

# On the Way to Fabric ...

Alexander Karl

# ABOUT ME



## Alexander Karl



SQL & BI Developer



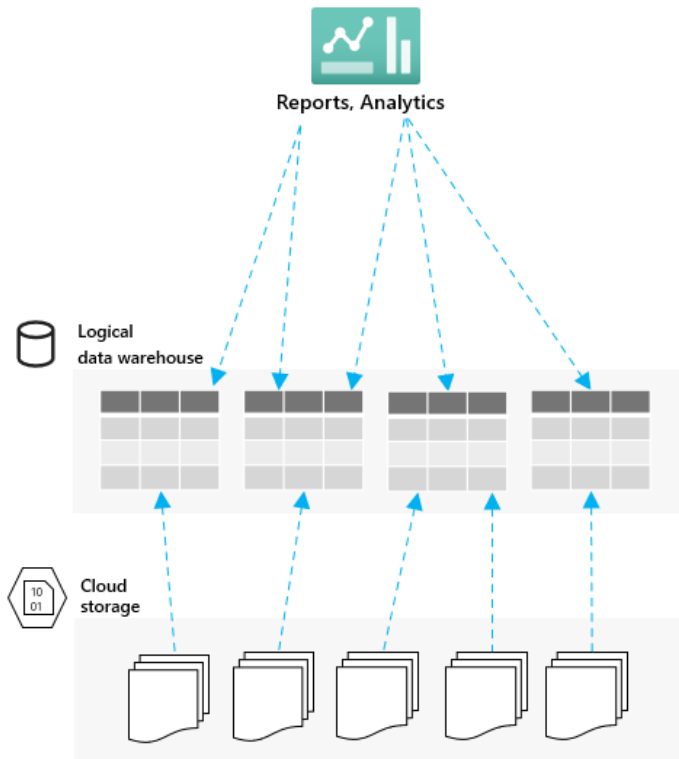
Stockstadt / PASS RG Rhein-Main



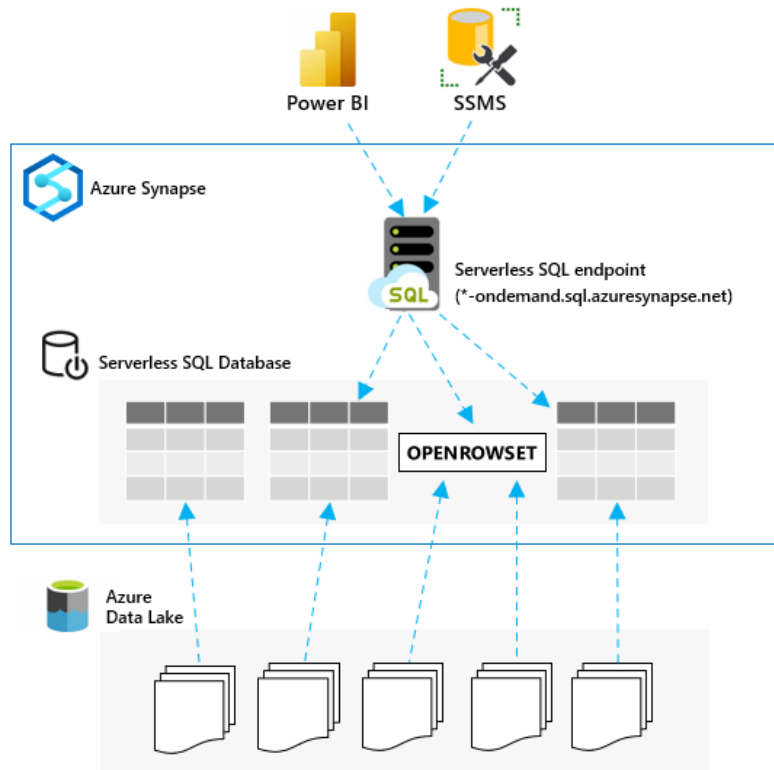
karl@net-CDE.de  
& aka@sqlpass.de



## Conceptual design



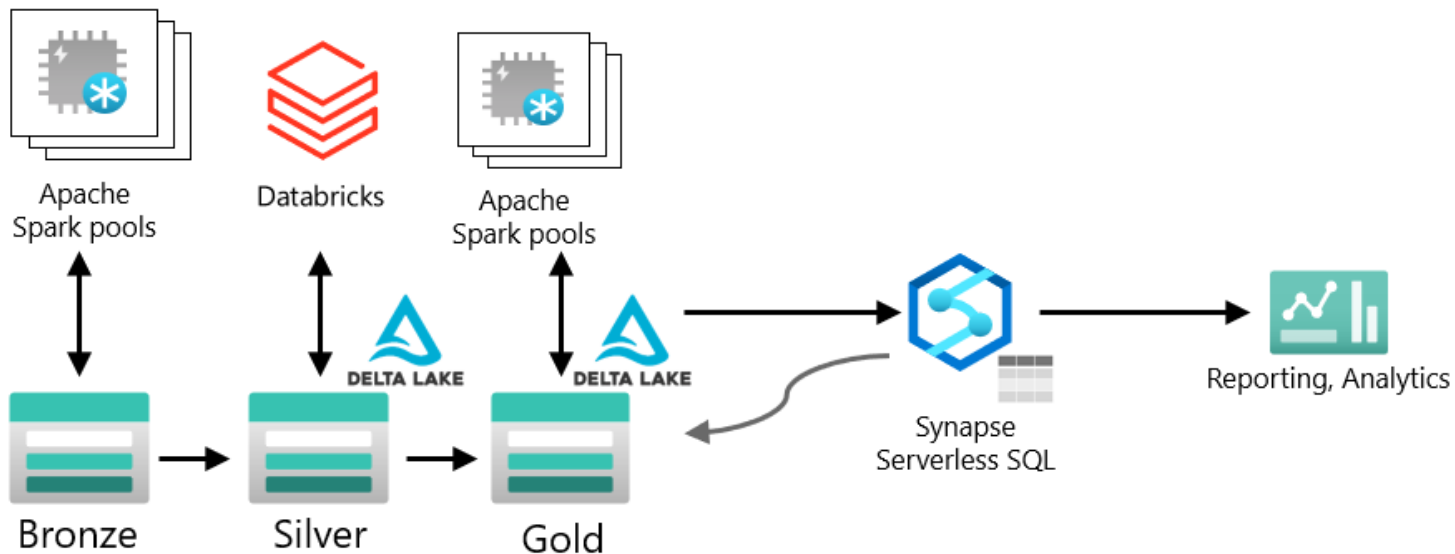
## Implementation with Azure Synapse

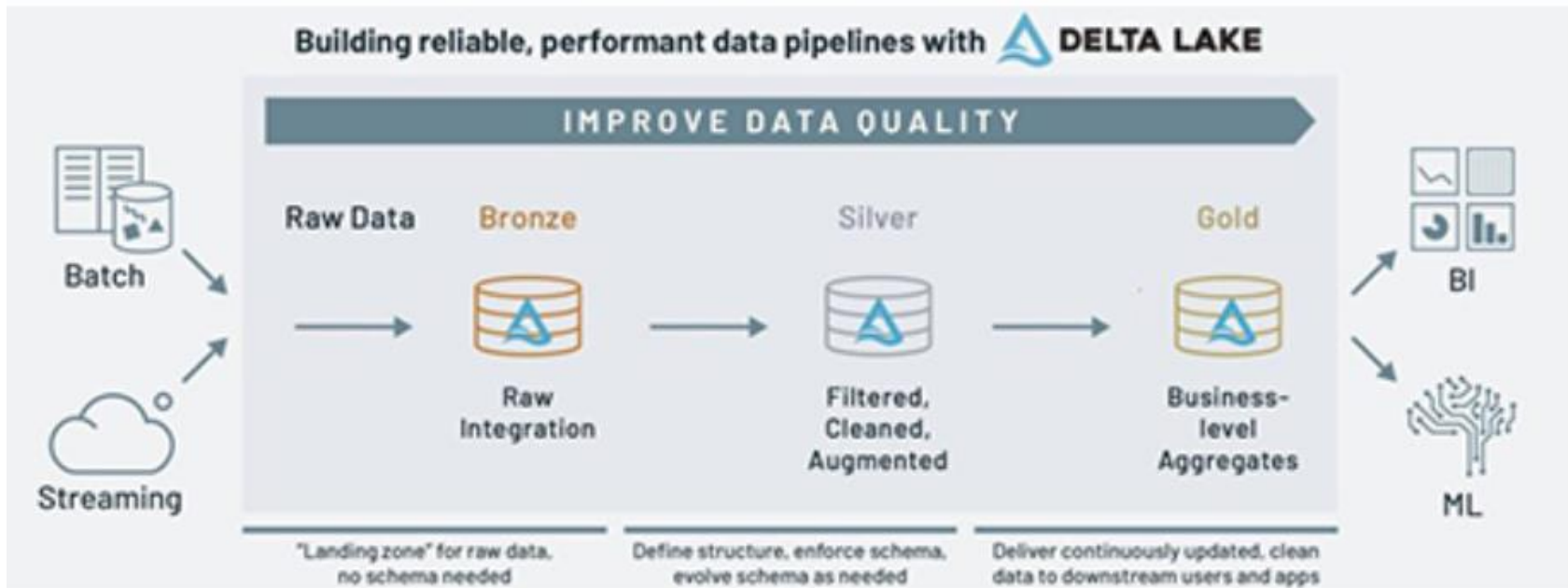


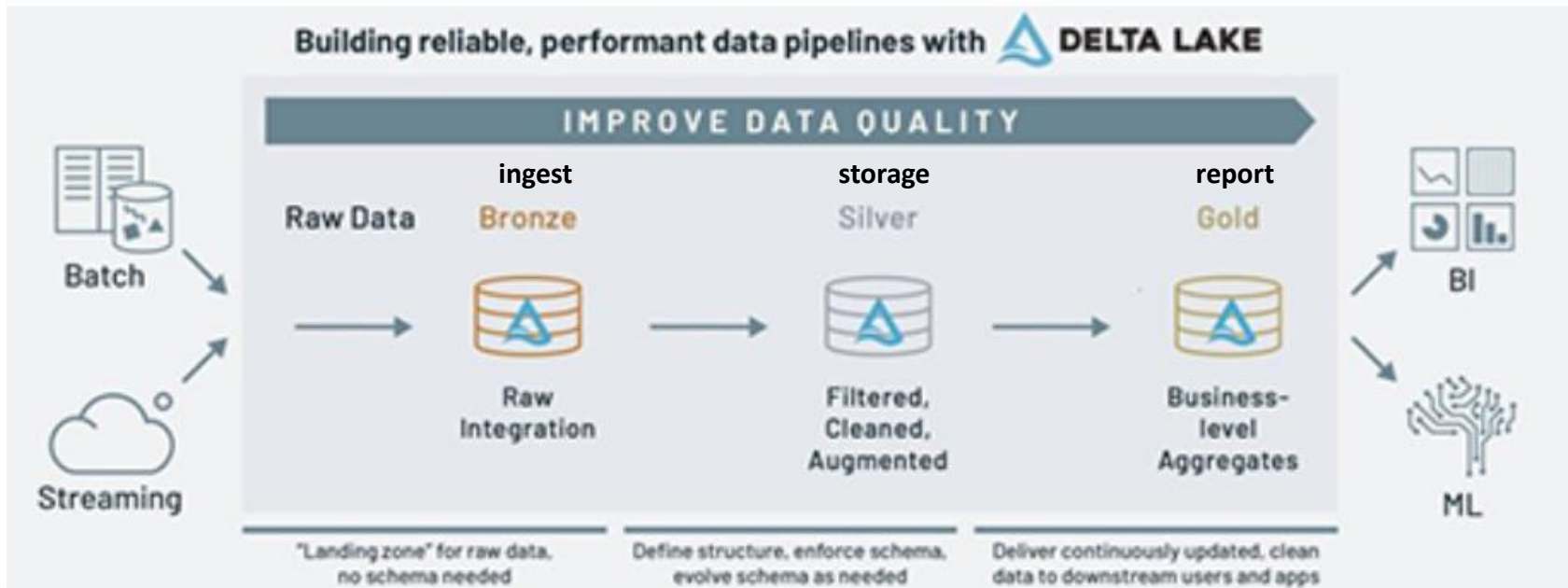
## Data processing

## Serving

## Reporting







# Blob Folders

Microsoft Azure

Home >

**StorageAccount**  
Storage account

Search

Overview

- Activity log
- Tags
- Diagnose and solve problems
- Access Control (IAM)
- Data migration
- Events
- Storage browser

Data storage

- Containers**
- File shares
- Queues
- Tables

Upload Open in Explorer Delete Move

Essentials

Resource group (move) : [resgroup-BWDE-cloud-iot](#)

Location : westeurope

Subscription (move) : [BWDE - IoT Data - PROD](#)

Subscription ID : 305d55fc-5c04-4948-a232-c37ee6689546

Disk state : Available

Tags (edit) : [Add tags](#)

Properties Monitoring Capabilities (5) Recommendations

**Data Lake Storage**

Hierarchical namespace	Enabled
Default access tier	Hot
Blob anonymous access	Enabled
Blob soft delete	Enabled (7 days)

**ContainerName**  
Container

Search

Upload Add Directory Refresh Rename Delete

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

Authentication method: Access key (Switch to Microsoft Entra user account)

Location: ContainerName / [zTest\\_delta](#) / [sampledata](#) / [2019](#)

Search blobs by prefix (case-sensitive)

Name

☐ [\[..\]](#)

☐ [01](#)

☐ [n2](#)

Upload Add Directory Refresh Rename Delete

Authentication method: Access key

Location: [/ zTest\\_delta / sampledata / 2019 / 03](#)

Name

☐ [\[..\]](#)

☐ [yellow\\_tripdata\\_2019\\_03.csv](#)

# Blob Folders SAS

ContainerName

Container

Overview

Diagnose and solve problems

Access Control (IAM)

Settings

Shared access tokens

Manage ACL

Access policy

Properties

Metadata

A shared access signature (SAS) is a URI that grants restricted access to an Azure Storage container. Use it when you want to gra  
[SAS](#)

Signing method

☒ Account key ☐ User delegation key

Signing key ⓘ

Key 1 ▾

Stored access policy

None ▾

Permissions \* ⓘ

7 selected ▾

Start and expiry date/time ⓘ

Start

12.03.2024

(UTC+01:00) Amsterdam, Berlin, Bern, Rome, Stockholm, Vienna ▾

Expiry

30.01.2025

(UTC+01:00) Amsterdam, Berlin, Bern, Rome, Stockholm, Vienna ▾

Allowed IP addresses ⓘ

for example, 168.1.5.65 or 168.1.5.65-168....

Allowed protocols ⓘ

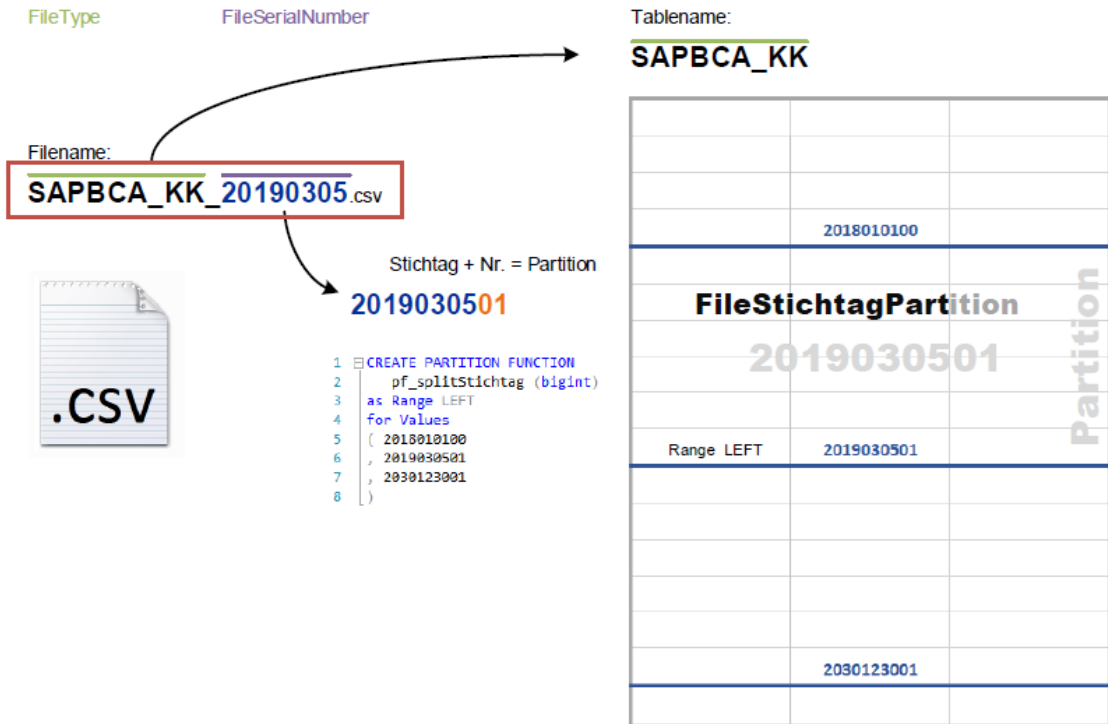
☒ HTTPS only ☐ HTTPS and HTTP

[Generate SAS token and URL](#)

Blob SAS token ⓘ

sp=racwIme&st=2024-03-12T07:26:57Z&se=2025-01-30T15:26:57Z&spr=https&sv=2022-11-02&sr=c&sig=EUnQjWyrq9Yz9qvMl



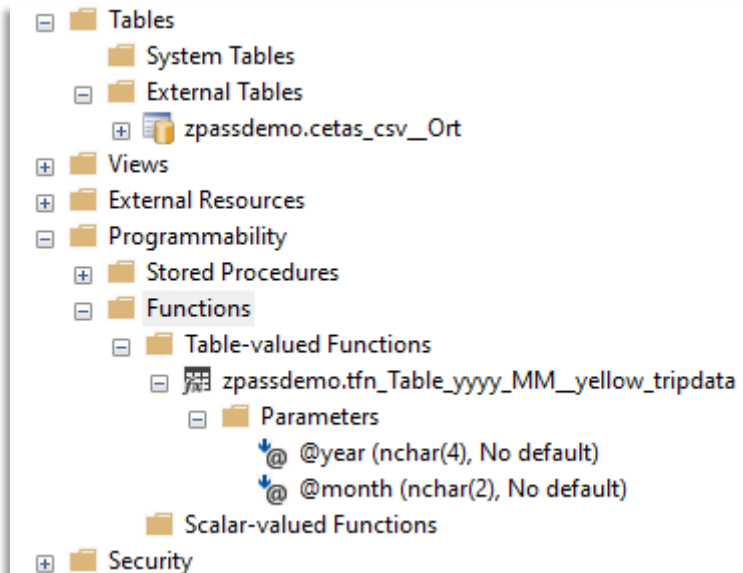
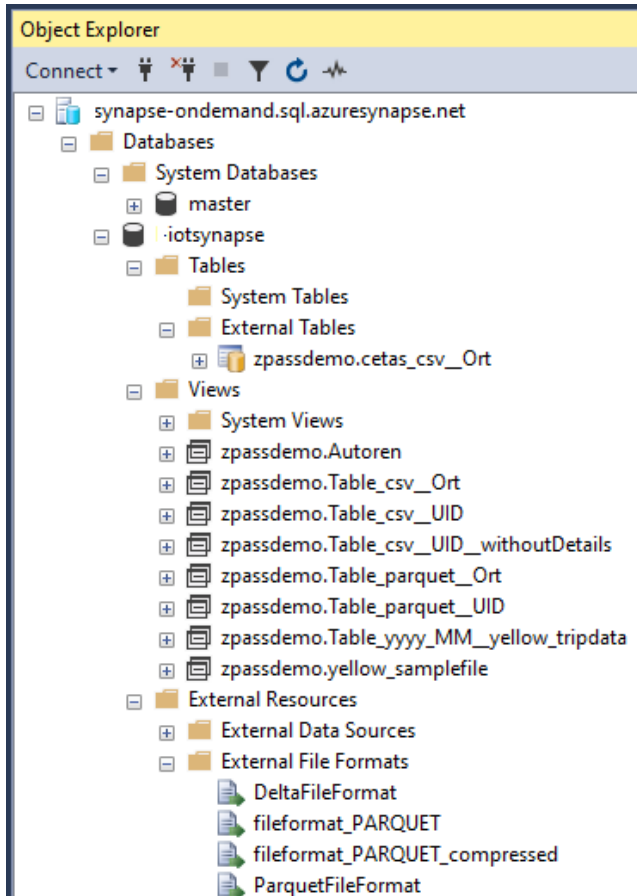


# DEMO

## Beispiel 1

tsql\_files 01 ... 04





```

1 CREATE OR ALTER FUNCTION [zpassdemo].[tfn_Table_yyyy_MM_yellow_tripdata]
2     ( @year  nchar(4)
3       , @month nchar(2)
4     )
5 RETURNS TABLE
6 AS RETURN
7 (
8     SELECT
9         count(*)                                as 'cnt_Rows'
10        , t.filepath(3)                          as 'blob_yyyy_MM'    ---- (3) = dritter ""
11        , t.filename()                          as 'blob_FileName'
12        , t.filepath()                          as 'blob_FilePath'
13    FROM
14        OPENROWSET( BULK '/zTest_delta/sampleddata/*/yellow_tripdata_*.csv'
15                   , DATA_SOURCE = 'eds_testfabrikpassdemo_testfabrikfilesystem'
16                   , FORMAT       = 'CSV'
17                   , FIELDTERMINATOR = ','
18                   --, ROWTERMINATOR = '\n'
19                   , PARSER_VERSION = '2.0'
20                   , HEADER_ROW = TRUE
21                   -- [ <reject_options> ]
22                 )
23          as [t]
24 WHERE
25     t.filepath(1) IN ( @year )
26 and t.filepath(2) IN ( @month )
27 GROUP by
28     t.filepath(3)
29     , t.filename()
30     , t.filepath()
31 )
32 -- END FUNCTION

```

```

1 CREATE OR ALTER FUNCTION [zpassdemo].[tfn_Table_yyyy_MM_yellow_tripdata]
2     ( @year  nchar(4)
3       , @month nchar(2)
4     )
5 RETURNS TABLE
6 AS RETURN
7 (
8     SELECT
9         count(*)                as 'cnt_Rows'
10        , t.filepath(3)          as 'blob_yyyy_MM'  ---- (3) = dritter "*"
11        , t.filename()           as 'blob_FileName'
12        , t.filepath()           as 'blob_FilePath'
13    FROM
14        OPENROWSET( BULK '/zTest_delta/sampleddata/*/yellow_tripdata_*.csv'
15                   , DATA_SOURCE = 'eds_testfabrikpassdemo_testfabrikfilesystem'
16                   , FORMAT       = 'CSV'
17                   , FIELDTERMINATOR = ','
18                   , ROWTERMINATOR = '\n'

```

```

1 SELECT *
2 FROM [zpassdemo].[tfn_Table_yyyy_MM_yellow_tripdata] ( '2019' , '08' )
3 GO

```

100 %

Results

Messages

	cnt_Rows	blob_yyyy_MM	blob_FileName	blob_FilePath
1	6073357	2019_08	yellow_tripdata_2019_08.csv	/zTest_delta/sampleddata/2019/08/yellow_tripdata_2019_08.csv

```

29     , t.filename()
30     , t.filepath()
31 )
32 -- END FUNCTION

```

# DEMO

## Beispiel 2

tsql\_admin files



# CETAS    CreateExternalTableAsSelect

?? Warum kein    SELECT ...    INTO

# CETAS CreateExternalTableAsSelect

```
1 CREATE EXTERNAL TABLE zpassdemo.cetas_csv__Ort
2 WITH
3 ( LOCATION      = '/report/2024/03/12/cetas_csv__Ort'
4   , DATA_SOURCE = eds_testfabrikpassdemo_testfabrikfilesystem
5   , FILE_FORMAT  = PARQUET
6 )
7 AS
8 SELECT *
9 FROM   zpassdemo.Table_csv__Ort;
```





```
1 CREATE OR ALTER PROC [dbo].[usp_cetas_from_VIEW]
2     @currDatetime smalldatetime
3     , @viewName sysname -- = 'zpassdemo.Table_csv__Ort'
4 AS
5 BEGIN
6     ----
7     DECLARE @StorageAccount sysname = 'testfabrikpassdemo'
8     , @Container sysname = 'testfabrikfilesystem'
9
10    DECLARE @sqlCmd nvarchar(4000)
11        --SELECT @currDatetime -- sys.time_zone_info --> UTC
12    ----
13    SET @sqlCmd = '
14    IF EXISTS (select * from sys.objects where object_id = OBJECT_ID(N''' + Replace( @viewName, 'Table', 'cetas' ) + ''') )
15        DROP EXTERNAL TABLE ' + Replace( @viewName, 'Table', 'cetas' ) + ';
16    ' --
17    --select @sqlCmd
18    EXEC sp_executesql @sqlCmd
19
20    SET @sqlCmd = N'
21    CREATE EXTERNAL TABLE ' + Replace( @viewName, 'Table', 'cetas' ) + '
22    WITH
23    (
24        LOCATION = '/report/' + Format( @currDatetime, 'yyyy/MM/dd/HHmm__' ) + Replace( Replace( @viewName, 'Table', 'cetas'
25                                                    , 'zpassdemo.', '' )
26                                                    + '.parquet' -- plus appropriate fileExtension
27    , DATA_SOURCE = eds_ ' + @StorageAccount + '_' + @Container + '
28    , FILE_FORMAT = ParquetFileFormat
29    )
30    AS
31    SELECT *
32    FROM ' + @viewName + ';
33    ' --
34    --select @sqlCmd
35    EXEC sp_executesql @sqlCmd
36
37    END -- end PROC
38    GO
```

# DEMO

## Beispiel 3

tsql\_cetas files



# .csv vs .parquet

Save Discard Download Refresh

Overview Versions Edit Generate SAS

```
1 1, Frankfurt-Main
2 2, Frankfurt-Main
3 3, Frankfurt-Main
4 4, Mönchengladbach
5 5, Mönchengladbach
6 6, Mönchengladbach
7 7, Mönchengladbach
8 8, München
9 9, München
10 10, Frankfurt-Main
11 11, Frankfurt-Main
12 12, Frankfurt-Main
13 13, Frankfurt-Main
14 14, Frankfurt-Main
15 15, Frankfurt-Main
16 16, Frankfurt-Main
```

Upload Add Directory ...

**Authentication method:** Access key ([Switch to Microsoft Entra user account](#))

**Location:** bwtde1cloud4iotfilesystem / zTest\_delta / sampledata / fromdatabricks

Search blobs by prefix (case...)

☐ Show deleted objects

Name

<input type="checkbox"/>	Folder	[.]	...
<input type="checkbox"/>	Folder	a1_Ort.csv	...
<input type="checkbox"/>	Folder	a2_UID.csv	...
<input type="checkbox"/>	File	a1_Ort.parquet	...
<input type="checkbox"/>	File	a2_UID.parquet	...

Save Discard Download Refresh Delete

Overview Versions Edit Generate SAS

⚠ The file 'zTest\_delta/sampledata/fromdatabricks/a1\_Ort.parquet' may not re

```
1 PAR1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
```

# .csv vs .parquet

Save Discard Download Refresh

Overview Versions Edit Generate SAS

```
1 1,Frankfurt-Main
2 2,Frankfurt-Main
3 3,Frankfurt-Main
4 4,Mönchengladbach
5 5,Mönchengladbach
6 6,Mönchengladbach
7 7,Mönchengladbach
8 8,München
9 9,München
10 10,Frankfurt-Main
11 11,Frankfurt-Main
12 12,Frankfurt-Main
13 13,Frankfurt-Main
14 14,Frankfurt-Main
15 15,Frankfurt-Main
16 16,Frankfurt-Main
```

Block blob

1.9 MiB

Upload Add Directory ...

**Authentication method:** Access key ([Switch to Microsoft Entra user account](#))

**Location:** bwtde1cloud4iotfilesystem / zTest\_delta / sampledata / fromdatabricks

Search blobs by prefix (case...)

☐ Show deleted objects

Name

<input type="checkbox"/>	Folder [..]	...
<input type="checkbox"/>	Folder a1_Ort.csv	...
<input type="checkbox"/>	Folder a2_UID.csv	...
<input type="checkbox"/>	File a1_Ort.parquet	...
<input type="checkbox"/>	File a2_UID.parquet	...

Save Discard Download Refresh Delete

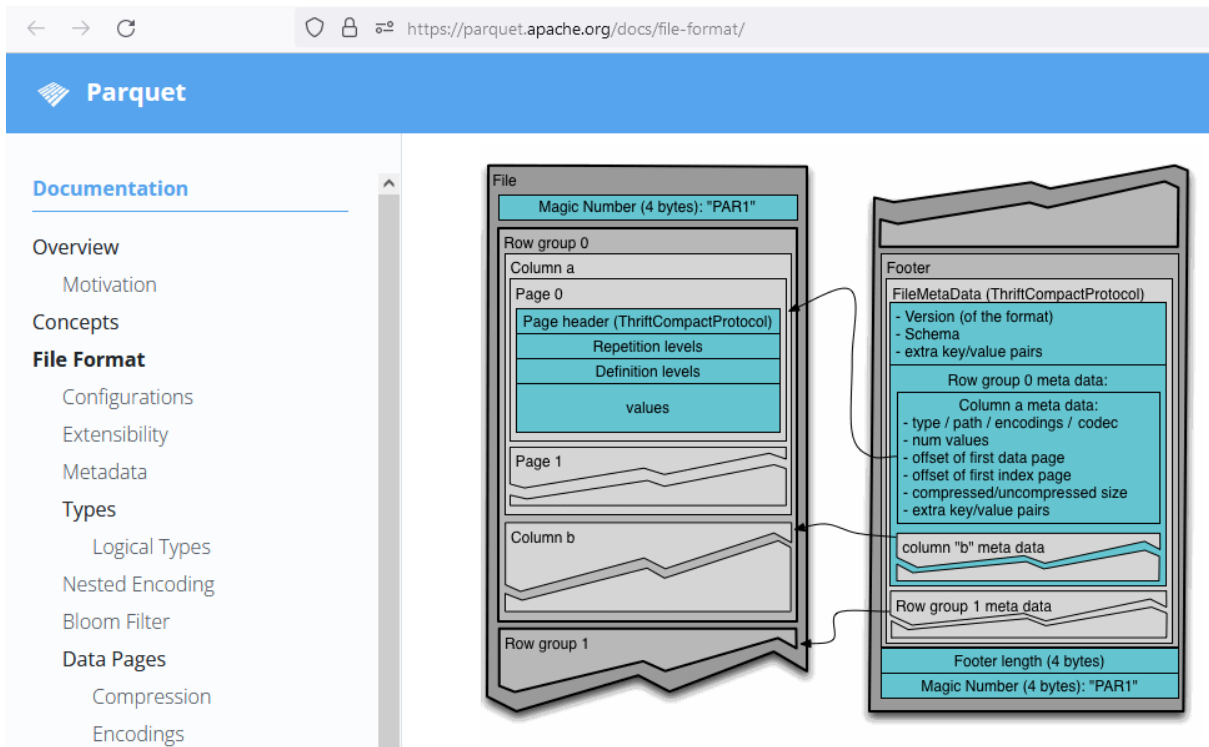
Overview Versions Edit Generate SAS

**Warning:** The file 'zTest\_delta/sampledata/fromdatabricks/a1\_Ort.parquet' may not re

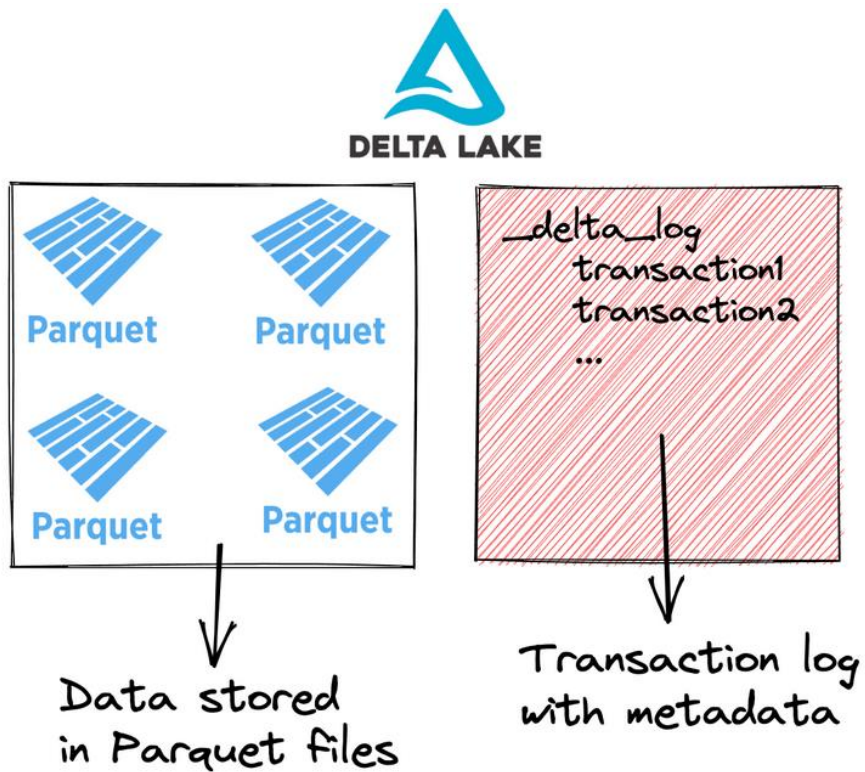
```
1 PAR1
2 
3 
4 
5 
6 
7 
8 
9 
10 
11 
12 
13 
14 
15 
16
```

Block blob

391.48 KiB



<https://parquet.apache.org/docs/file-format/>



# DEMO

## Beispiel 4

.parquet vs. .delta-Files



## LINKS

- [https://github.com/sql-fabrik/AzureSynapse\\_gettingStarted](https://github.com/sql-fabrik/AzureSynapse_gettingStarted)
- >> README.md
- <https://www.databricks.com/de/glossary/medallion-architecture>
- <https://learn.microsoft.com/de-de/azure/synapse-analytics/sql/develop-tables-cetas>
- <https://parquet.apache.org/docs/file-format/>
- <https://www.databricks.com/de/glossary/what-is-parquet>
- <https://delta.io/blog/delta-lake-vs-parquet-comparison/>





# THANK YOU TO OUR SPONSORS



Mainzer Datenfabrik



TIMEXTENDER



ITGAIN 





Thank you very much for your attention.  
Vielen Dank für Eure Aufmerksamkeit.