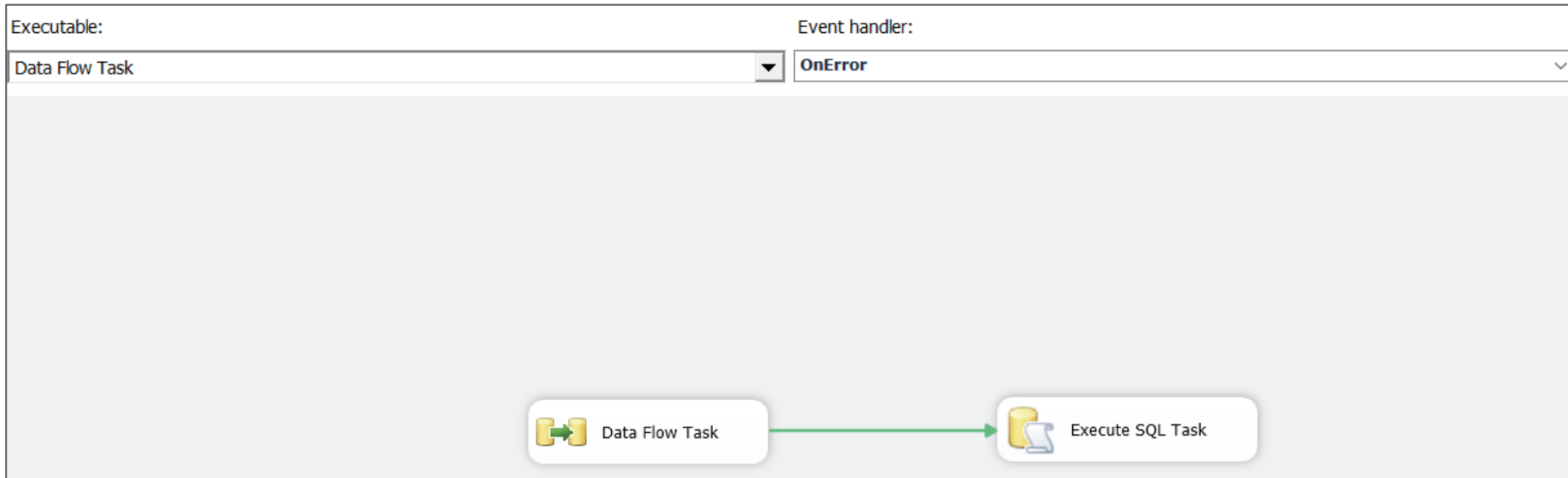


# Event Handlers

Executable: Data Flow Task

Event handler:

- OnError
- OnError
- OnExecStatusChanged
- OnInformation
- OnPostExecute
- OnPostValidate
- OnPreExecute
- OnPreValidate
- OnProgress
- OnQueryCancel
- OnTaskFailed
- OnVariableValueChanged
- OnWarning



# DEMO

## SSIS Log e Event Handlers



# Não esqueça...



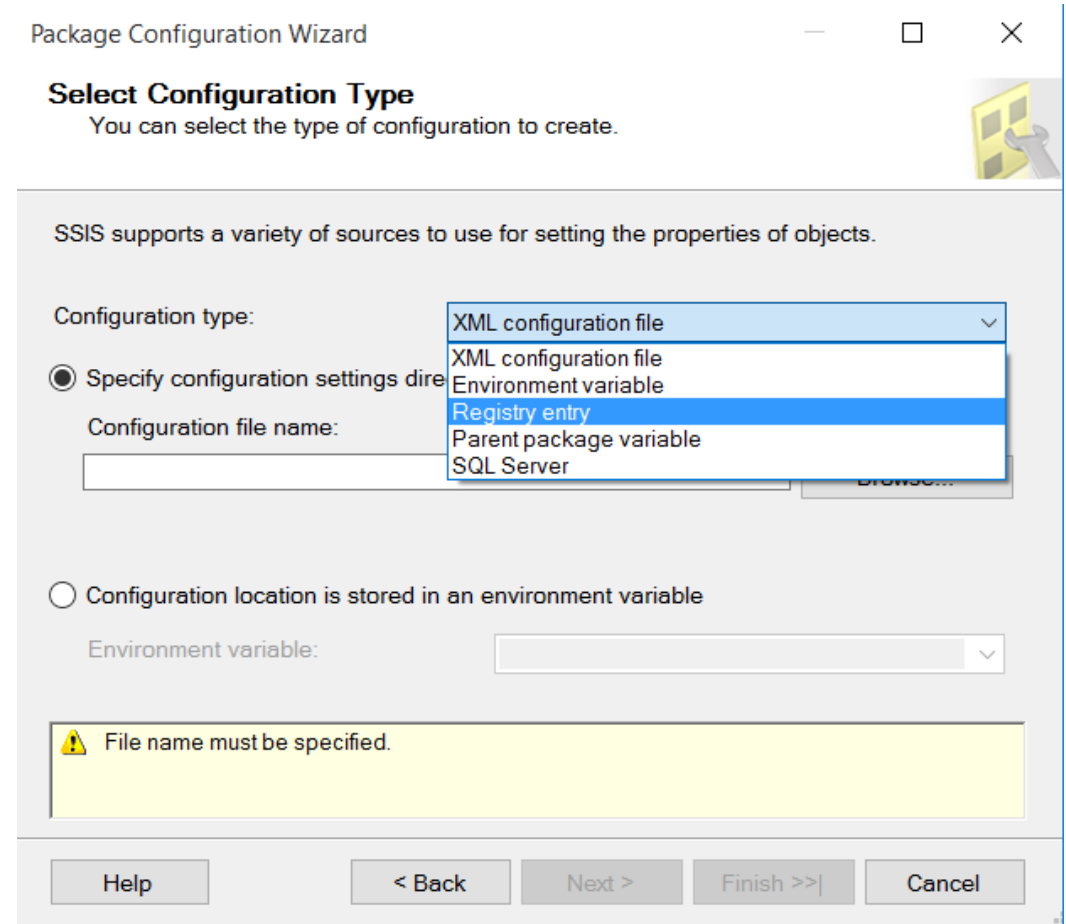
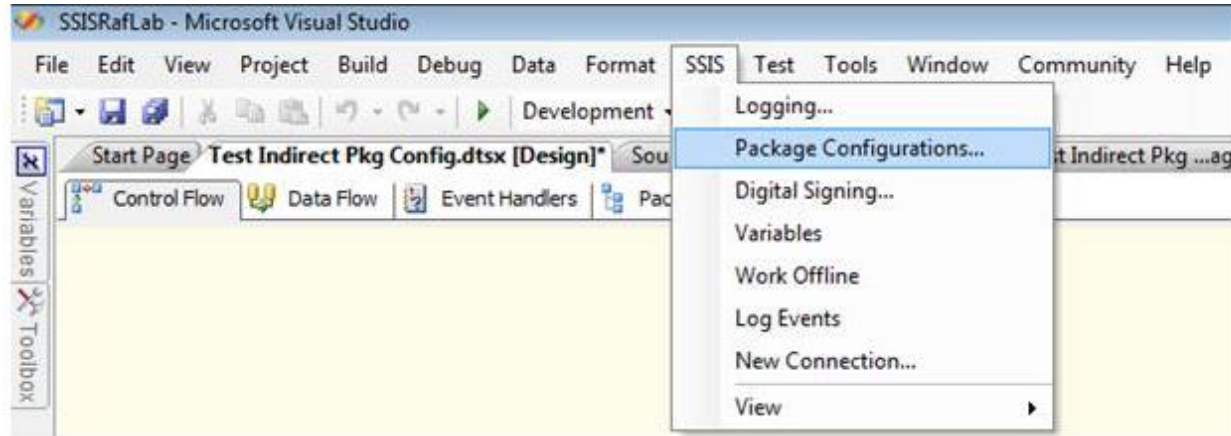
- Configure o log de maneira eficiente e eficaz.

SSIS Logging

Event Handlers

Logs Customizados

# Configuração do pacote SSIS (2008 ou >)



# Configuração do pacote SSIS (2008 ou >)



# Configuração do pacote SSIS (2008 ou >)

- XML File Configuration

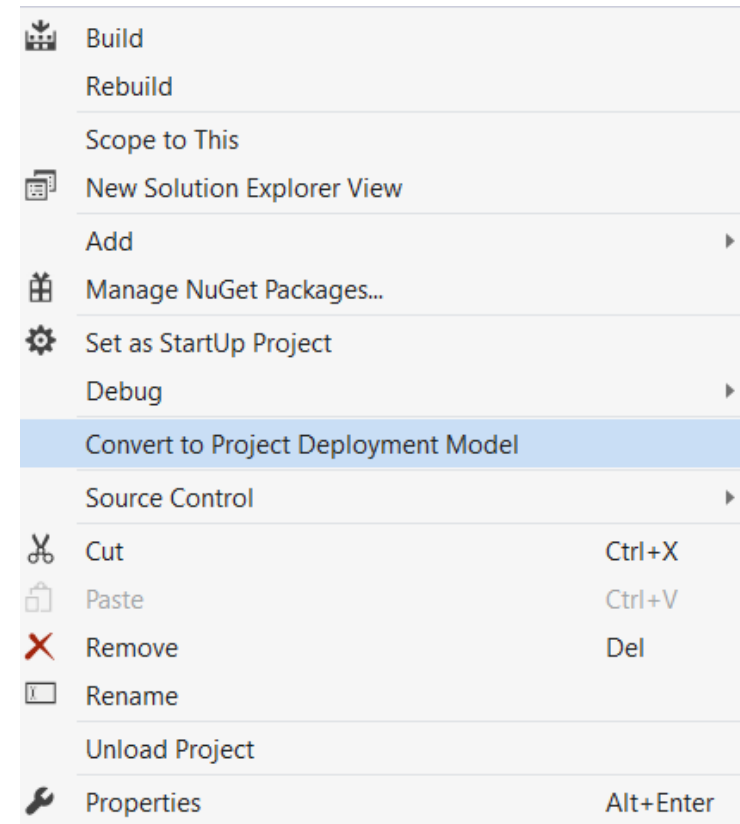
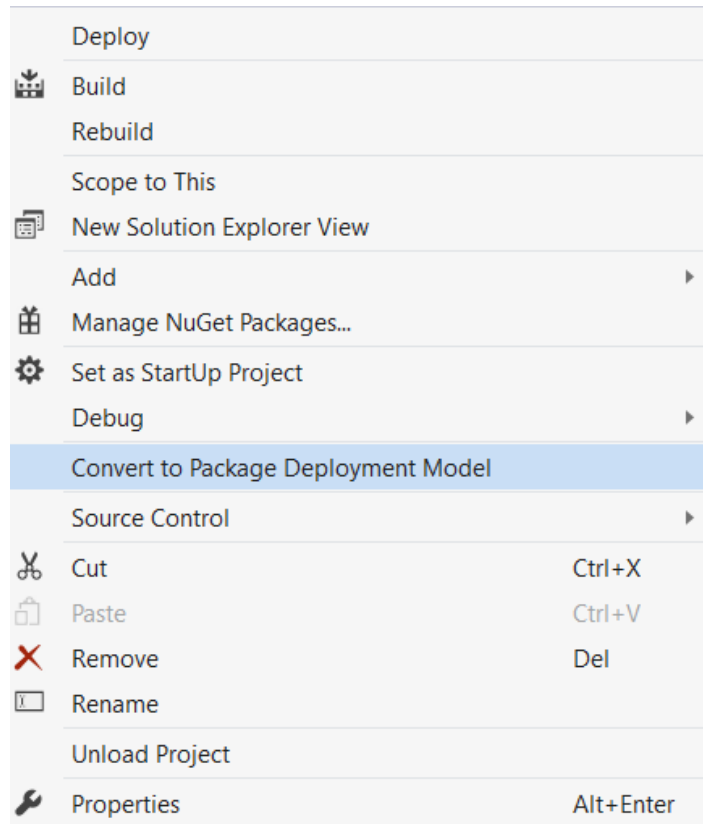
```
<?xml version="1.0" ?>
- <DTSTConfiguration>
- <DTSTConfigurationHeading>
  <DTSTConfigurationFileInfo GeneratedBy="MARINER\rsalas" GeneratedFromPackageName="Test XML Config - MSDN Article"
  </DTSTConfigurationHeading>
- <Configuration ConfiguredType="Property" Path="\Package.Variables[User::ConfigTarget].Properties[Value]" ValueType
  <ConfiguredValue>Run time Value from config file - direct Method</ConfiguredValue>
  </Configuration>
-
  </DTSTConfiguration>
```

- Environment Variables
- Registry Entry
- Parent Package



# Configuração do pacote SSIS (2012 ou >)

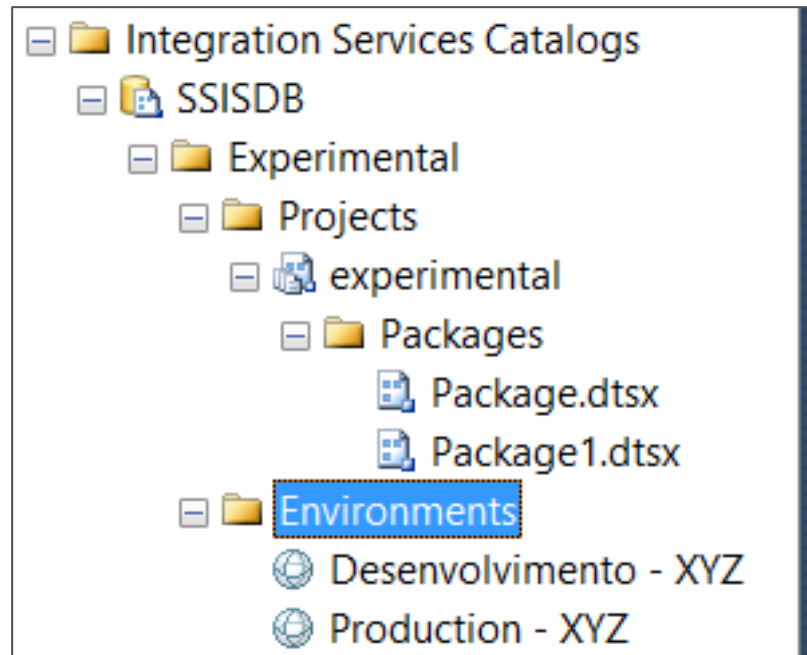
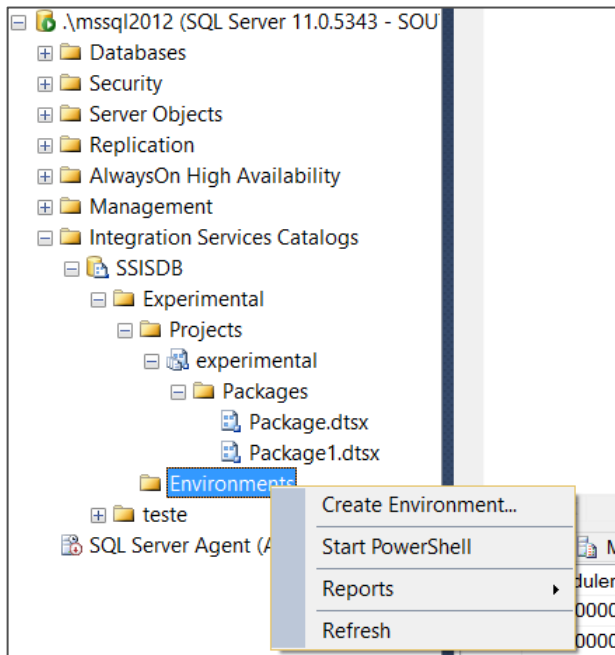
- Deployment Model



# Configuração do pacote SSIS (2012 ou >)

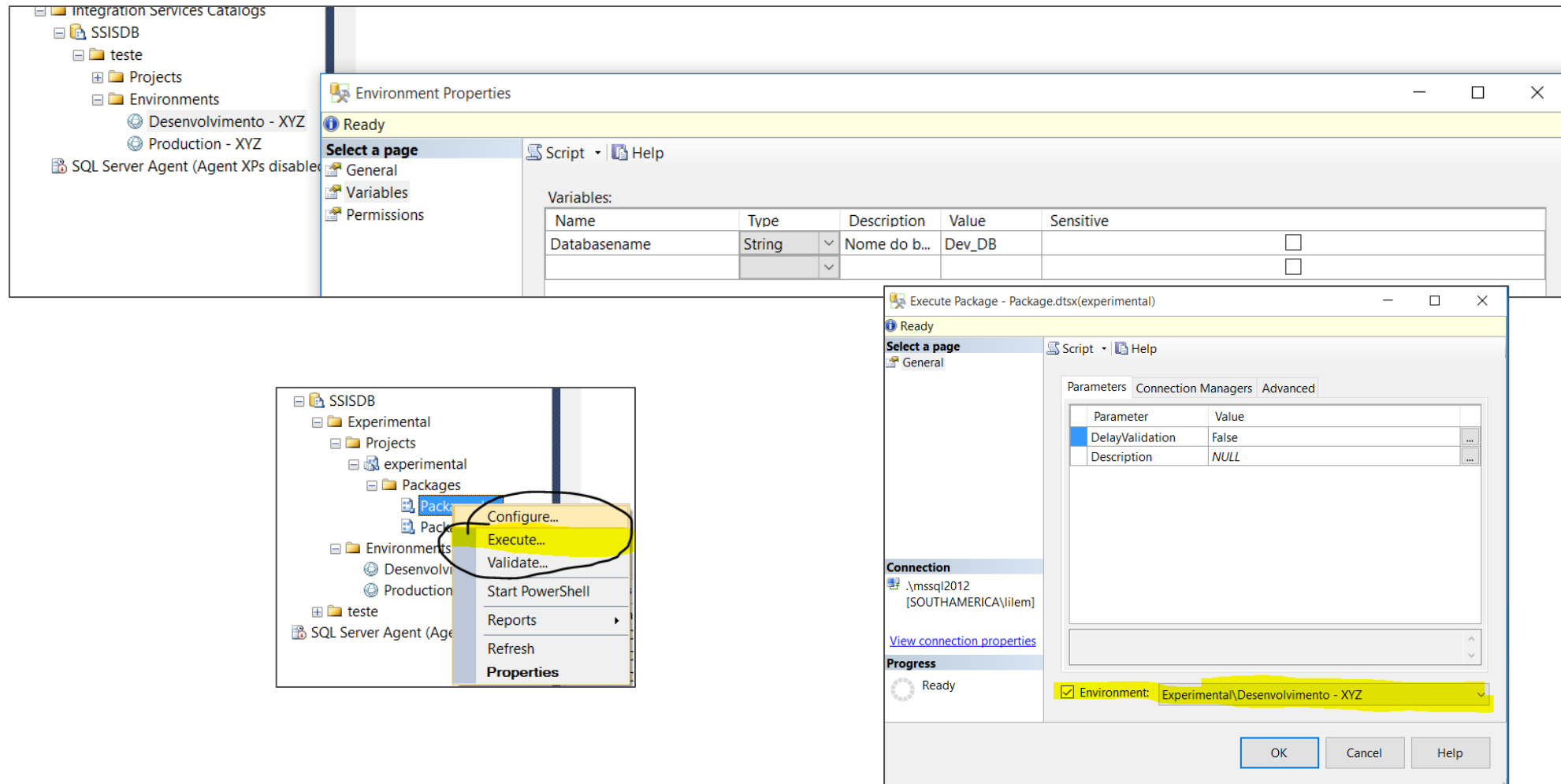
- Parameters e Enviroments

Package1.dtsx [Design] x Package.dtsx [Design] Package.dtsx [XML]					
Control Flow Data Flow Parameters Event Handlers Package Explo...					
Name	Data type	Value	Sensitive	Required	Description
Databasename	String	0	False	False	

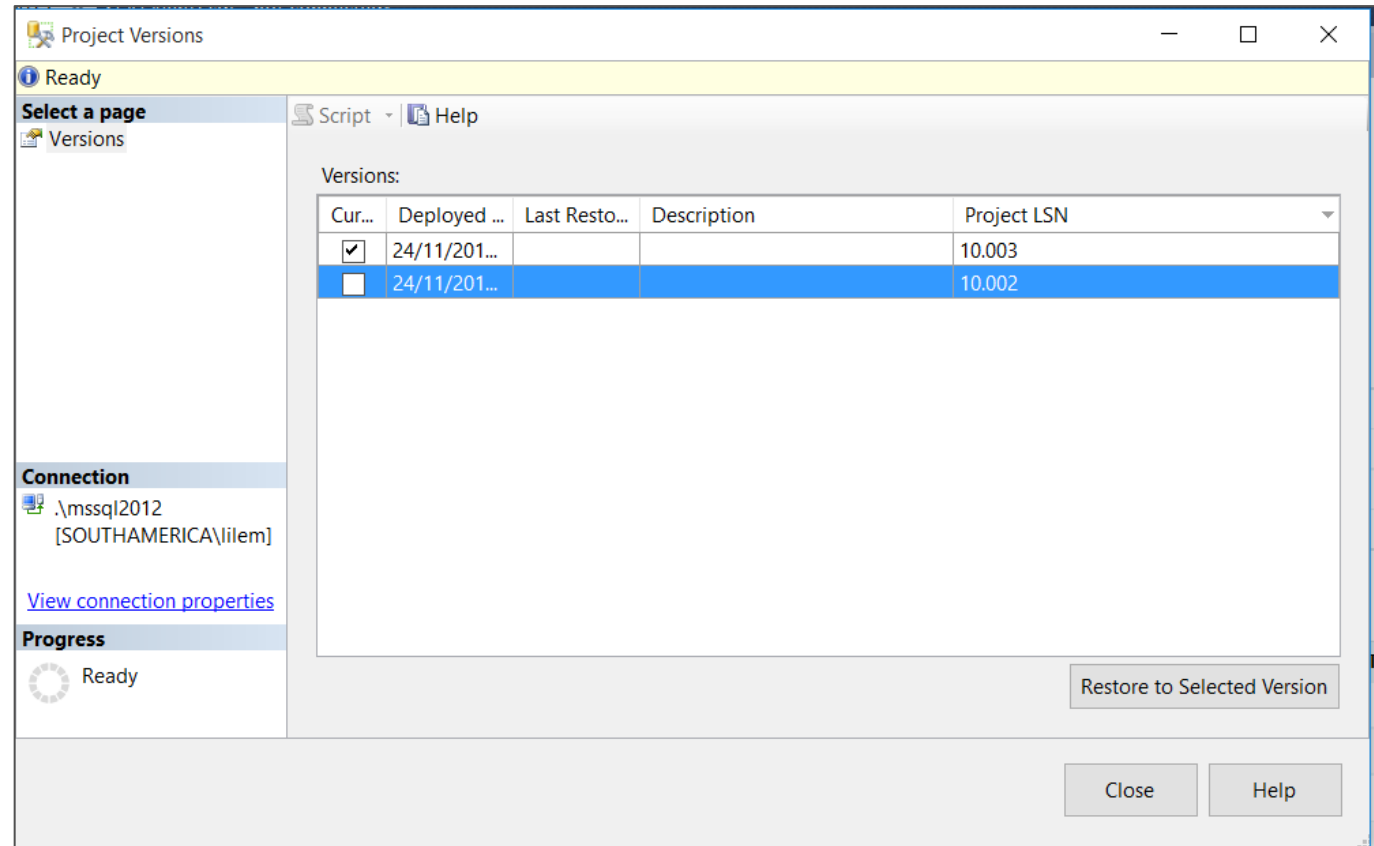
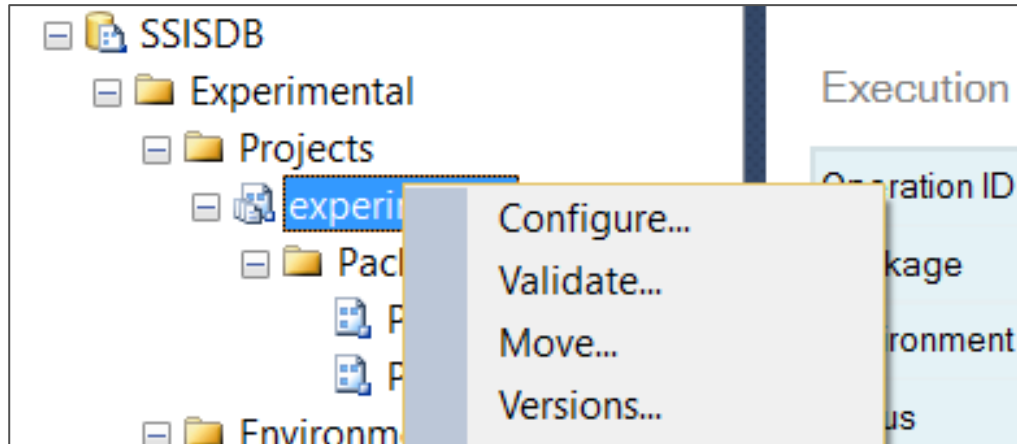


# Configuração do pacote SSIS (2012 ou >)

- Parameters e Enviroments



# Versionamento

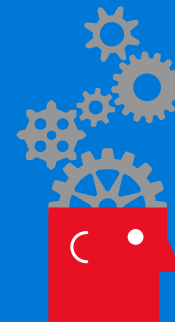


# DEMO

Package Config 2008 / 2012 e Versionamento

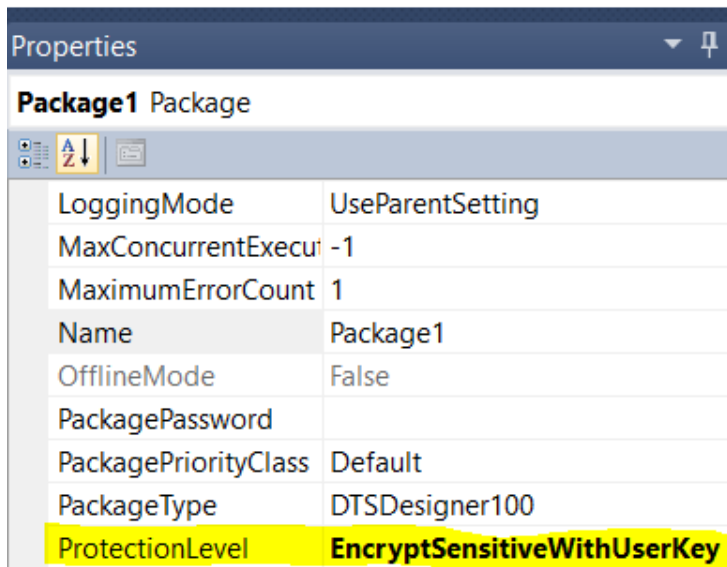
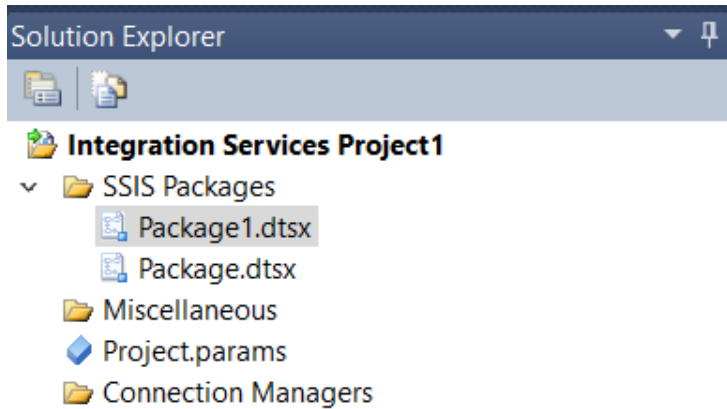


# Não esqueça...



Use o **Project deployment model** para ter um uso mais flexível da configuração do SSIS.

# Protection Level



Dont Save  
Sensitive

Encrypt  
Sensitive With  
User Key

Encrypt  
Sensitive With  
Password

Encrypt All  
With  
Password

Encrypt All  
With User Key

Server  
Storage

# DEMO

Protection Level





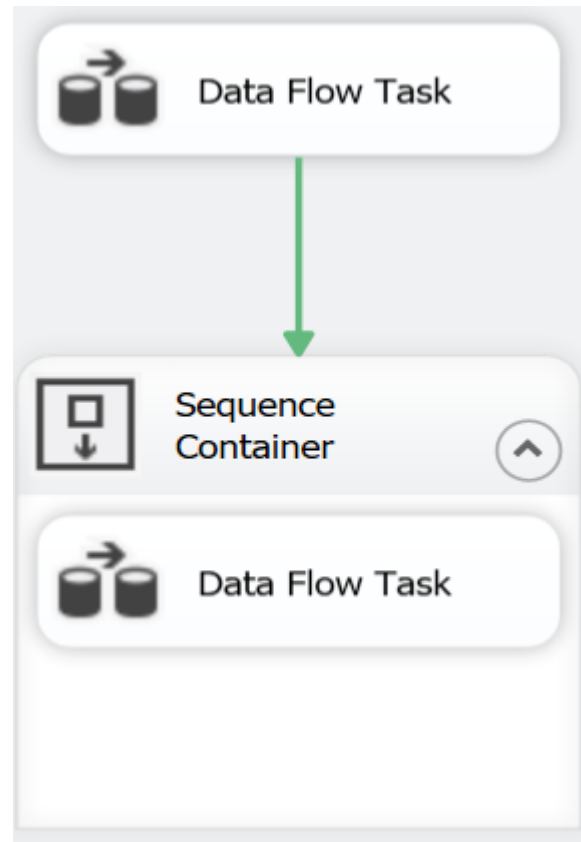
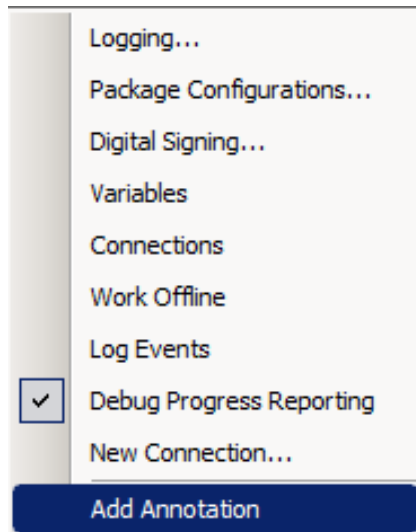
# Não esqueça...



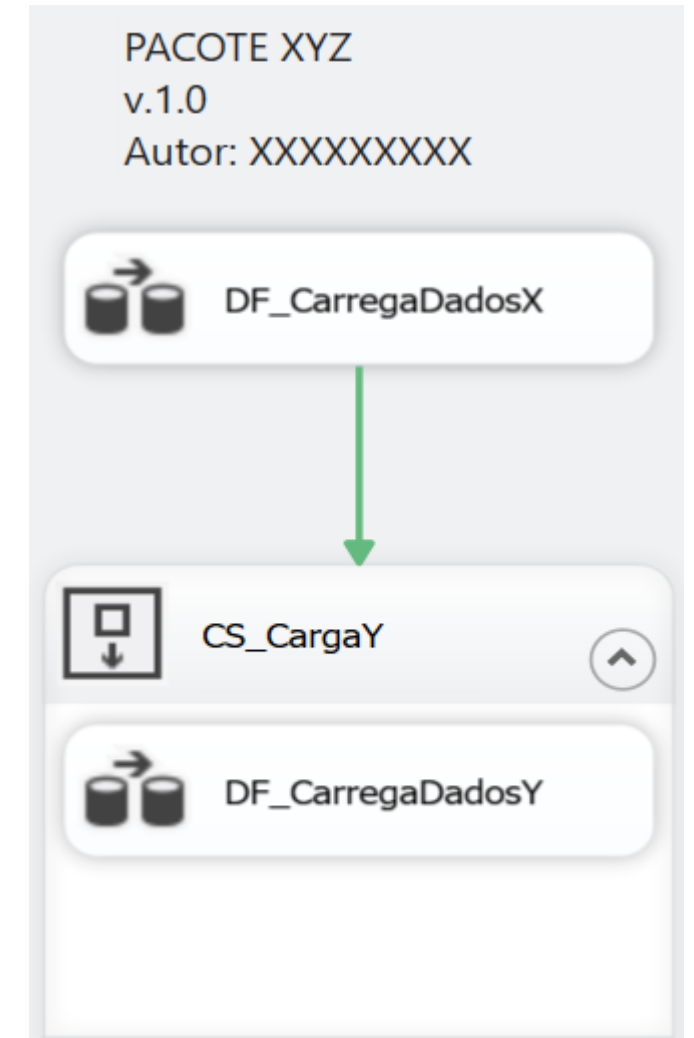
Configure as packages tornando flexível o processo de deploy em ambientes diferentes

Adeque o protection level para o ambiente de produção

# Documentação



VS



# DEMO

## Documentação



# Não esqueça...

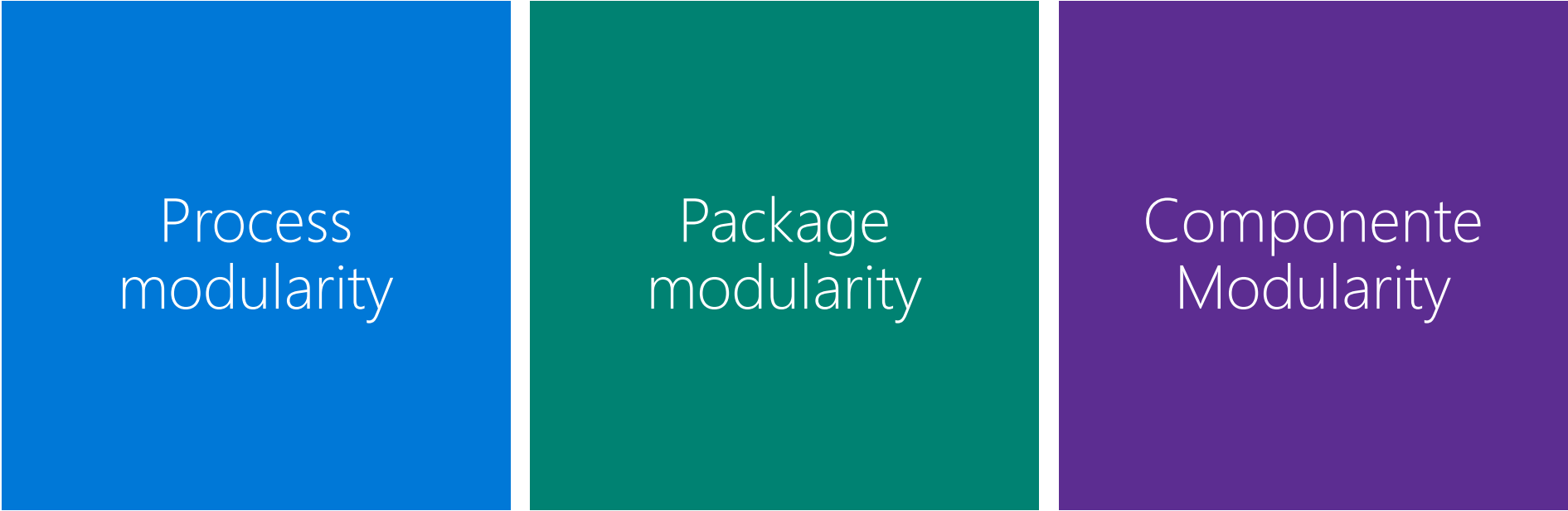


Comente o código

Documente as tasks

Outras pessoas irão  
dar manutenção  
neste código

# Design

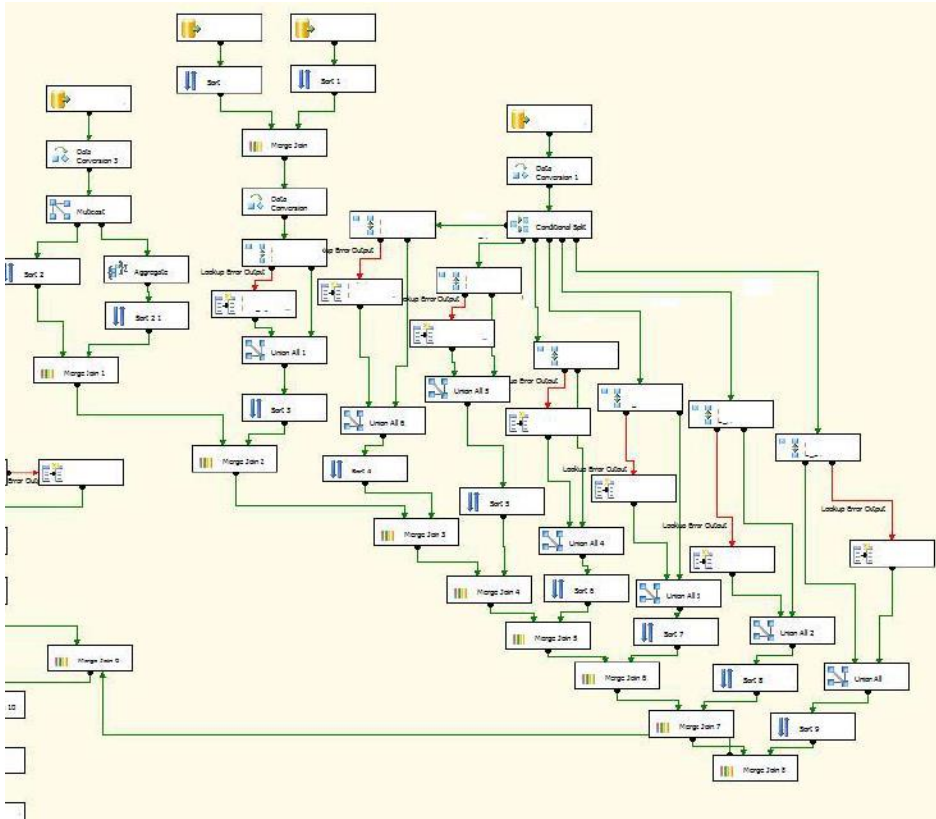


Process  
modularity

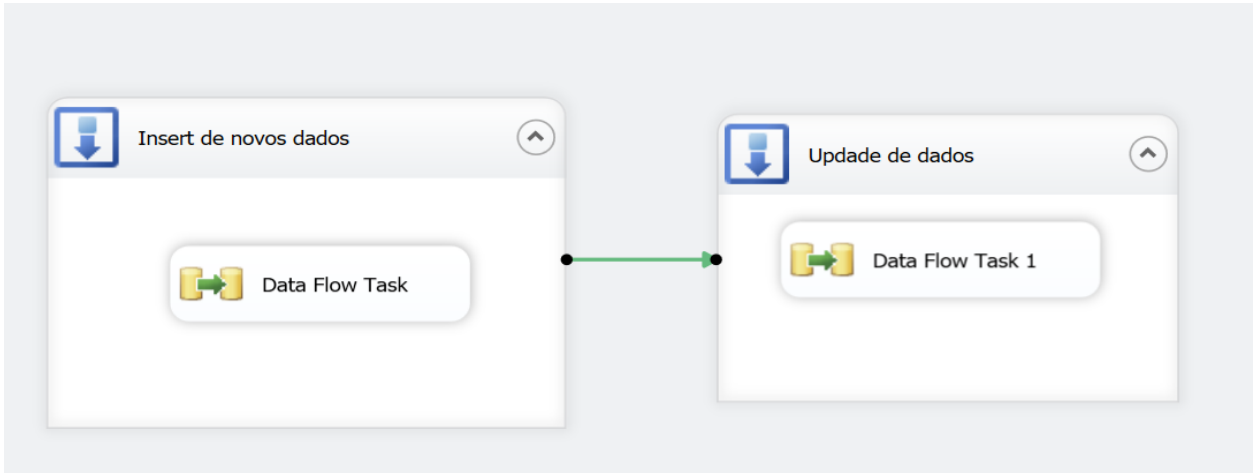
Package  
modularity

Componente  
Modularity

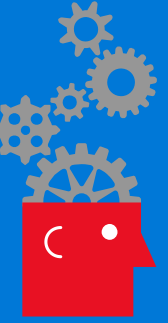
# Design



VS



# Não esqueça...



Simplifique a lógica  
do pacote

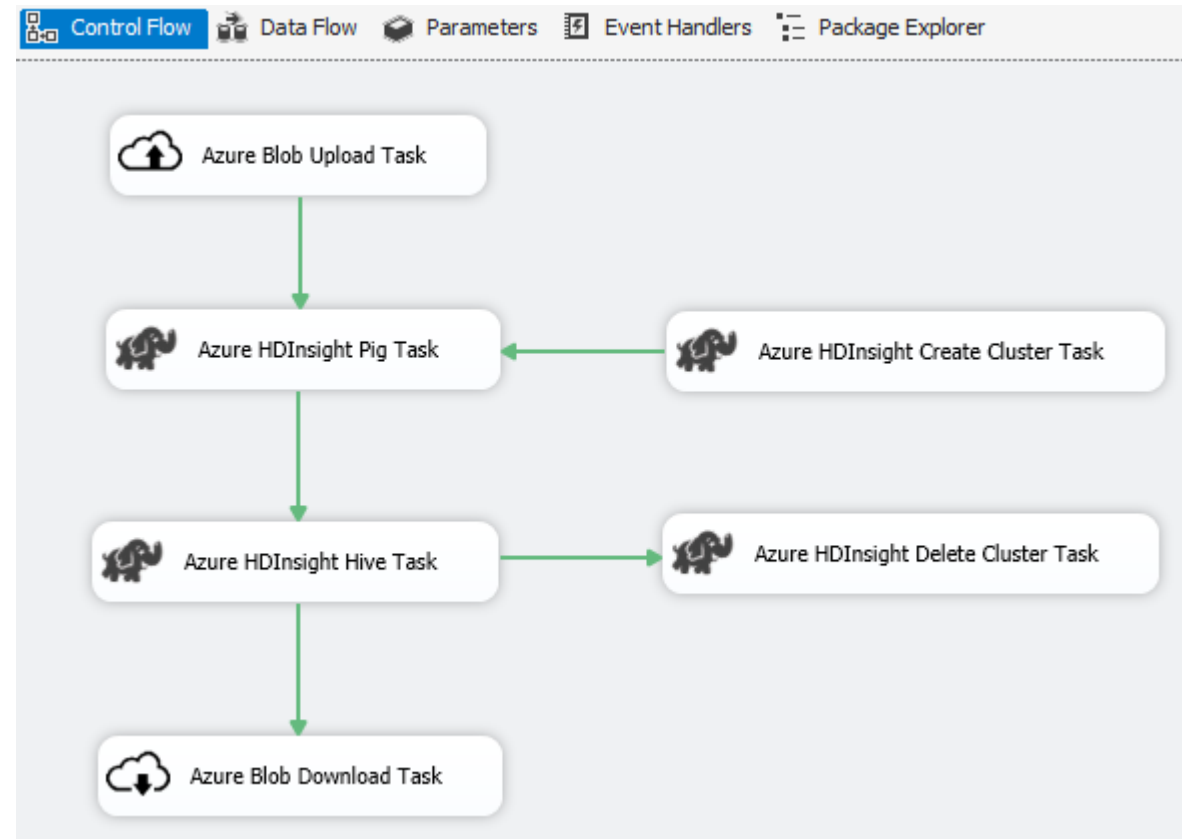
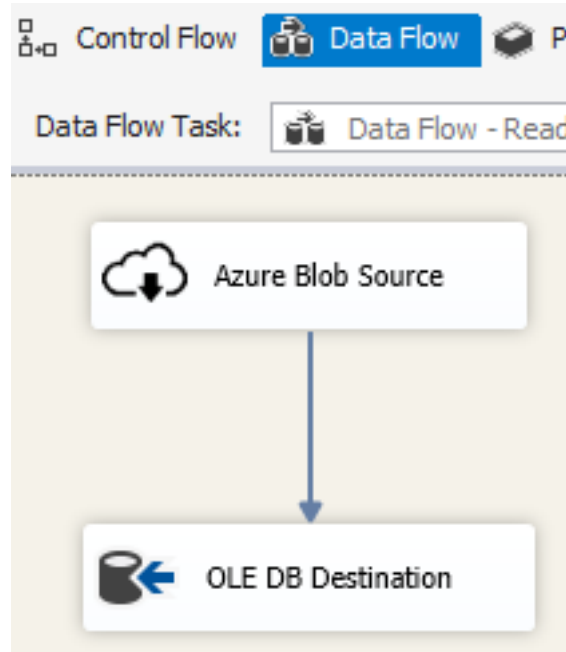
Agrupe tasks através  
do uso de containers

# SSIS e Azure





# SSIS Azure Feature Pack



# DEMO

## SSIS Azure Pack



