



MANUAL DO AMBIENTE DE BI

Criado em 02/05/2018

Ádan Santos Yoshimoto

Atualizado em 04/01/2022

Edirlan Carlos Soares

OBJETIVO

Descrever o funcionamento e características do ambiente de BI, ferramentas de Integração de dados, as formas de armazenamento e ferramentas de Visualização de Dados como Power BI e Reporting Services.

Também é abordado as formas de atualização e formas de capturar os dados.

PÚBLICO

Arquiteto de dados

ARQUITETURA

A arquitetura representada abaixo na imagem é bem típica da literatura sobre ETL ou ELT, onde favorece a centralização de todas as informações relevantes para a empresa no processo de tomada de decisão.

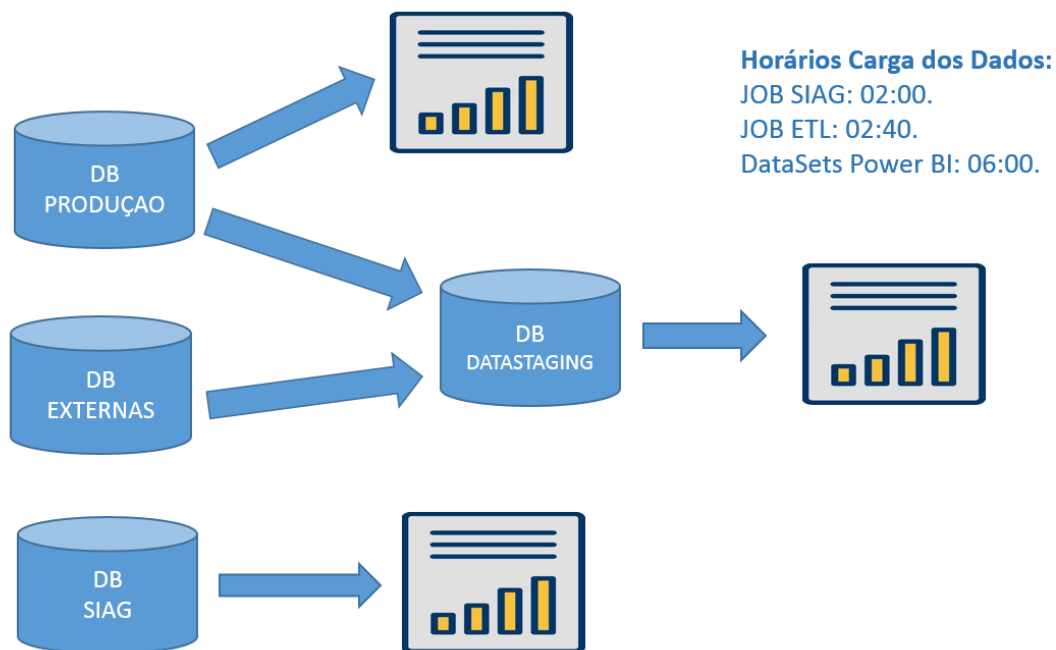
Esta centralização minimiza as chances de divergências de conceitos e de números diferentes de um setor para o outro.

Banco de Dados: MS SQL Server 2016

Instancia: VSERVERSQLSEC

Bases:

- dbCIEE_DataStaging



MODELO DE DADOS

A estrutura de dados é baseada no data lake, para agilizar e acelerar o processamento de dados, não foi criada uma estrutura de datawarehouse, como o dw amarra muito os dados, foi definido manter um data lake para quando for utilizar as ferramentas de dataviz, o processamento ser mais ágil.

O data lake é um repositório que centraliza e armazena todos os tipos de dados gerados pela e para a empresa. Eles são depositados ali ainda em estado bruto, sem o processamento e análise e até mesmo sem uma governança.

A ideia é manter na organização dados que podem ser estrategicamente úteis, mesmo que eles, na realidade, não sejam requeridos em nenhum momento posterior.

DATASETS POWER BI

Armazenado no "TFS\CIEEPR.BI\Repositorio\producao\Power BI\DataSets

Também localizado em uma tabela no database dbCIEE_DataStaging : DocBI.DataSetsPowerBI

A Carga é Feita no Integration Services (SSIS).

Projeto: TFS\CIEEPR.BI\Producao\Projetos\Data Mining\Docs

DATA STAGING

Contém os dados extraídos diretamente dos sistemas de origens, a ideia é obter os dados de forma que traga o mínimo de impacto no desempenho do sistema de origem. Geralmente esta base fica no mesmo servidor da origem dos dados, desta maneira evita-se o uso de tráfego pela rede através de linked servers.

PROCEDIMENTOS CARGA DE DADOS

Base DataStaging

A Carga é Feita no Integration Services (SSIS).

Projeto: TFS\CIEEPR.BI\Producao\Projetos\Data Mining\ETL

A Documentação do projeto está toda descrita em cada etapa do mesmo.

ESQUEMAS

Foram definidos Schemas dentro da base para separar os processos de negócio das informações de controle da própria base.

Schema dbo – Armazena os dados referente aos processos de negócio do CIEE-PR.

Schema DocBI – Armazena os dados referentes a meta dados e controles internos e documentação.

POWER BI

Ferramenta de visualização e exploração de dados da Microsoft.

A licença adquirida é a conta PRO.

O ponto positivo do licenciamento é o seu baixo custo, como ponto negativo podemos citar que a funcionalidade de publicação na Web torna o painel público, ou seja, se alguém externamente obter o endereço URL, poderá ver os dados.

GATEWAY

O gateway configura a autenticação do serviço do Power BI com a base de dados do CIEE.

Atualização:

GATEWAY:

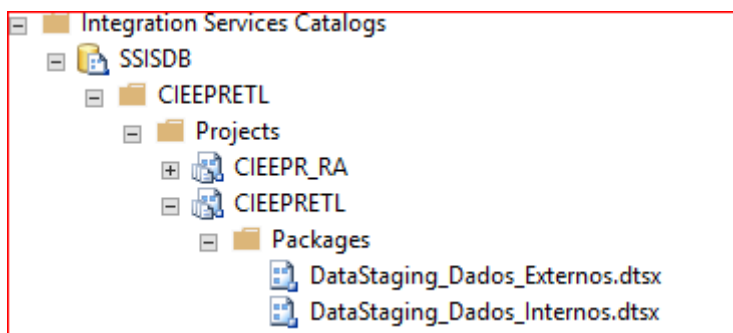


AGENDAMENTOS

O agendamento da execução da carga diária está configurado para ser executado de Segunda a Sábado as 02:40 hrs na instancia VSERVERSQLSEC.

Arquivos Packages para carregar os dados: Estancia VSERVERSQLSEC.

Processo de Atualização feito no projeto SSIS via Deploy.



Geral:

Job Properties - Carga Dados - BI

Select a page

- General
- Steps
- Schedules
- Notifications
- Targets

Script Help

Name: Carga Dados - BI

Owner: CIEEPR\sql_user

Category: [Uncategorized (Local)]

Description: Carga de Dados DataStaging e DataWarehouse.

Connection

Server: SQLCLUSTER

Connection: CIEEPR\vedirfan.soares

[View connection properties](#)

Progress

Ready

☒ Enabled

Source:

Created: 30/10/2018 15:15:40

Last modified: 30/09/2021 08:42:37

Last executed: 08/11/2021 02:40:00

[View Job History](#)

OK Cancel

Passos do Agendamento:

Select a page

General

Steps

Schedules


Notifications

Targets


Connection

Server:
SQLCLUSTER

Connection:
CIEEPR\edirlan.soares

 [View connection properties](#)

Progress

 Ready

Script ? Help

Job step list:

St...	Name	Type	On Success	On Failure
1	DataStaging_Dados_Internos	SQL Serv...	Go to ste...	Quit the job...
2	Crescimento Base de dados BI (DataStaging ...	Transact-...	Quit the jo...	Quit the job...

Move step:

↑

↓

Start step:
1:DataStaging_Dados_Internos

New...

Insert...

View

Delete

OK

Cancel

Horário do Agendamento:

Job Schedule Properties - Carga de Dados - BI

Name: Jobs in Schedule

Schedule type: ☒ Enabled

One-time occurrence

Date: Time:

Frequency

Occurs:

Recurs every: day(s)

Daily frequency

☒ Occurs once at:

☐ Occurs every: hour(s) Starting at: Ending at:

Duration

Start date:

☐ End date:

☒ No end date:

Summary

Description:

MONITORAMENTO

Monitorar diariamente a execução dos Jobs relativos ao BI.

Os logs de execução do job do SQL podem ajudar na solução de algum possível problema.

As tabelas de controle de controle também podem ajudar na solução do problema, nelas há os campos de data de início e final da execução.

Log file summary: No filter applied

Date	Step ID	Server	Job Name	Step Name	Notifications	Message	Duration
19/11/2018 02:40:00	✓	SQLCLUSTER	Carga Dados - BI			The job succeeded. The Job was invoked by Schedule 405 (Carga de Dados - BI). The last step to run was step 3 (Fatos).	00:09:39
18/11/2018 02:40:00	✓	SQLCLUSTER	Carga Dados - BI			The job succeeded. The Job was invoked by Schedule 405 (Carga de Dados - BI). The last step to run was step 3 (Fatos).	00:11:16
17/11/2018 02:40:00	✓	SQLCLUSTER	Carga Dados - BI			The job succeeded. The Job was invoked by Schedule 405 (Carga de Dados - BI). The last step to run was step 3 (Fatos).	00:10:05
16/11/2018 02:40:00	✓	SQLCLUSTER	Carga Dados - BI			The job succeeded. The Job was invoked by Schedule 405 (Carga de Dados - BI). The last step to run was step 3 (Fatos).	00:09:52
15/11/2018 02:40:00	✓	SQLCLUSTER	Carga Dados - BI			The job succeeded. The Job was invoked by Schedule 405 (Carga de Dados - BI). The last step to run was step 3 (Fatos).	00:09:49
14/11/2018 03:40:00	✓	SQLCLUSTER	Carga Dados - BI			The job succeeded. The Job was invoked by Schedule 405 (Carga de Dados - BI). The last step to run was step 3 (Fatos).	00:12:12

SENHAS E INFORMAÇÕES

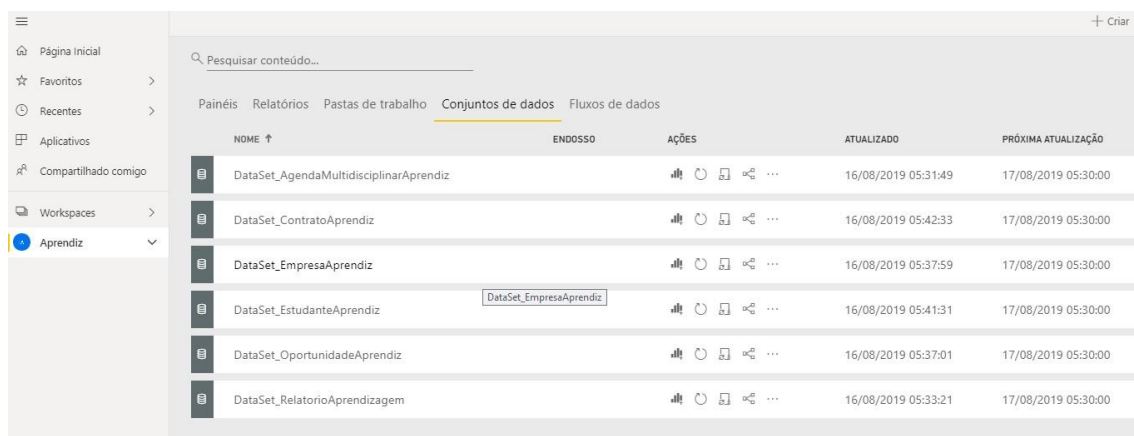
Fica localizado no C:\TFS\CIEEPR.BI\Documentação\Processos\Informações e Senhas BI.txt

























POSSIVEIS FALHAS

- Shedule do Job Carga Dados – BI.

- DataSets

Que Ficam localizado no WorkSpace de Cada Categoria.



Painéis	Relatórios	Pastas de trabalho	Conjuntos de dados	Fluxos de dados
NOME ↑	ENDOSSO	AÇÕES	ATUALIZADO	PRÓXIMA ATUALIZAÇÃO
DataSet_AgendaMultidisciplinarAprendiz		   	16/08/2019 05:31:49	17/08/2019 05:30:00
DataSet_ContratoAprendiz		   	16/08/2019 05:42:33	17/08/2019 05:30:00
DataSet_EmpresaAprendiz		   	16/08/2019 05:37:59	17/08/2019 05:30:00
DataSet_EstudanteAprendiz	DataSet_EmpresaAprendiz	   	16/08/2019 05:41:31	17/08/2019 05:30:00
DataSet_OportunidadeAprendiz		   	16/08/2019 05:37:01	17/08/2019 05:30:00
DataSet_RelatorioAprendizagem		   	16/08/2019 05:33:21	17/08/2019 05:30:00

-Gateway de Dados

Pode acontecer de ficar Off-line (VserverBI Delisgado).



ADICIONAR FONTE DE DADOS

CLUSTERS DE GATEWAY

- ✓ CIEE_GatewayPadrao
 - Diretório IBGE/CAGED
 - dbCIEE_SCE_Temp
 - dbCIEE_RFB
 - dbCIEE_DataWarehouse
 - dbCIEE_SCE
 - dbCIEE_Trace

Testar todas as conexões

Configurações do Cluster de Gateway Administradores

✓ Online: tudo pronto.

Nome do Cluster de Gateway

CIEE_GatewayPadrao

Departamento

Descrição

Informações de Contato

powerbi@cieepr.org.br

☒ Permitir que as fontes de dados de nuvem do usuário sejam atualizadas por meio deste cluster de gateway. Essas fontes de dados de nuvem não precisam ser configuradas neste cluster de gateway. [Saiba mais](#)

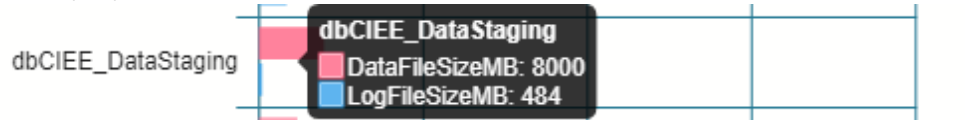
☒ Permitir que os conectores de dados personalizados do usuário atualizem por este cluster do gateway (versão prévia). [Saiba mais](#)

☒ Distribua solicitações em todos os gateways ativos neste cluster. [Saiba mais](#)

Aplicar Descartar

ARMAZENAMENTO

Data:19/11/2018



Pode ser verificado através da ferramenta Data Tools ou através do report de governança dos dados, o link segue a baixo.

Relatório detalhado de Governança de Dados:

<https://app.powerbi.com/view?r=eyJrljoiOWVINTNmODctNDRmNS00OWIzLWUwY2ItYjI4YzBhZmU0ZmRlIiwidCI6ImZmOTU4YzEzLTljZmQtNDAzZC1iInja0LWQxNWQ0MTQ1OWE5NCJ9>

Links Manual e Projetos que vai no relatório de Governança de dados estão no Sharepoint:

<https://cieepr.sharepoint.com/portalprocessos/Indicadores/Forms/AllItems.aspx>

FERRAMENTAS BI:

- SQL SERVER
- POWERBI
- INTEGRATION SERVICES
- REPORT SERVICES
- POWER ARCHITECT
- AZURE APIS

FONTE DE DADOS:

- BASE PRODUÇÃO
 - Base Principal.
- SIAG
 - Base Histórica de dados.
- DataStaging
 - Base BI.