

Direct Lake mode in Power BI

MsBIP #77

June 15th 2023



ORANGEMAN

Just Blindbæk

- ▶ Self-employed BI consultant in justB
 - ▶ Trainer at Orange Man
 - ▶ Founder
 - Danish Microsoft BI Community (MsBIP.dk)
 - Power BI UG Denmark (PowerBI.dk)
 - Power BI Cruise, Power BI Next Step and Data Platform Next Step
 - ▶ Strong focus on
 - Azure BI architecture
 - Analysis Services
 - Reporting Services
 - Power BI
- ▶ just@justB.dk / blog.justB.dk / [@justblindbaek](https://twitter.com/justblindbaek) / youtube.com/c/justblindbaek



Microsoft
CERTIFIED
Trainer

Microsoft
CERTIFIED
IT Professional

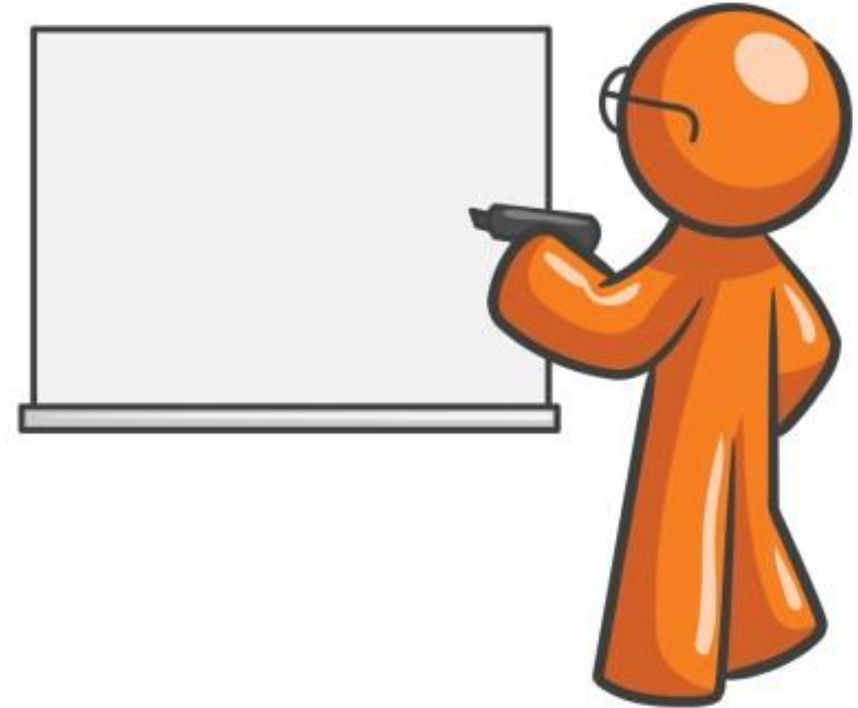
SSAS
MAESTRO
by Microsoft

MVP **Microsoft**
Most Valuable
Professional

ORANGEMAN

Agenda

- ▶ What is Direct Lake mode?
- ▶ V-ORDER compression
- ▶ Compared to Import and Direct Query
- ▶ Swapping into memory
- ▶ Analyze query processing
- ▶ Fall back to Direct Query
- ▶ Current limitations

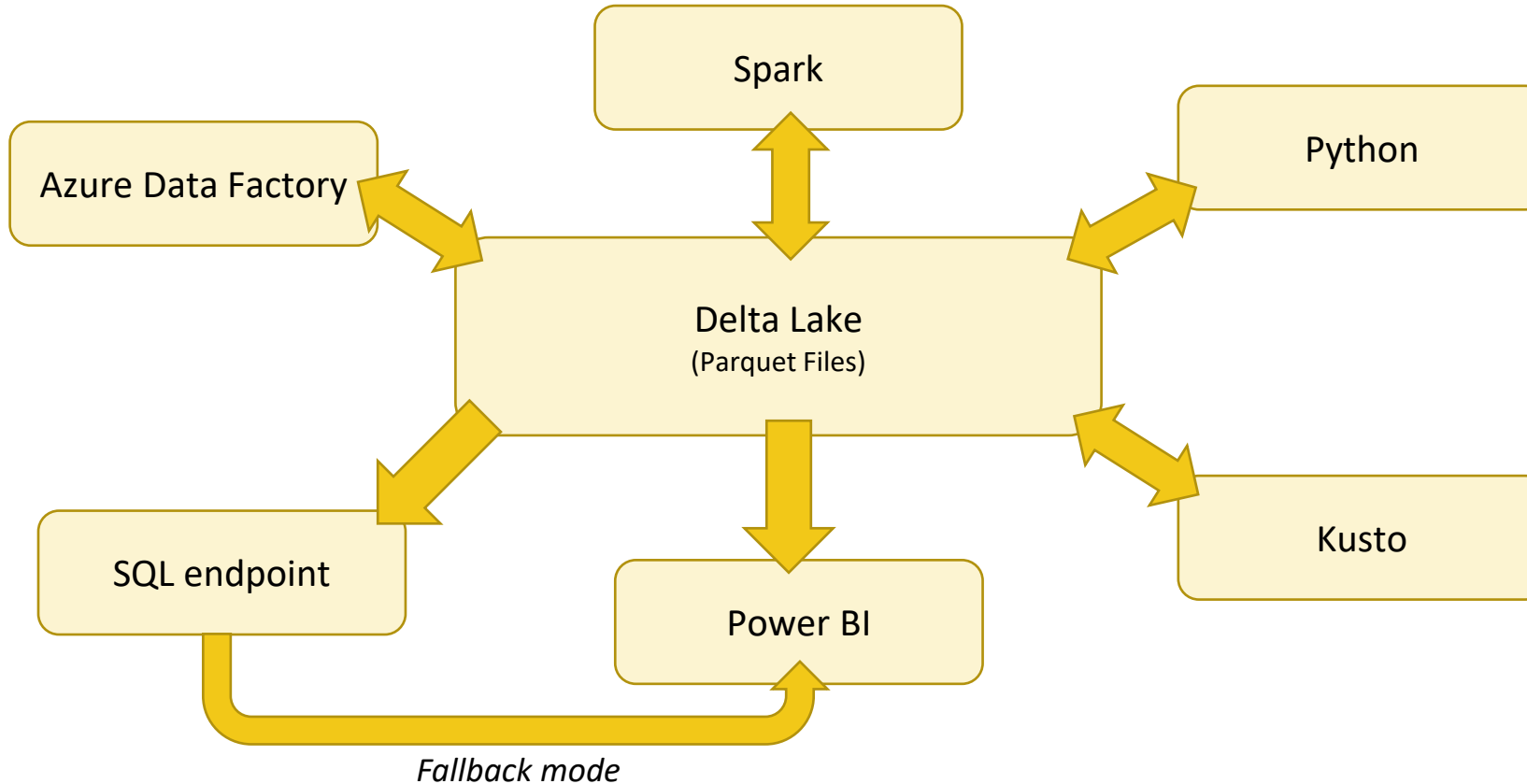


Direct Lake mode in Power BI

Using the open standard Delta Lake/Parquet format for all Fabric artefacts, users enjoy incredible performance for interactive analysis directly against Fabric OneLake, without any data movement.

- ▶ Groundbreaking new dataset capability
 - ▶ Fast-path to load the data from the lake straight into the Power BI engine
 - ▶ Eliminates the import requirement
- ▶ Amir Netz: *Direct Lake - the killer feature of Power BI*

Delta Lake for all Fabric artifacts



V-ORDER compressed Parquet

Microsoft internal Sales
Data (162 tables)



880GB



268GB



84GB

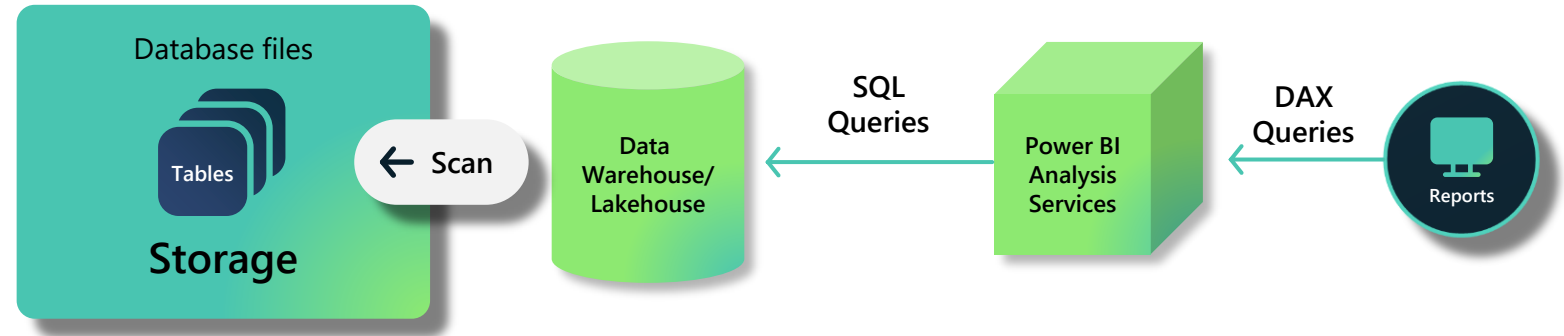
x3.2

Less I/O for all
workloads

Current storage modes

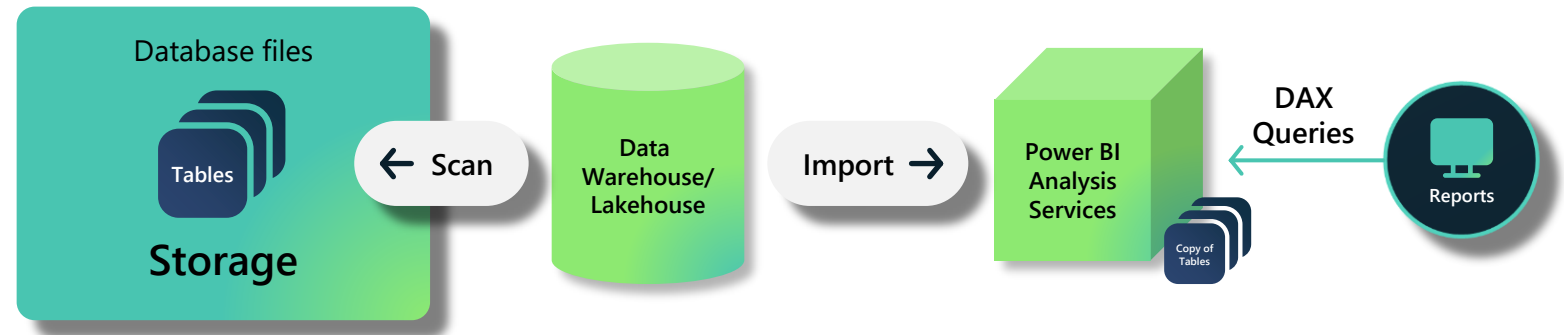
"Direct Query Mode"

Slow, but real time



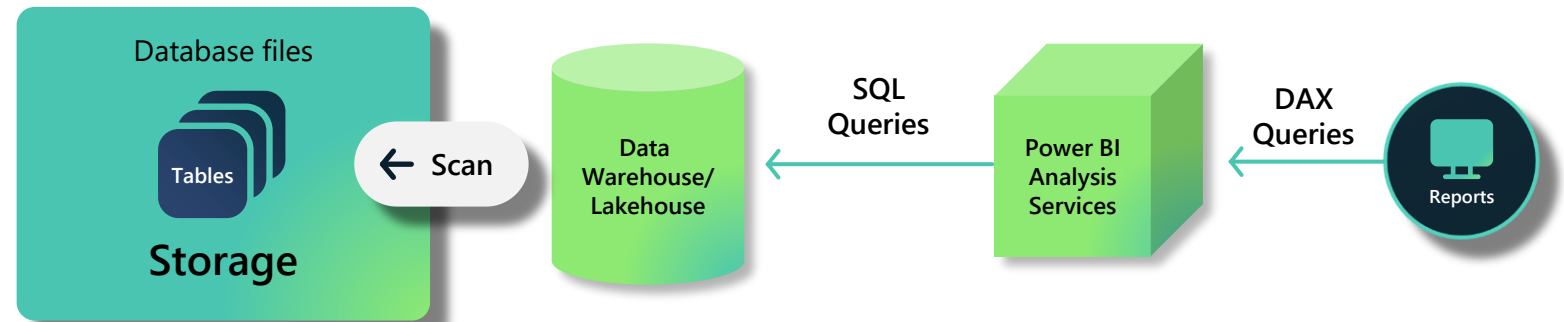
"Import Mode"

Latent & duplicative but fast



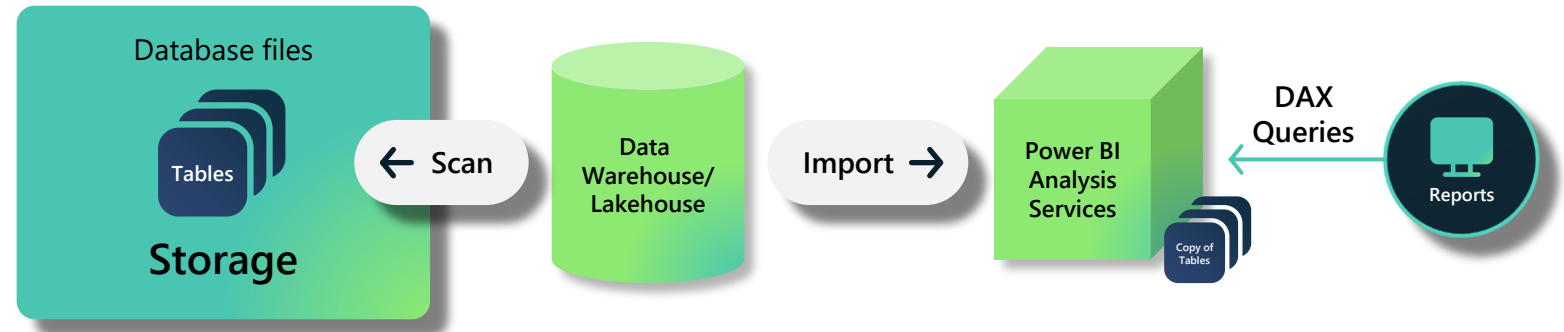
"Direct Query Mode"

Slow, but real time



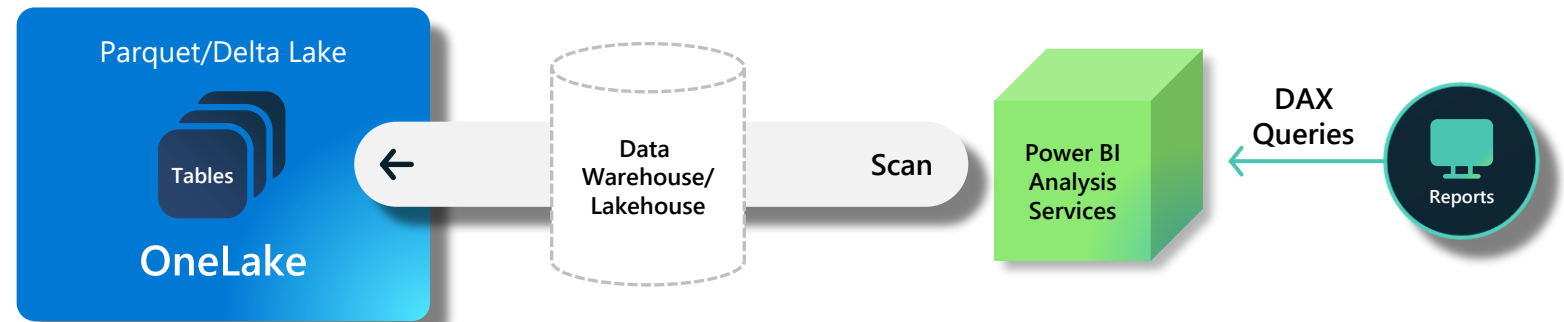
"Import Mode"

Latent & duplicative but fast



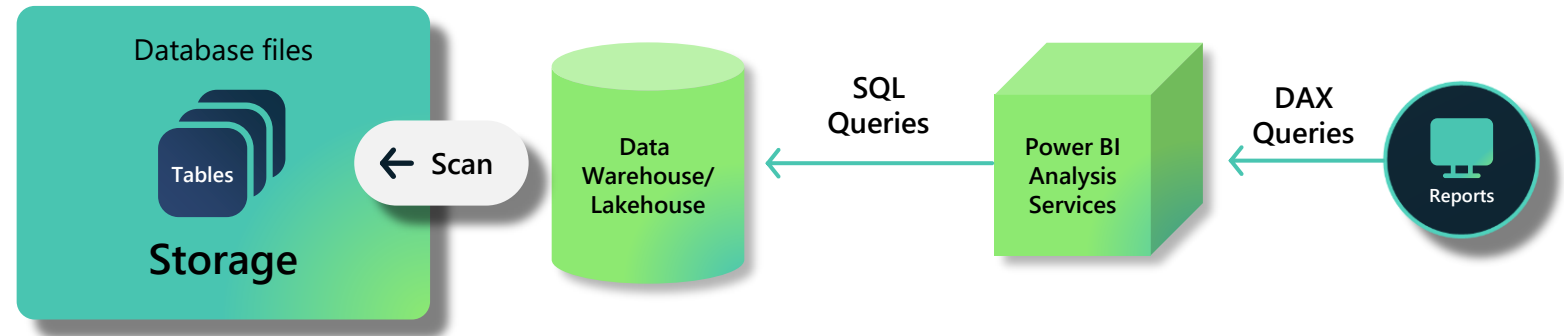
"Direct Lake Mode"

Perfect!



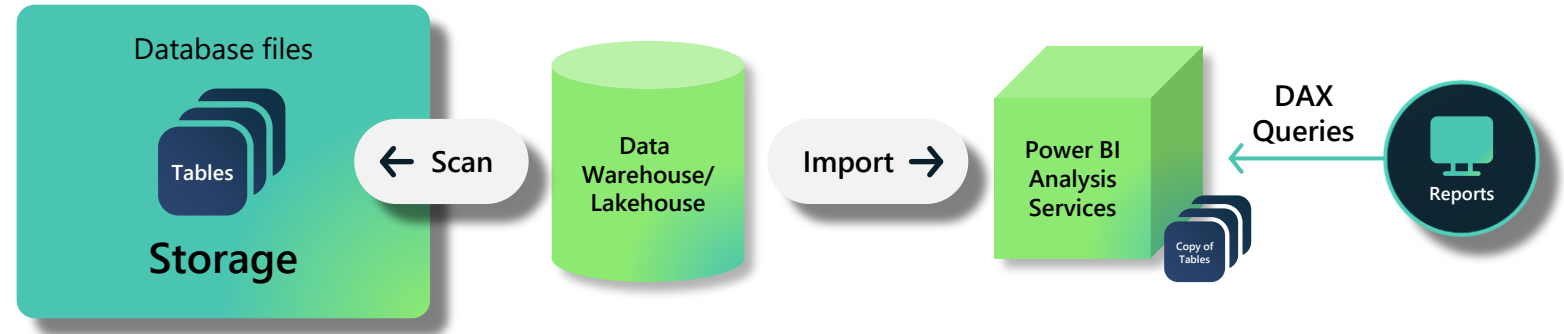
"Direct Query Mode"

Slow, but real time



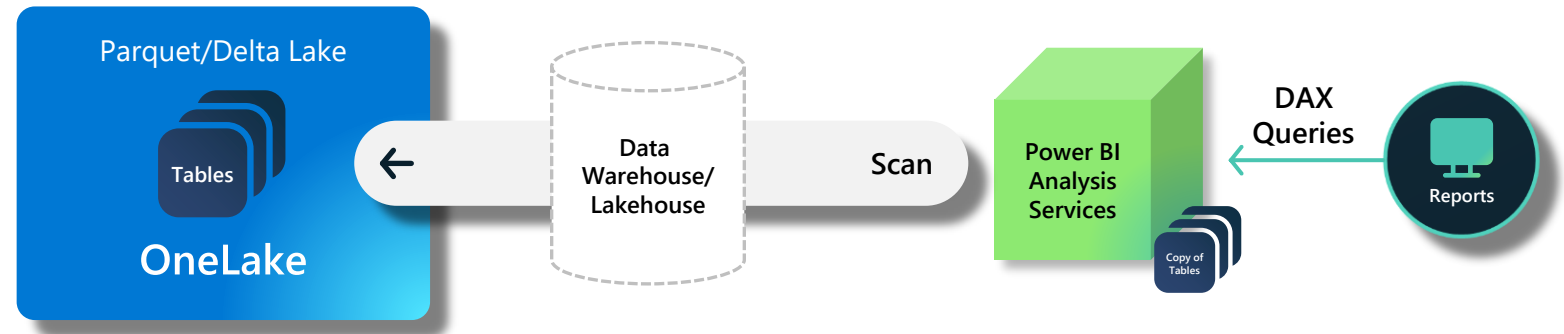
"Import Mode"

Latent & duplicative but fast



"Direct Lake Mode"

Perfect!



VertiPaquet as the Storage format

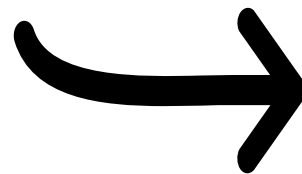
SSAS, AAS, Power BI large models

C:\Program Files\Microsoft SQL Server\MSAS15.MSSQLSERVER\OLAP\Data\AdventureWorksTabular.0.db\DimCustomer (10).tbl\238 prt				
Name	Date modified	Type	Size	
1.DimCustomer (10).AddressLine1 (78).0.idf	1/29/2020 6:36 PM	IDF File	37 KB	
1.DimCustomer (10).AddressLine1 (78).0.idfmeta	1/29/2020 6:36 PM	IDFMETA File	1 KB	
1.DimCustomer (10).AddressLine2 (79).0.idf	1/29/2020 6:36 PM			
1.DimCustomer (10).AddressLine2 (79).0.idfmeta	1/29/2020 6:36 PM			
1.DimCustomer (10).BirthDate (66).0.idf	1/29/2020 6:36 PM			
1.DimCustomer (10).BirthDate (66).0.idfmeta	1/29/2020 6:36 PM			
1.DimCustomer (10).CommuteDistance (82).0.idf	1/29/2020 6:36 PM			
1.DimCustomer (10).CommuteDistance (82).0.idfmeta	1/29/2020 6:36 PM			

Delta Lake/Parquet

Finance		
Lake view	Table view	
Tables		
> ACR Adjustment Type		
> Adjustment Type		
> Budget		
> Business		
> Calendar		
> Consumed Revenue		
> Forecast		
> Forecast Type		
> Future Flag		
> ...		

Tables > Revenue		
Name ↑	Size	Type
Folder _delta_log	2 items	Folder
part-00000-87858576-90b7-4aff-8c9e-69dcc52dbt	8.4 GB	PARQUET
part-00001-631fb085-0591-46b8-a0b5-0fec8f2255	8.4 GB	PARQUET
part-00002-0469bb29-daa4-4ecd-a3ee-bb90331a6	8.4 GB	PARQUET
part-00003-27e6062b-4d55-4469-b285-7cb2a2f32	8.4 GB	PARQUET
part-00004-b12eea8e-f255-41fa-a943-a78def57ce	8.4 GB	PARQUET



Swapping in and out of memory

- ▶ Also known as framing
- ▶ It's nothing new
- ▶ Dataset evictions and dynamic memory management is a core concept on the architecture behind Power BI Premium Gen2
- ▶ Relevant data pages are loaded on-demand (paged into memory)

Refresh

Keep your Direct Lake data up to date

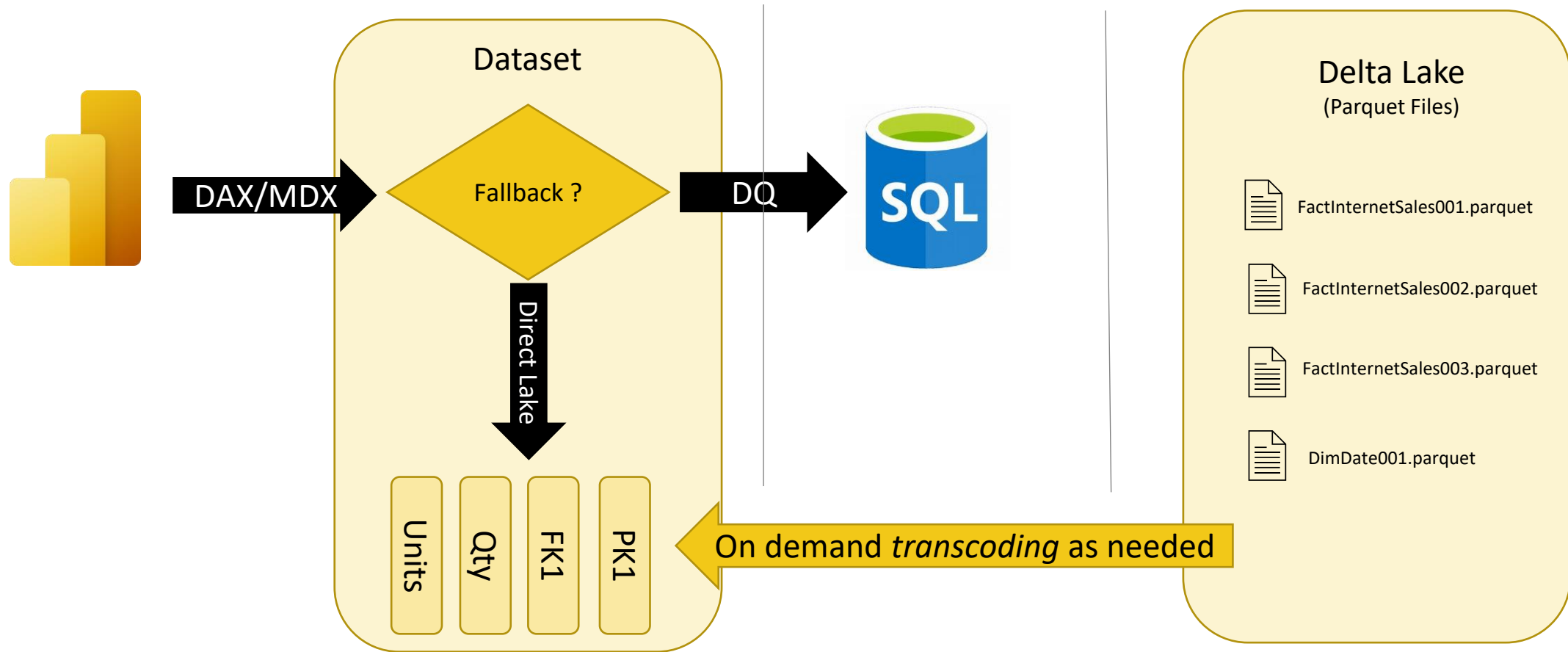
Configure Power BI to detect changes to the data in OneLake and automatically update the Direct Lake tables that are included in this dataset. [Learn more](#)



Analyze query processing

- ▶ Performance Analyzer in Desktop
- ▶ See DMV data in DAX Studio
- ▶ Parquet file storage
- ▶ Discover Dictionary
 - IsPageable
 - IsResident
 - Temperature?

Fall back to DirectQuery



Limitations

- ▶ **PREVIEW !**
- ▶ No XMLA end-point – only web authoring
- ▶ Direct Lake size limits – if reached, then DirectQuery mode
- ▶ Calculated columns and calculated tables are not yet supported
- ▶ Some data types may not be supported
- ▶ Only Single Sign-On (SSO) is supported
- ▶ No Role Level Security... Waiting for OneSecurity 😊

Useful ressources



- ▶ Direct Lake: <https://learn.microsoft.com/en-us/power-bi/enterprise/overview>
- ▶ How direct lake mode works: <https://learn.microsoft.com/en-us/fabric/data-engineering/lakehouse-pbi-reporting>
- ▶ Direct Lake - the killer feature of Power BI: <https://twitter.com/AmirNetz/status/1661045933256364032>
- ▶ Power BI Direct Lake Mode : Frequently Asked Questions: <https://fabric.guru/power-bi-direct-lake-mode-frequently-asked-questions>
- ▶ Delta Lake table optimization and V-Order: <https://learn.microsoft.com/en-us/fabric/data-engineering/delta-optimization-and-v-order?tabs=sparksql>

Questions



ORANGEMAN