# Lab 03 - SQL DW Data Loading Scenarios and Best Practices

Contents

[Lab 03 - SQL DW Data Loading Scenarios and Best Practices 1](#_Toc19365782)

[Overview 2](#_Toc19365783)

[Pre-requisites: 2](#_Toc19365784)

[Setup - Create your loading objects 2](#_Toc19365785)

[Part 1 - Connect to your SQL Data Warehouse as loading user 5](#_Toc19365786)

[Impact of file format on loading 6](#_Toc19365787)

[First - Delimited Text 6](#_Toc19365788)

[Second - Parquet 8](#_Toc19365789)

[Third - GZIP Compressed Delimited Text 10](#_Toc19365790)

[Part 4 - Impact of Single File Compression 12](#_Toc19365791)

[Part 5 - Impact of Table Distribution on loading 13](#_Toc19365792)

[Let’s take a look again at the DMVs and compare our load speeds 15](#_Toc19365793)

[Part 6 - Impact of CTAS vs Insert into Select 15](#_Toc19365794)

[Closing 16](#_Toc19365795)

## Overview

This module will go over the internals of loading data with PolyBase and provide a deeper understanding of factors that can impact loading performance.

To simplify this experience, the entire lab is conducted using SSMS and T-SQL. In the real world, you would automate everything that you see in this lab with a tool like Azure Data Factory, SSIS, or another 3rd party tool; however, the learnings presented here are still valuable as you move into production scenarios.

The TSQL Script in the GitHub directory,

* **02\_Lab03\_ConnectivityArchive.sql**
* **03\_Lab03\_DataLoadingAsLoadingUser.sql**
* **03\_Lab03DataLoadingAdminUser.sql**

contains the scripts located in the document below. For ease of use, all the source files are in a blob store with the key provided in the script directory, **blobkey.txt**.

**Note**: The following example creates a database scoped credential that can be used to create an [external data source](https://docs.microsoft.com/en-us/sql/t-sql/statements/create-external-data-source-transact-sql?view=sql-server-2017), which can do bulk operations, such as [BULK INSERT](https://docs.microsoft.com/en-us/sql/t-sql/statements/bulk-insert-transact-sql?view=sql-server-2017) and [OPENROWSET](https://docs.microsoft.com/en-us/sql/t-sql/functions/openrowset-transact-sql?view=sql-server-2017). Shared Access Signatures cannot be used with PolyBase in SQL Server, APS or SQL DW.

Reference: <https://docs.microsoft.com/en-us/sql/t-sql/statements/create-database-scoped-credential-transact-sql?view=sql-server-2017>

## Pre-requisites:

* Existing Azure SQL Data Warehouse
* Existing Azure Storage Instance with datasets
* Azure subscription
* SQL USER in SQL DW with CONTROL Permissions (From Lab 01)

## Setup - Create your loading objects

When loading data into SQL DW it is a best practice to use PolyBase. PolyBase allows you to load files (Delimited Text, ORC, RC, and Parquet) from Azure Blob Storage, Azure Data Lake Store and Azure Data Lake Store Gen2.

To use PolyBase there are four objects that must be created:

* **Database Scoped Credential** – Credential used to authenticate to External Data Source.
* **External Data Source** – Describes where the data is located at a coarse grain.
* **External File** – Describes how the files are written.
* **External Table** – Describes the fine grain location and how the data is written in the file.

Let’s take a look at the code!

Select New Query and verify that AdventureWorksDW is the selected database.

Execute the following TSQLs as admin user (sqladmin). This code is also in the file, 02\_Lab03\_ConnectivityArchive.sql. Remember to connect to the Master Database. **You may get error message that the “key” is already created. You still need to run the Master Certificate in the Master Database.**

--CREATE MASTER KEY;

IF NOT EXISTS (SELECT \* FROM sys.symmetric\_keys WHERE symmetric\_key\_id = 101 ) --name = '##MS\_ServiceMasterKey##')

BEGIN

PRINT 'Creating Database Master Key'

CREATE MASTER KEY ENCRYPTION BY PASSWORD = 'PutYourDemoEncryptionPassword@!!'

END

ELSE

BEGIN

PRINT 'Database Master Key Alread Exists'

END

GO

-- ===================================================================

-- Master Certificate - Run In Master

-- ===================================================================

--drop certificate MyDemoDataSecurityCertificate?

--create our certificate.

IF NOT EXISTS(SELECT \*

FROM sys.certificates

WHERE name = 'AzureDemoDataSecurityCertificate')

BEGIN

CREATE CERTIFICATE AzureDemoDataSecurityCertificate WITH SUBJECT = 'AZURE DataSecurity Certificate', EXPIRY\_DATE = '12/31/2024'

PRINT 'AzureDemoDataSecurityCertificate Created'

END

ELSE

BEGIN

PRINT 'AzureDemoDataSecurityCertificate Already Exists.'

END

-- ===================================================================

-- Database Scoped Credential/Data Source & TextFileFormat - Run In AdventureWorksDW

-- ===================================================================

IF EXISTS (SELECT \* FROM sys.external\_data\_sources WHERE name = 'AzureBlobStore')

DROP EXTERNAL DATA SOURCE AzureBlobStore;

go

--Drop if already created

--DROP DATABASE SCOPED CREDENTIAL BLOBCredentialDEV;

CREATE DATABASE SCOPED CREDENTIAL BLOBCredentialDEV

WITH

IDENTITY = 'myuser',

SECRET = 'dRIZB1nQbsIa2pi+o45h1pFRbrNiiAbjMoT3WVlskJf7e9nCwgP9imVFwXm8oG7YYHZyTvWkCZQ/Qxj7UzsdbA=='

CREATE EXTERNAL DATA SOURCE USGSWeatherEvents with (

TYPE = HADOOP,

LOCATION ='wasbs://usgsdata@dwdatdstorage.blob.core.windows.net',

CREDENTIAL = BLOBCredentialDEV

);

CREATE EXTERNAL FILE FORMAT TextFileFormat

WITH

(

FORMAT\_TYPE = DELIMITEDTEXT,

FORMAT\_OPTIONS

(

FIELD\_TERMINATOR = '|'

)

)

GO

We now have the External Table and connectivity.

The next thing that we need to do is create an external table over the files so that we can access the data from SQL DW. We can do this with the objects that we created above. Think of an External Table as a pointer to the files on the storage system.

Execute the following TSQL to create an external table:

-- Create schema for external tables over Azure Data Lake Store Gen2

CREATE SCHEMA [EXT]

GO

-- Create schema for staging tables after loading

CREATE SCHEMA [STG]

GO

-- Create schema for cleaned production tables

CREATE SCHEMA [PROD]

GO

-- Create schema for cleaned production tables

CREATE SCHEMA [staging]

GO

-- Create the External Table

CREATE EXTERNAL TABLE [staging].[STG\_factWeatherMeasurements\_text]

(

[StationId] [nvarchar](12) NOT NULL,

[ObservationTypeCode] [nvarchar](4) NOT NULL,

[ObservationValueCorrected] [real] NOT NULL,

[ObservationValue] [real] NOT NULL,

[ObservationDate] [date] NOT NULL,

[ObservationSourceFlag] [nvarchar](2) NULL,

[fipscountycode] [varchar](5) NULL

)

WITH

(

DATA\_SOURCE = USGSWeatherEvents,

LOCATION = '/usgsdata/weatherdata/factWeatherMeasurements/',

FILE\_FORMAT = TextFileFormat

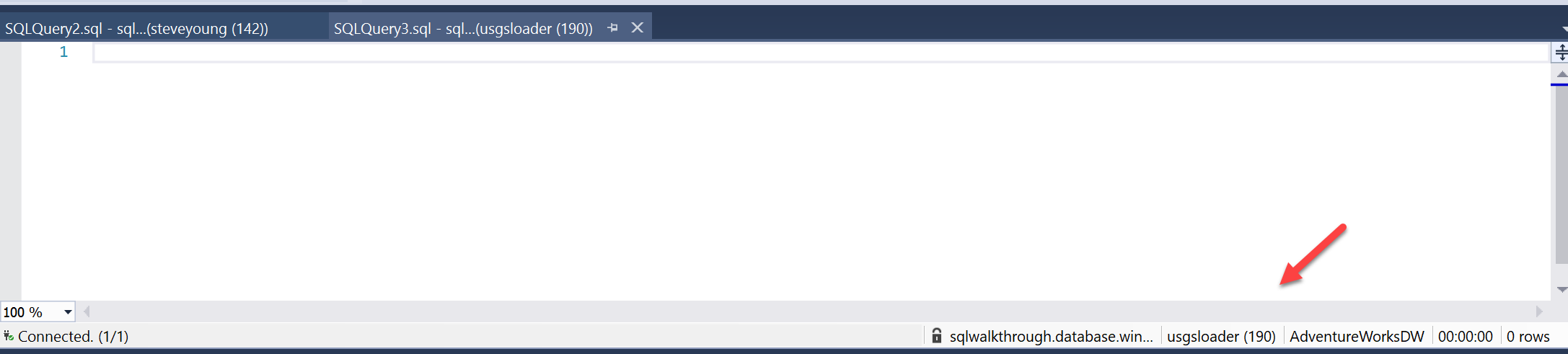
);

select count(\*) from [staging].[STG\_factWeatherMeasurements\_text]

The rest of the 02\_Lab03\_ConnectivityArchive.sql script can be run to create the rest of the Weather Events DW, however the rest of the labs only use the TABLE [staging].[STG\_factWeatherMeasurements\_text]

## Part 1 - Connect to your SQL Data Warehouse as loading user

You will have 2 Query windows open for this section. One will load data, the other will run operational queries to see various performance views. The bottom right of the screen will show the currently logged in user.



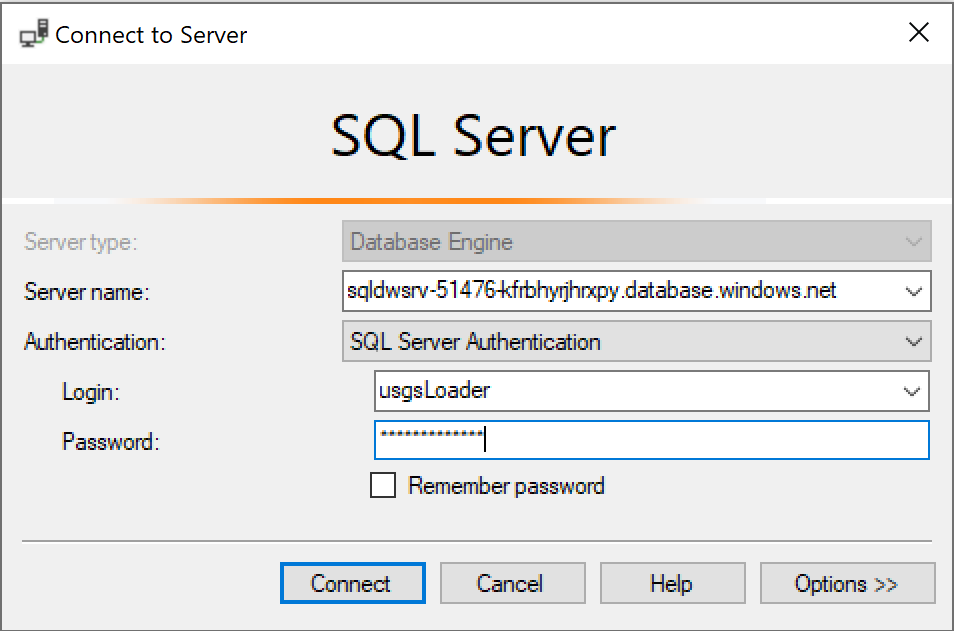
Before you start loading data into your SQL Data Warehouse, **you need to login into the DW with your loading user**. We advise having a separate user for loading because it gives you greater control over security, resource allocation, and workload management by separating loading jobs from analytical queries.

For this portion of the lab, you **should also have another Query Window logged in as your Admin user.** This will allow you to see loading statistics using Data Management Views (DMVs.)

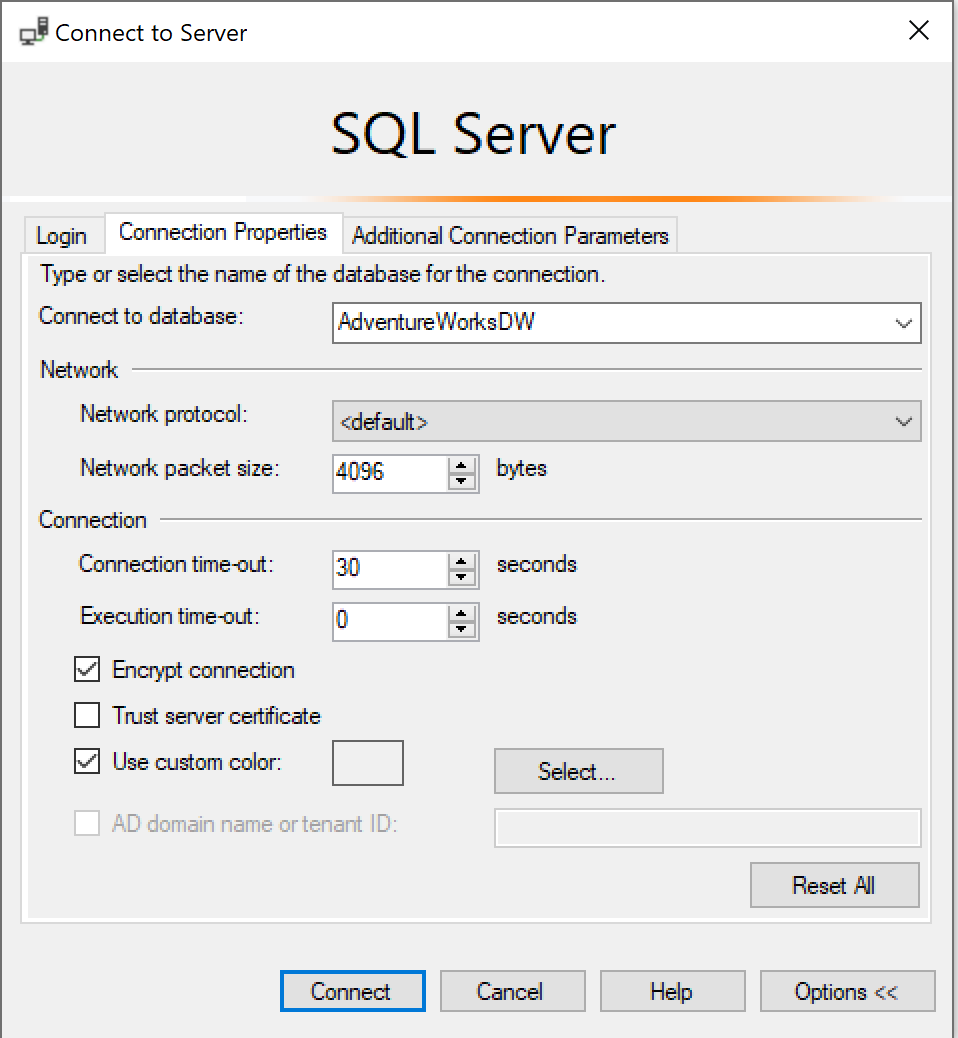
These Scripts are in **03\_DataLoadingAsLoadingUser.sql**

In SQL Server Management Studio:

1. Log in as ‘usgsLoader’



1. Click on ‘Options’, change the database name to ‘AdventureWorksDW’ and click ‘Connect’.



## Impact of file format on loading

Choosing the right file format for loading is often an overlooked choice when customers are designing their load jobs into SQL DW.

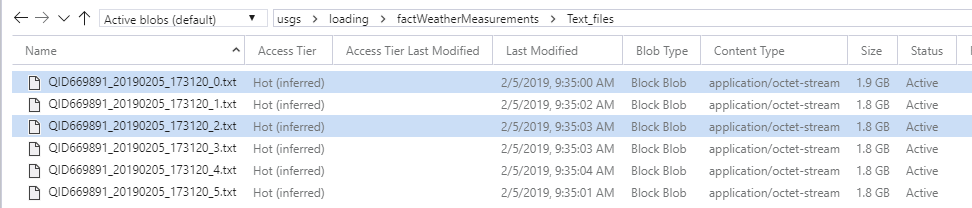
By far, the most popular format used to load data into SQL DW is the delimited text file. It is human readable, generic and the default output of many on-prem systems. However, the delimited text file can cause loading issues because it is essentially a collection of strings which are open to interpretation by the loading tool. Additionally, several delimited text generation tools do not enforce data quality during writes, which means you when the data is written is can no longer be interpreted as the same data. For example, the difference between a null value and an empty string is subtle in a delimited text file. If quotes are turned off, then all empty strings will be interpreted as a null! This is seemingly innocent form of data loss that can have large consequences.

Parquet is another very popular format that is gaining popularity in the big data ecosystem because of its binary data representation, strongly typed data, and compression capabilities. This format is popularized by the spark and hive ecosystem and we are seeing the increase in adoption as the Modern Data Warehouse Architecture continues to grow in popularity.

For this section of the lab, the same data is stored on Azure Storage in Delimited Text, Parquet and Gzip Compressed Delimited text. We are going to load each in turn and look at storage size and load performance.

### First - Delimited Text

The data below was generated via a PolyBase export of 134,333,511 records. As you can see below, PolyBase tries to evenly distribute the data across multiple files. This is a best practice when writing data because it effectively parallelizes the DW resources to load each file simultaneously.



One should note that these are 1.8GB files because the files are uncompressed. This is raw data size which we will see change as we change file formats and compressions later in the lab.

This statement points to the directory location of the files. PolyBase will traverse all files in a directory and all subdirectories (recursive directory traversal) up to 33k files.

The next step is to run a CREATE TABLE AS SELECT (CTAS) to create a new table from the external data. Let’s load some data!

CREATE TABLE [staging].[STG\_text\_load]

WITH

(

DISTRIBUTION = ROUND\_ROBIN,

HEAP

)

AS

SELECT \* FROM [staging].[STG\_factWeatherMeasurements\_text] option(label = 'STG\_text\_load')

While this is running you can view the steps with the Dynamic Management Views (DMVs)

**Open another Query windows as admin user** and execute the following script with the AdventureWorksDW selected.

SELECT s.\*

FROM

sys.dm\_pdw\_exec\_requests r

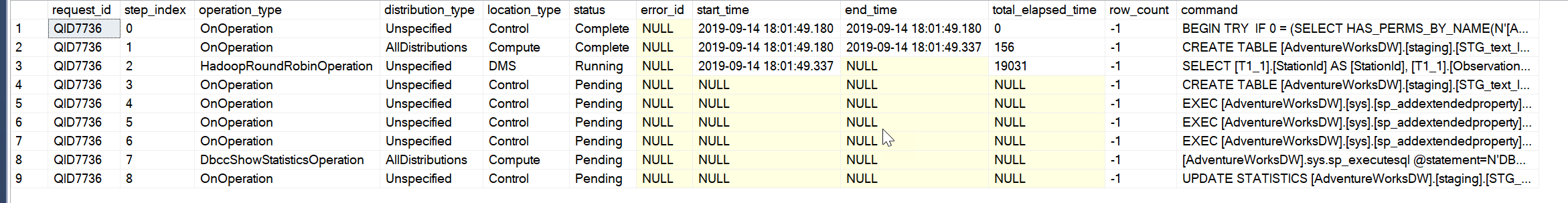
JOIN

Sys.dm\_pdw\_request\_steps s

ON r.request\_id = s.request\_id

WHERE r.[label] = 'STG\_text\_load'

You should see results like this while running. Note the End times as the query is not complete yet;



You’ll probably notice that the operation is in the HadoopRoundRobinOperation operation\_type for quite a while. This is the operation the reads the files from the data source into SQL DW. Notice that this step has a DMS location\_type. DMS stands for Data Movement Service. This component of the DW architecture is responsible for data movement between nodes and data movement from outside sources into SQL DW.

We can see more detailed information about what is happening in this step with the following.

SELECT ew.\*

FROM[sys].[dm\_pdw\_dms\_external\_work] ew

JOIN sys.dm\_pdw\_exec\_requests r

ON r.request\_id = ew.request\_id

JOIN Sys.dm\_pdw\_request\_steps s

ON r.request\_id = s.request\_id

WHERE r.[label] = 'STG\_text\_load'

ORDER BY input\_name, read\_location

This is particularly interesting because it shows how we effectively load uncompressed Delimited Text files in parallel. For large files >512 MB, we split the file into multiple chunks so it can be read more efficiently.

### Second - Parquet

For Parquet, let’s kick off the code and then talk about what’s happening while the data loads.

Execute the following statements as **admin user (sqladmin)**

CREATE EXTERNAL FILE FORMAT [Parquet] WITH (FORMAT\_TYPE = PARQUET)

GO

-- DROP EXTERNAL TABLE [staging].[STG\_factWeatherMeasurements\_parquet]

CREATE EXTERNAL TABLE [staging].[STG\_factWeatherMeasurements\_parquet]

(

[SiteName] [varchar](52) NOT NULL,

[ReadinType] [varchar](52) NOT NULL,

[ReadingTypeID] [real] NOT NULL,

[ReadingValue] [real] NOT NULL,

[ReadingTimestamp] [datetime2](7) NOT NULL,

[ReadingUnit] [varchar](52) NOT NULL,

[fpscode] [int] NOT NULL

)

WITH (

DATA\_SOURCE = USGSWeatherEvents,

LOCATION = N'/usgsdata/Parquet\_files/',

FILE\_FORMAT = [Parquet],

REJECT\_TYPE = VALUE,

REJECT\_VALUE = 0)

GO

Execute the following CTAS as **load user** (**usgsLoader**) (in the query tab loggin as the load user

CREATE TABLE [staging].[STG\_parquet\_load]

WITH

(

DISTRIBUTION = ROUND\_ROBIN,

HEAP

)

AS

SELECT \*

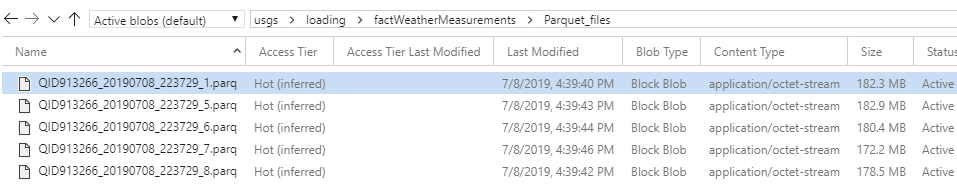
FROM [staging].[STG\_factWeatherMeasurements\_parquet]

OPTION (label = 'STG\_parquet\_load')

GO

The first thing that we did above was create a new External File Format named Parquet. Notice that this didn’t require a lot of information because the Parquet file format is predefined unlike delimited text. This helps keep data consistent when it moves around an open big data ecosystem.

If we look at the files on the storage layer, we can see that the data size is much smaller than delimited text files. This is because Parquet is a columnar format and has compression built into the file format.



Let’s take a look at the DMVs again to see what happens with Parquet load.

SELECT DISTINCT ew.\*

FROM[sys].[dm\_pdw\_dms\_external\_work] ew

JOIN sys.dm\_pdw\_exec\_requests r

ON r.request\_id = ew.request\_id

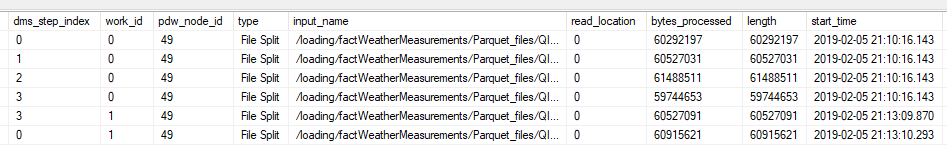
JOIN Sys.dm\_pdw\_request\_steps s

ON r.request\_id = s.request\_id

WHERE r.[label] = 'STG\_parquet\_load'

ORDER BY start\_time ASC, dms\_step\_index

Parquet is not split-able today in SQL DW. This is shown by each file only appearing once. Also, notice that DMS\_Step\_Index values show up more than once. This means that the reader read one file through to its entirety then went back and started reading the next one.



### Third - GZIP Compressed Delimited Text

Again, let’s go ahead and execute the following statements before discussing

**As admin user:**

CREATE EXTERNAL FILE FORMAT compressed\_TextFileFormat\_Ready

WITH

(

FORMAT\_TYPE = DELIMITEDTEXT,

FORMAT\_OPTIONS

(

FIELD\_TERMINATOR = '|'

), DATA\_COMPRESSION = N'org.apache.hadoop.io.compress.GzipCodec'

)

GO

CREATE EXTERNAL TABLE [staging].[STG\_factWeatherMeasurements\_CompressedText]

(

[SiteName] [varchar](52) NOT NULL,

[ReadinType] [varchar](52) NOT NULL,

[ReadingTypeID] [real] NOT NULL,

[ReadingValue] [real] NOT NULL,

[ReadingTimestamp] [datetime2](7) NOT NULL,

[ReadingUnit] [varchar](52) NOT NULL,

[fpscode] [int] NOT NULL

)

WITH

(

DATA\_SOURCE = USGSWeatherEvents,

LOCATION = N'/usgsdata/Compressed\_Text\_files',

FILE\_FORMAT = [compressed\_TextFileFormat\_Ready],

REJECT\_TYPE = VALUE,

REJECT\_VALUE = 0

)

GO

**As load user (usgsLoader)**

CREATE TABLE [staging].[STG\_compressed\_text\_load]

WITH

(

DISTRIBUTION = ROUND\_ROBIN,

HEAP

)

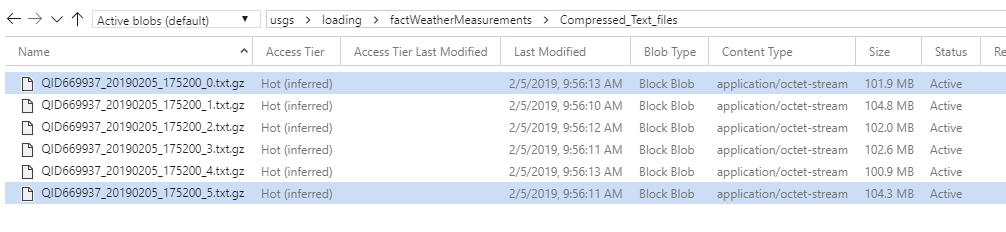
AS

SELECT \*

FROM [staging].[STG\_factWeatherMeasurements\_CompressedText]

OPTION (label = 'STG\_compressed\_text\_load')

GO



Again, you will notice that SQL DW does not split GZIP compressed files. This is because the compression algorithm itself is not split-able. More on this problem later!

Let’s take a moment and compare performance of the three file formats that we used. (NOTE: Wait until the CTAS is completed before you execute the statement below)

SELECT AVG(total\_elapsed\_time) AS [avg\_loadTime\_ms], [label]

FROM sys.dm\_pdw\_exec\_requests

WHERE [label] IS NOT NULL

AND [label] <> 'System'

AND Status = 'Completed'

GROUP BY [label]

|  |  |
| --- | --- |
| **avg\_loadTime\_ms** | **label** |
| 253246 | STG\_compressed\_text\_load |
| 263624 | STG\_parquet\_load |
| 274567 | STG\_text\_load |

It’s clear that for this particular table and data volume that compression helped. Gzip compression was the fastest, then Parquet and finally uncompressed text files. Now there are drawbacks with this. Both Gzip and Parquet files are not splittable in SQL DW. If not done correctly, this will severely impact load performance. Additionally, we have found that for wider tables with lots of strings, the CPU overhead of converting the strings slows down the overall performance. Finally, neither of the fastest two file formats are human readable. There is no easy way to visually inspect the data which can make them hard to debug in the case that anything happens during the load.

## Part 4 - Impact of Single File Compression

Let’s take the same data and try to load it from a single compressed Gzip file.

**As admin user (sqladmin)**

CREATE EXTERNAL TABLE [staging].[STG\_factWeatherMeasurements\_CompressedText\_single\_file]

(

[SiteName] [varchar](52) NOT NULL,

[ReadinType] [varchar](52) NOT NULL,

[ReadingTypeID] [real] NOT NULL,

[ReadingValue] [real] NOT NULL,

[ReadingTimestamp] [datetime2](7) NOT NULL,

[ReadingUnit] [varchar](52) NOT NULL,

[fpscode] [int] NOT NULL

)

WITH

(

DATA\_SOURCE = USGSWeatherEvents,

LOCATION = N'/usgsdata/singlefile/Compressed\_Text\_files',

FILE\_FORMAT = [compressed\_TextFileFormat\_Ready],

REJECT\_TYPE = VALUE,

REJECT\_VALUE = 0)

GO

**As loading user (usgsloader)**

CREATE TABLE [staging].[STG\_CompressedText\_single\_file]

WITH

(

DISTRIBUTION = ROUND\_ROBIN, HEAP

)

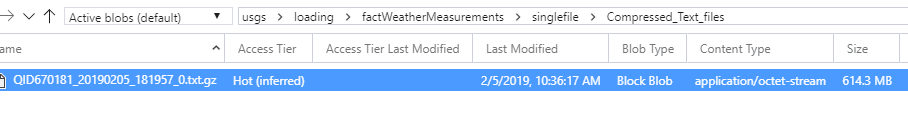
AS

SELECT \*

FROM [staging].[STG\_factWeatherMeasurements\_CompressedText\_single\_file]

OPTION(label = 'STG\_single\_compressed\_load')

GO



You’ve probably notice that this is taking quite a bit a time. While the load is running, we would like for you to think about the following questions and discuss with the person next to you.

-- Rerun this command as Admin when the Loader query is complete

SELECT AVG(total\_elapsed\_time) AS [avg\_loadTime\_ms], [label]

FROM sys.dm\_pdw\_exec\_requests

WHERE [label] IS NOT NULL

AND [label] <> 'System'

AND Status = 'Completed'

GROUP BY [label]

## Part 5 - Impact of Table Distribution on loading

So far, we have run Create Table as Select (CTAS) statements with Round\_Robin distributions and Heap indexes. This is the theoretical maximum speed that you can load a set of data into SQL DW. Fixing file format issues as describe above or scaling the DW are the only ways to improve performance. However, there are ways to *degrade* performance off of this standard.

The most common problem is data skew. This occurs when the cardinality of the distribution column is low or skewed to a certain set of values. The canonical case is a key that has a lot of NULL or 0 values.

Let’s suppose that we wanted to load that same data again but hash the data on the readingUnit column. Run this command as the USGSLoader (See Part 5 of script) (may take 5-7 minutes depending on load)

CREATE TABLE [staging].[STG\_Hash\_ReadingUnit]

WITH

(

DISTRIBUTION = HASH(ReadingUnit),

HEAP

)

AS

SELECT \*

FROM [staging].[STG\_factWeatherMeasurements\_CompressedText]

OPTION (label = 'STG\_Hash\_ReadingUnit')

While this is running, run the following in another query window as admin user (sqldmin):

Select \*

FROM sys.dm\_pdw\_dms\_workers dw

JOIN sys.dm\_pdw\_exec\_requests r

ON r.request\_id = dw.request\_id

WHERE r.[label] = 'STG\_Hash\_ReadingUnit'

This shows how the data is being read from files, how that data is being converted to SQL data type format and how it is being written to SQL DW distributions. The steps are the following:

* EXTERNAL\_READER = Each line is an individual thread reading data from files
* HASH\_CONVERTER = Thread responsible for converting the data generated by the EXTERNAL\_READER
* WRITER = Thread responsible for writing to the correct distribution based on sink table definition.

Notice how many writers are sitting around with 0 bytes\_processed. This is because of the table skew! Table skew occurs when certain distributions receive more data than others. Distributions receive data based on the hash value of the distribution.

If we did a simple check on how many rows per readingUnit, we would see that there are only 17 unique readingUnit values. When writing, these 17 distributions receive all of the load causing a bottleneck. Now, some skew is natural, but this is extreme! The table would be better suited to a ROUND\_ROBIN distribution.

|  |  |
| --- | --- |
| **RowCount** | **readingUnit** |
| 6658050 | T |
| 47216 | D |
| 11394 | A |
| 134061 | B |
| 10871562 | 7 |
| 36655 | M |
| 9902661 | W |
| 75345491 | 0 |
| 16481 | 6 |
| 126562 | S |
| 615070 | H |
| 13002585 | N |
| 9373275 | U |
| 755951 | K |
| 91195 | Z |
| 293697 | R |
| 7051605 | X |

## Let’s take a look again at the DMVs and compare our load speeds

Execute the following query as admin user (wait until the previous load is completed):

Select avg(total\_elapsed\_time) as [avg\_loadTime\_ms], [label]

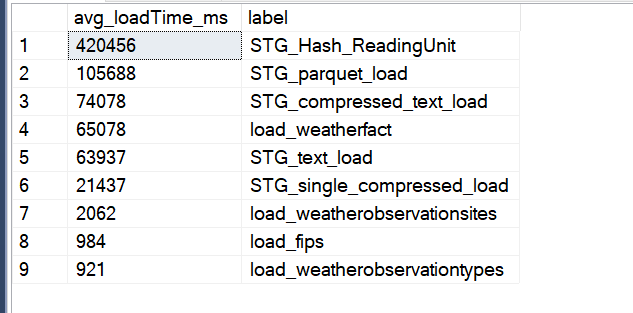
FROM sys.dm\_pdw\_exec\_requests

where [label] is not null

and [label] <> 'System'

and Status = 'Completed'

GROUP BY [label] order by 1 desc



As you can see, table structure had the biggest impact on my loading performance! NOTE!!! For the labs we may only be using 1 node in order to limit cost.

## Part 6 - Impact of CTAS vs Insert into Select

Depending on your ELT pattern, you can either use a CTAS to create a brand-new table or an INSERT INTO SELECT from external table to append records to an existing table. Both methods have their purposes, but the two syntaxes are not functionally equivalent.

Let’s take a look and see the performance differences. Since we have already done several CTAS statements lets focus on INSERT INTO SELECT and compare. Run as USGLoader

CREATE TABLE [staging].[STG\_factWeatherMeasurements\_Insert\_Into]

(

[SiteName] [varchar](52) NOT NULL,

[ReadinType] [varchar](52) NOT NULL,

[ReadingTypeID] [real] NOT NULL,

[ReadingValue] [real] NOT NULL,

[ReadingTimestamp] [datetime2](7) NOT NULL,

[ReadingUnit] [varchar](52) NOT NULL,

[fpscode] [int] NOT NULL

)

WITH

(

DISTRIBUTION = ROUND\_ROBIN,

HEAP

)

GO

INSERT INTO [staging].[STG\_factWeatherMeasurements\_Insert\_Into] SELECT \* FROM [staging].[STG\_factWeatherMeasurements\_text] option (label = 'STG\_insertInto')

GO

INSERT INTO [staging].[STG\_factWeatherMeasurements\_Insert\_Into]

SELECT \* FROM [staging].[STG\_factWeatherMeasurements\_text] option (label = 'STG\_insertInto\_two')

GO

Wait until the above inserts are completed and execute the following query as admin user:

SELECT avg(total\_elapsed\_time) as [avg\_loadTime\_ms], [label]

FROM sys.dm\_pdw\_exec\_requests

WHERE [label] is not null

and [label] <> 'System'

and Status = 'Completed'

GROUP BY [label] ORDER BY 1 DESC

As you can see CTAS is slightly faster (~10%) than writing directly into a table via insert into select, but that isn’t why we suggest loading to a staging table via CTAS statements. **CTAS is a fully parallelize**, **minimally logged operation while Insert INTO has normal logging enabled**. This functional difference accounts for the difference in load speeds.

## Closing

You’ve reached the end of the DW loading lab. Hopefully you are now comfortable with the concepts behind loading data into SQL DW and better understand some of the levers you can utilize to optimize your data loads

