THE INFOTRUST QIK ...

Managing your data assets in Qlik environment

May 20-22 Qlik Customer Day 2025 Theinfotrust.com



Milda Vitalyte-Masaite Qlik Architect, The Infotrust

Qlik Enterprise Architect and Qlik certified trainings instructor.

Delivering the most complex projects for the largest Baltics and international customers. 16+ years of experience with Qlik.



Simas Baranauskas

Business Intelligence Solutions Architect, The Infotrust

Data management evangelist and Microsoft data platform architect. Success record of 50 projects within 20 years of experience.



Variety of data architectures

LATE 1980'S

Data Warehouse





2011

2020

Data Lake



Structured, Semi-Structured and Unstructured Data

Lakehouse



Structured, Semi-Structured and Unstructured Data

Should I stay (on qvds) or should I go (data platform)?

Decision point: stay on-prem and Move to Cloud

Move to Cloud

- Alerting, Reporting, Automate, Qlik Predict (ex. AutoML), Qlik Answers (gen2), data lineage, data flow
- Capacity based:
 - unlimited users (data is value metrics)

• •

- capacity in cloud will mirror your on-prem resources
- 3Y of parallel usage: enough time to migrate even the most complex landscapes
- Migration tool (available for partners and customers Q2)
- Hybrid usage (data prep on-prem, consumption in Qlik Cloud): drive user adoption, data prep migration at your pace

Stay on-prem

Less and less reasons to stay

All usual benefits of cloud: Qlik takes care of HW resources, most of administration, upgrades. Major innovations only in cloud.



LOCK-IN: DATA STORAGE

Recap: 1.5 year ago, in Qlik Customer Day



LOCK-IN: DATA STORAGE

PARQUET VS QVD

File size:

- Parquet (Brotli Compr.): up to 121% smaller than equivalent QVD files
- Parquet (Snappy Compr.): up to 81% smaller than QVD files

Load Performance in Qlik:

- Parquet: Faster in almost every scenario
- **QVD Files**: faster only when full loading all rows and columns of large datasets

File Size comparison									
Table Name	Source Type	Parquet (BROTLI)	Parquet (GZIP)	Parquet (HADOOP)	Parquet (LZ4)	Parquet (SNAPPY)	Parquet (UNCOMPRESSED)	Parquet (ZSTD)	
CUSTOMER		84,345,839	90,704,767	141,952,132	141,948,772	134,617,491	255,309,411	96,152,572	
LINEITEM		1,862,649,877	1,972,080,892	2,633,180,140	2,633,015,997	2,600,742,142	4,163,351,636	2,184,179,203	
NATION		2,220	2,381	2,866	2,802	2,737	3,735	2,383	
ORDERS		446,217,809	473,853,252	704,552,715	704,526,262	681,663,521	1,326,943,725	540,944,585	
PART		43,683,886	48,608,972	74,713,612	74,710,092	68,330,514	139,175,340	50,205,063	
PARTSUPP		269,205,591	290,485,906	502,585,503	502,572,063	460,322,114	1,193,426,726	333,194,541	
REGION		1,157	1,260	1,305	1,257	1,268	1,340	1,223	
SUPPLIER		5,449,336	5,815,106	9,126,052	9,125,852	8,579,432	15,976,584	6,195,265	

Green = Smallest | Red = Biggest





Reload Duration Comparison (Median) - LINEITEM Table ~60M rows



LOCK-IN: DATA STORAGE

LAYERING OF QLIK'S QVD

- "Layers" a.k.a "Medallion" architecture.
- Almost lakehouse approach, but not open until QVDs are replaced by Parque files.

Source	Э	
PostgreSQL	:	
MySQL	:	
PostgreSQL	1	->
MySQL	: -	





Should I stay (on qvds) or should I go (data platform)?

The case of staying on qvd/parquet even in Cloud

- Limited number of data (and sources)
- No other drivers for data lakehouse, data warehouse (not all data is "big data", no IOT,....)
- Qlik to be used for most of consumption use cases (Analytical dashboards, Reporting, Alerting, AutoML/Predict..)
- Data Flows better support now full self service
- Qlik Talend with it's data move (RT) are included in Qlik Cloud Analytics
- Data platforms are still evolving > Iceberg ahead.



Go data platform?

Recommended options

- Qlik Talend (data movement, transformations, data quality, data products) -> any cloud data platform
- Qlik Open lakehouse
- Snowflake, dbt, ADF

Data Lakehouse

Combining the best of data lakes and data warehouses



Data Lakehouse

- Unified Workloads
- Governance
- Interoperable
- storage
- Schema-on-read or

Fast, governed access for all data and users

Data Warehouses

SQL

- Analytics
- Performance
- Reliable transactions
 - Schema-on-write

Fast access, many users

Why Apache Iceberg?









WHY INDUSTRY FOCUSES ON APACHE ICEBERG?

ΟΡΕΝΝΕSS

- Compatible across multiple analytics engines
- Cloud-agnostic seamlessly works with AWS, Azure, Google Cloud, or hybrid setups.
- Supported widely, reducing risk of vendor lock-in.

• •

PERFORMANCE AND SCALABILITY

 Incremental changes handled efficiently, ideal for large data sets.

- Concurrency Control optimistic concurrency ensures smooth simultaneous access.
- Robust Metadata Management

INDUSTRY ADOPTION

- Open-source Leadership governed transparently by Apache Software Foundation
- Rapidly increasing adoption and contributions from industry leaders.







 Community-driven innovations, ensuring futureproofing.

PARQUET & ICEBERG



• Efficiently store your data





 Manages these Parquet files as organized, transactional tables



ICEBERG UNDER THE COVERS



ICEBERG TABLES

Open standard for creating, updating, optimizing and querying large analytics tables on object stores.

- **Open format for managing table metadata** (schema, partitions, types, etc.)
- **Insert / Update / Delete** \rightarrow work with data just like in a database
- **ACID transactions** \rightarrow consistency guarantees needed for reliability
- **Schema evolution** \rightarrow automatically adapt to changing data models
- **Dynamic partitions** \rightarrow dynamically adapt to changing query patterns
- **Pluggable data store format** (Parquet, Avro, ORC) \rightarrow analytics today, transactional or vector tomorrow
- **Storage optimizations** \rightarrow continuously optimize data for best performance



Build a Snowflake-managed Iceberg Lakehouse with Qlik Talend Cloud Pipelines



Supported Sources

Qlik...

 \checkmark





COST AND GOVERNANCE ADVANTAGES

- Avoid Duplicate Storage
- Optimized Spending
- Unified Governance open, self hosted catalog or in Snowflake
- Performance & Scalability





STRATEGIC ARCHITECTURE OUTLOOK

Agility & Future-Proofing

hostage, so you can evolve your architecture as needed.

Multi-Cloud & Hybrid

clouds using the same data. This flexibility is key for resilience and optimization.

Accelerating Innovation

analytics since data isn't locked in silos – more innovation and insights from the same data.

Vendor Leverage

vendors to provide value (since you're not irreversibly locked to them).



Open, portable data means you can adopt new technologies or switch platforms with minimal friction. No single vendor can hold your data

An open data layer can span clouds and on-prem seamlessly. You avoid cloud vendor lock-in and leverage best-of-breed services on different

When any team can access data with the tools of their choice, experimentation flourishes. Open data fosters AI/ML and cross-domain

Strategically, using open standards increases your leverage in vendor negotiations. You can opt in or out of services more freely, pushing



Lean architecture: Snowflake, dbt, ADF, Qlik

















DBT



Project "E Database	
= Group	RANDVEL_RIVILE.I06_PARH source table
Sources	Details Description Columns Referenced By SQL
CONSTRO_APP	
CONSTRO_RIVILE	Referenced By
RANDVEL_RIVILE	
DB DB	Models
I02_DKH	
I03_DKD	Randvel DimDokumentas
E 104_ATH	 Randvel_FaktaiDK
I06_PARH	Randvel_Faktai
I07_PARD	Randvel_Pardavimai
E 127 FIFO	Randvel_DimPardavimuDok
□ 144 SKOL	Randvei_Pardavimari
	Code
	Sample SOL
	2 DB.
	3 I06_KODAS_PO,
	4 I06_OP_TIP,
	5 I06_VAL_POZ,
	6 106_PVM_TIP, 7 106 OP STORNO.
E] N23_PRMD	8 I06_DOK_NR,
E N25_KOD	9 I06_OP_DATA,
N26_KOMP	10 I06_DOK_DATA,
N37_PMAT	11 106_KODAS_MS,

Description	Columns	Referenced By	SQL
iced By			

	← → C C cloud.getdbt.com/accounts/232890/jobs/653634/docs/#!/model/mod							
	★ Bookmarks 🛛 🗱 Worksheets - Snowf 💧 ebooks							
DBT	dbt	Search	Search for models					
		•						
	E N09_VIET	Rand	vel Parc					
	N15_MEN		_					
	N17_PROD	Details	Description					
	N18_MAT							
	N19_PGR	54	case					
	N21_SKAI	55	wr el					
	N22_PRMH	50	case					
	N23 PRMD	58	wł					
		59	el					
		60	FROM RANDV					
		61	left j					
	E N37_PMAT	62	Left j					
	■ N40_ABAR	64	or					
	N45_VAL	65	Pr					
	N47_PPAR	66	left j					
		67	or					
	Projects	68	left j					
	my_new_project	70	left f					
	🚔 macros	71	or					
	generate_schema_name	72	left j					
	🗁 models	73	or					
	Constro	74	left j					
	🖨 Randvel	75	or left f					
		76	or					
		78	left j					
	Facts	79	or					
	Randvel_Faktai	80	left j					
	Randvel_FaktaiCF	81	or					
	Randvel_FaktaiDK	82	where "I06					
	Randvel_Pardavimai	4						
	🗀 Test							

-

davimai table

and the second second

Columns Depends On Code

```
hen sh."I06_OP_TIP" in (52) then sl.I07_KIEKIS/nullif(sl.I07_FRAKCIJA,0) *(-1) * mat.N37_BRUTTO
lse sl.I07_KIEKIS/nullif(sl.I07_FRAKCIJA,0) * mat.N37_BRUTTO end
                                                                          as "L_Brutto",
hen sh."I06_OP_TIP" in (52) then sl.I07_KIEKIS/nullif(sl.I07_FRAKCIJA,0) *(-1) * mat.N37_NETTO
lse sl.I07_KIEKIS/nullif(sl.I07_FRAKCIJA,0) * mat.N37_NETTO end
                                                                          as "L_Netto"
                                              as sl
VEL.STG.I07_PARD
join RANDVEL.STG.I06_PARH
                                               as sh
n sh.I06_KODAS_PO = sl.I07_KODAS_PO
join RANDVEL.STG.N08_KLIJ
                                              as k
on sh."I06_KODAS_KS" = k."N08_KODAS_KS"
rekes
join RANDVEL.PBI.DimPreke
                                                               as pr
on sl."I07_KODAS" = pr."Prek kodas" and sl.I07_TIPAS = pr."Prek tipo kodas"
join RANDVEL.STG.N18_MAT
                                                   as mt
n sl."I07_KODAS_US" = mt."N18_KODAS_US"
join RANDVEL.STG.N07_IMON
                                              as im
n im.N07_KODAS_IS = sl.I07_KODAS_IS
join RANDVEL.STG.N06_OBJ
                                               as os
on os."N06_KODAS_OS" = sl.I07_KODAS_OS
join RANDVEL.STG.N06_OBJ
                                               as osc
on osc."N06_KODAS_OS" = sl.I07_KODAS_OS_C
join RANDVEL.STG.N01_ACCT
                                               as ac
n sl.I07_KODAS = ac."N01_KODAS_SS"
join RANDVEL.STG.N37 PMAT
                                               as mat
n sl.I07_KODAS = mat.N37_KODAS_PS and sl.I07_KODAS_US = mat.N37_KODAS_US
join RANDVEL.STG.N15_MEN as men
on sh.I06_KODAS_MS = men."N15_KODAS_MS"
6_OP_TIP" in (51,52,53,54,55)
```

Generic advice

- Use the momentum "Move to Cloud" -> move "as is" (capitalize on business value, unlimited user, quick and easy migration)
- selected data platform.

Deal with challenges of transformation, data volume, data sharing, ingestion, cloud sources my moving data ingestion and transformation to



THE INFOTRUS 1

Thank You

Milda Vitalyte Simas Baranauskas

The Infotrust vilnius@theinfotrust.com



Qlik CUSTOMER DAY 2025

10 M 10 M 10 M 10 M

STATES AND THE STATES

CONTRACTOR OF

-

