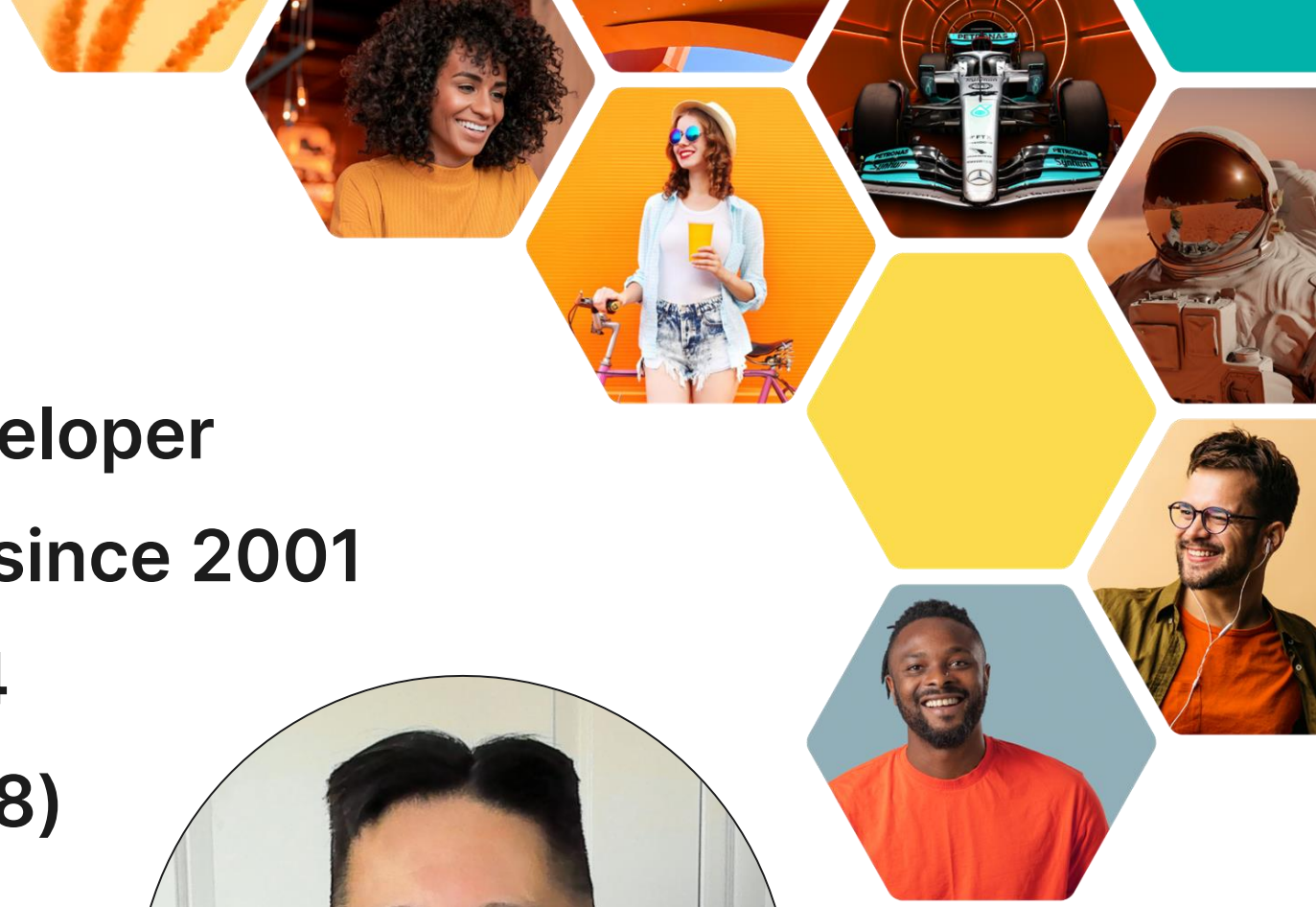
PURESTORAGE®

A New Solution for Data Archiving Using Data Virtualization

Andy Yun
Field Solution Architect
ayun@purestorage.com

Andy Yun

Field Solution Architect



- SQL Server DBA & DB Developer
- Working with SQL Server since 2001
- Speaking since Early 2014
- Microsoft MVP (2017-2018)

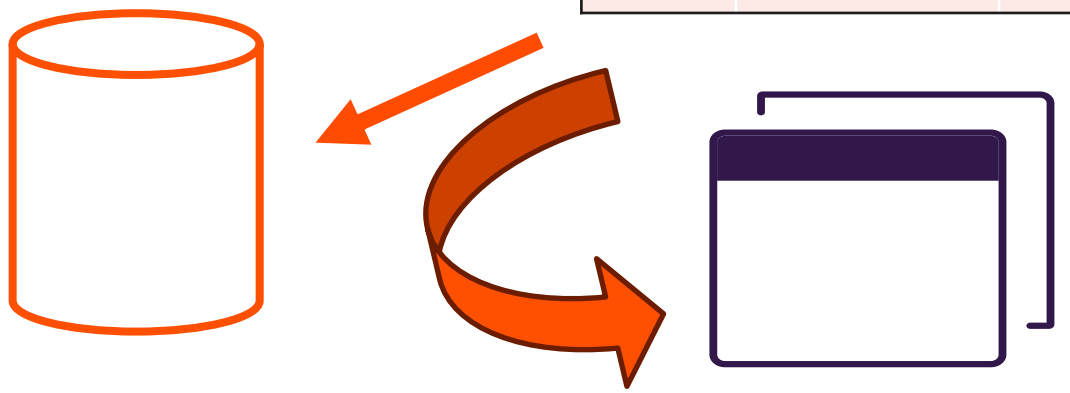
ayun@purestorage.com
sqlbek@gmail.com – sqlbek.bsky.social
https://sqlbek.wordpress.com/
https://www.github.com/sqlbek/

©2023 Pure Storage

Give Me Some Data

```
SELECT *
FROM dbo.SalesOrders
```

ID	ProductID	Date
1	AK303	2023-05-06
2	MC302	2022-11-22



Files in S3 Object Storage

©2023 Pure Storage

When You Hear…

S3 Object Storage

≠

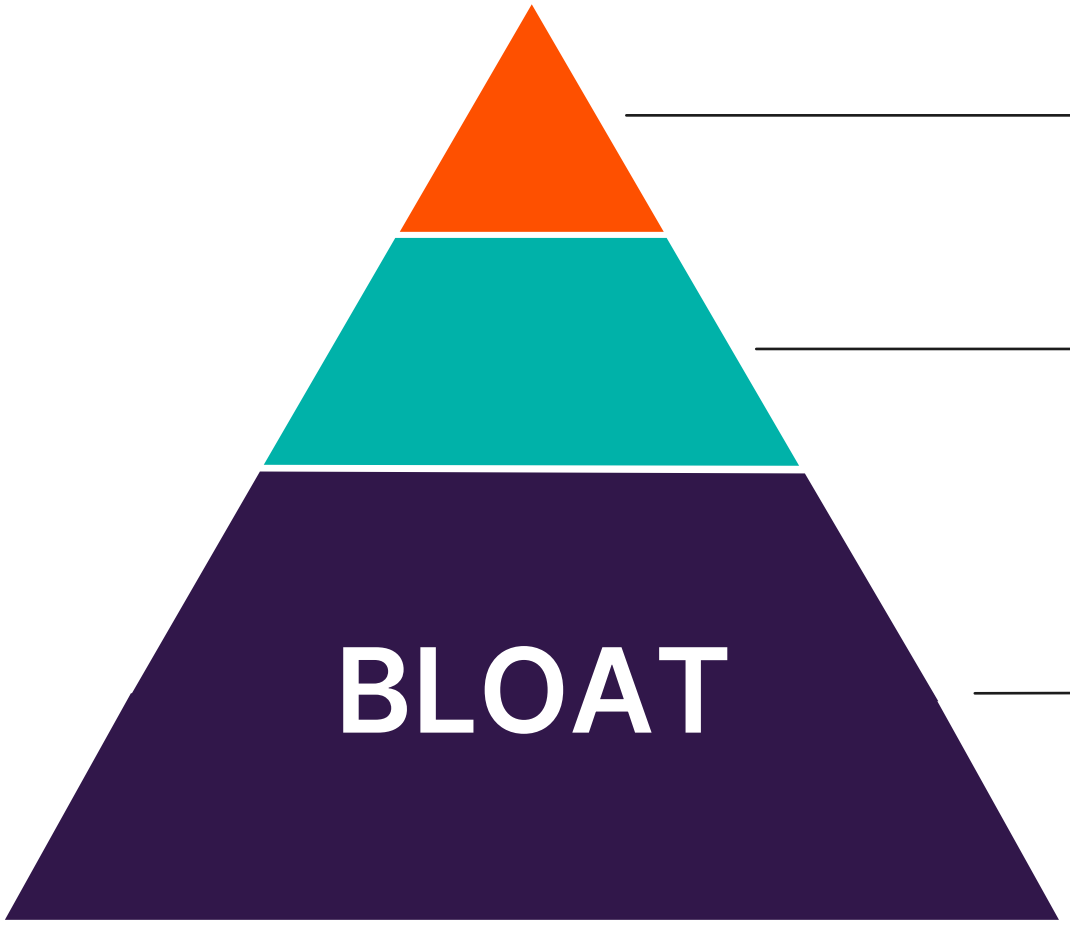
©2023 Pure Storage

Why Should You Care?

Who has a VLDB with data that will never change again?

And that data must be kept available for occasional reporting or “just in case?”

Data Distribution of a VLDB



Fresh Data

Regularly Modified Data

Static Unchanging Data

©2023 Pure Storage

Benefits of Reducing VLDB Size

- Backups, CHECKDB, Index Maintenance, HA/DR
- Full table/index scans
- Anything & everything impacted by table size

Better Performance = Happier Users

©2023 Pure Storage

SQL Server 2022 Data Virtualization

- Extract static, unchanging data into CSV or Parquet
- Park that data outside of your database
- Use External Tables to seamlessly query that data

Offload that data bloat from your VLDBs

©2023 Pure Storage

PURESTORAGE®

Demo

Parquet File Benefits

- Parquet features ROW and COLUMN elimination
- Data is compressed

Caveats

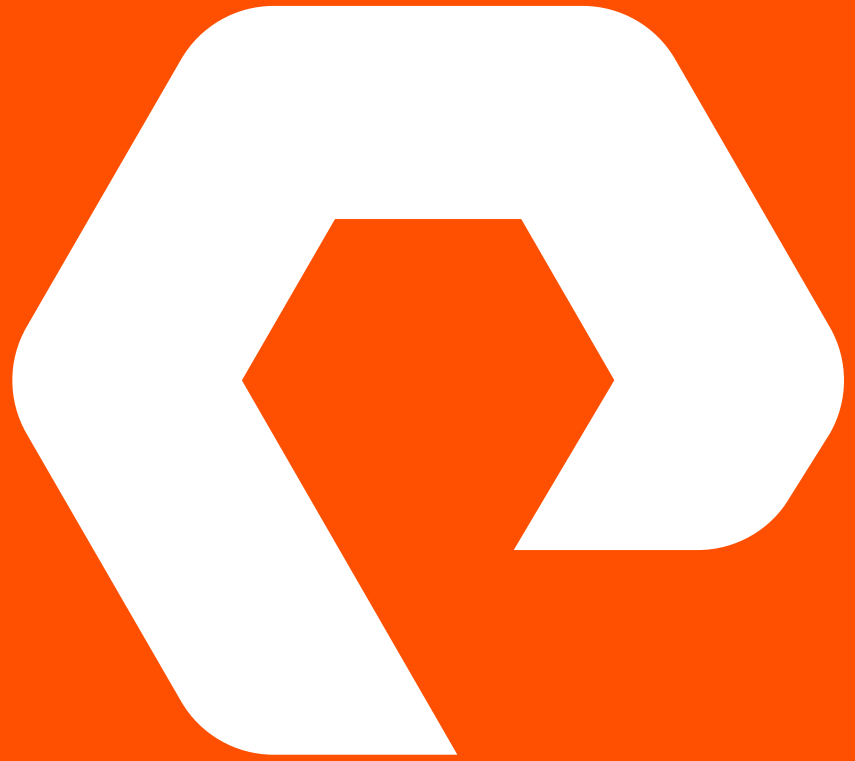
- Data is effectively static
- Strings need a collation change to support predicate elimination

©2023 Pure Storage

Remember:

S3 Object Storage

©2023 Pure Storage



Uncomplicate Data Storage, Forever

Andy Yun

ayun@purestorage.com