PURESTORAGE®

Modern Storage Platforms for SQL Server

@purestorage

September 2022

Anthony Nocentino

Principal Field Solution Architect @ Pure Storage

- Specializes in system architecture, performance, SQL Server, Kubernetes, Containers, Microsoft Azure, and VMware
- Master's degree in Computer Science

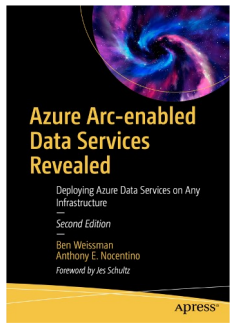
Email: anocentino@purestorage.com

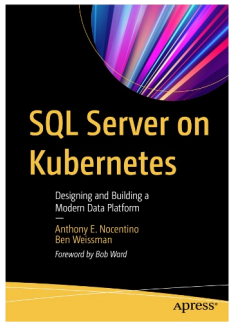
Blog: www.nocentino.com

Twitter: @nocentino

GitHub: https://github.com/nocentino/

Pluralsight Author: www.pluralsight.com





©2022 Pure Storage, Inc.

Hands on Lab Information

SQL Server 2022 and Pure Storage

0 **https://bit.ly/pssqlworkshops**

1 **https://purestorage.skytap-portal.com/**

2 Enter your passcode and open a desktop to **Windows1**

©2022 Pure Storage, Inc.

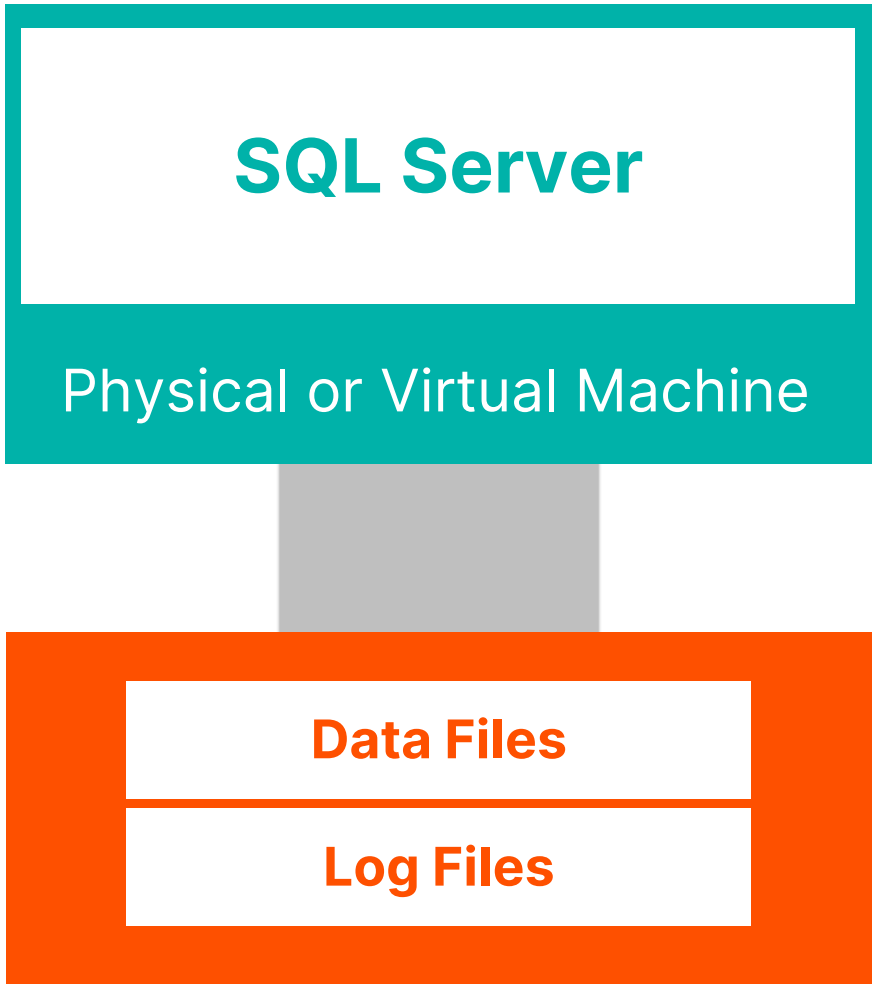
PURESTORAGE®

Storage Fundamentals for DBAs

Where data lives in a computer system...

Storage Is Where Data Lives

- SQL Server stores data...
- On disks
  - In files
  - Local
    - SCSI
    - PCIe
  - **Remote**
    - Fibre Channel
    - iSCSI
    - NVMe-oF
    - S3 compatible object storage



©2022 Pure Storage, Inc.

Key Performance Metrics

Latency	Throughput	IOPs
How long a request takes	Amount of data moved	Number of requests
Transaction log I/Os	Function of the storage interconnect	Size depends on the application
Saturation leads to queuing	Type of storage used	

©2022 Pure Storage, Inc.

Why Database Systems Care about These…

Latency	Throughput	IOPs
Does your system sell things?	Does users need reports?	Larger requests take longer
Is your system time sensitive?	Backups?	Often governed in cloud and on premises
OLTP	OLAP	

©2022 Pure Storage, Inc.

Basics of SQL Server I/O

Operation	IO Block Size
Transaction log write	512 bytes - 60 KB
Checkpoint/Lazywriter	8KB - 1MB
Read-Ahead Scans	128KB - 512KB
Bulk Loads	256KB
Backup/Restore	1MB
ColumnStore Read-Ahead	8MB
File Initialization	8MB
In-Memory OLTP Checkpoint	1MB

https://www.nocentino.com/posts/2021-12-10-sqlserver-io-size/

©2022 Pure Storage, Inc.

Let's Look Closer…at Latency

- Monitor Latency
- **`sys.dm_io_virtual_file_stats`**
- Average per file since instance startup
- Have a monitoring tool

https://www.nocentino.com/posts/2021-10-06-sql-server-file-latency/

©2022 Pure Storage, Inc.

Storage Devices

This is where your data really lives...

Block IO

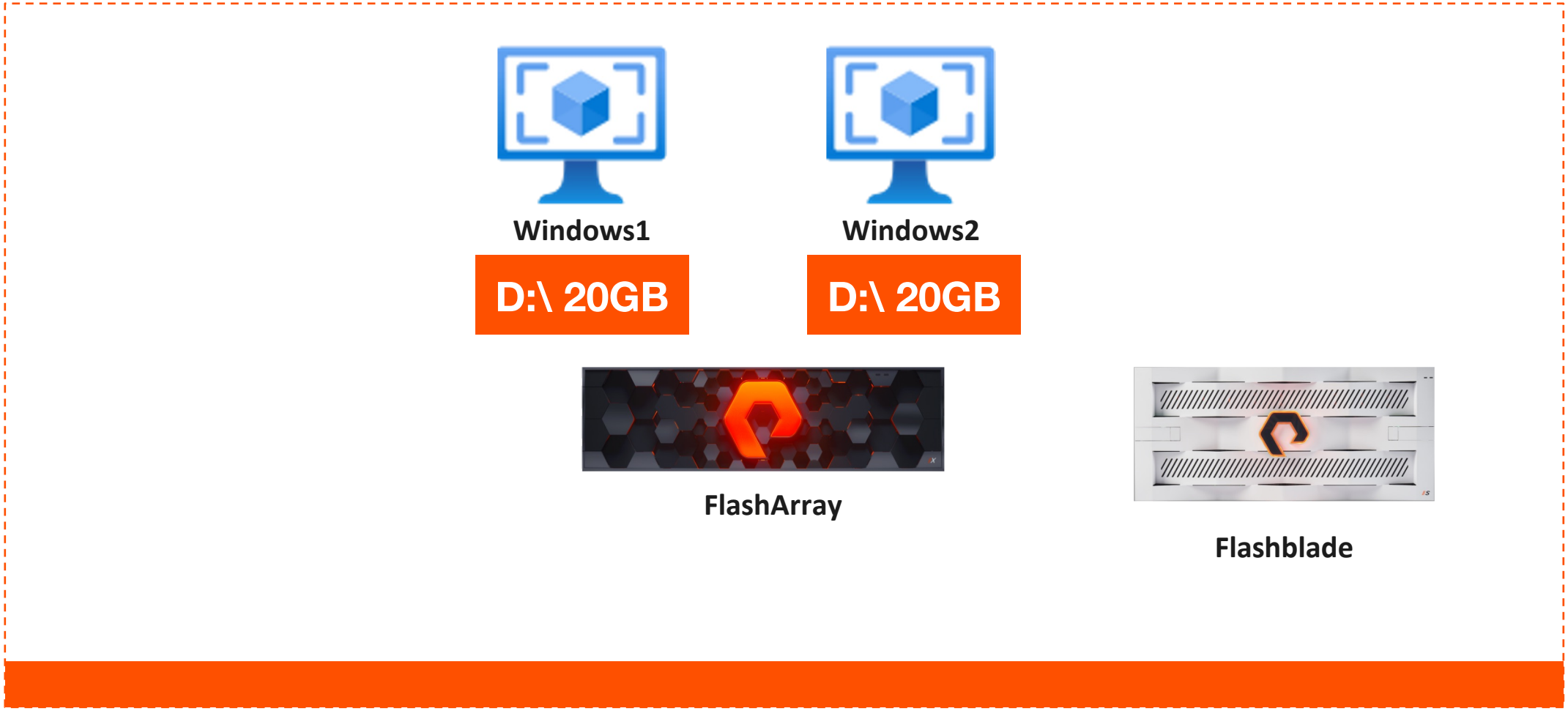
Object Storage

Encryption at Rest

©2022 Pure Storage, Inc.

Hands On Lab Architecture

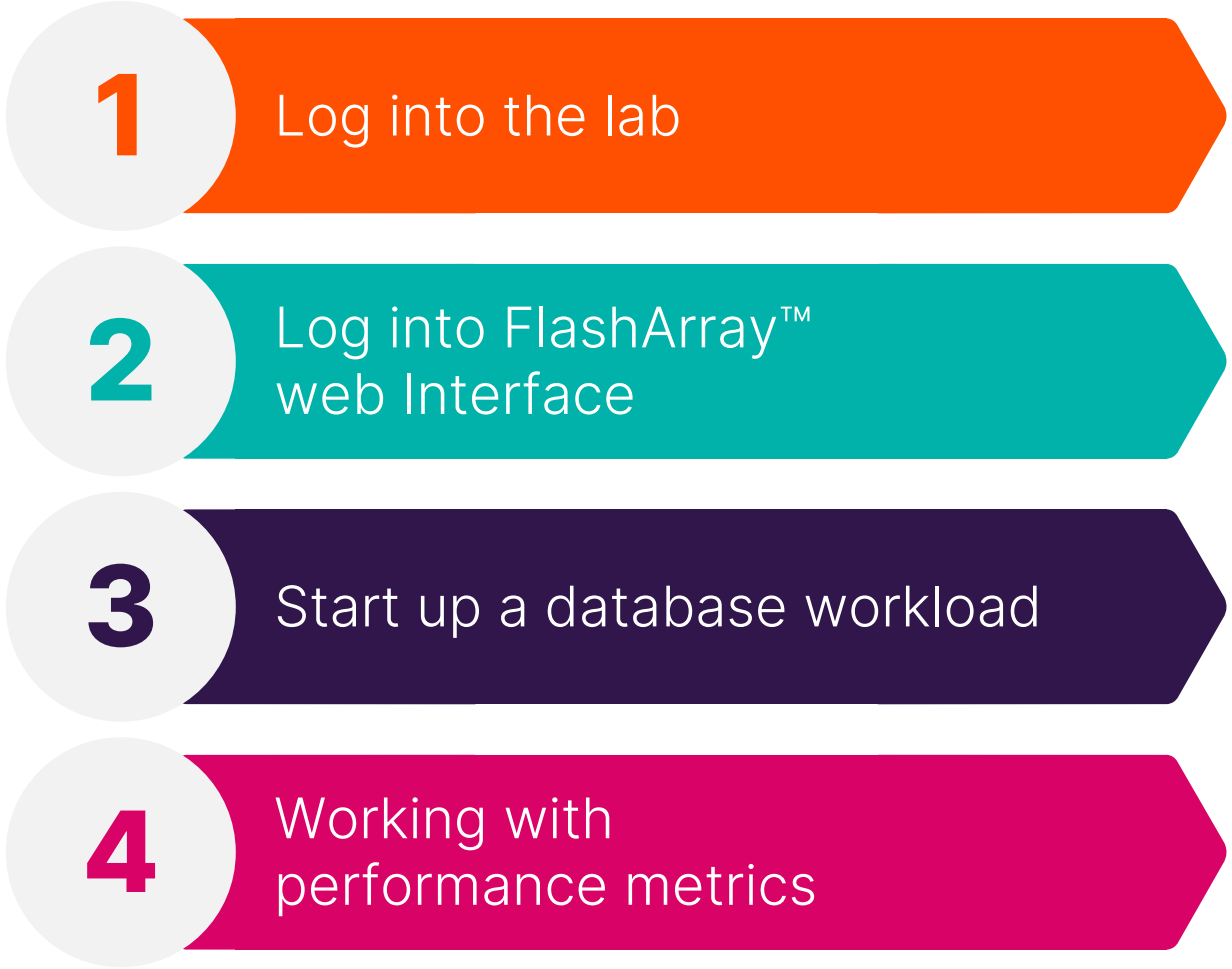
SQL Server 2022 and Pure Storage



©2022 Pure Storage, Inc.

Hands On Lab Information

SQL Server 2022 and Pure Storage



©2022 Pure Storage, Inc.

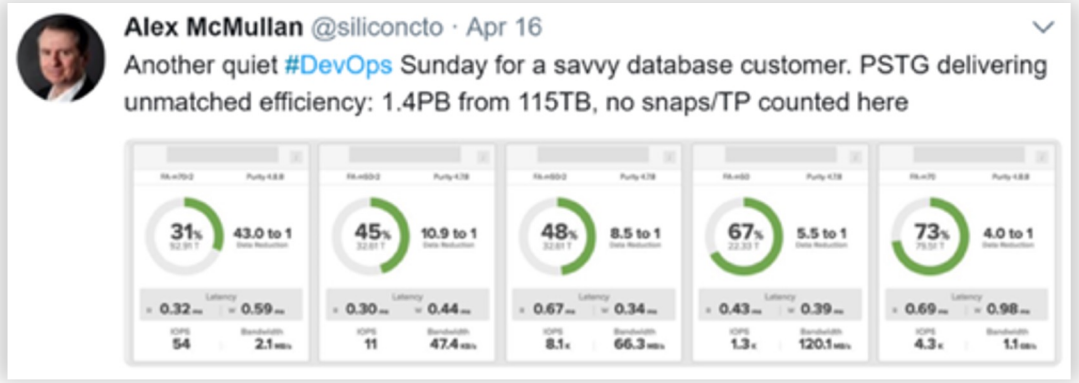
PURESTORAGE®

Storage-based Snapshots and SQL Server

Let’s bend space and time when it comes to working with large databases

Data Reduction

- In-line deduplication and compression
- Performance over data reduction
- Deep background deduplication and compression
- Data reduction that's typically 2x better than retro-fit architecture-based arrays
- Deduplication down to 512 bytes block level
- Purity delivered by an engineering team that's constantly looking to improve data reduction
- But my database is encrypted with TDE!



©2022 Pure Storage, Inc.

Snapshots of Volumes

A Snapshot is a point in time representation of data on Volumes

Why leverage Snapshots for your databases:

- Instant data + ransomware protection + protecting audit and ledger files
- Dev/Test refreshes in seconds
- In-place application and database upgrades
- Intra-Instance ETL
- Offload database maintenance

Snapshots can significantly reduce compute, networking and storage, and overhead.

Snapshots consume very little space.

Clones will not impact performance of your source volumes.

Traditional restores are expensive for your company.

Choose your tooling – PowerShell, Python, REST, and more...

Snapshots are immutable and can be protected with SafeMode.

©2022 Pure Storage, Inc.

SQL Server 2022 – Database Engine Features

Protecting Large Databases

Crash consistent vs. application consistent snapshots

- Enables point in time recovery of a database using snapshots

TSQL Based Snapshot (Cross Platform Snapshot)

- Enables point in time recovery without VSS on FlashArray™
- Availability Group Replica Seeding from snapshot
  - Build Availability Groups faster
  - Get back into a high availability posture faster after a failure or failover

Let's talk about Intel Quick Assist Technology (QAT)

https://www.nocentino.com/posts/2022-05-26-seed-ag-replica-from-snapshot/

©2022 Pure Storage, Inc.

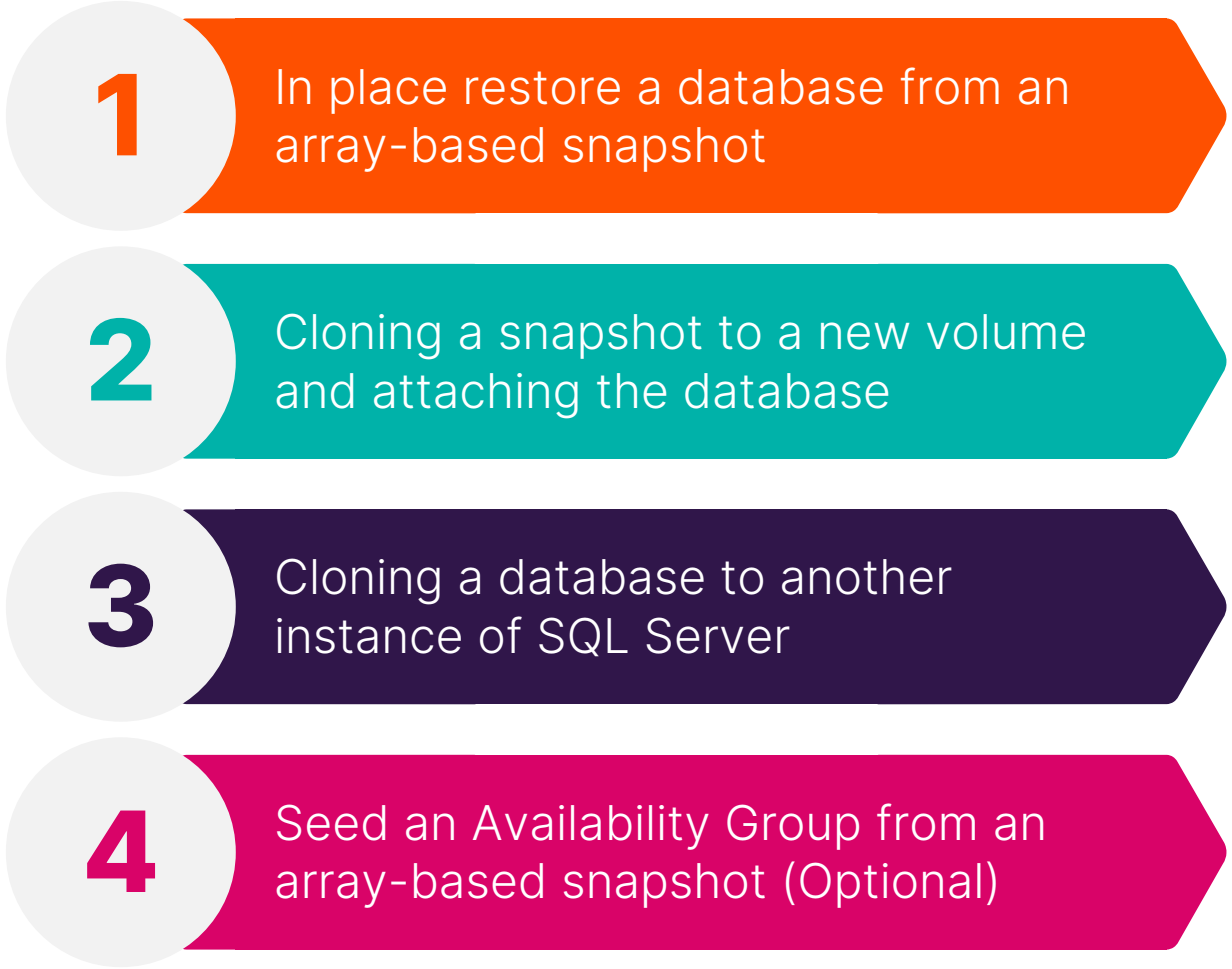
A Layered Approach to Backup and Recovery

- Replication of Snapshots to:
  - Another FlashArray™
  - FlashBlade™
  - NFS target
  - Public cloud like Amazon S3, Azure, GCP, or Cloud Block Store
- Replication of the reduced data (DR/Seeding)
- Ultra-fast backup and restore with FlashBlade™
- Use native or third-party backup tools
- Compliment your existing backup regime with snapshots of FlashArray™ Volumes

©2022 Pure Storage, Inc.

Hands on Lab Information

SQL Server 2022 and Pure Storage



Take a break after the lab – 15 minutes

©2022 Pure Storage, Inc.

PURESTORAGE®

SQL Server Object Integration: Backup and Restore

Backup and restore to s3 and performance tuning considerations

S3 – 101

Modern Object Storage and Data

- AWS Simple Storage Service (S3)
  - Storage service in the cloud
- API is open and available
- Has become the “standard” for object storage
- Companies have built their own s3 compatible object storage platforms
- Means you can get access to s3 anywhere
  - Pure Storage FlashBlade
  - MinIO
  - Many others

©2022 Pure Storage, Inc.

S3 Object Integration – Backup and Restore

Modern Object Storage and Data

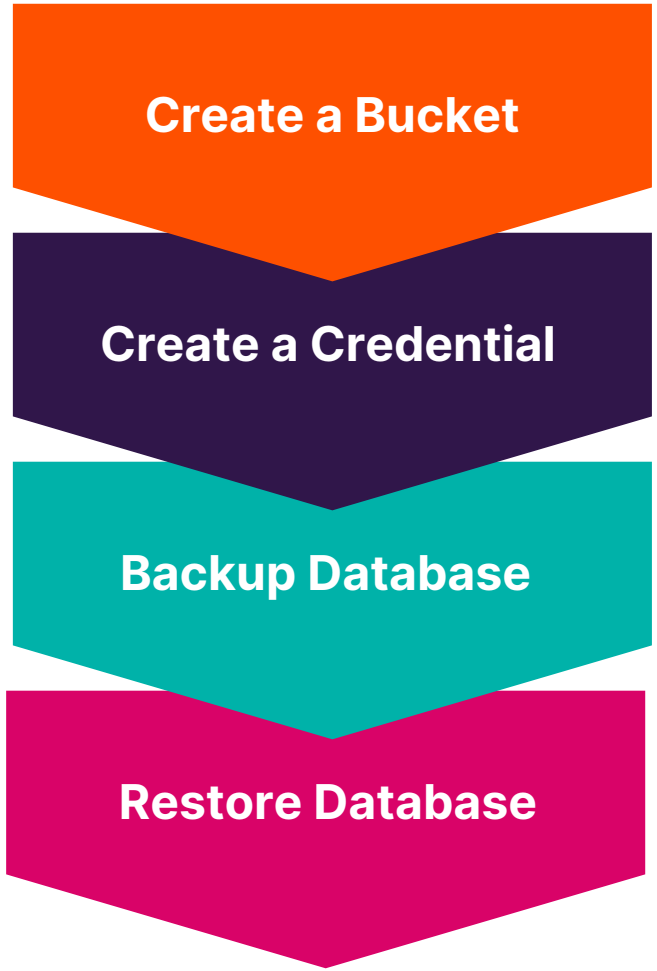
- Scale out rather than scale up
  - Single database high throughput
  - Concurrent backups
- Large environments
- Single Namespace
- Easy and native replication
- DBAs have one job
  - Get backups off the primary storage
  - Get them out of the data center as fast as possible…

https://www.nocentino.com/posts/2022-06-06-backing-up-to-s3-storage-with-sqlserver/

©2022 Pure Storage, Inc.

S3 Object Integration – Backup and Restore

Modern Object Storage and Data




```
CREATE CREDENTIAL [s3://s3.example.com/sqlbackups]
WITH IDENTITY = 'S3 Access Key', SECRET = ACCESSKEYID:SECRETKEY';
```

```
BACKUP DATABASE TestDB1
TO URL = 's3://s3.example.com/sqlbackups/TestDB1.bak'
WITH COMPRESSION, STATS = 10, FORMAT, INIT
```

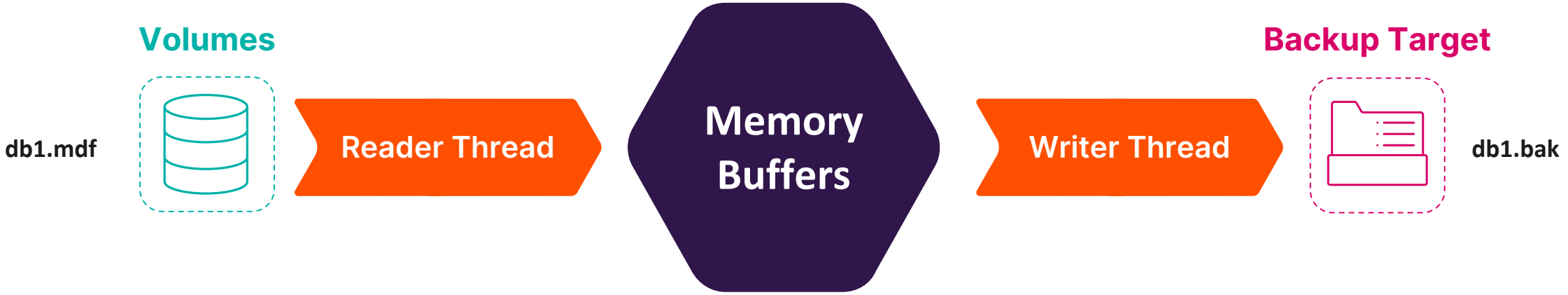
```
RESTORE DATABASE TestDB1
FROM URL = 's3://s3.example.com/sqlbackups/TestDB1.bak'
WITH STATS = 10
```

©2022 Pure Storage, Inc.

SQL Server Backup Architecture

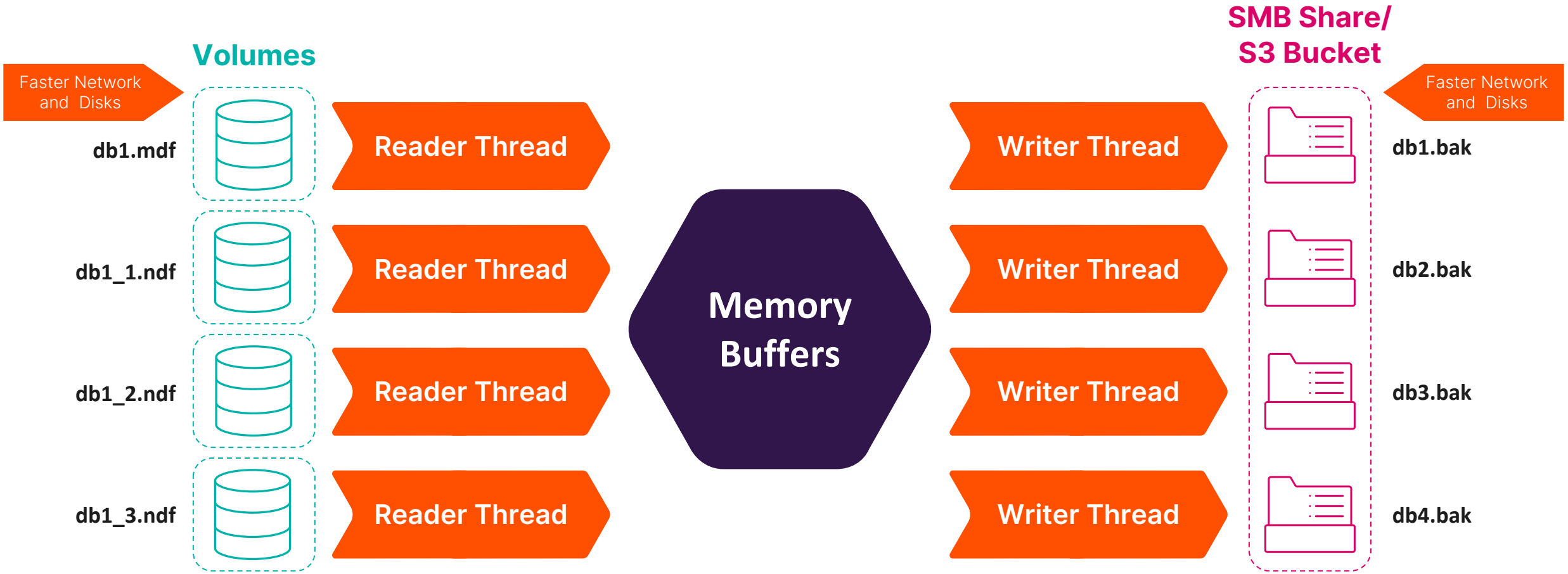
Database Files are read by the **Reader Thread** into **Memory Backups Buffers**.

The **Writer Thread** reads from backup buffers and is written to the **Backup File.**



©2022 Pure Storage, Inc.

Performance Tuning Backups



How many database files and shares/buckets should you use?

©2022 Pure Storage, Inc.

S3 Object Integration – Backup and Restore

Modern Object Storage and Data

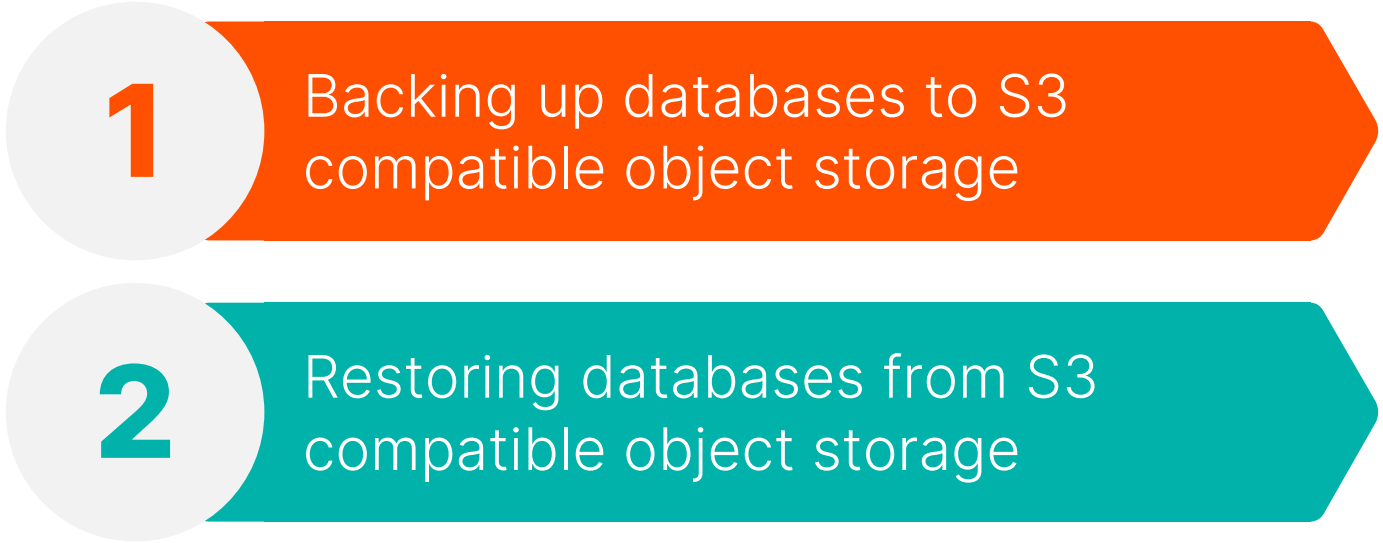
```
BACKUP DATABASE TestDB1
   TO  URL = 's3://s3.example.com/sqlbackups/TestDB1.bak',
       URL = 's3://s3.example.com/sqlbackups/TestDB2.bak',
       URL = 's3://s3.example.com/sqlbackups/TestDB3.bak',
       URL = 's3://s3.example.com/sqlbackups/TestDB4.bak',
       URL = 's3://s3.example.com/sqlbackups/TestDB5.bak',
       URL = 's3://s3.example.com/sqlbackups/TestDB6.bak'
WITH COMPRESSION, STATS = 10
```

- Each URL statement is a separate TCP stream.
- S3 has a max object size, so when working with backup files > 100GB, you must do a little tuning.

©2022 Pure Storage, Inc.

Hands-on Lab Information

SQL Server 2022 and Pure Storage



Take a break after the lab – 15 minutes

©2022 Pure Storage, Inc.

PURESTORAGE®

SQL Server Object Integration: Data Virtualization

Modern object storage data and SQL Server

S3 Object Integration – Data Virtualization

Modern Object Storage and Data

Why Data Virtualization?

- Access object storage directly from SQL Server engine
- Minimize overhead to get access to data
- Access data where it lives
- Backup restore / partitioning / index tuning not needed

Supported external file types

- Parquet/CSV/Delta

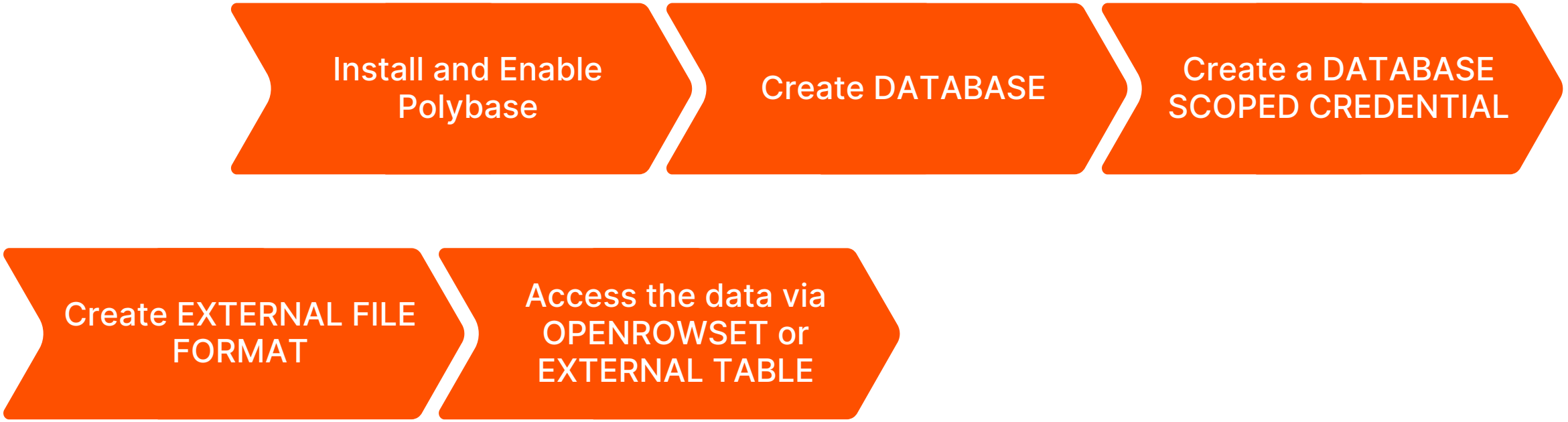
How to access external object data

- OPENROWSET
- EXTERNAL TABLE
- CREATE EXTERNAL TABLE AS SELECT

©2022 Pure Storage, Inc.

SQL Server 2022 – S3 Object Integration

Modern Object Storage and Data



©2022 Pure Storage, Inc.

Hands on Lab Information

SQL Server 2022 and Pure Storage

1. Query data on S3 compatible object storage with OPENROWSET
2. Query data on S3 compatible object storage with EXTERNAL TABLE
3. Wrap-up after the lab

©2022 Pure Storage, Inc.

PURESTORAGE®

Uncomplicate Data Storage, Forever