

什么是 Azure Synapse Analytics?

Azure Synapse 是一项企业分析服务，汇集了企业数据仓库中所用 SQL 技术的精华、用于大数据的 Spark 技术、用于数据集成和 ETL/ELT 的 Pipelines，以及与其他 Azure 服务（Power BI、CosmosDB 和 AzureML）的深度集成。

- T-SQL 的分布式查询系统，行业领先的SQL行业标准的 `Apache Spark`
- 数据湖
- 内置数据集成
- 统一的监控、管理和安全性

[通过在 Azure Synapse Analytics 中 \(以前的 SQL DW\) 升级专用 SQL 池来优化性能](#)

1. 采用@wanlu 之前 spike 的从还原点升级到 Gen2

[链接](#)

按照步骤升级在 QA 完成升级

注意：在 IP 全部复制进去之后，还会会有那么一小会时间 `cannot access`，等一会就可以了，为什么可以目前还未知，早上操作的时候是这样的，什么操作也没有做，只是过了一会就连上了。

2. 检查需要修改的代码

- 代码：onedata_platform/python_script/util/db_connection/config.py
 - ETL 链接DW `DATA_PLATFORM_URI`
 - BLOB `BLOB_CONFIG` ?????
- CICD：不需要修改，SP 的更新是在 azkanban中完成的，调度python 脚本，更新的时候调用了 config.py 中的配置。运行在 `prepare_env.flow` 依赖中
- azkanban，不需要修改，azkanban主要是调度代码，DW 的连接都在代码中做修改，理解 azkanban不需要该配置

本地测试新DW

3. Gen2 的部分新功能

- 开启查询缓存
- 开启结果集缓存
- 查看一个复杂查询是否能够命中缓存

```
-- 为数据库启用查询存储
ALTER DATABASE [onedataqagen2]
SET QUERY_STORE = ON;
```

```

SELECT name, is_query_store_on
FROM sys.databases;

-- Run this command when connecting to the MASTER databas
-- 需要先链接到 master 数据库进行设置
ALTER DATABASE [onedaqagen2]
SET RESULT_SET_CACHING ON;

SELECT name, is_result_set_caching_on
FROM sys.databases;

select top 10 * from raw.t_dim_distributor_database_new

SELECT request_id, command, submit_time,result_cache_hit FROM
sys.dm_pdw_exec_requests order by submit_time desc
-- 可以查看到上面的查询是缓存命中了的, (result_cache_hit > 0)
-- QID5753948,select top 10 * from raw.t_dim_distributor_database_new,2021-04-
21 03:26:27.120,1

```

4. 测试代码功能完备,运行代码, 是否正常工作, 写入到新 dw

qa 环境执行etl, 查看是否运行正常, 新数据库中是否写入数据

5. 添加用户

添加前需要先删除从别的库中同步过来的用户

LoaderRC60、analyst、readonlyreport、reportreader、performancetester

这些事从 Dev 库中同步过来的用户

问题还是挺多。

```

SELECT name, sid FROM sys.sql_logins

-- 登陆master
CREATE LOGIN LoaderRC80 WITH PASSWORD = 'Sh3lldN5&Z5';
SELECT name, sid FROM sys.sql_logins

select * from sys.database_principals order by type_desc desc
-- dbo,SQL_USER
-- guest,SQL_USER
-- INFORMATION_SCHEMA,SQL_USER
-- sys,SQL_USER
-- fly.hong,SQL_USER

--
-- CREATE USER LoaderRC80 FOR LOGIN LoaderRC80;

```

```

-- GRANT CONTROL ON DATABASE::[onedataqagen2] to LoaderRC80;
--
-- CREATE USER analyst FOR LOGIN analyst;
-- GRANT CONTROL ON DATABASE::[onedataqagen2] to analyst;

select * from sys.database_principals order by type_desc desc

-- 这三个可以直接删除
DROP USER readonlyreport
DROP USER reportreader
DROP USER performancetester

select * from sys.database_principals order by type_desc desc

DROP USER LoaderRC60
DROP USER analyst
-- The database principal owns a schema in the database, and cannot be dropped.
-- 找到这些用户的拥有的 shcema 权限, 并改变这些schema 的权限到 dbo 用户, 它是sqlSeerver
的默认用户

SELECT s.name
FROM sys.schemas s
WHERE s.principal_id = DATABASE_PRINCIPAL_ID('LoaderRC60');
-- dpal_ext
-- zxwang
-- temp
-- aggregation
-- mengcheng
-- crawl_staging
-- crawl_raw
-- wanlu

ALTER AUTHORIZATION ON SCHEMA::dpal_ext TO dbo;
ALTER AUTHORIZATION ON SCHEMA::zxwang TO dbo;
ALTER AUTHORIZATION ON SCHEMA::temp TO dbo;
ALTER AUTHORIZATION ON SCHEMA::aggregation TO dbo;
ALTER AUTHORIZATION ON SCHEMA::mengcheng TO dbo;
ALTER AUTHORIZATION ON SCHEMA::crawl_staging TO dbo;
ALTER AUTHORIZATION ON SCHEMA::crawl_raw TO dbo;
ALTER AUTHORIZATION ON SCHEMA::wanlu TO dbo;

DROP USER LoaderRC60
-- The database principal owns an database scoped credential and cannot be
dropped.

-- DROP DATABASE SCOPED CREDENTIAL credential_name

select * from sys.database_scoped_credentials
-- onapgedev_blob

```

```

-- OneDataDWLoaderQa
-- chaofandashi

DROP DATABASE SCOPED CREDENTIAL onapgedev_blob
-- Cannot drop the credential 'onapgedev_blob' because it is used by an
external data source.

SELECT s.name
FROM sys.schemas s
WHERE s.principal_id = DATABASE_PRINCIPAL_ID('analyst');
-- Chains
-- staging
-- raw
-- onedata_sys
-- adhoc
-- xiao
-- analyst
-- base_report
-- xiaoqiang
-- temp_staging
-- temp_adhoc
-- temp_raw
-- base
-- agg

ALTER AUTHORIZATION ON SCHEMA::Chains TO dbo
ALTER AUTHORIZATION ON SCHEMA::staging TO dbo
ALTER AUTHORIZATION ON SCHEMA::raw TO dbo
ALTER AUTHORIZATION ON SCHEMA::onedata_sys TO dbo
ALTER AUTHORIZATION ON SCHEMA::adhoc TO dbo
ALTER AUTHORIZATION ON SCHEMA::xiao TO dbo
ALTER AUTHORIZATION ON SCHEMA::analyst TO dbo
ALTER AUTHORIZATION ON SCHEMA::base_report TO dbo
ALTER AUTHORIZATION ON SCHEMA::xiaoqiang TO dbo
ALTER AUTHORIZATION ON SCHEMA::temp_staging TO dbo
ALTER AUTHORIZATION ON SCHEMA::temp_adhoc TO dbo
ALTER AUTHORIZATION ON SCHEMA::temp_raw TO dbo
ALTER AUTHORIZATION ON SCHEMA::base TO dbo
ALTER AUTHORIZATION ON SCHEMA::agg TO dbo

DROP USER analyst
-- The database principal owns an database scoped credential and cannot be
dropped.
-- 同样删除不掉

select * from sys.database_scoped_credentials
-- onapgedev_blob
-- OneDataDWLoaderQa

```

```
-- chaofandashi

DROP DATABASE SCOPED CREDENTIAL onapgedev_blob
-- Cannot drop the credential 'onapgedev_blob' because it is used by an
external data source.
```

将 Azure 资源文档上的用户、密码以及资源规格等写入新DW。

```
-- 创建login：必须连接到 master 数据库(onedataqagen2)
CREATE LOGIN LoaderRC80 WITH PASSWORD = 'Sh3lldN5&z5'; (经测试可以链接
onedataqagen2)
CREATE LOGIN analyst WITH PASSWORD = 'Sh3ll$9M6@u';
CREATE LOGIN reportreader WITH PASSWORD = 'On3Datar3Ad!@#Qa'; (经测试可以链接
onedataqagen2)
CREATE LOGIN performancetester WITH PASSWORD = 'Sh3llOn6Perfor&Tes@t!er'; (经测
试可以链接 onedataqagen2)
-- SELECT name, sid FROM sys.sql_logins WHERE name = 'LoaderRC80';

-- 创建user：登录到指定的数据库
-- 授予该user权限：这步操作需要登录到指定的数据库
-- USE onedataqagen2;
CREATE USER LoaderRC80 FOR LOGIN LoaderRC80;
GRANT CONTROL ON DATABASE::[onedataqagen2] to LoaderRC80;

CREATE USER analyst FOR LOGIN analyst;
GRANT CONTROL ON DATABASE::[onedataqagen2] to analyst;

CREATE USER reportreader FOR LOGIN reportreader;
GRANT CONTROL ON DATABASE::[onedataqagen2] to reportreader;

CREATE USER performancetester FOR LOGIN performancetester;
GRANT CONTROL ON DATABASE::[onedataqagen2] to performancetester;

-- 授予该 user 资源等级：这步操作需要登录到指定的数据库
-- USE onedataqagen2;
-- 为用户分配Resource Class 等级，不同等级的用户在Query时会被分配不同的资源
-- 如果要重新分配用户的 resource class，需要先删除当前的resource class，
-- 再分配新的resource class。例如，要为analyst 用户重新分配resource class
'staticrc60'
-- EXEC sp_droprolemember 'mediumrc', 'analyst'
-- EXEC sp_addrolemember 'staticrc60', 'analyst'

-- smallrc 这个规格没有了
-- analyst 由于一值都存在，
-- 所以创建的时候一值失败
-- 也无法登陆
```

```
EXEC sp_addrolemember 'mediumrc', 'analyst'
EXEC sp_addrolemember 'mediumrc', 'LoaderRC80'
EXEC sp_addrolemember 'smallrc', 'reportreader'
-- 这里出错 smallrc 是没有的, (官方文档有说, 这是一个默认的规格)
EXEC sp_addrolemember 'staticrc80', 'performancetester'

-- 为这个账户创建专门的特点shcema 上的权限
GRANT CREATE TABLE to reportreader;
GRANT DROP TABLE to reportreader;
GRANT INSERT, SELECT ON SCHEMA::report TO reportreader;
GRANT CREATE VIEW on SCHEMA::test_report TO reportreader;
GRANT SELECT on SCHEMA::report to reportreader;

EXEC sp_addrolemember 'staticrc80', 'performancetester'
```

查看用户角色

```
SELECT r.*,m.*, rm.*
FROM sys.database_role_members rm
JOIN sys.database_principals AS r ON rm.role_principal_id = r.principal_id
JOIN sys.database_principals AS m ON rm.member_principal_id = m.principal_id
WHERE r.name IN
('smallrc','mediumrc','largerc','xlargerc','staticrc10','staticrc20','staticrc30',
'staticrc40','staticrc50','staticrc60','staticrc70','staticrc80');

select * from sys.database_role_members
select * from sys.database_principals
```

[创建login](#)

[创建user](#)

[授予该user权限](#)

[赋予该user指定的资源类](#)

```
-- 创建login: 必须连接到 master 数据库
CREATE LOGIN <login_name> WITH PASSWORD = '<enterStrongPasswordHere>';
SELECT name, sid FROM sys.sql_logins WHERE name = 'xxx';

-- 创建user: 登录到指定的数据库
USE <database_name>;
CREATE USER <user_name> FOR LOGIN <login_name>;

-- 授予该user权限: 这步操作需要登录到指定的数据库
USE <database_name>;
```

```
GRANT CONTROL TO <user_name>;
```

-- 赋予该user指定的资源类：如果<user_name>不是成员或任何其他资源类，则会转到具有 3% 内存授予的默认 `smallrc` 资源类。

5. 后续观察期功能是否一切正常，如推上线之后，QA/dev/prod 的 host 等链接信息都是在 `azure key vault` 里面放着，这个需要改的。