Integration Services SSIS
Guia de boas práticas

Liliam Leme
Sergio Fonseca
Renata Festa

Microsoft

AGENDA

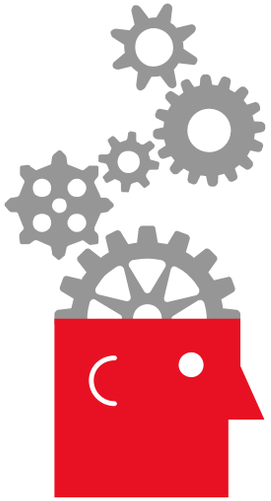
Otimizando Inserts

Componentes

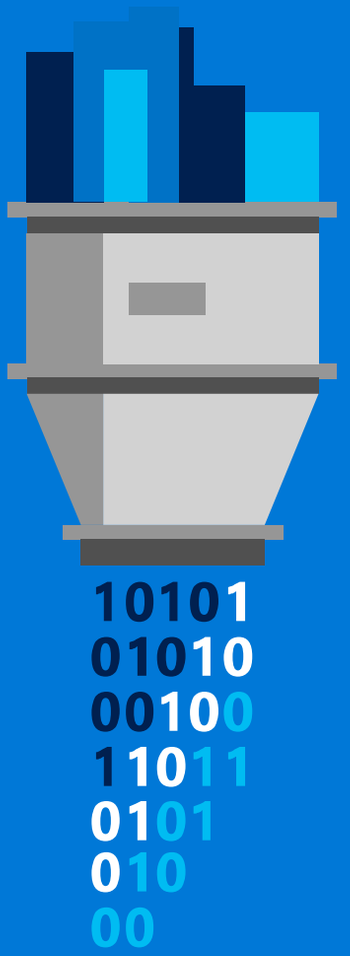
SSIS e o Azure

SSIS

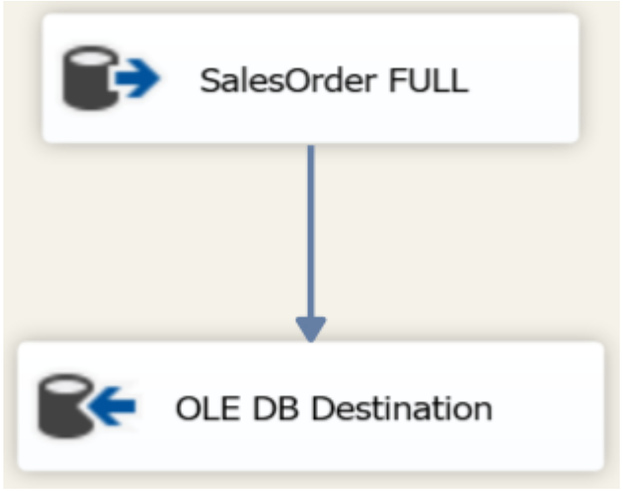
O SQL Server Integration Services (**SSIS**) é uma ferramenta de **ETL**: Extract Transformation Load e pode ser usado para resolver problemas complexos de negócio, permitindo a **transferência de dados**, **copiar arquivos**, **enviar email** em resposta a eventos, **tratar a informação** e **administrar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

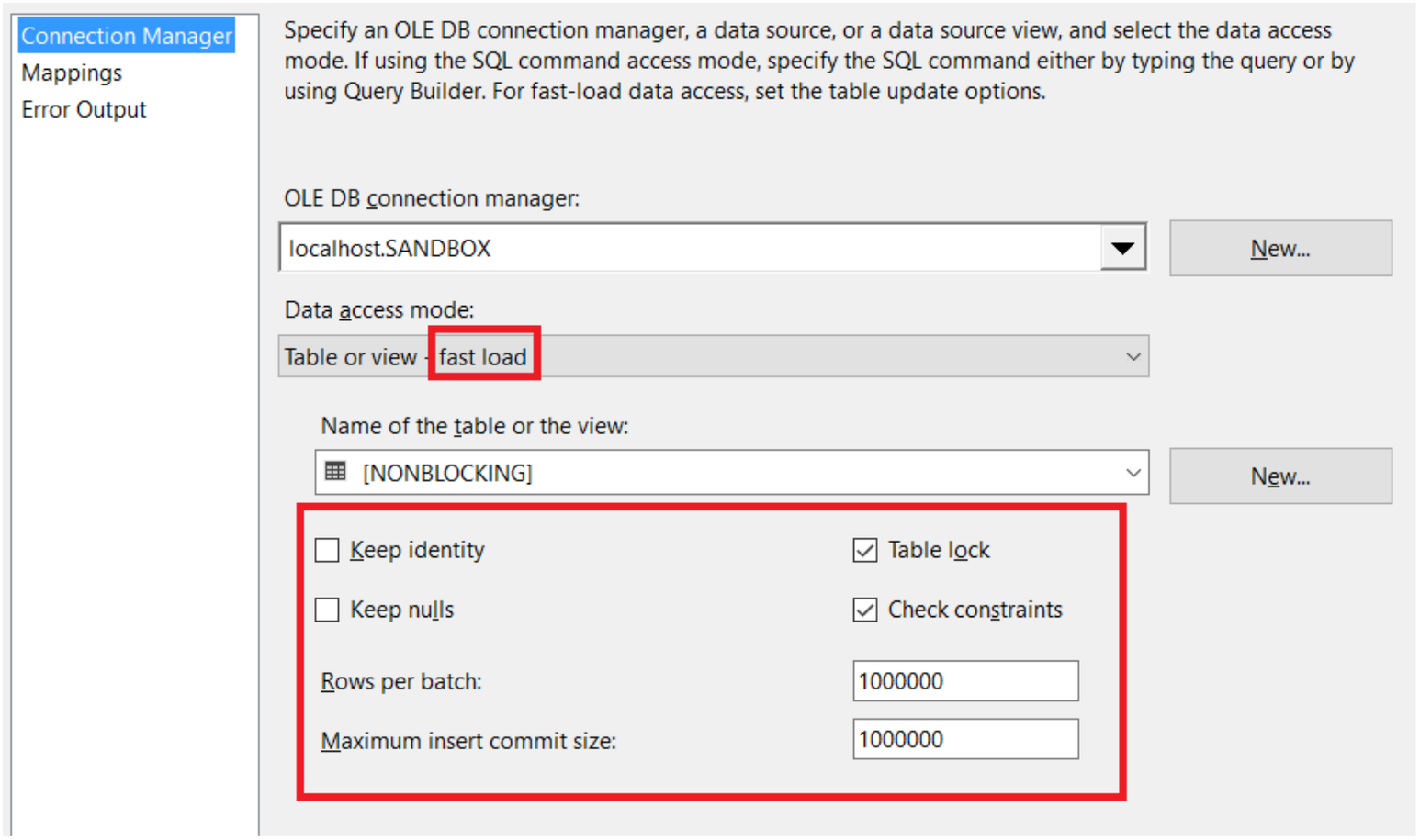


Otimizando Inserts



OLEDB Destination





Não esqueça...

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

Use Tablelock se o cenário for aplicável

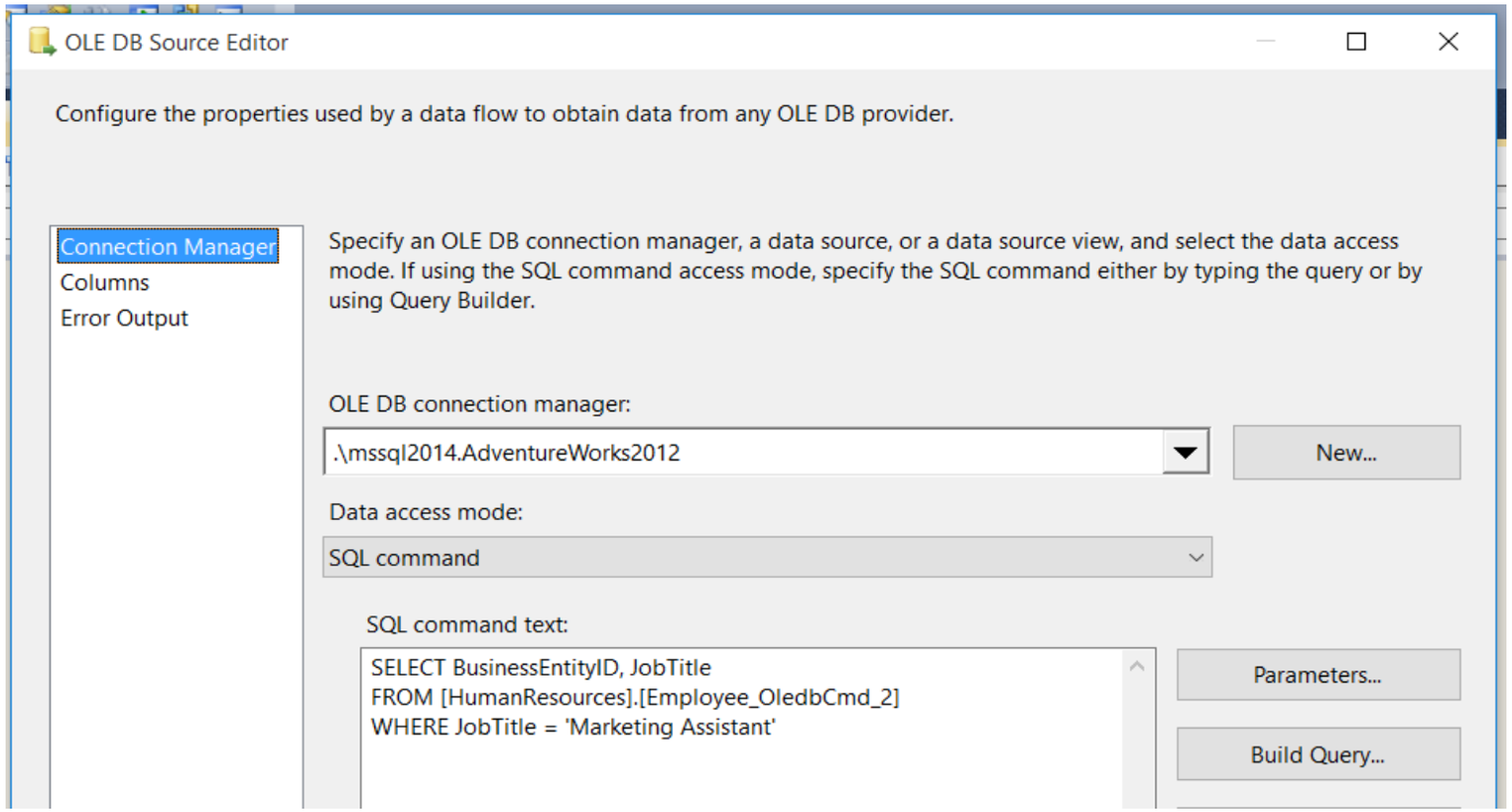
Avalie commit size e rows por batch

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* remover todos os dados da tabela e apenas informa isso no log.
- Use *Switch partition* para tabelas particionadas

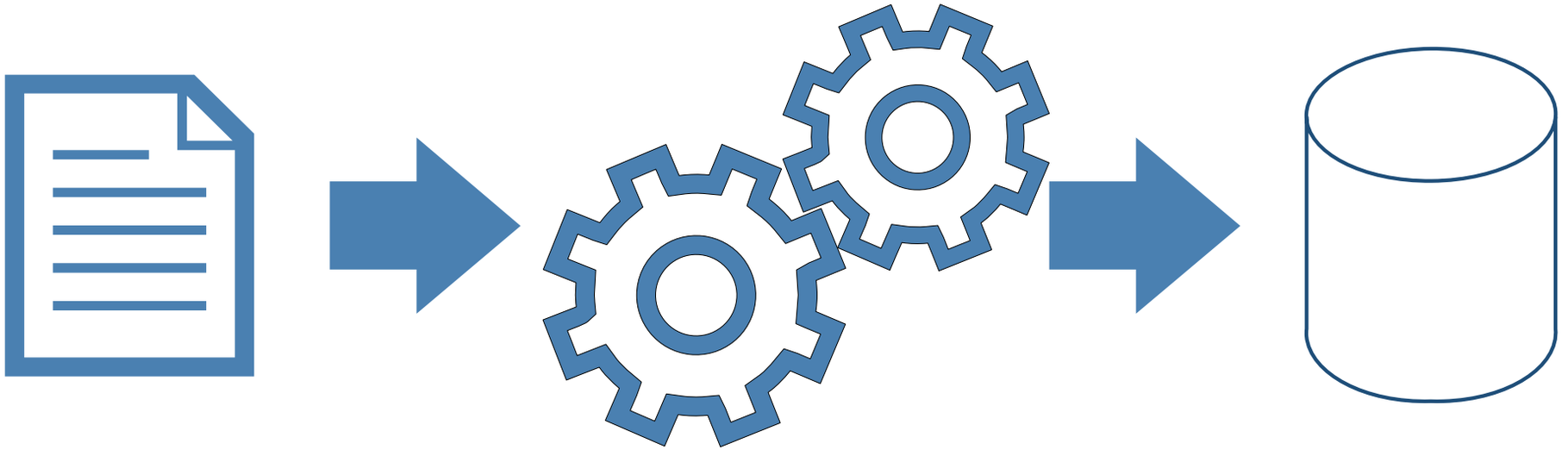
Otimize sua transferência de dados

- Selecione apenas as colunas necessárias.



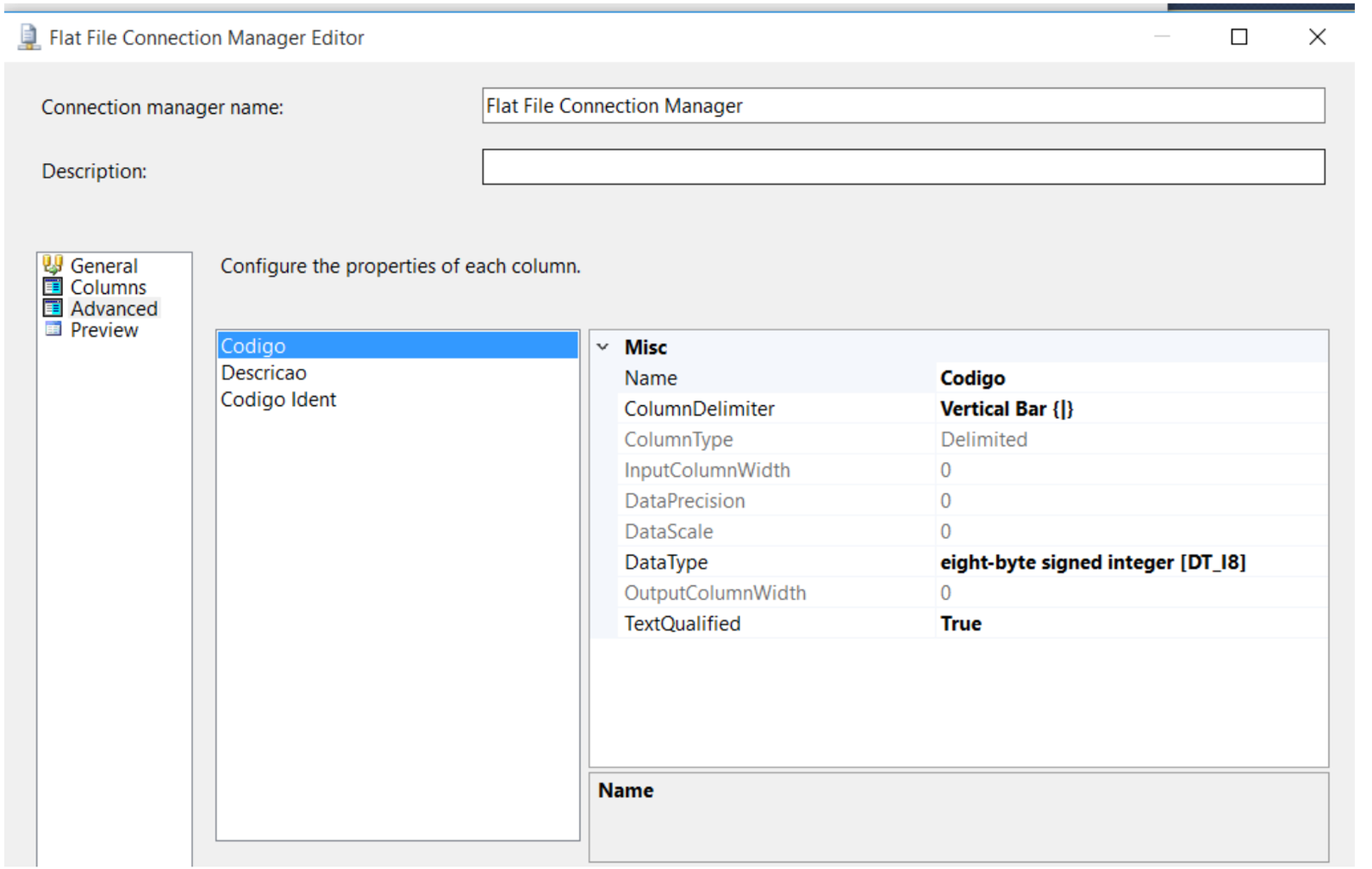
Carga para Arquivos

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



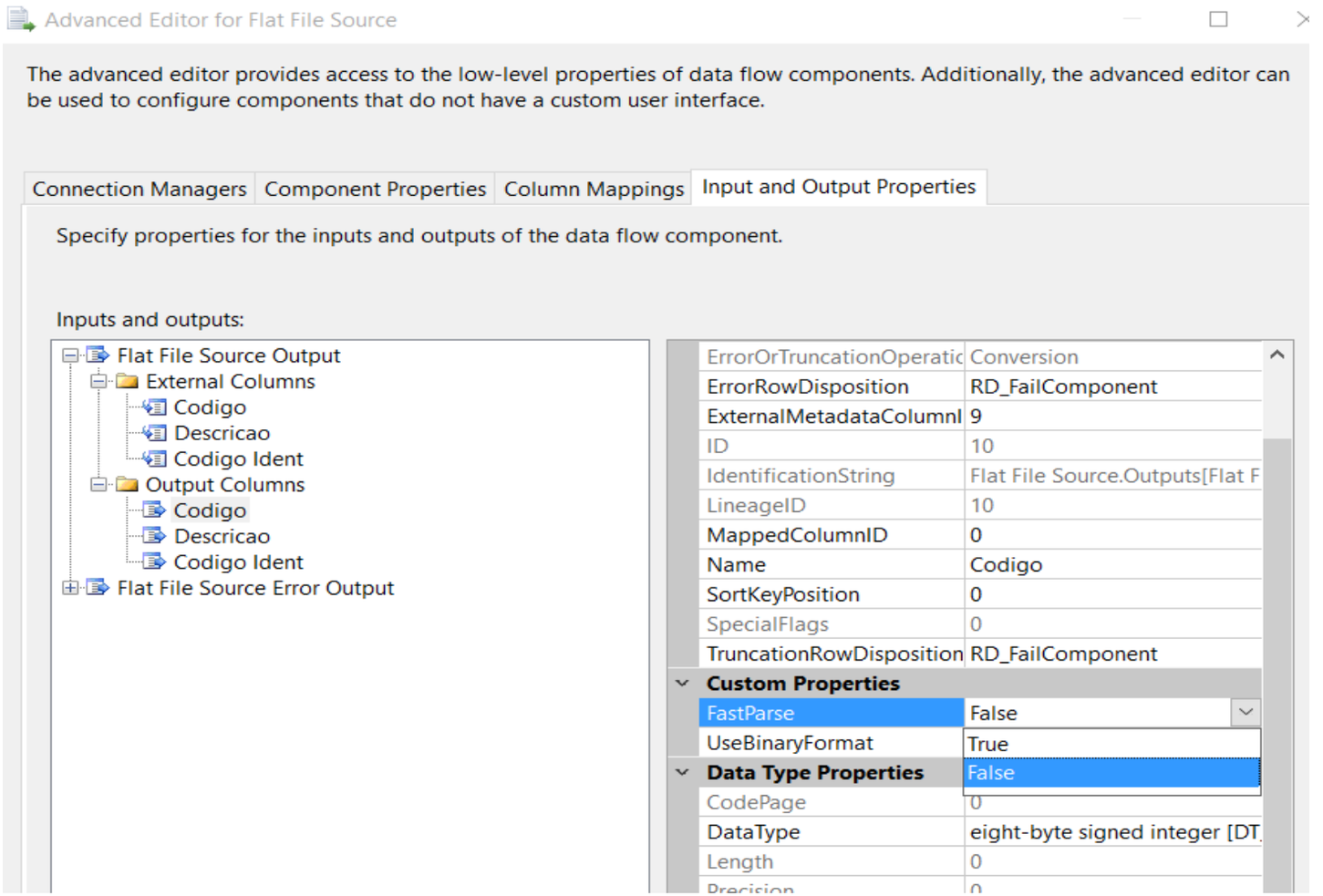
Carga para Arquivos

- Tipos de dados



Carga para Arquivos

- Int e datetime



Não esqueça...

Dê preferência a operações minimamente logadas.

Não traga mais colunas ou mais dados que o necessário

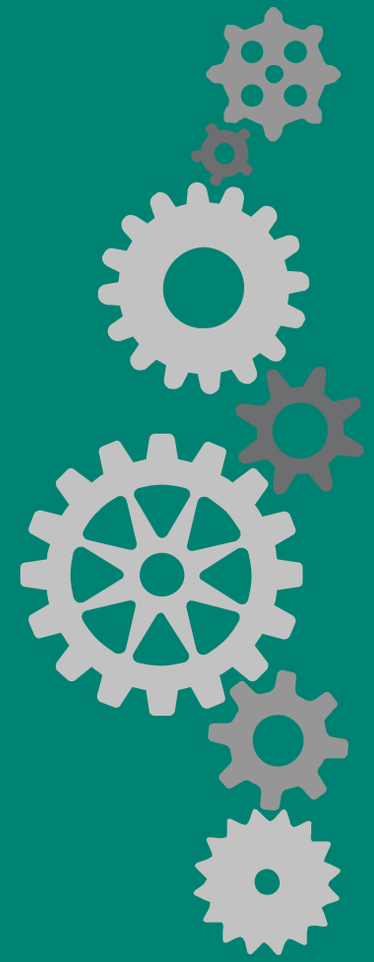
Configure adequadamente os tipos de dados do arquivo.

Use fast parse para flat file

DEMO

Inserts / Fast Parse

Componentes



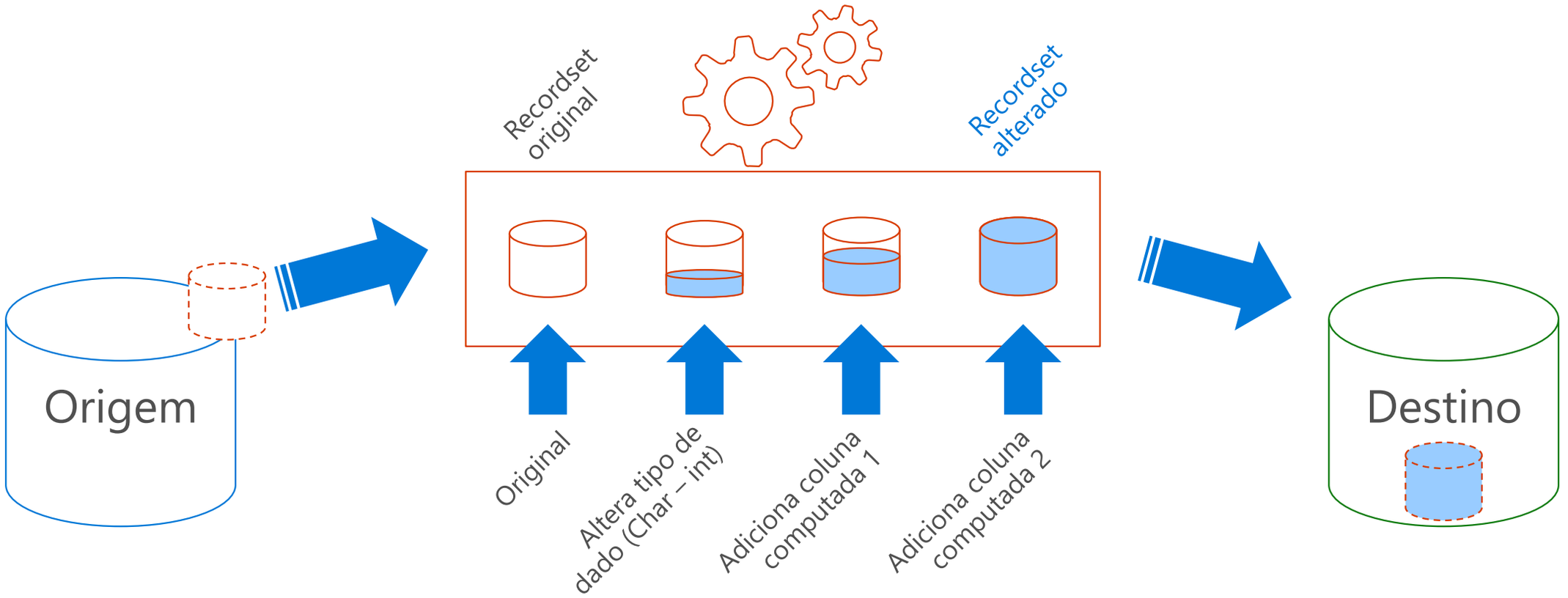
Full, Partially ou Non Blocking operations

Nonblocking ou row transformations

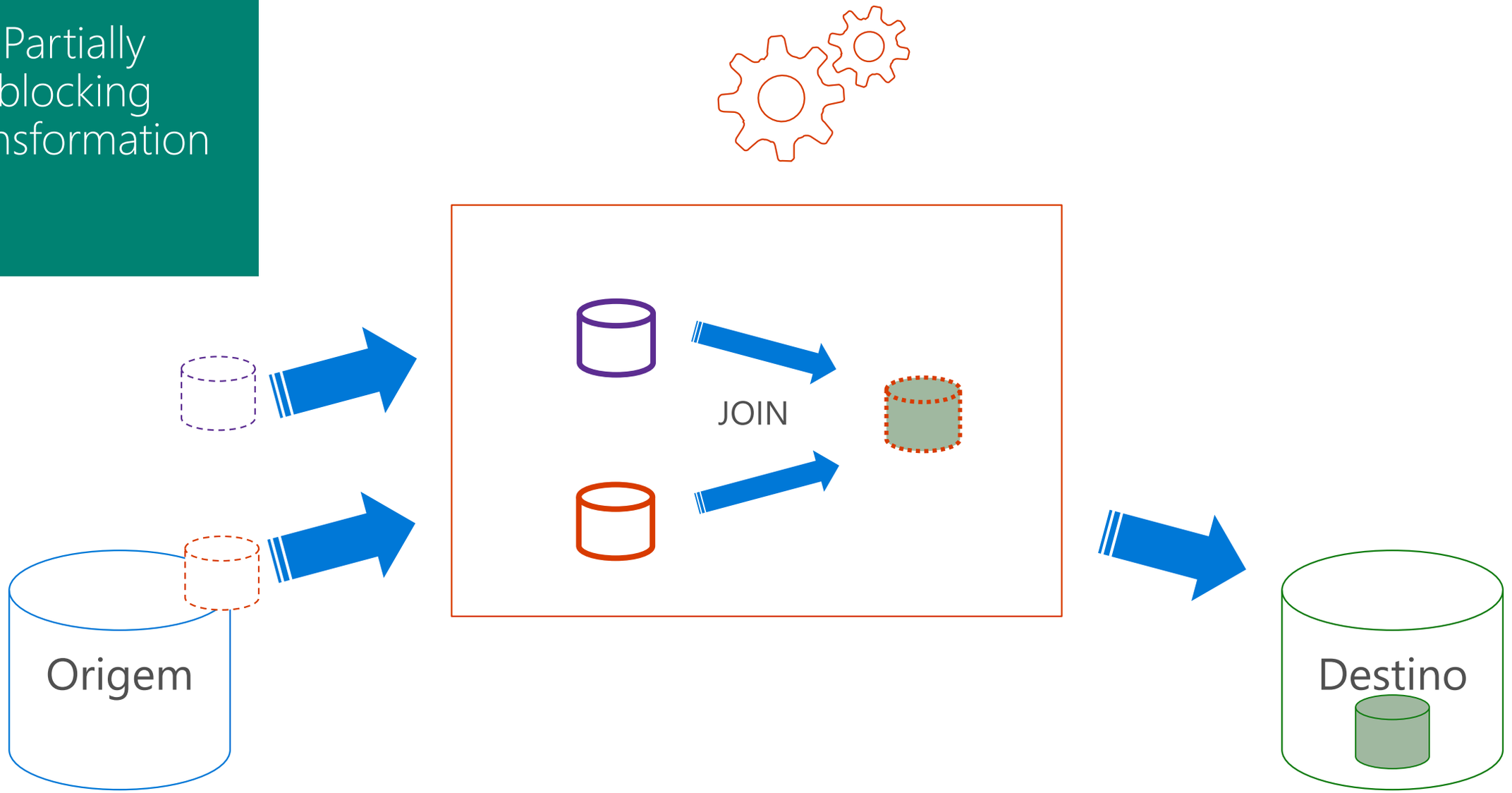
Partially blocking transformation

Full blocking transformations

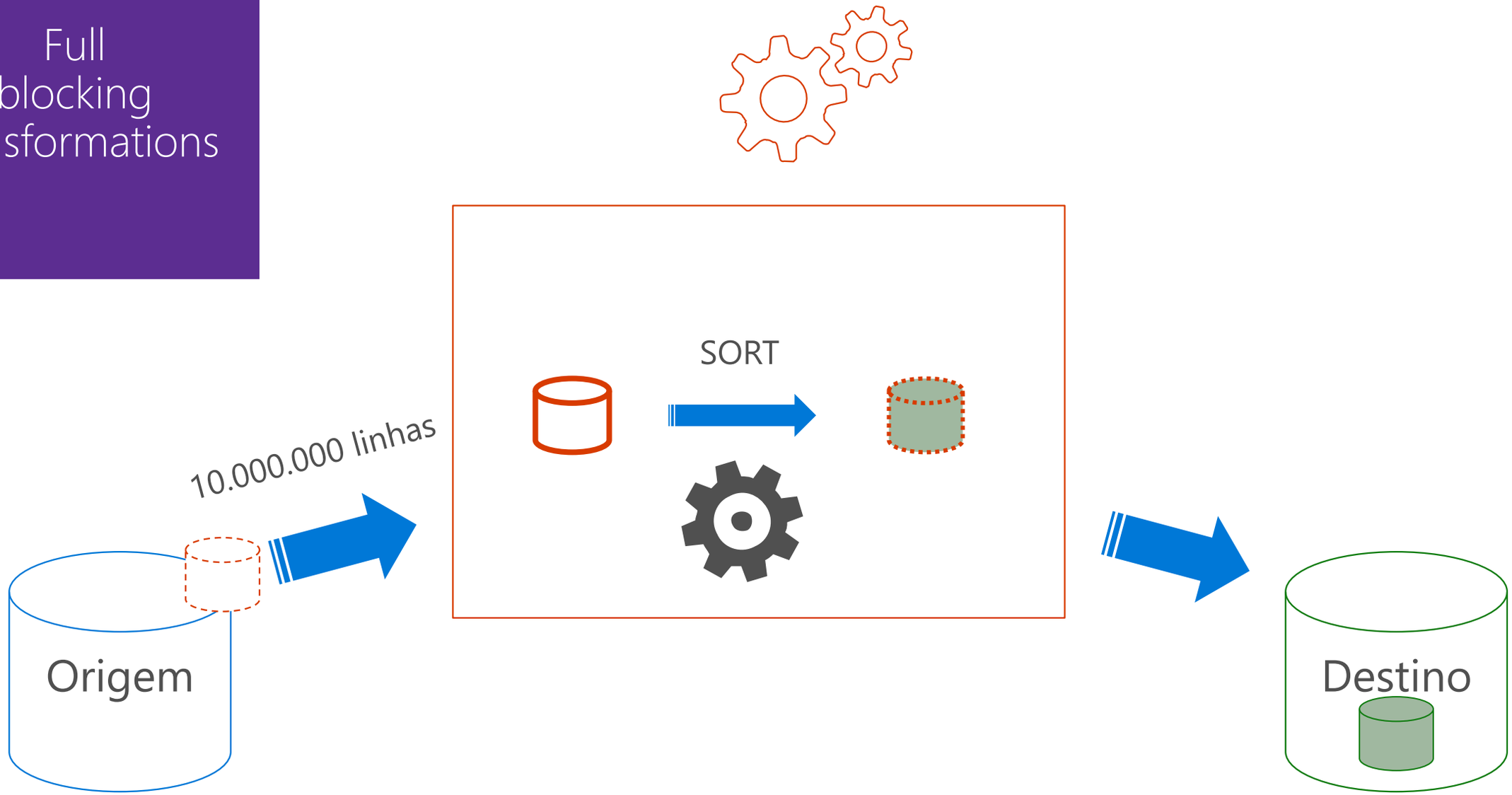
Nonblocking ou row transformations



Partially blocking transformation

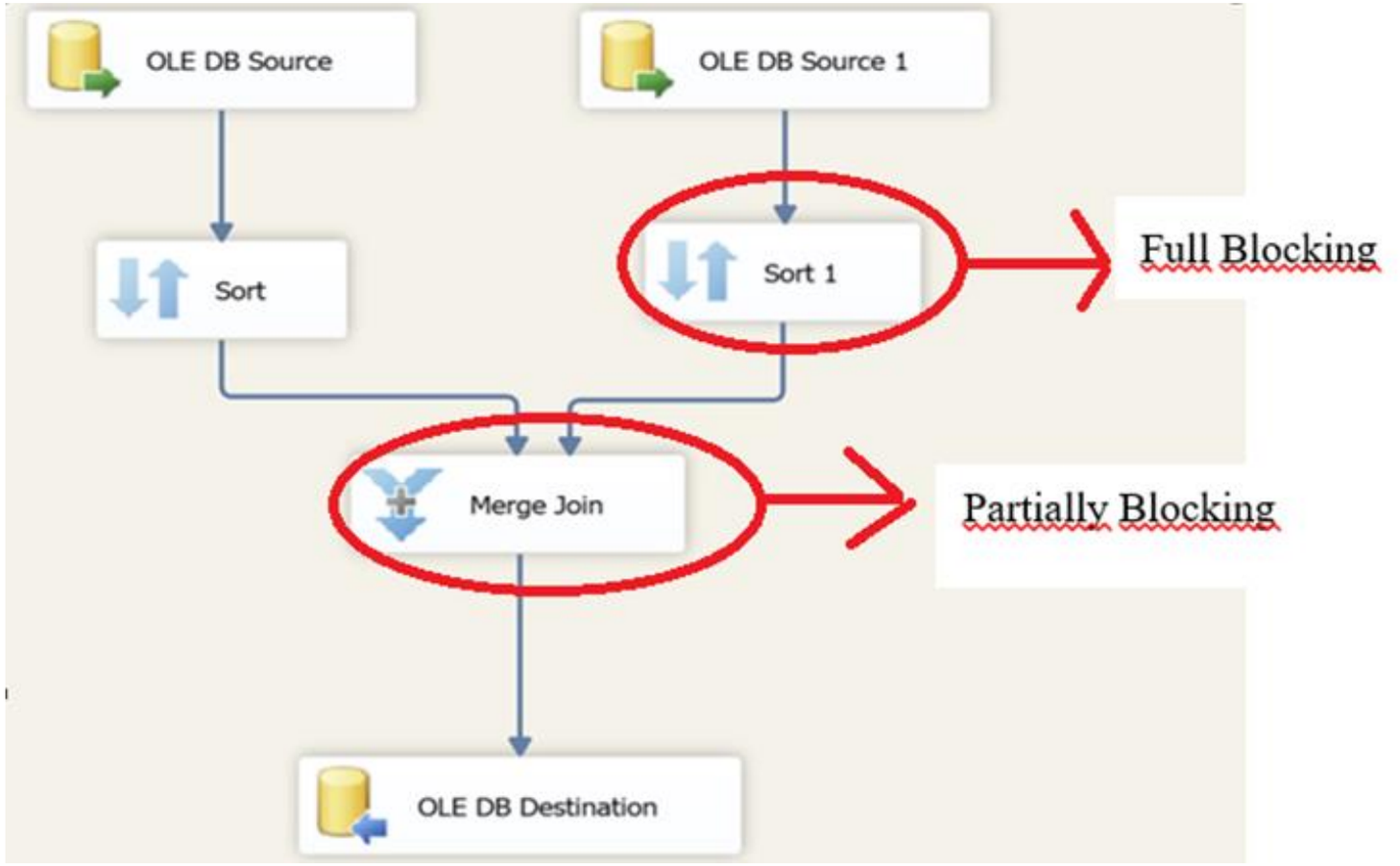


Full blocking transformations

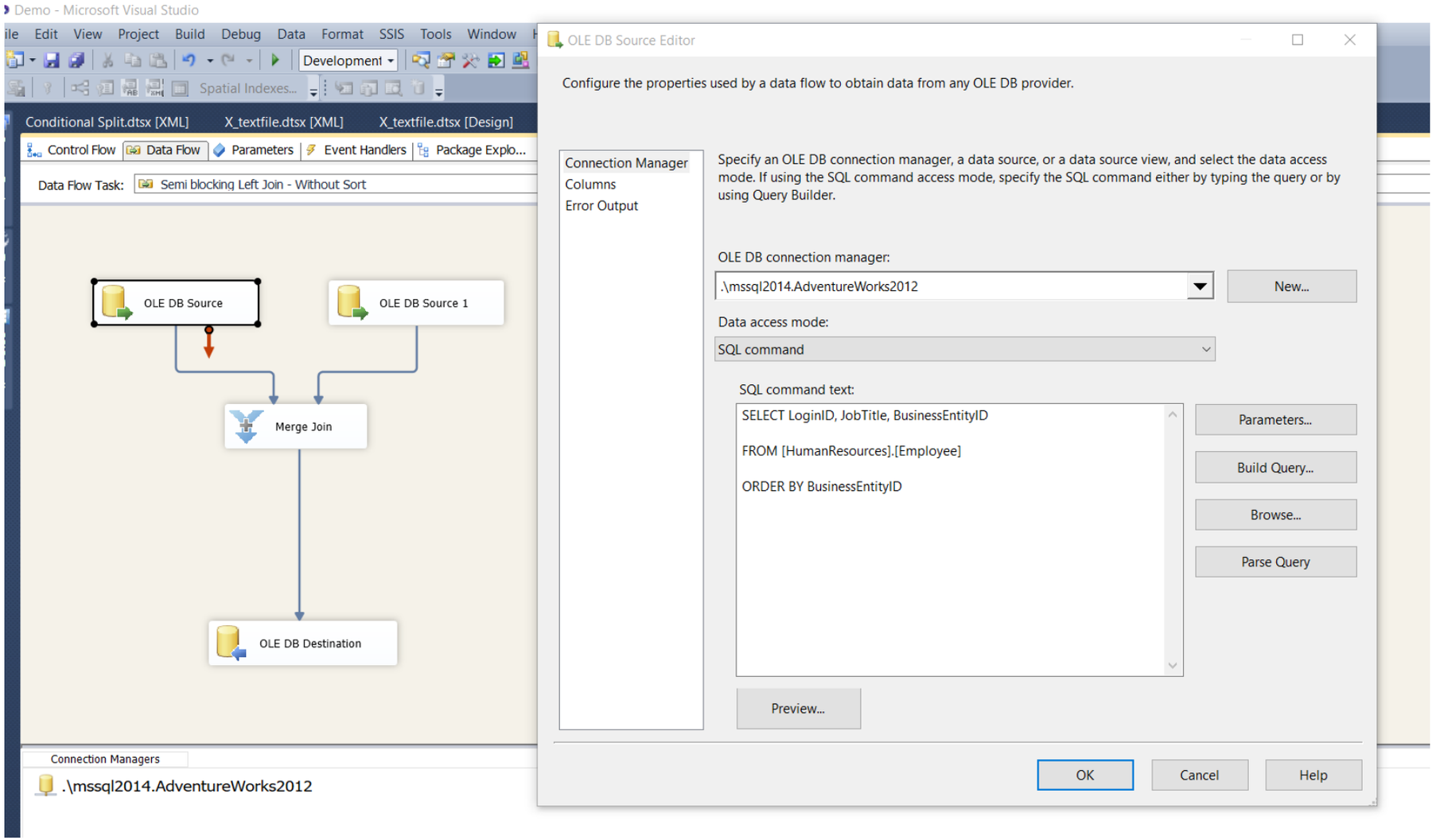


Full, Partially ou Non Blocking operations

- ✓ 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?

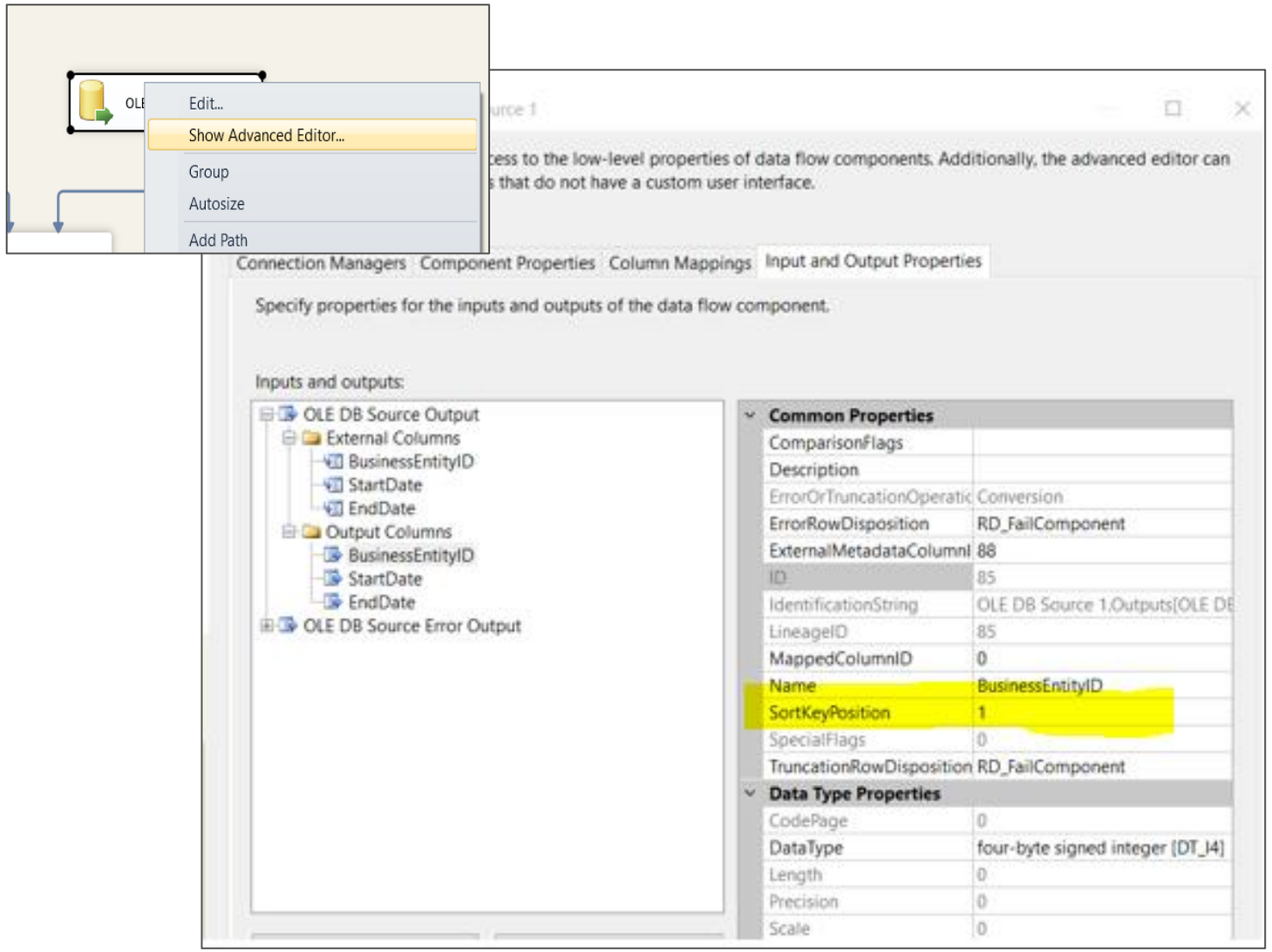


SELECT com ORDER BY

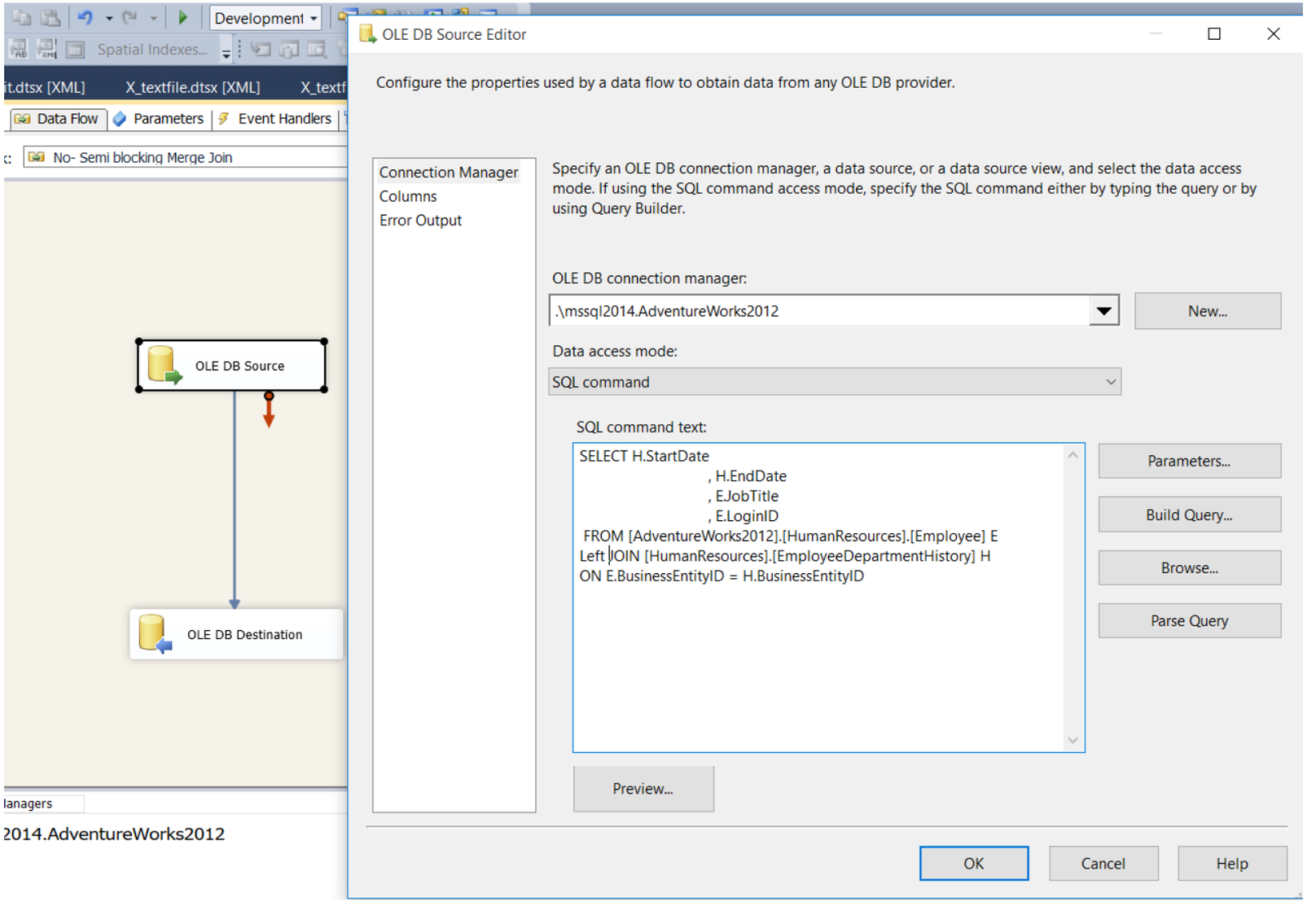


SELECT com ORDER BY

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 commom 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

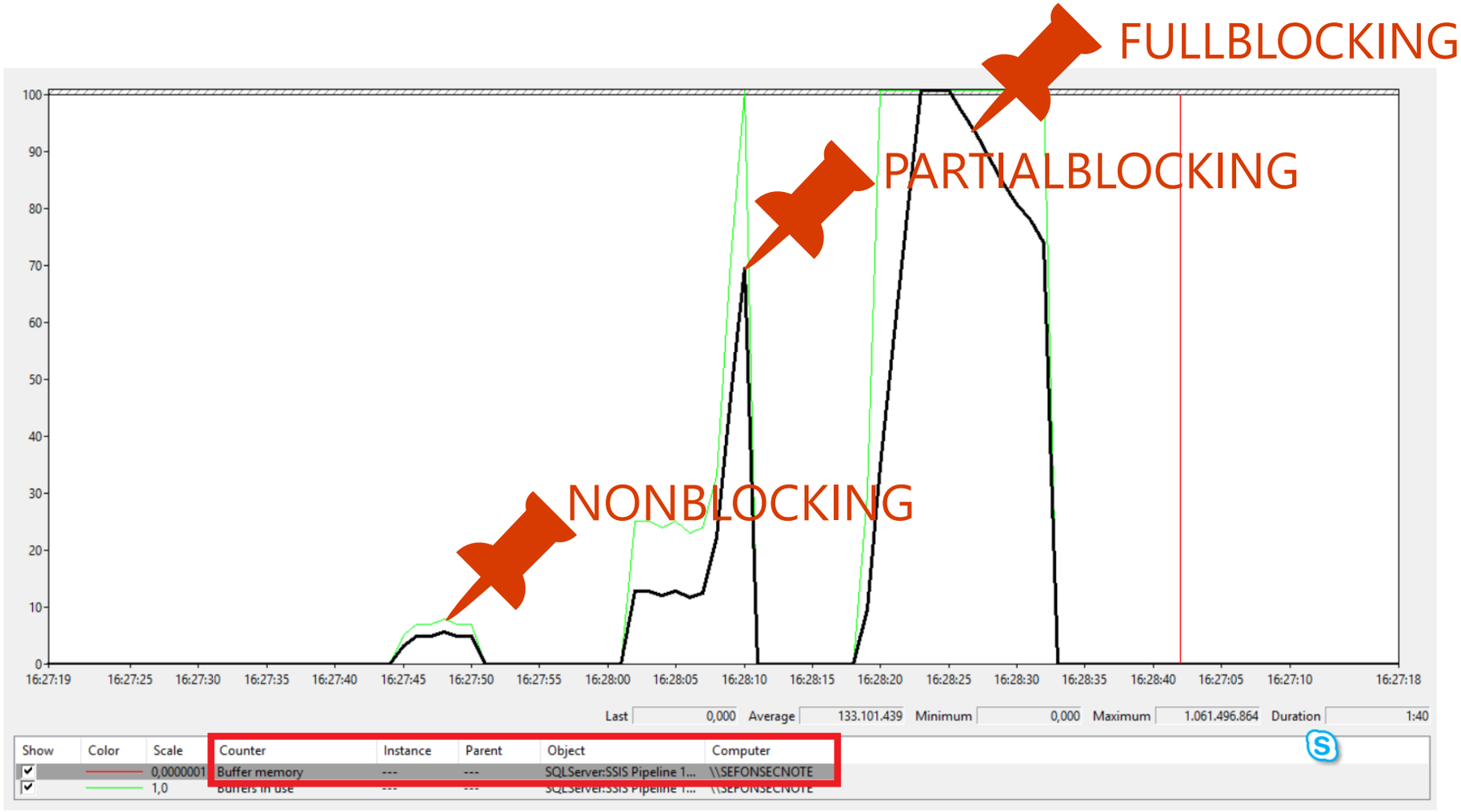
	Non-blocking	Semi-blocking	Fully-blocking
Synchronous/asynchronous	Síncrono	Assincrono	Assincrono
Número de linhas de saída igual a entrada	Sim	Geralmente não	Geralmente não
Precisa ler todos os dados antes de processar	Não	Não	Sim
Novo buffer é criado?	Não	Sim	Sim
Nova thread se faz necessária?	Não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

Full, Partially ou 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

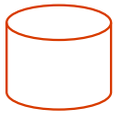
Evite, se possível, transformações full blocking ou partially blocking

OLEDB Command

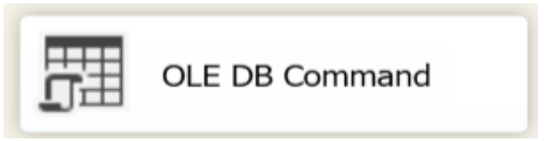
BUFFER

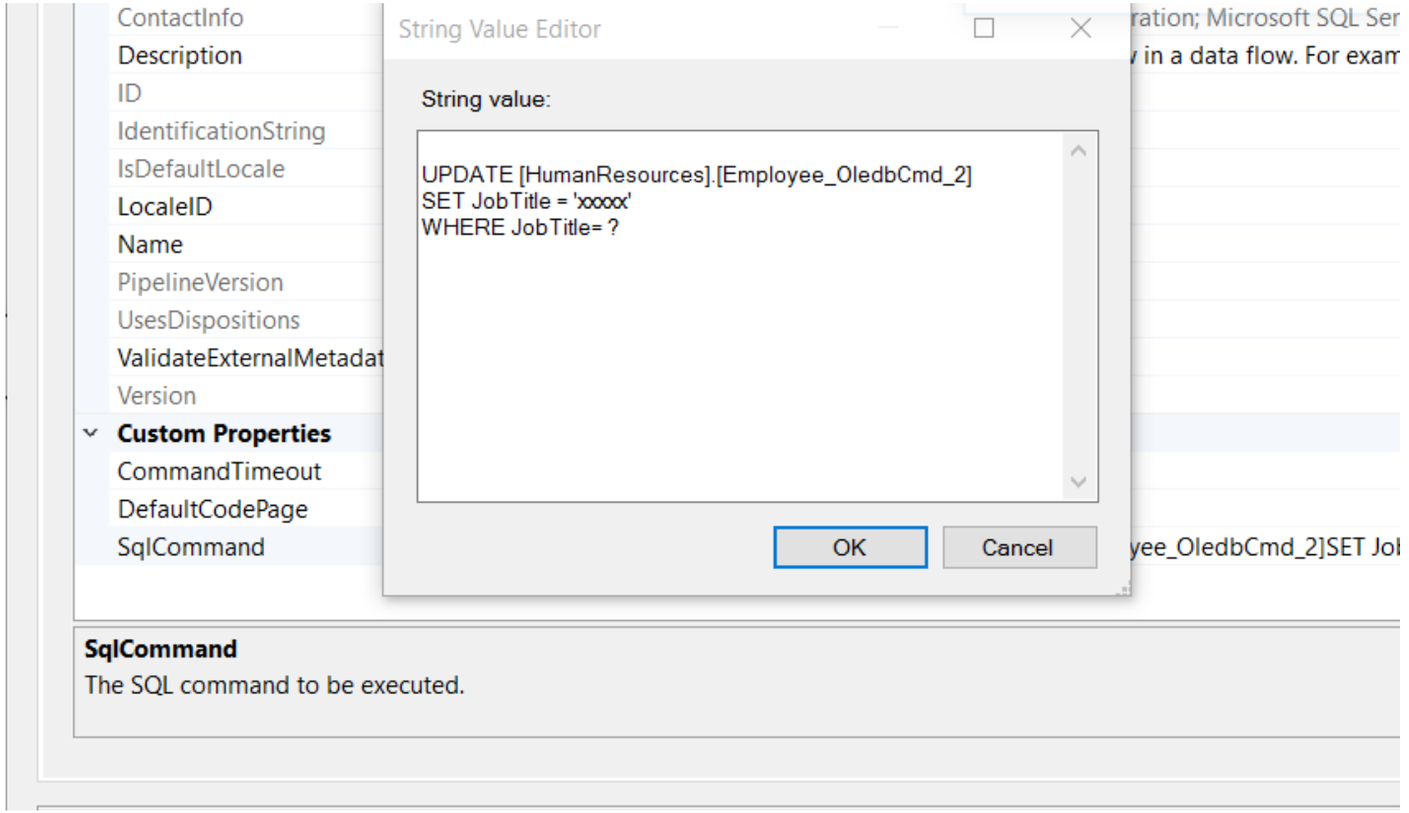
UPDATE TB_Destino

(Linha a Linha)



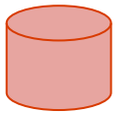
TB_Destino





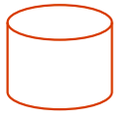
OLEDB Command -> Execute SQL Task

BUFFER

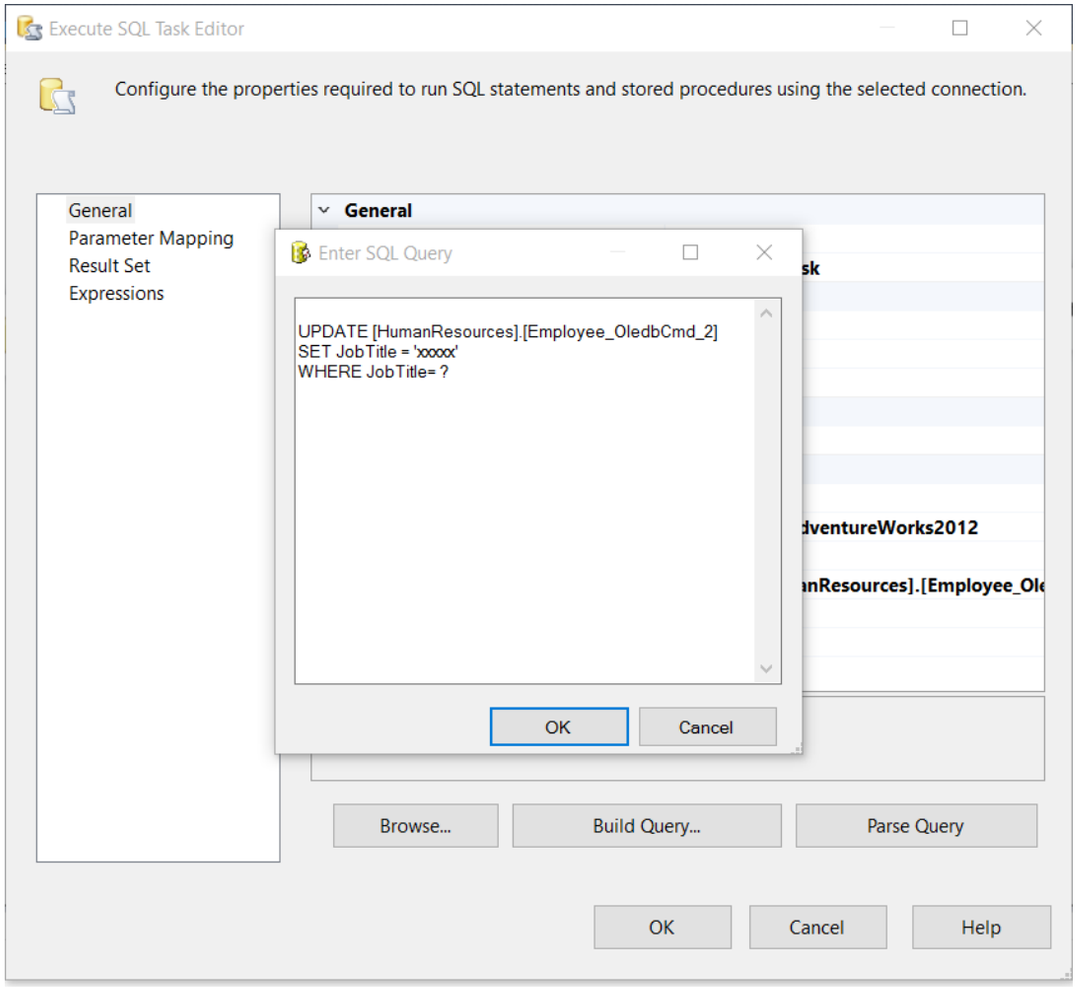


TB_Staging

UPDATE TB_Destino
JOIN TB_Staging



TB_Destino



DEMO

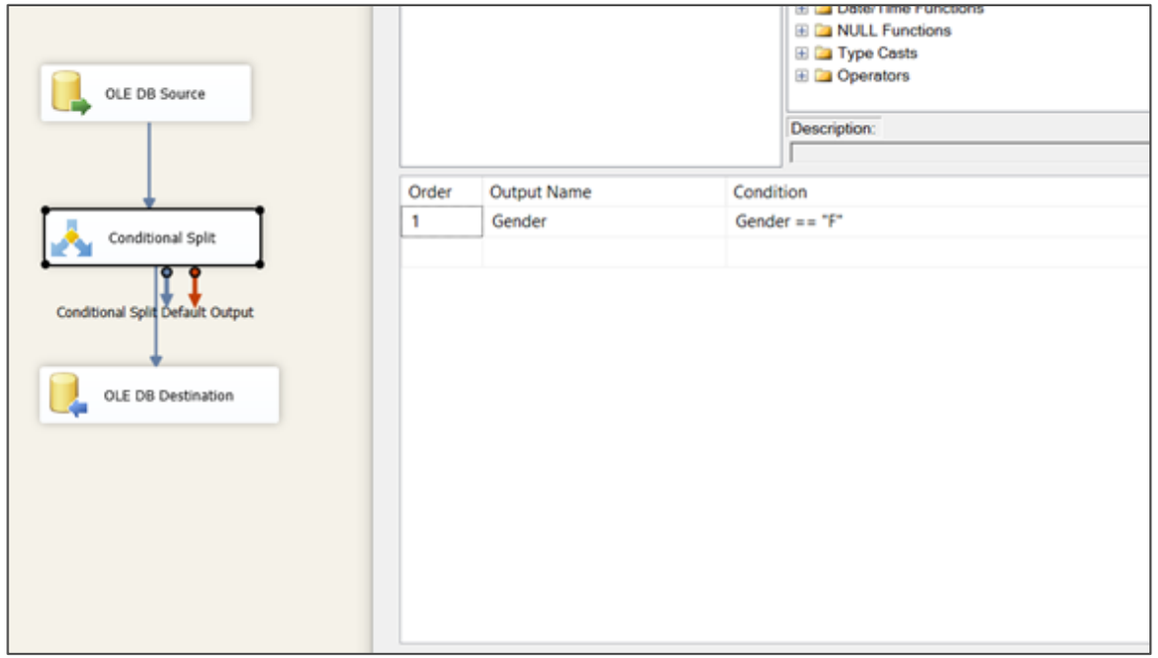
OLEDB Command

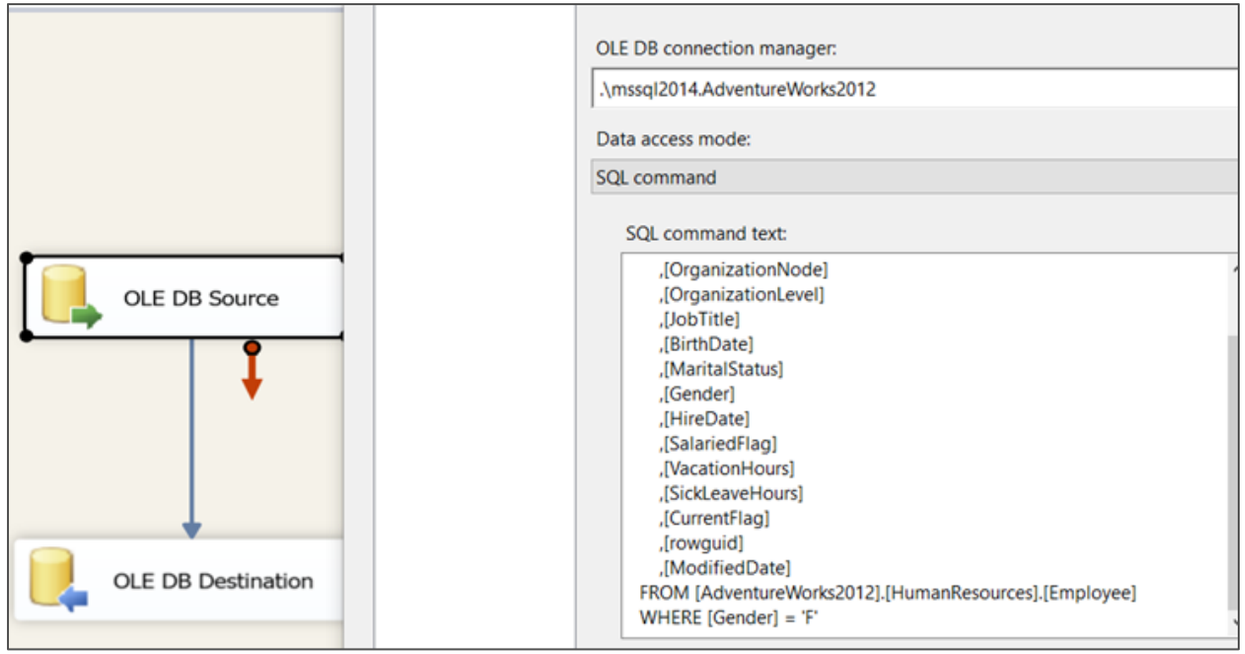
Não esqueça...

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.

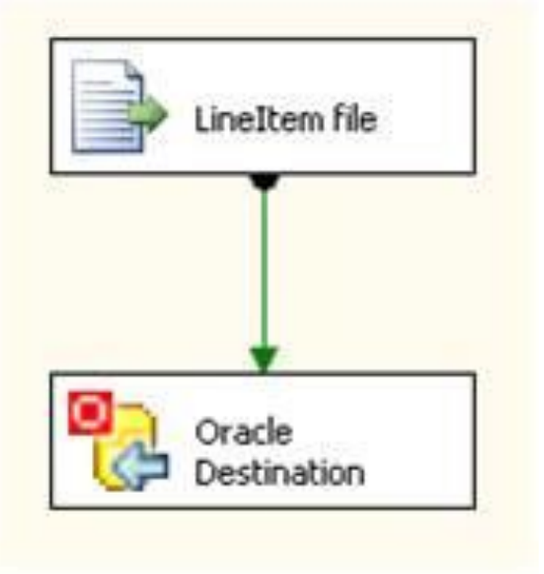
Se for possível filtre na origem dos dados

Oracle para SQL lento

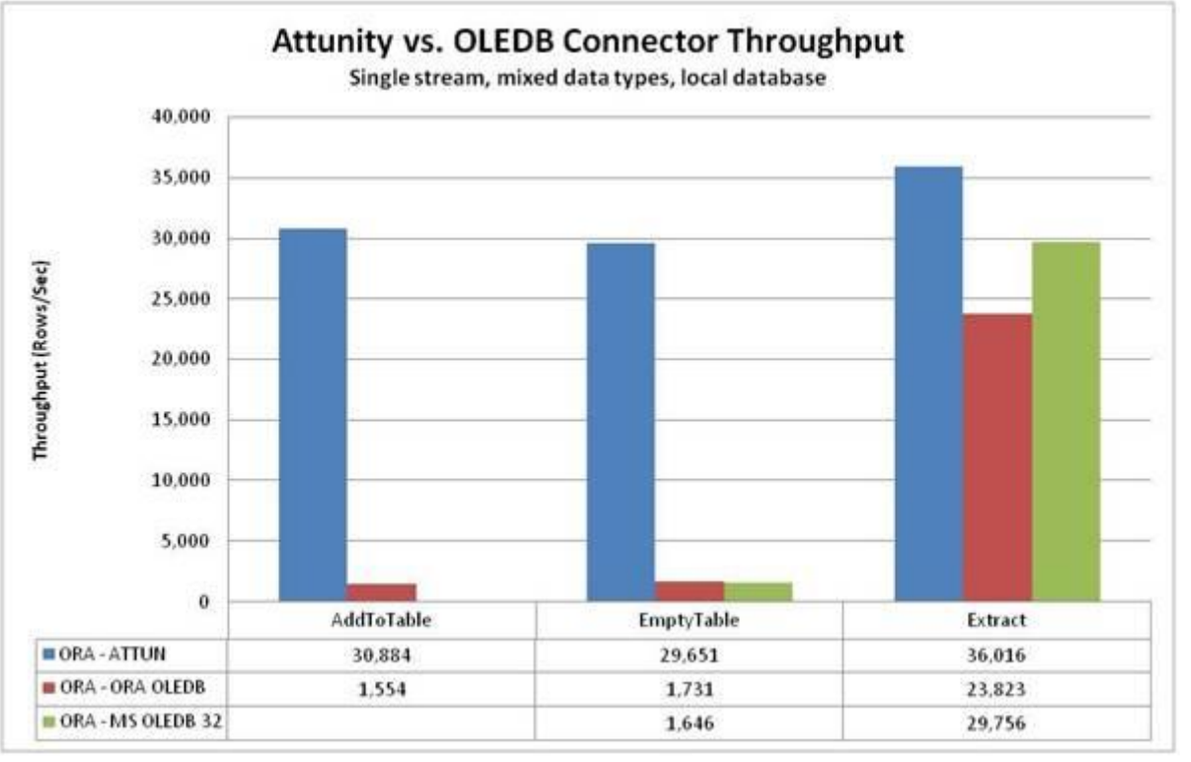
Usar Microsoft Connector for Oracle by Attunity (API Oracle)

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

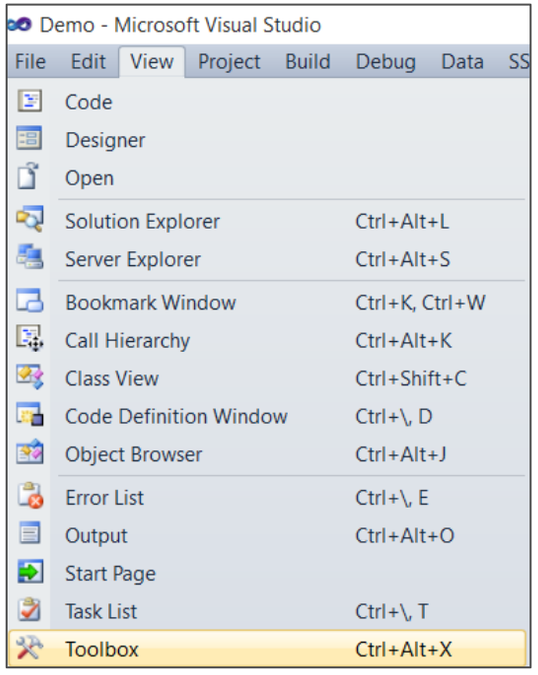


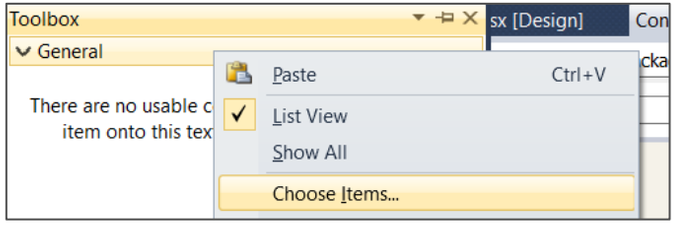
Ref: https://technet.microsoft.com/en-us/library/ee470675(v=sql.100).aspx

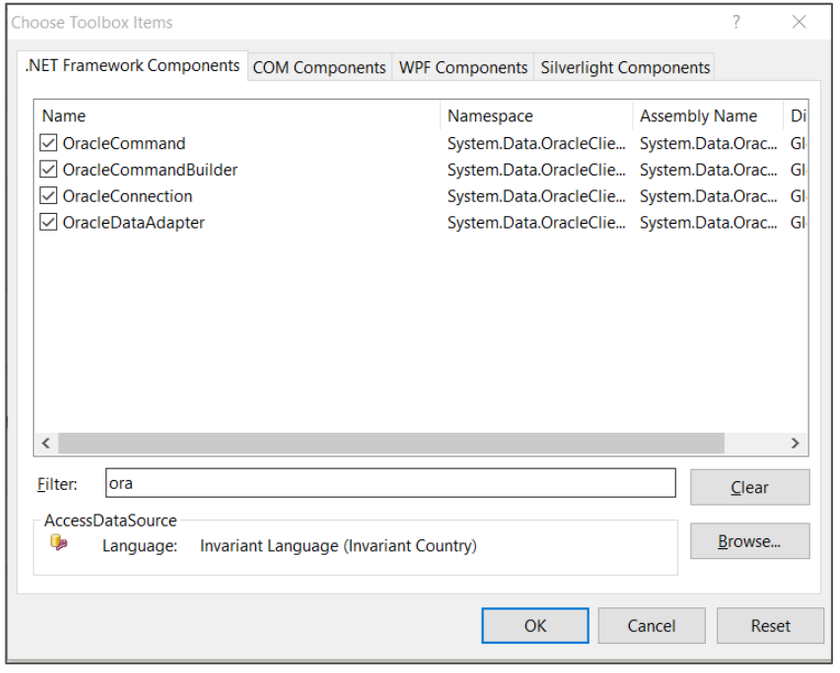
Oracle Attunity - Passos da instalação

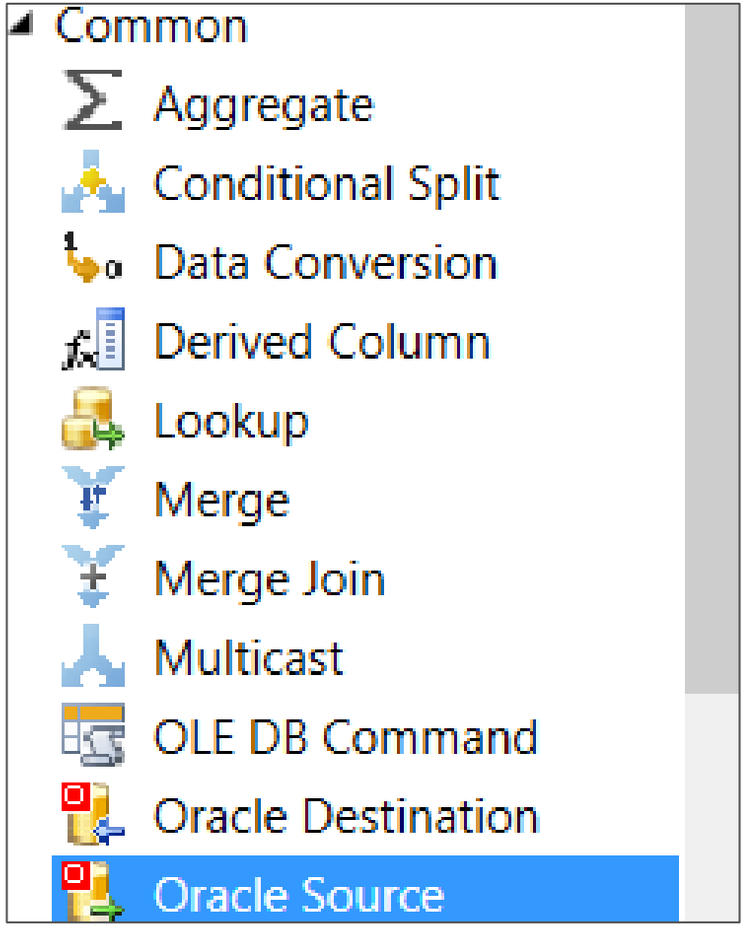
- Instale o Oracle Attunity 64 bit connector
- Instale o client do Oracle 32 bits
  - É necessário para funcionar no Visual Studio que roda em 32 bits
- Instale o client do Oracle 64 bits
  - O Serviço do SSIS roda em modo 64 bits
- Reinicie o computador

Oracle Attunity - Configuração









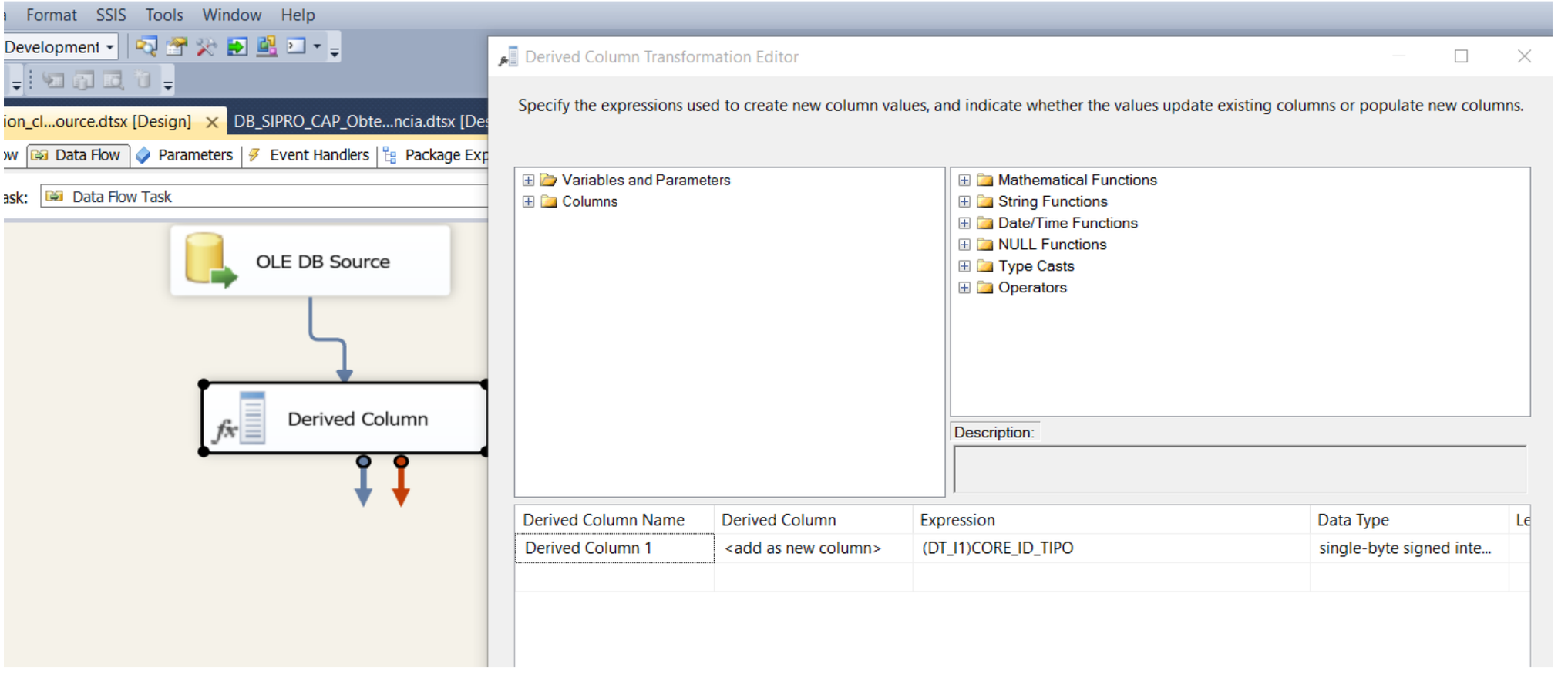
DEMO

Oracle Attunity

Não esqueça...

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

T-SQL vs Coluna Derivada



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

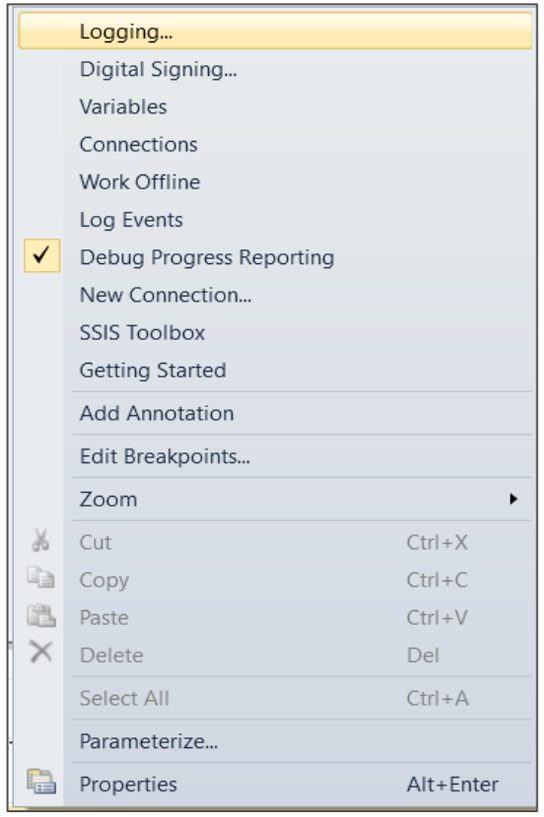
Não esqueça...

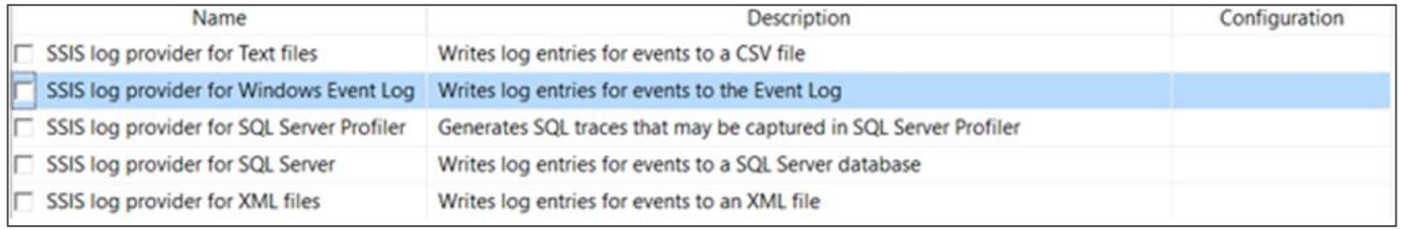
- Se estiver usando Script Componente ou Script Task:

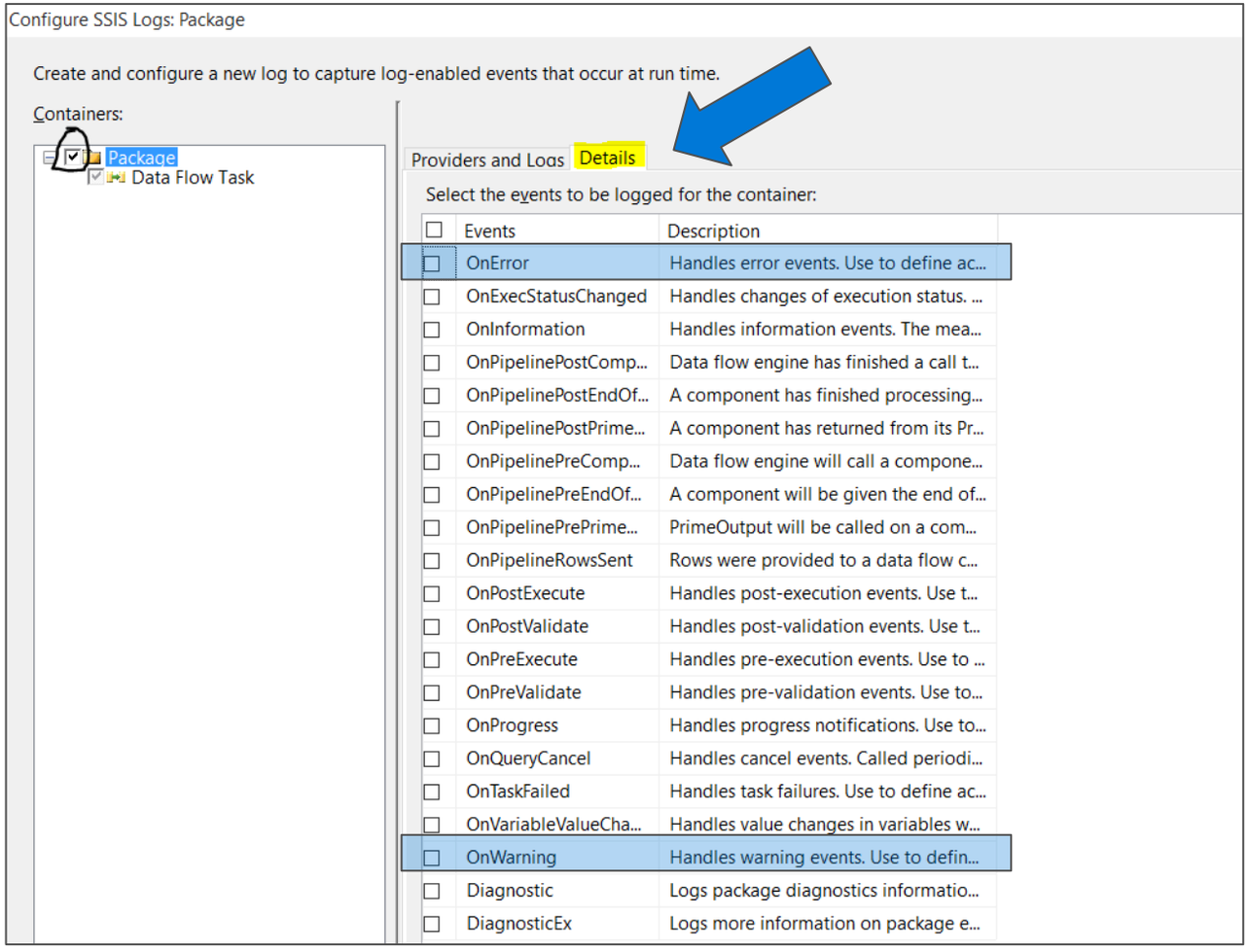
Libere as conexões através do comando release connection ao adquiri-las no pacote por meio do acquire connection.

Use Try catch para tratar erros

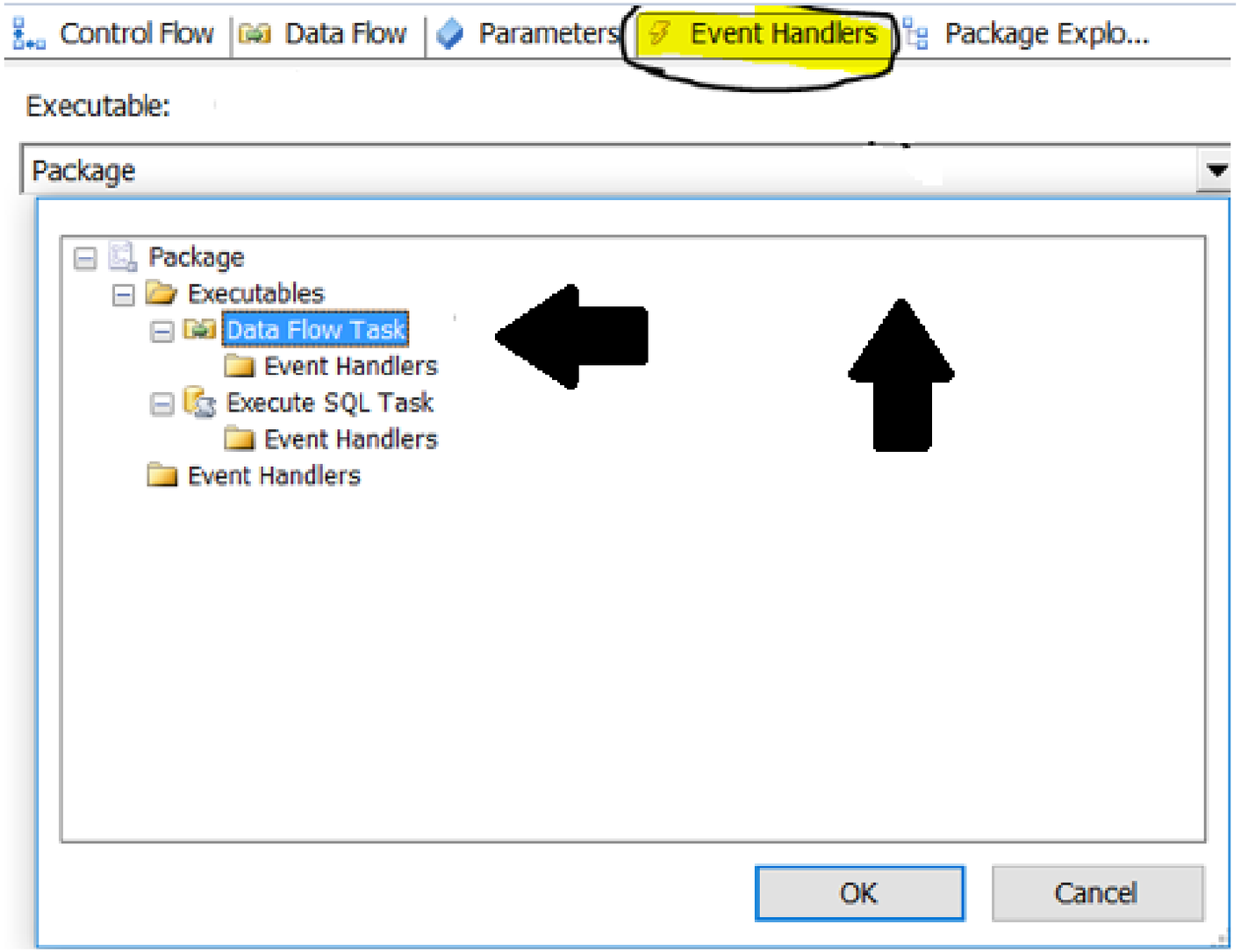
SSIS LOG


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

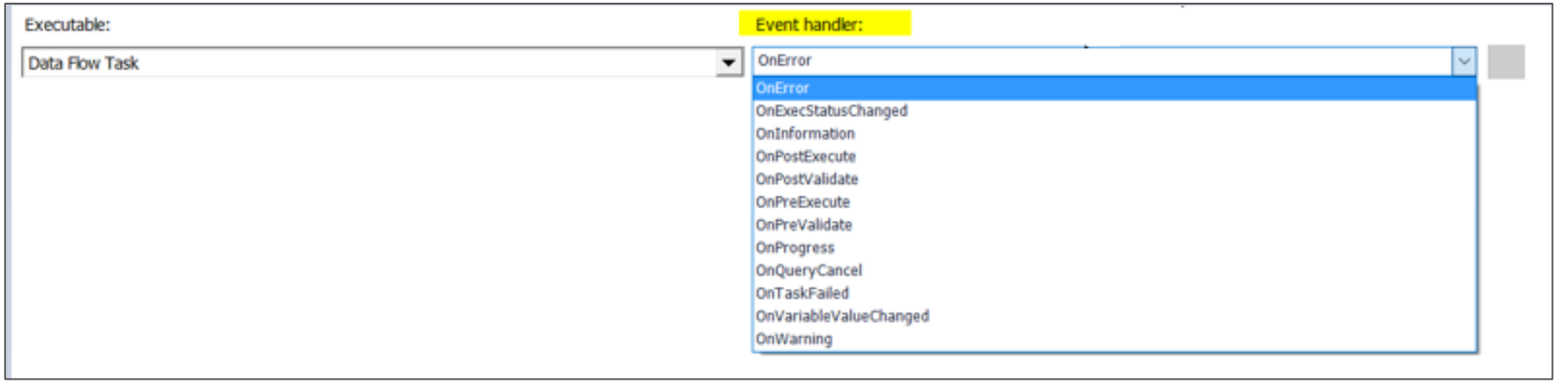

Name
Description
Configuration
SSIS log provider for Text files
Writes log entries for events to a CSV file
SSIS log provider for Windows Event Log
Writes log entries for events to the Event Log
SSIS log provider for SQL Server Profiler
Generates SQL traces that may be captured in SQL Server Profiler
SSIS log provider for SQL Server
Writes log entries for events to a SQL Server database
SSIS log provider for XML files
Writes log entries for events to an XML file

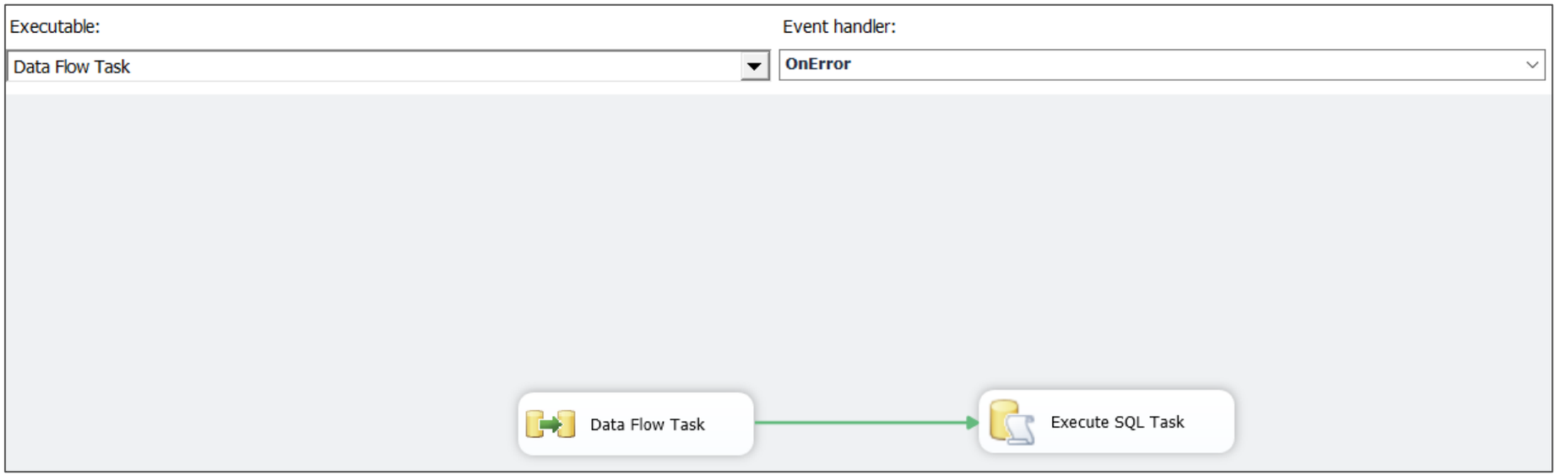

Configure SSIS Logs: Package
Create and configure a new log to capture log-enabled events that occur at run time.
Containers:
Package
Data Flow Task
Providers and Logs
Details
Select the events to be logged for the container:
Events
Description
OnError
Handles error events. Use to define ac...
OnExecStatusChanged
Handles changes of execution status. ...
OnInformation
Handles information events. The mea...
OnPipelinePostComp...
Data flow engine has finished a call t...
OnPipelinePostEndOf...
A component has finished processing...
OnPipelinePostPrime...
A component has returned from its Pr...
OnPipelinePreComp...
Data flow engine will call a compone...
OnPipelinePreEndOf...
A component will be given the end of...
OnPipelinePrePrime...
PrimeOutput will be called on a com...
OnPipelineRowsSent
Rows were provided to a data flow c...
OnPostExecute
Handles post-execution events. Use t...
OnPostValidate
Handles post-validation events. Use t...
OnPreExecute
Handles pre-execution events. Use to ...
OnPreValidate
Handles pre-validation events. Use to...
OnProgress
Handles progress notifications. Use to...
OnQueryCancel
Handles cancel events. Called periodi...
OnTaskFailed
Handles task failures. Use to define ac...
OnVariableValueCha...
Handles value changes in variables w...
OnWarning
Handles warning events. Use to defin...
Diagnostic
Logs package diagnostics informatio...
DiagnosticEx
Logs more information on package e...

Event Handlers



Event Handlers





DEMO

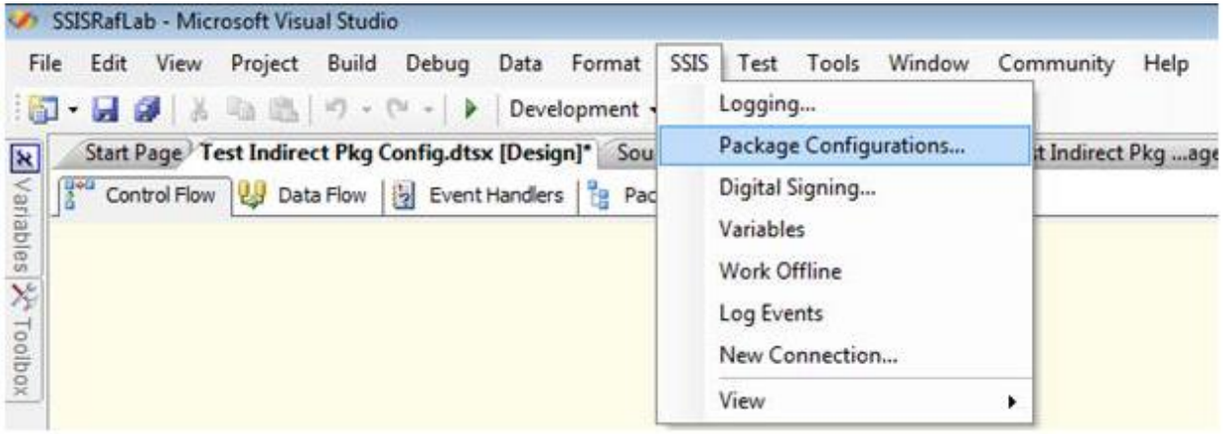
SSIS Log e Event Handlers

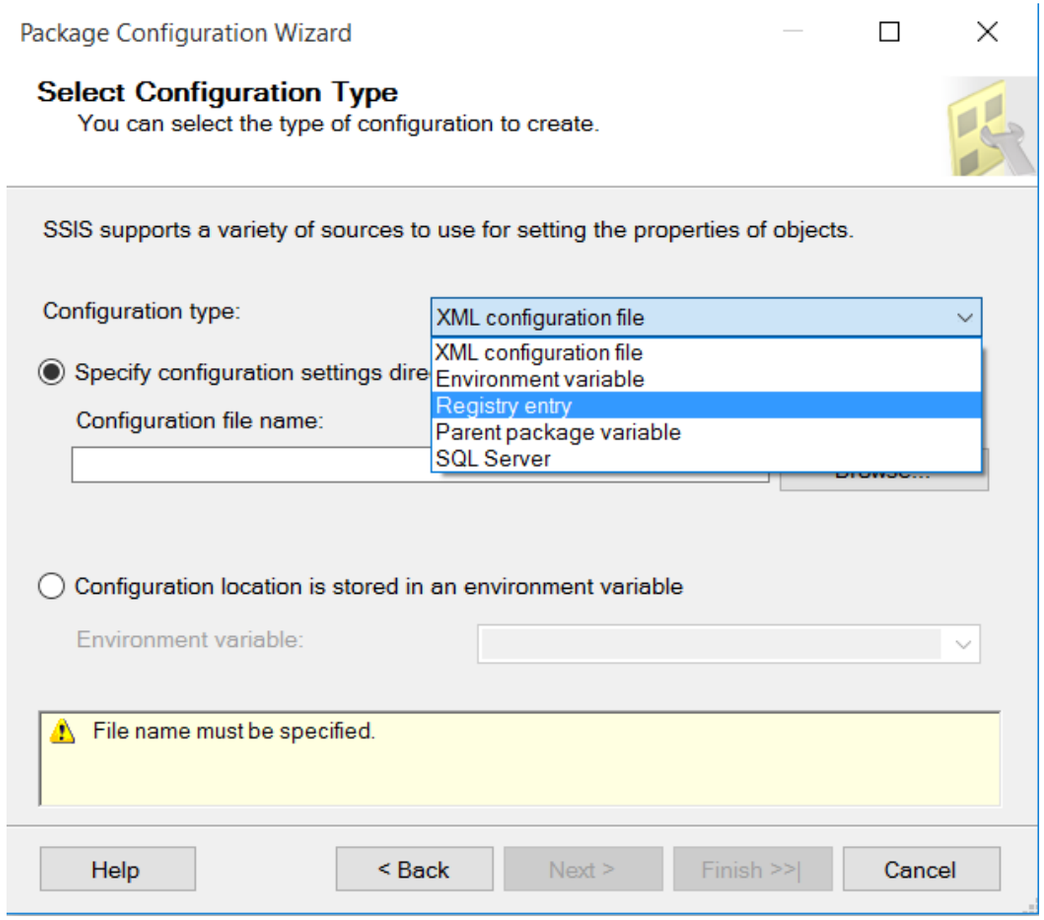
Não esqueça...

- 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 >)

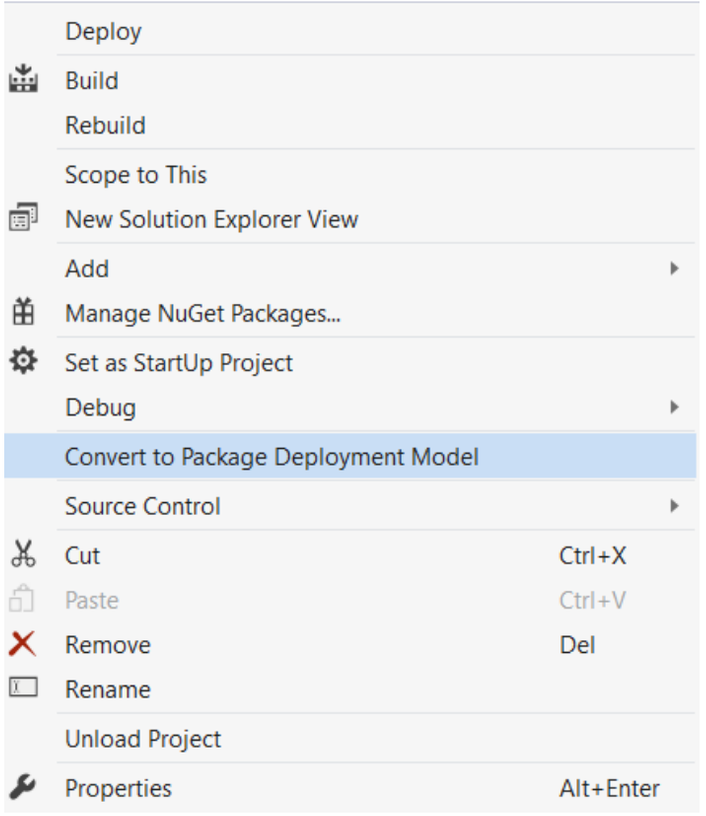
- XML File Configuration

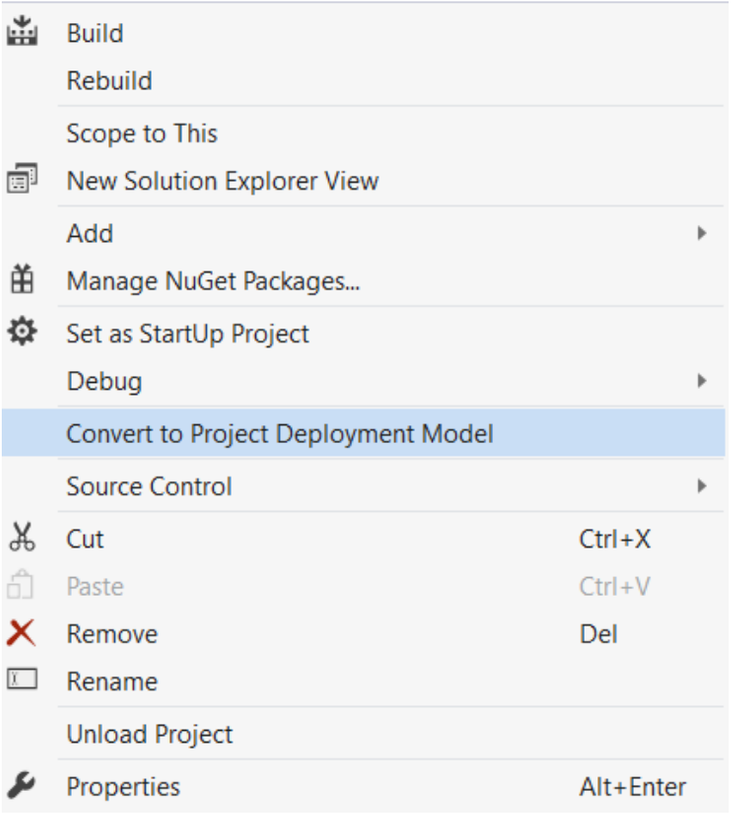
```
<?xml version="1.0" ?>
- <DTSConfiguration>
- <DTSConfigurationHeading>
<DTSConfigurationFileInfo GeneratedBy="MARINER\rsalas" GeneratedFromPackageName="Test XML Config - MSDN Article"
</DTSConfigurationHeading>
- <Configuration ConfiguredType="Property" Path="\Package.Variables[User::ConfigTarget].Properties[Value]"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

- Environment Variables
- Registry Entry
- Parent Package

Configuração do pacote SSIS (2012 ou >)

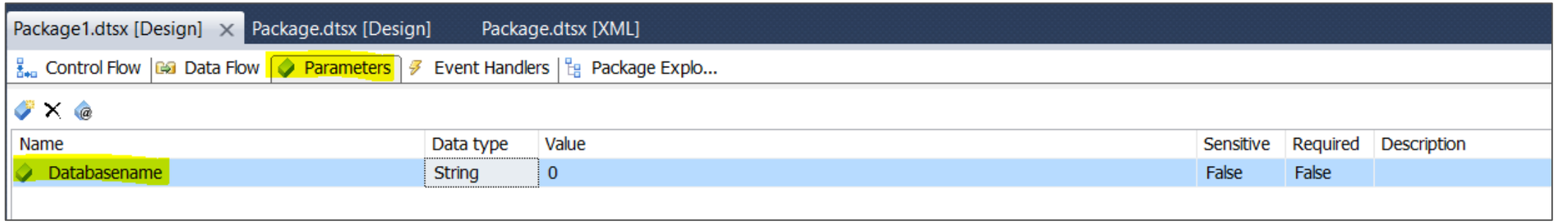
- Deployment Model

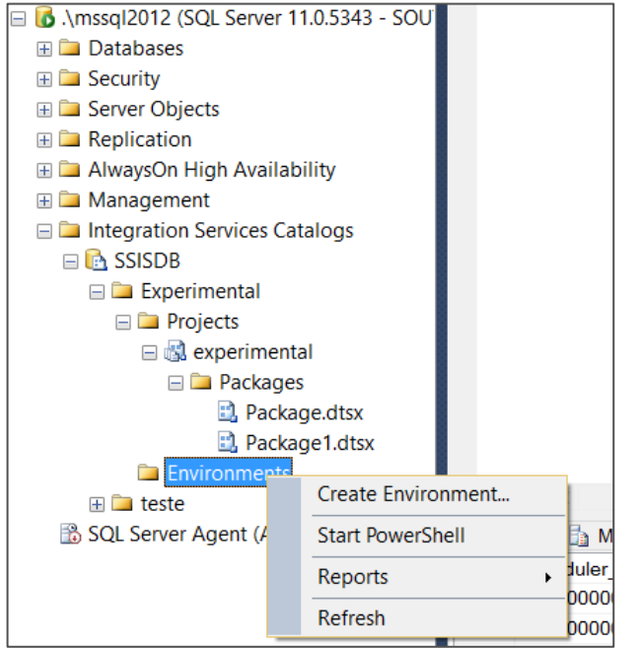


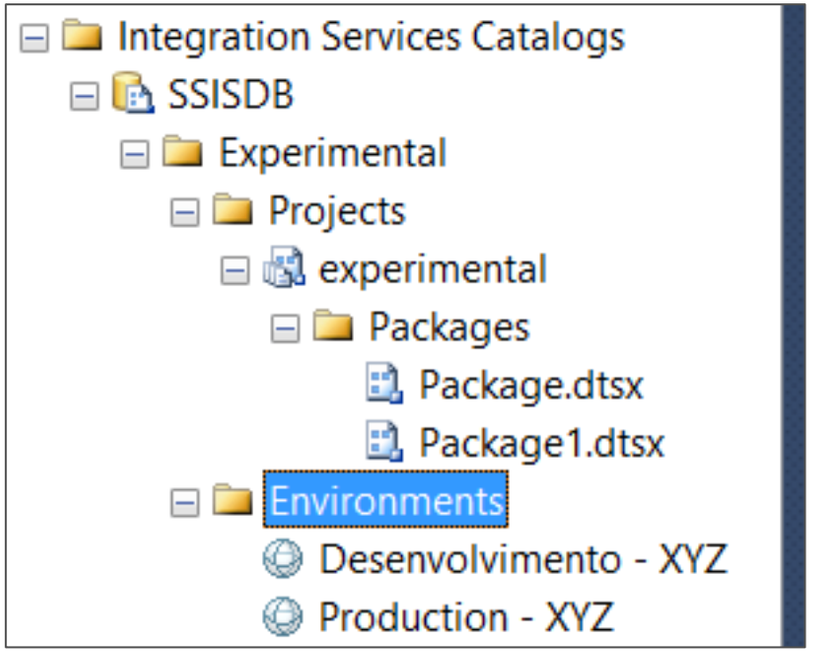


Configuração do pacote SSIS (2012 ou >)

- Parameters e Enviroments

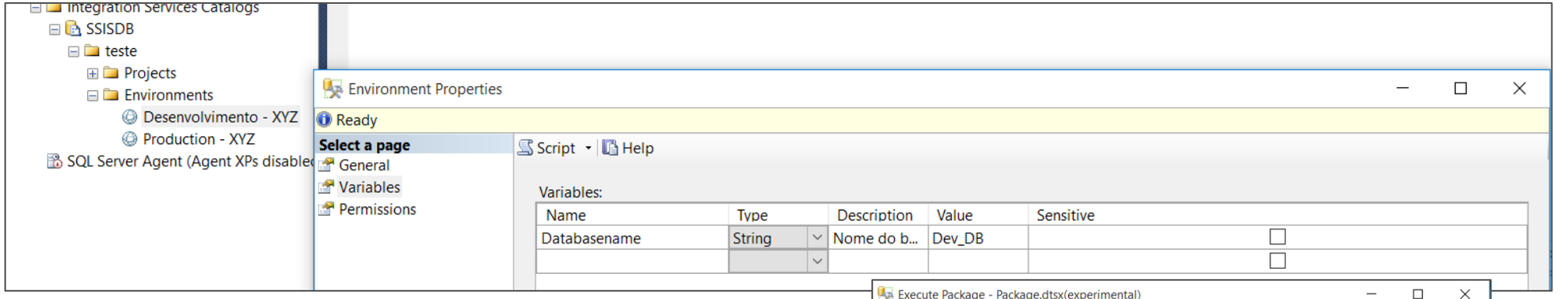


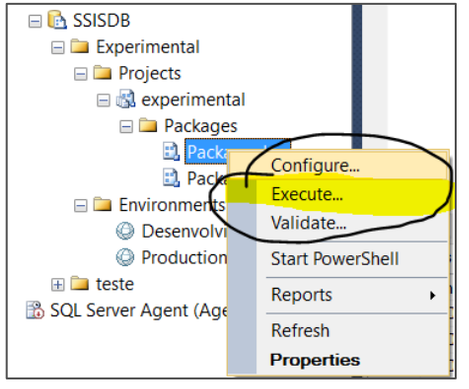


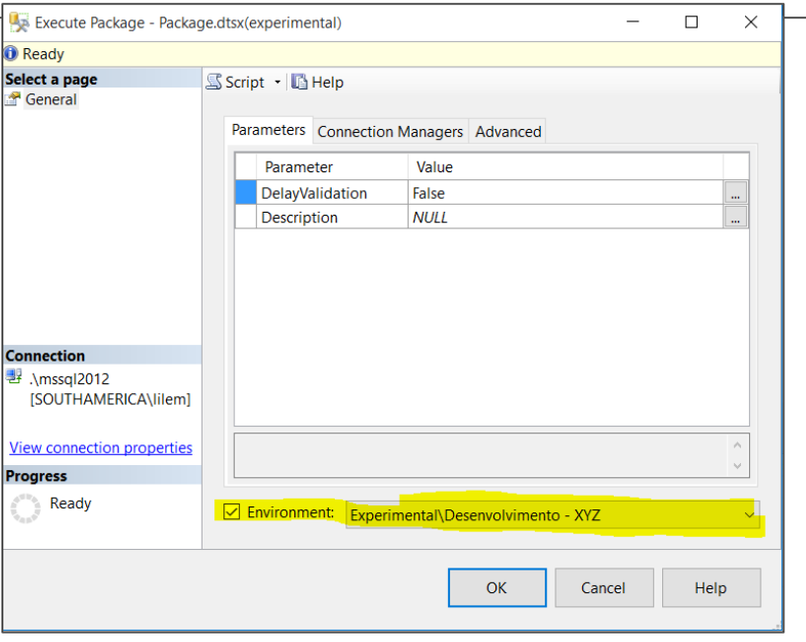


Configuração do pacote SSIS (2012 ou >)

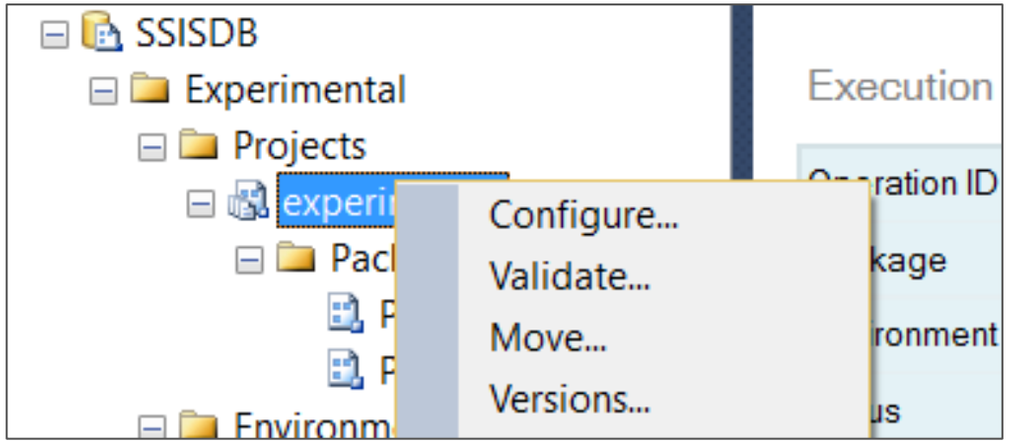
- Parameters e Enviroments

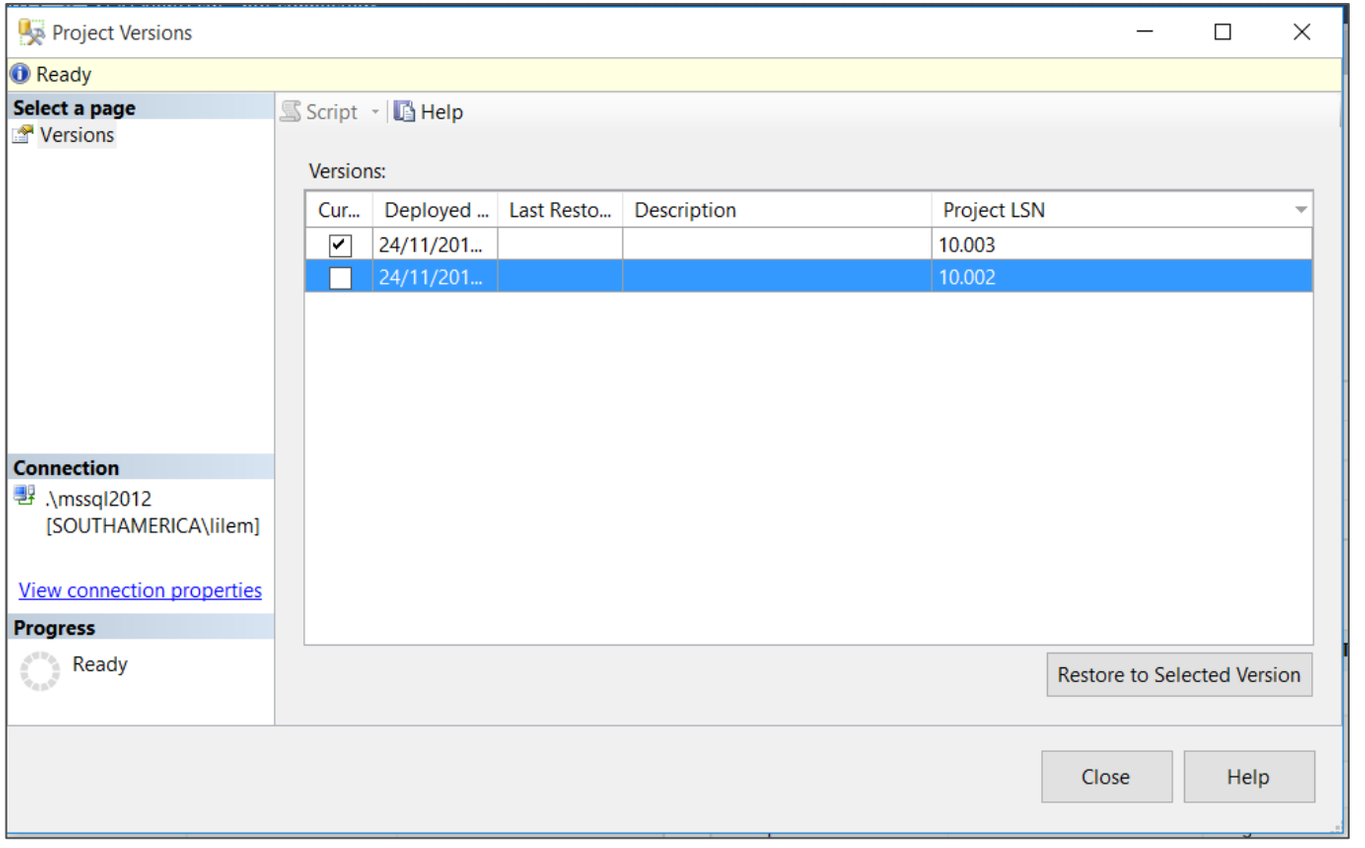






Versionamento





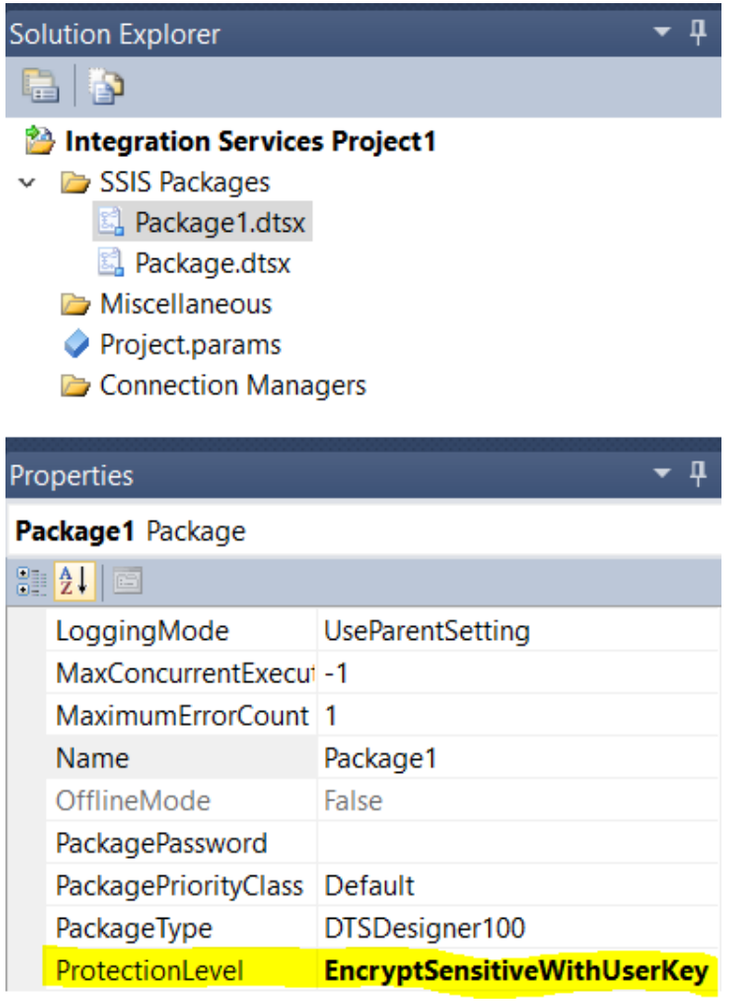
DEMO

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

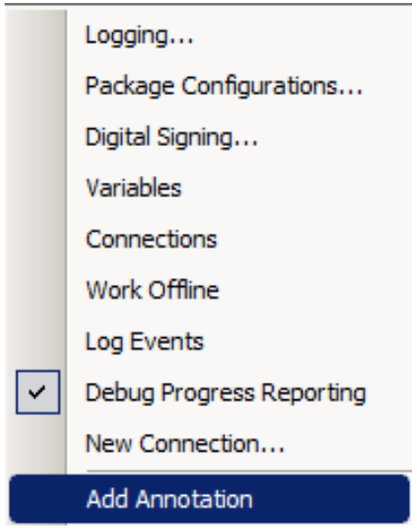
Protection Level

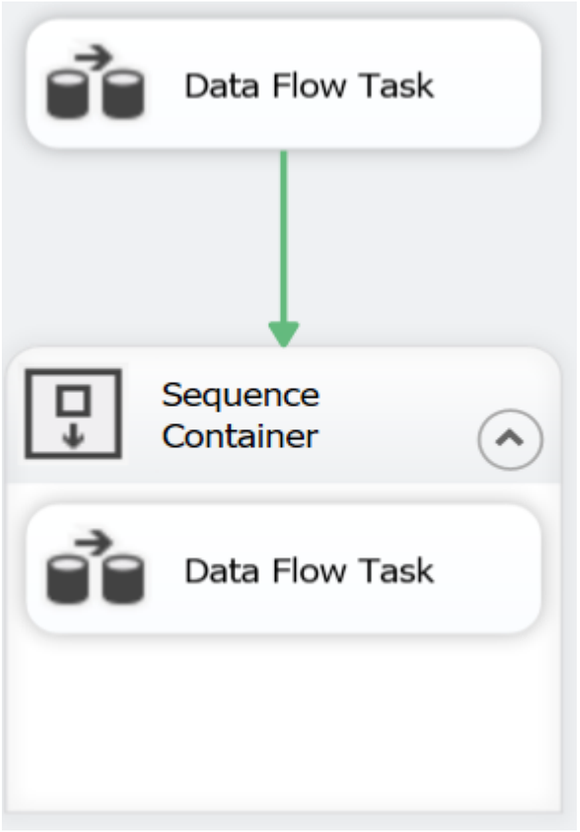
Não esqueça...

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

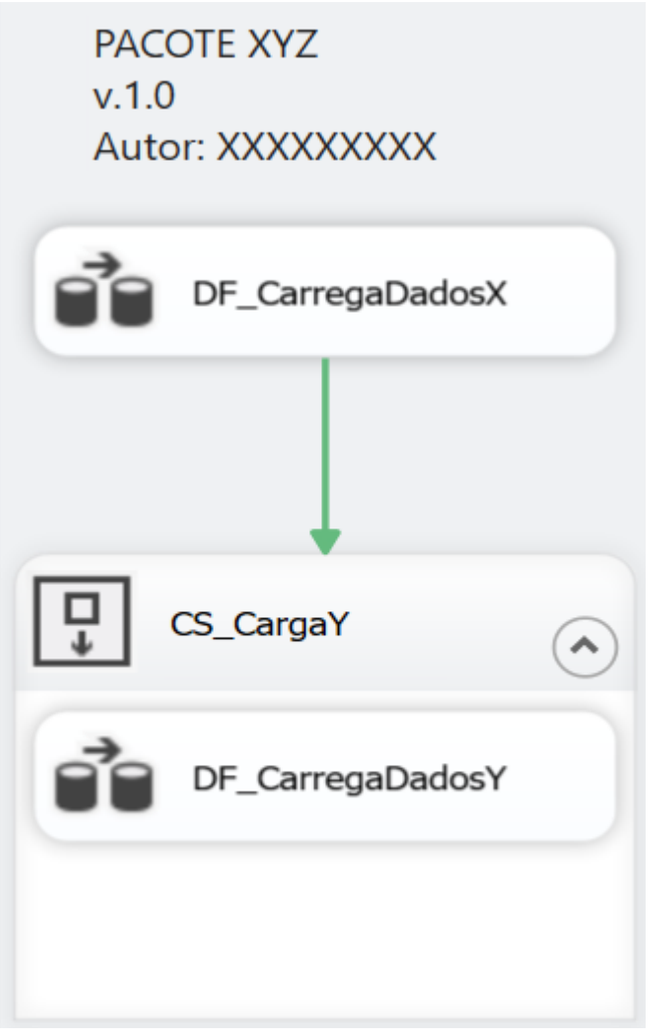
Adeque o protection level para o ambiente de produção

Documentação





VS



DEMO
Documentação

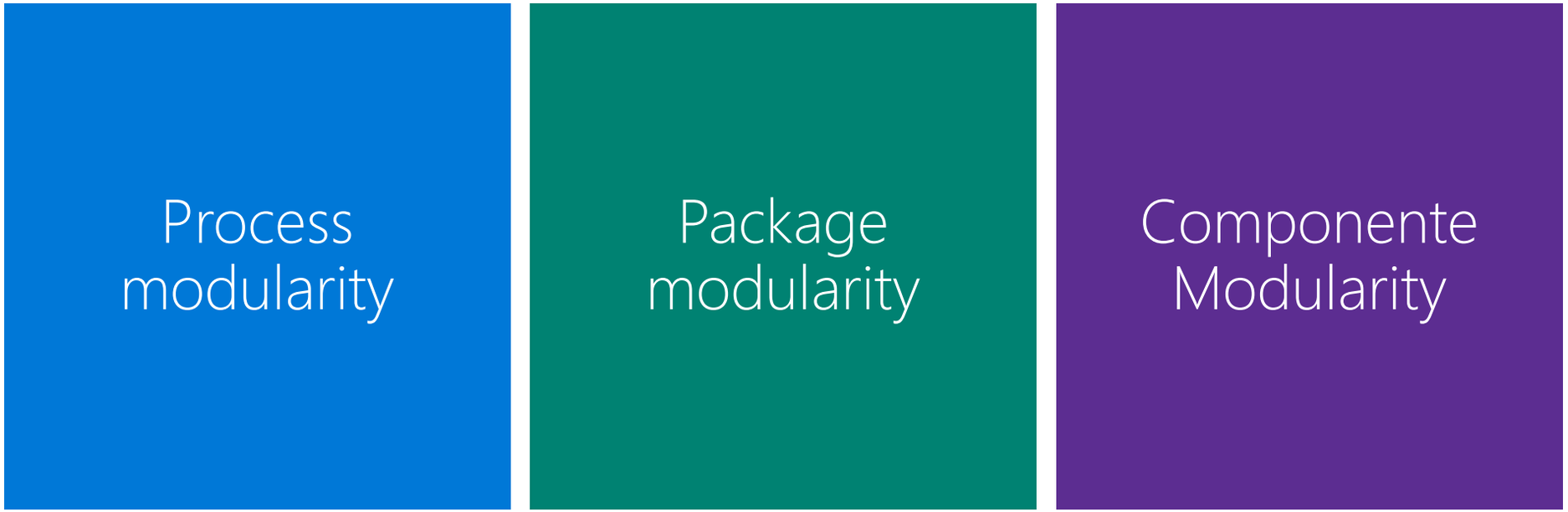
Não esqueça...

Comente o código

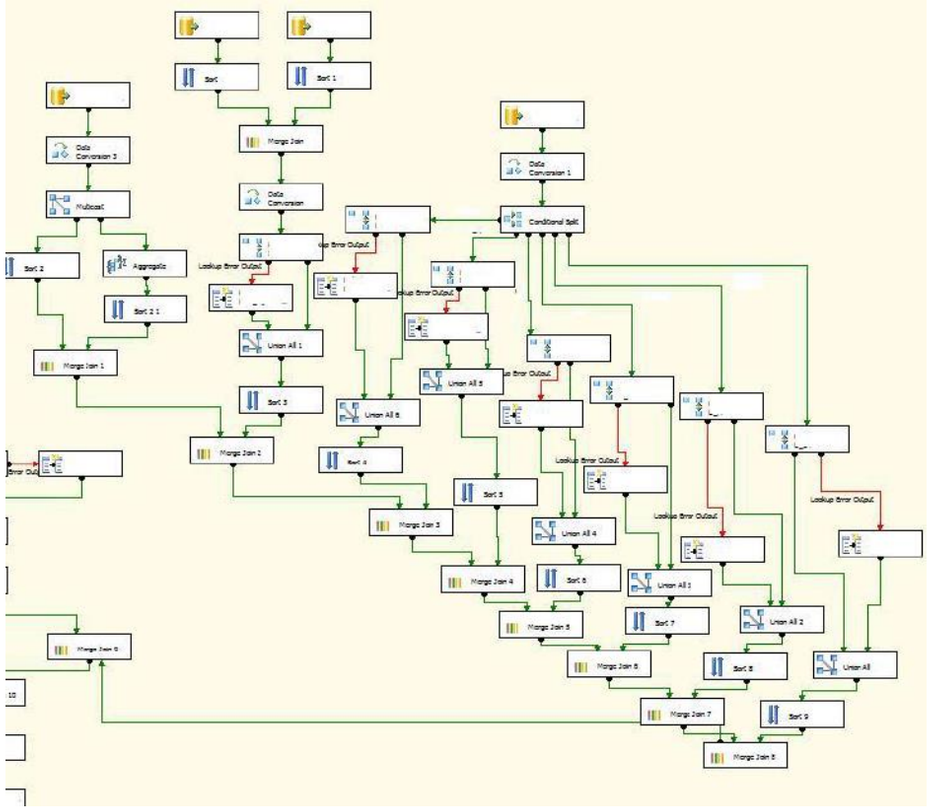
Documente as tasks

Outras pessoas irão dar manutenção neste código

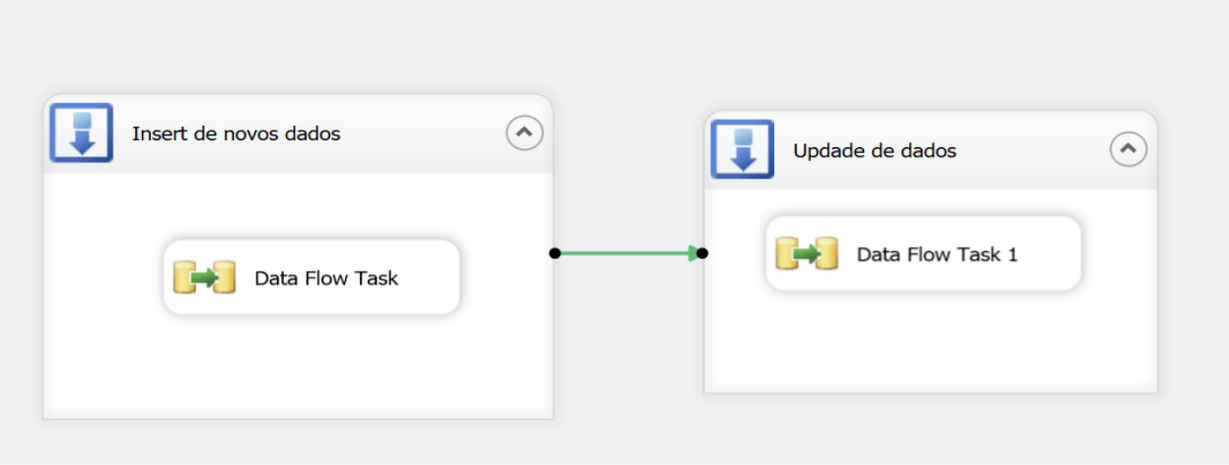
Design


Process modularity
Package modularity
Componente Modularity

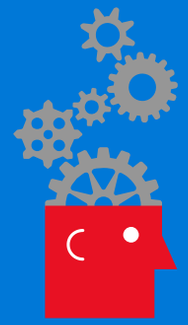
Design



VS


Insert de novos dados
Data Flow Task
Updade de dados
Data Flow Task 1

Não esqueça...

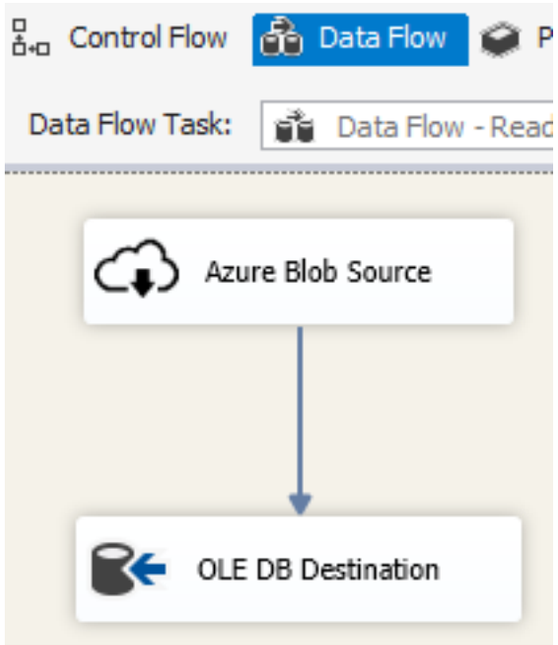


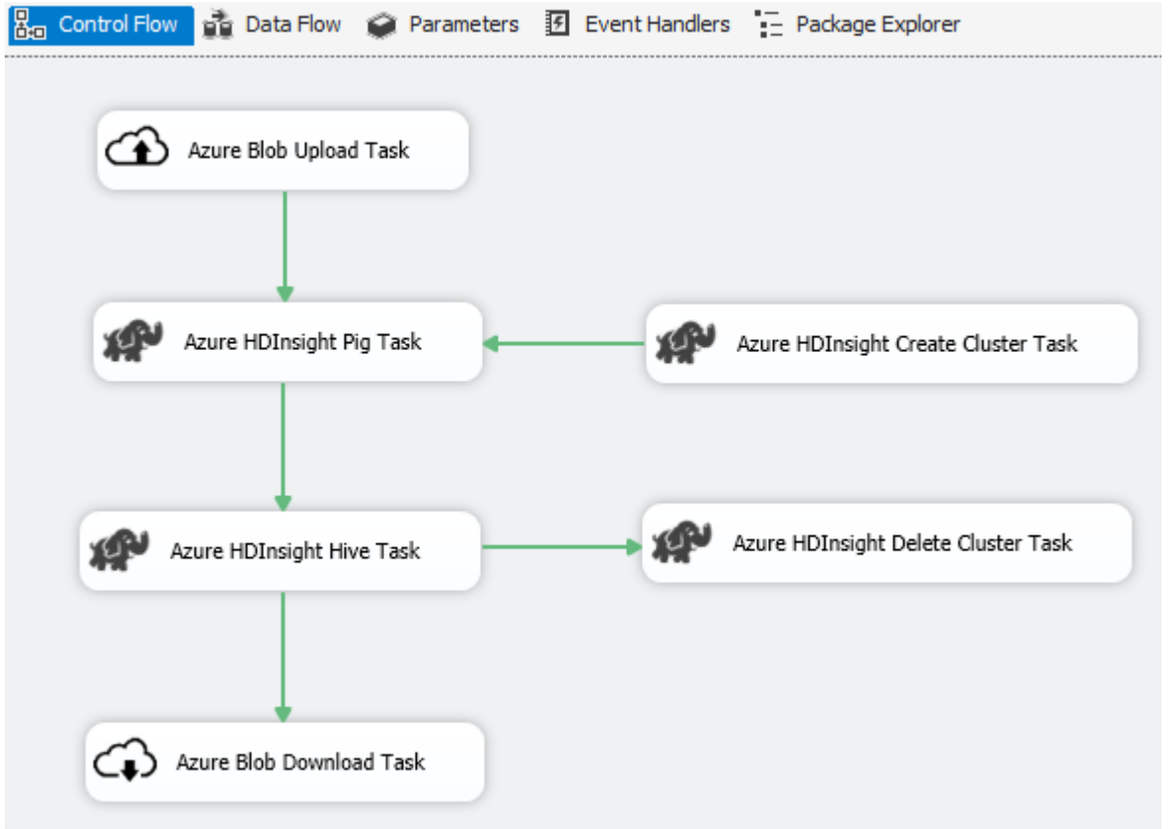
Simplifique a lógica do pacote

Agrupe tasks através do uso de containers

SSIS e Azure

SSIS Azure Feature Pack





DEMO

SSIS Azure Pack