SNIA DEVELOPER CONFERENCE

SDC 24

BY Developers FOR Developers

September 16-18, 2024
Santa Clara, CA

Azure Files

Design challenges for the biggest file server in the world

Utsav Mohata, Principal Software Engineer, Microsoft

Rena Shah, Senior Product Manager, Microsoft

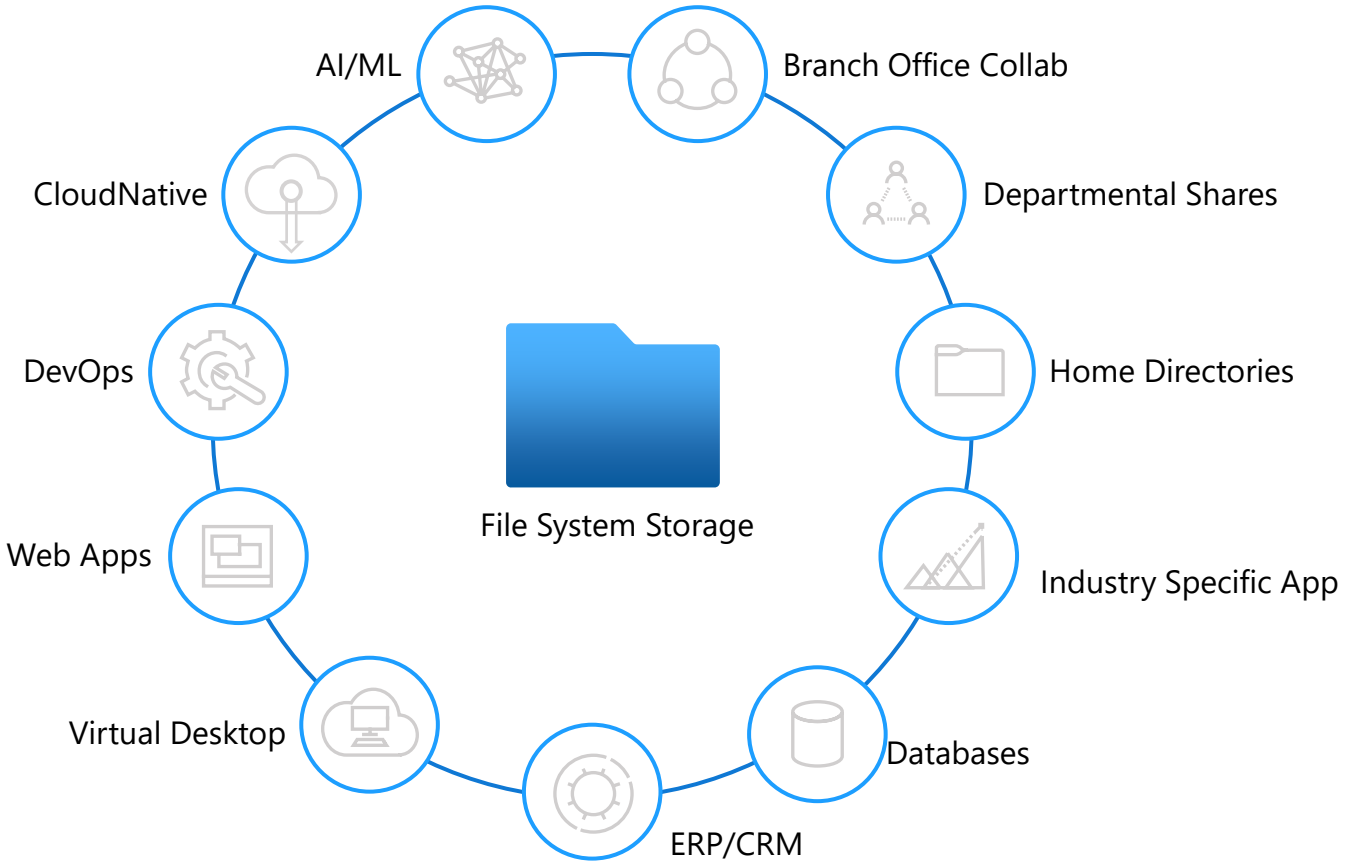
Agenda

- File storage workloads in Cloud
- Deep dive into some unique problems solved for
  - High Availability
  - Performance
  - Security
- Understanding how REST interop with SMB/NFS helps

File Storage Workloads in Cloud

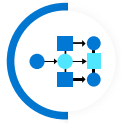
mount to the cloud

File Storage Workloads



Unique expectations from Hyperscalers

Multi-Tenancy and high availability

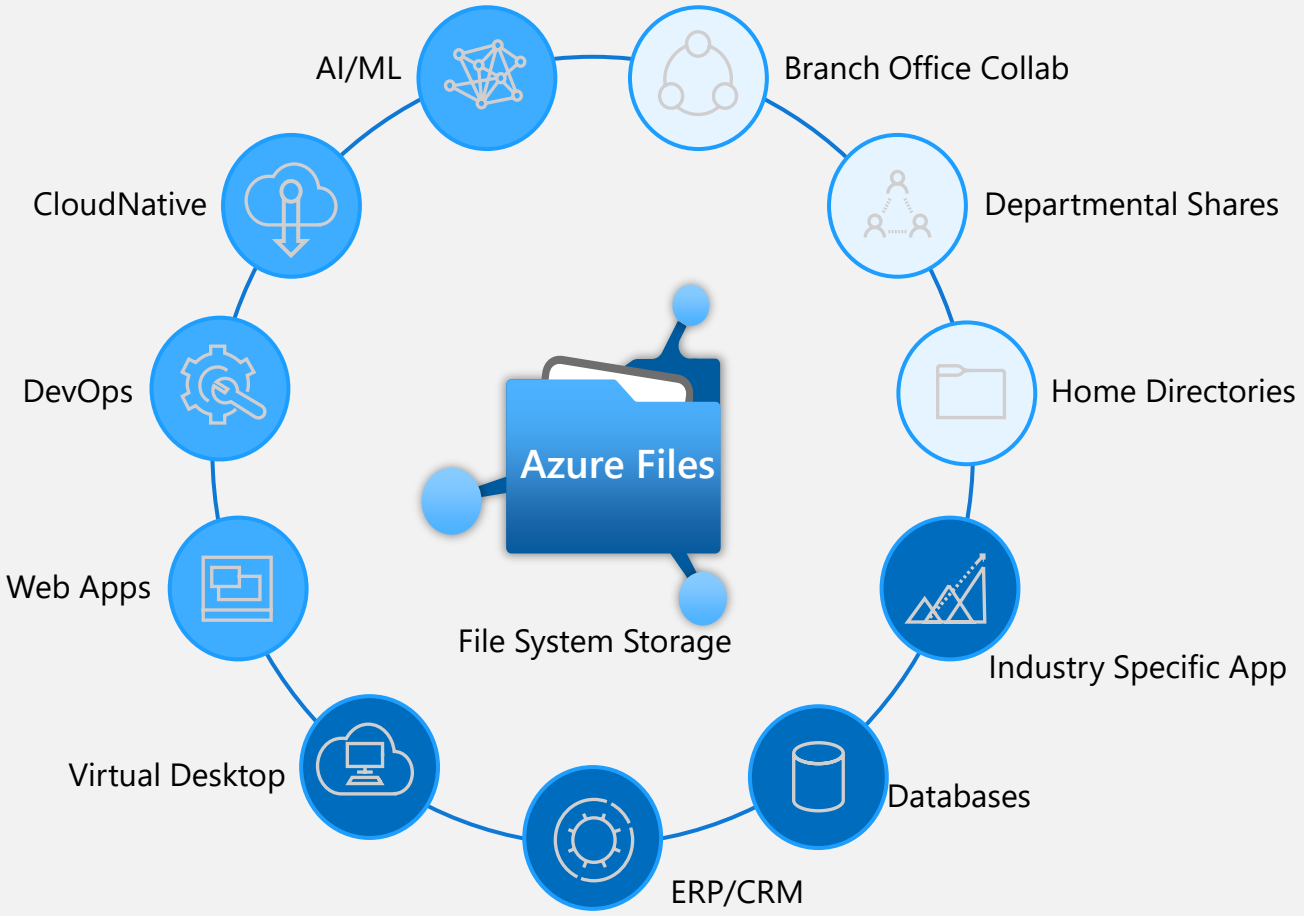


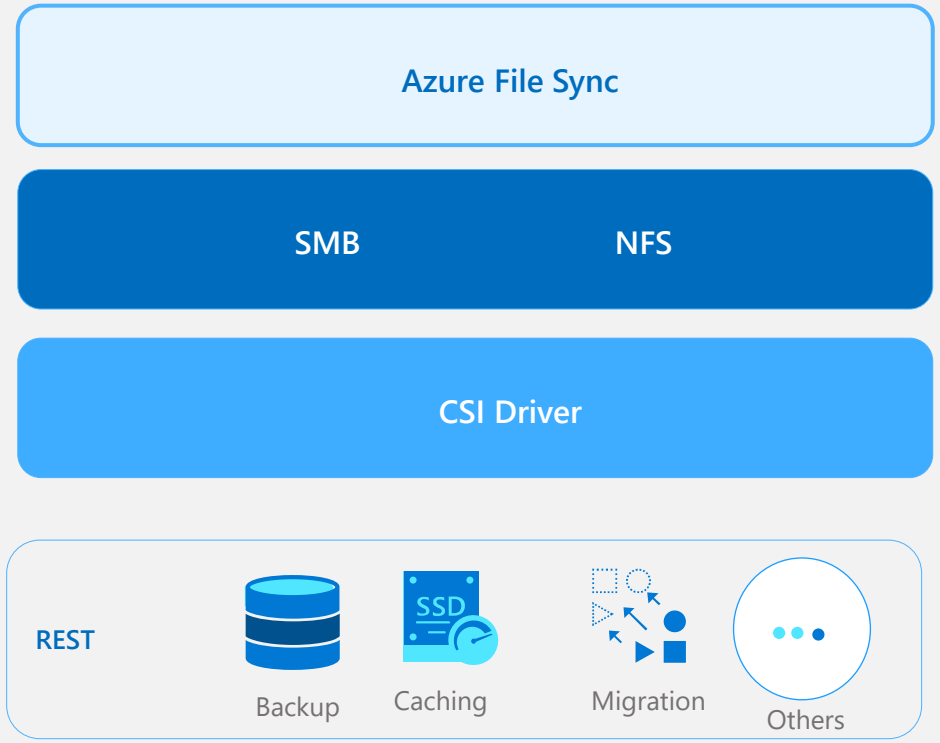
Workloads with Hybrid requirements

Traditional application workload rehosted or re-platformed to cloud

Leveraging the economies of cloud scale efficiently

Workloads Enabled by Azure Files



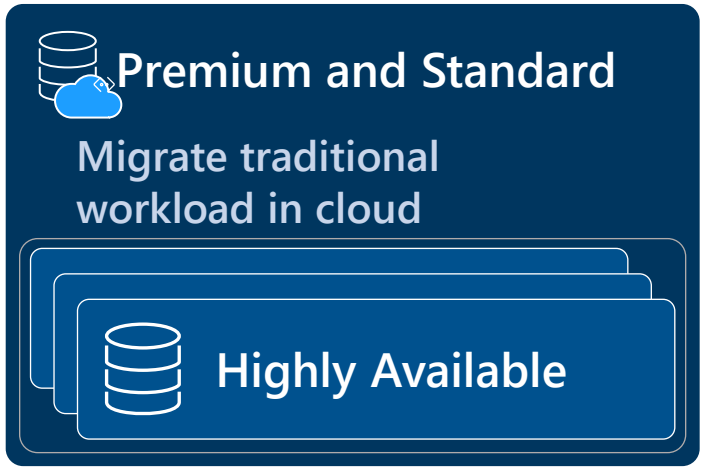


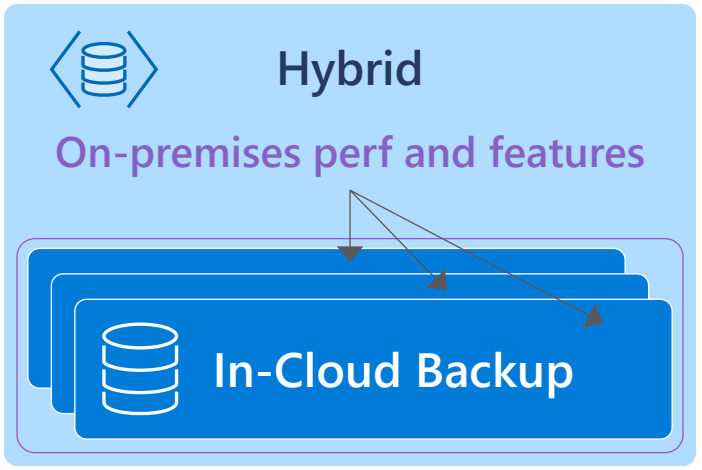
Capabilities addressing workload variety

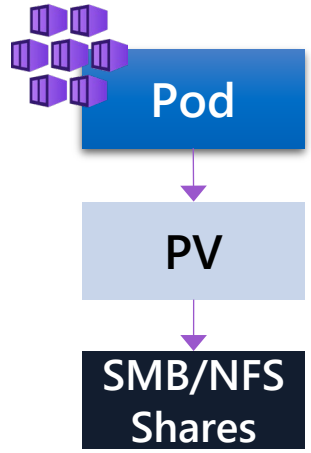
SMB/NFS

Azure File Sync

Cloud Native







CSI Driver

REST- Value Added Services

*preview expected later this year

Azure Files

Deep Dive into a Distributed Cloud File System

Azure Files

- A cloud native scalable implementation for file storage-built ground up
- **Not** a simple rehost of a Windows or Linux file server
- Fully Managed File Share
- Allows customers to seamlessly migrate mission critical applications to the cloud.

Azure Storage Architecture Basics

Front-end Layer

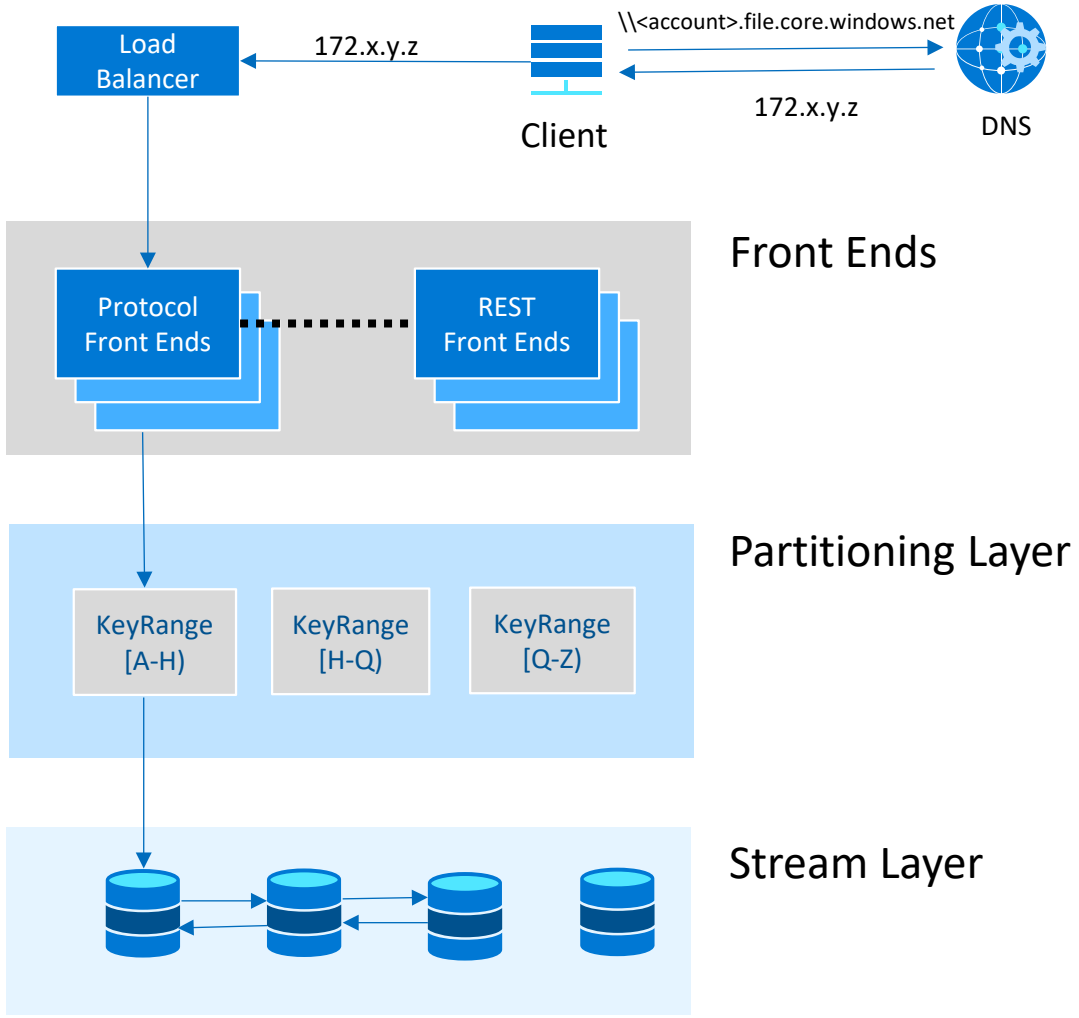
- Protocol endpoint
- Authentication/Authorization
- Metrics / Logging

Partition Layer

- Massively scalable key/value store
- Key ranges assigned to a server
- Understands and manages our data abstractions

Stream Layer

- Data persistence and replication
- Append-only file system
- 3 copies



Leveraging the Azure Storage Stack to build Azure Files

- Reuse Azure Table and Blobs as the backing store
  - Azure Table for Metadata
  - Azure Page Blobs for Data
- Azure Tables allow to associate a set of tables as a group for the same Key Range
- Metadata is stored in these set of tables
  - Entity – All files and directories existing in the share
  - Data – Allocated file ranges and pointers to the data
  - Handle – All open handles to files and directories
- Atomic update can be made across the group of tables for the same Key Range
- Tier Data by durability requirement to help with reconnect/reclaim

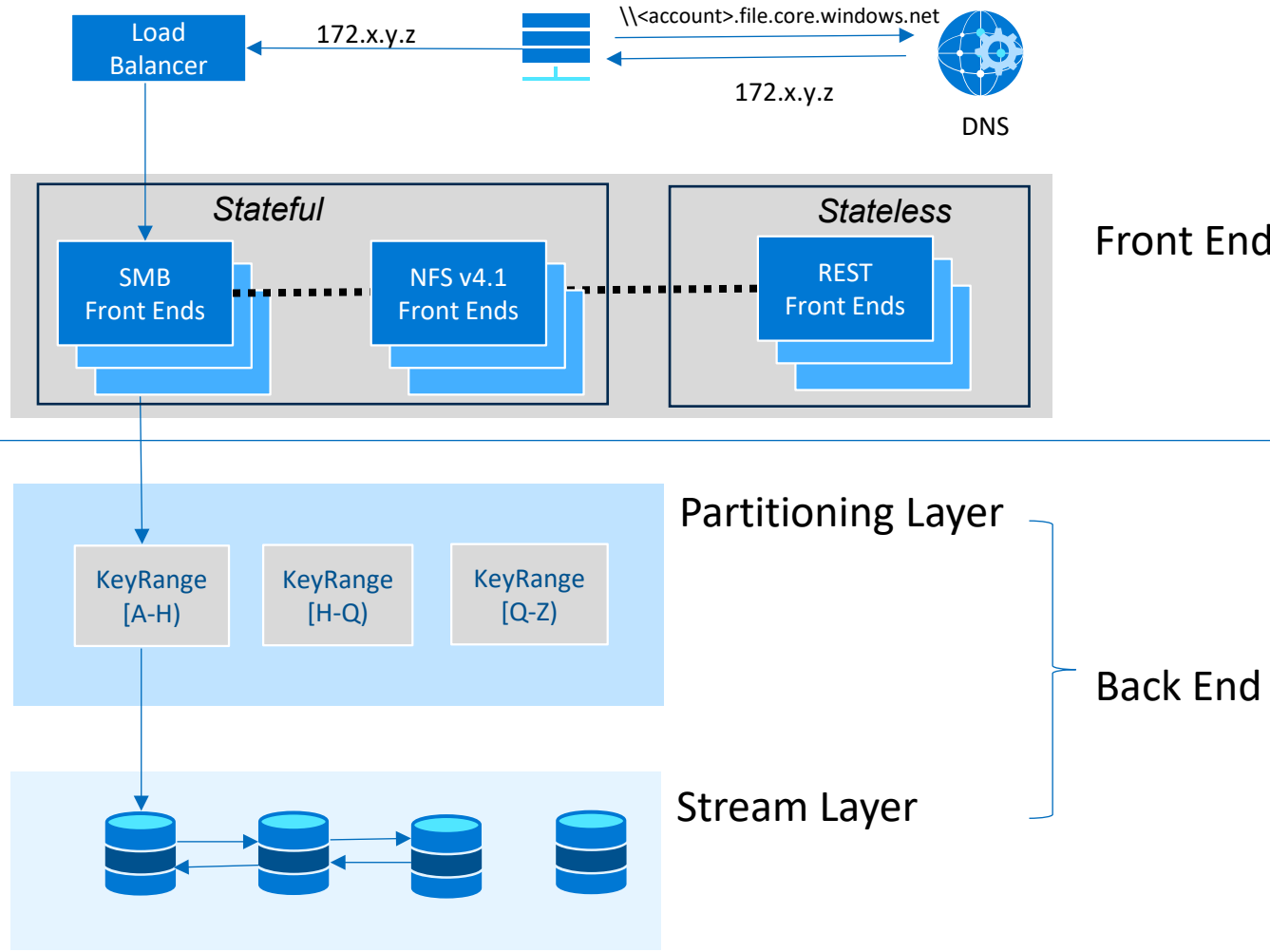
Azure Files - Architecture

Ephemeral state

- *SMB2_FILEID.Volatile*
- *credits*
- *TCP socket details*
- *StateId [nfs]*

Durable state

- *SMB2_FILEID.Persistent*
- *SessionId*
- *HandleId*
- *Open Counts, File Name, size, lease levels*
- *Byte Range locks [nfs]*



Azure Files Deep Dive

- Focus Areas
  - High Availability
  - Perf
  - Security

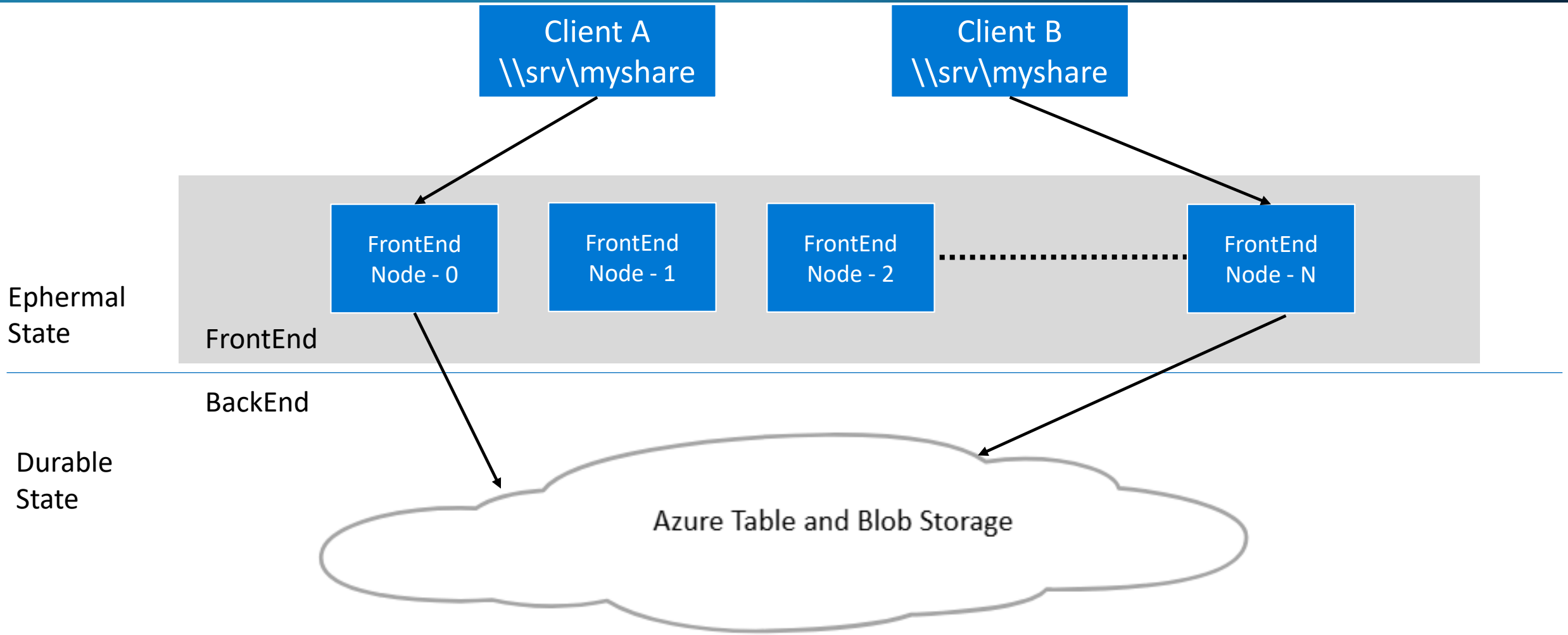
High Availability

HA for the continuously available share

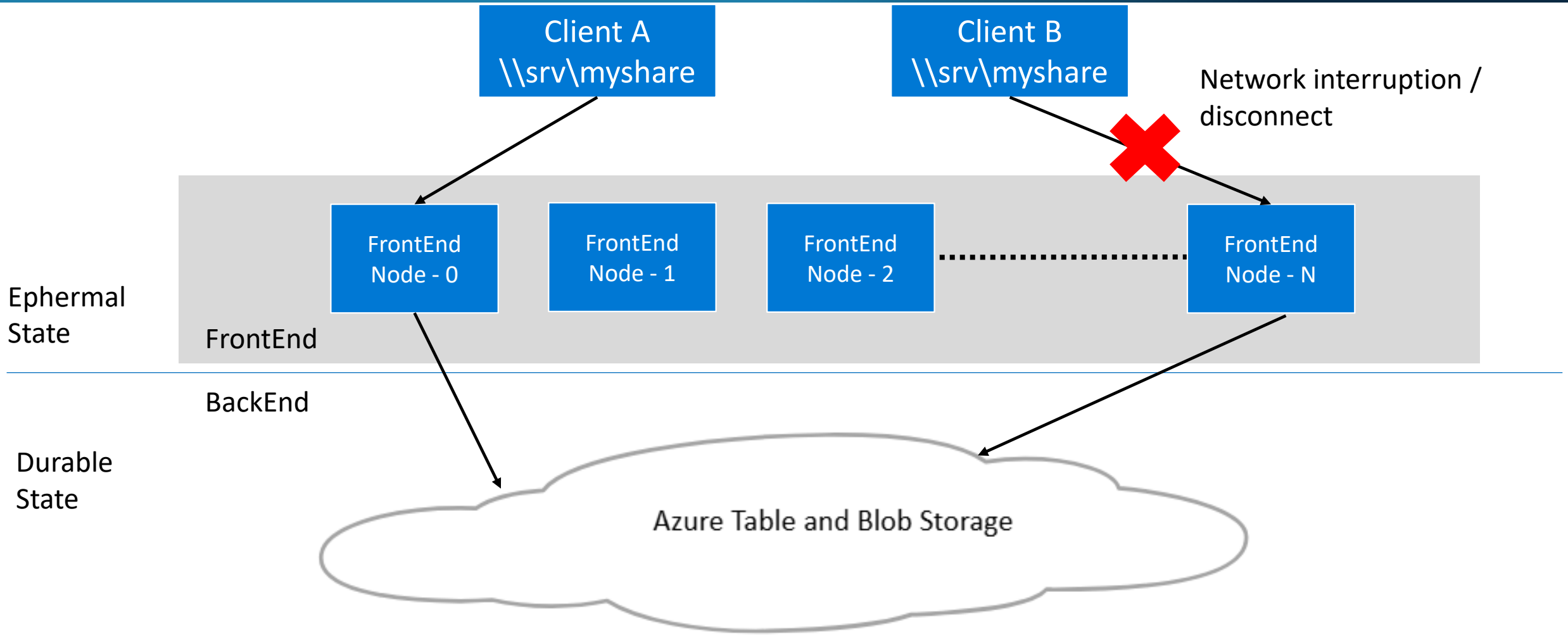
Restarts and Reconnects

- Unique requirement
  - Storage cluster is made up of 100s of nodes running microservices
  - Azure File Service restarts due to upgrades / deployments
- Solution
  - Tier Data by durability requirement to allow for transparent failover
    - Reconnect: Persistent Handles support >= SMB 3.x protocol onwards

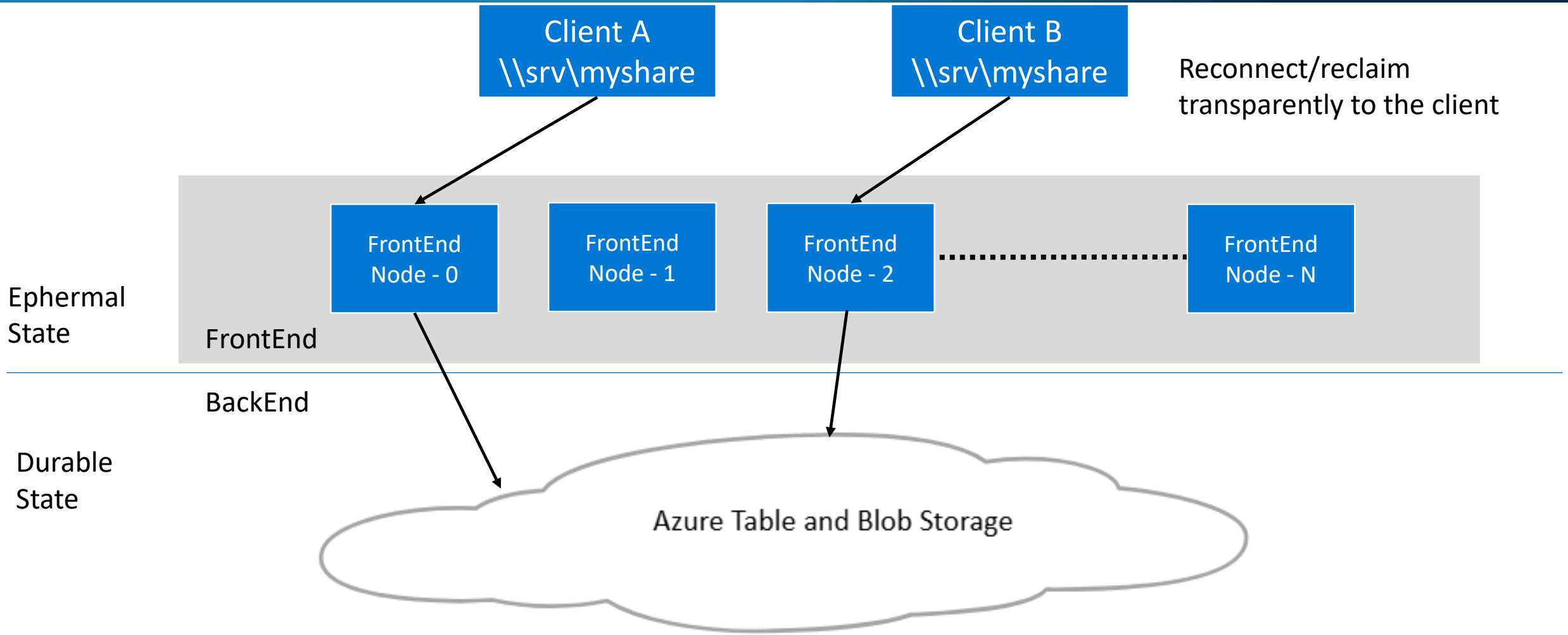
Availability



Availability



Availability

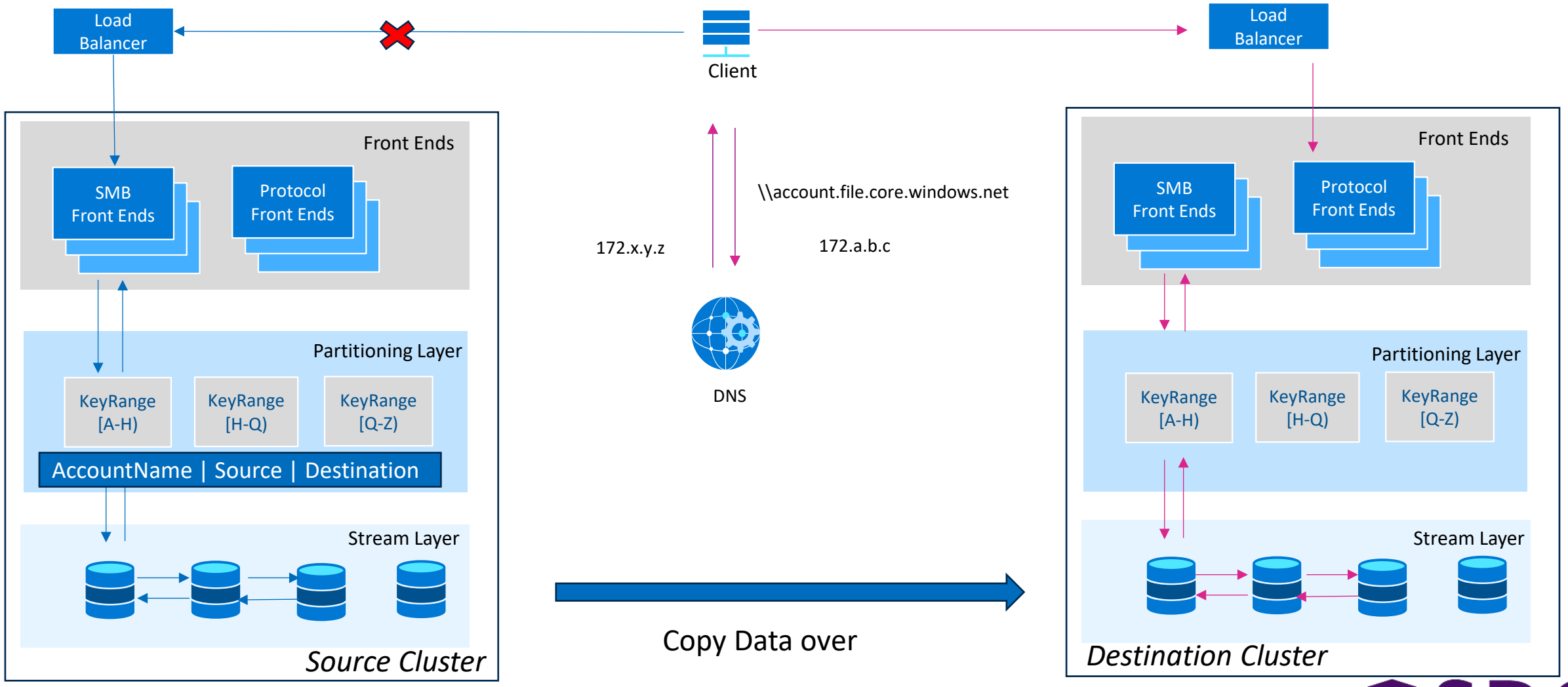


Live Migration

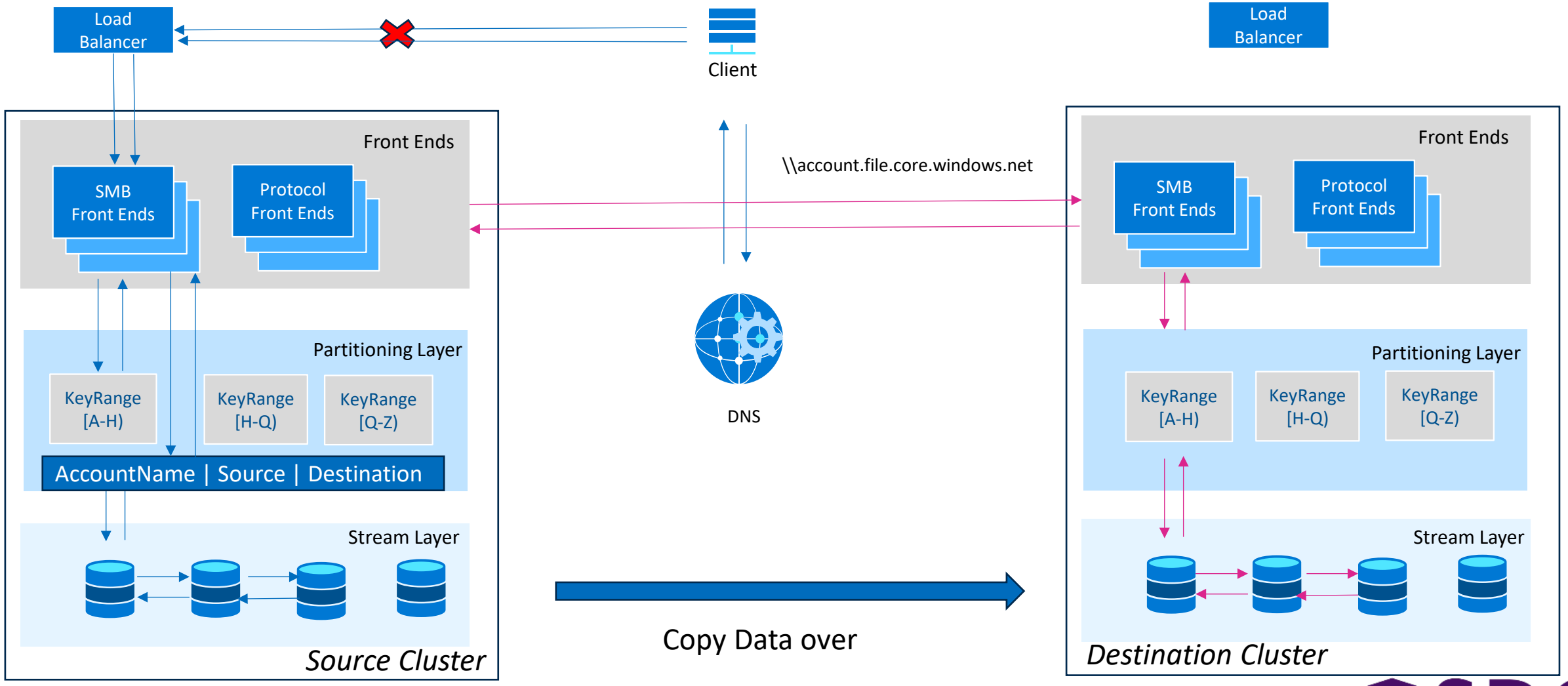
- Scenario :
  - A customer account is hosted on a Storage Cluster
  - Multiple Storage Clusters within a region.
  - Move storage accounts between Storage Clusters for Load Balancing
  - Migration has multiple phases
    - Copy data from source to destination cluster
    - Updating DNS mapping
    - Disconnect connection with client
    - Redirection
  - Transparent to the end user

- Issue :
  - Client resolving DNS only at mount time
  - File I/Os start giving "Host is down" or "Permission denied" errors

Live Migration - Reconnected



Live Migration - Redirected



Live Migration

- Issue (recap) :
  - File I/Os start giving "Host is down" or "Permission denied" errors on old linux clients
- Solution :
  - Server Side : Increase the window for which the service will continue to be in redirection mode
  - SMB Client Side : Fixes submitted to the mainline Linux kernel (Upgrade to client OS with fixes)
    - Benefits any file system provider
- Benefit :
  - Reduced customer reported incidents by 99 %
- Future :
  - Contribute to the community for the same improvement for the NFS v4.1 client

Performance

Client as well as Server side

Performance

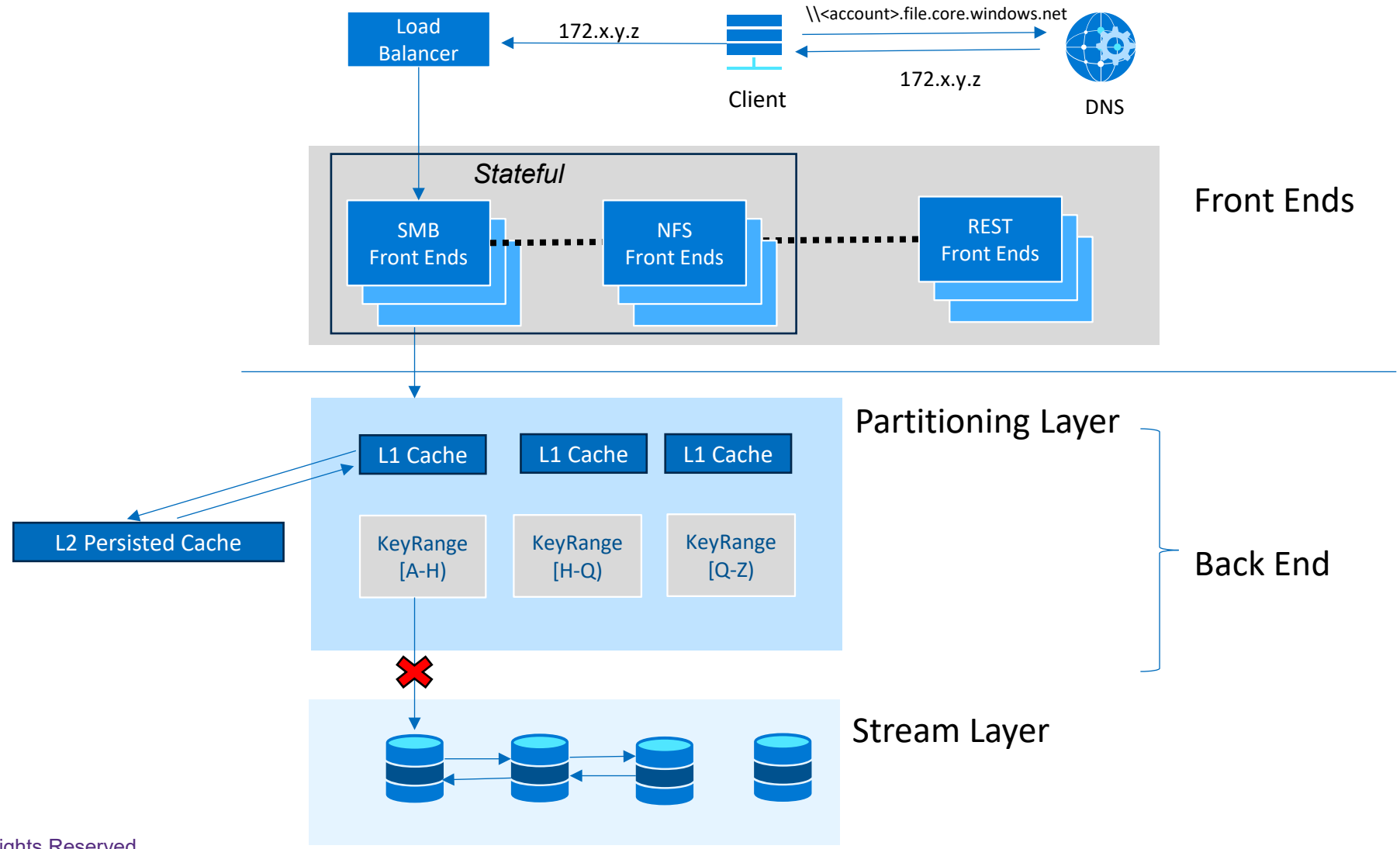
- Focus Areas
  - Improve latency of metadata operations
    - Advanced caching for Metadata
  - Increase client throughput
    - Multichannel / Nconnect
  - Reduce roundtrips
    - Deferred close for file handles
    - Lease Break Optimization

Advanced Caching for Metadata

- Workloads that perform a high volume of metadata operations (creating/opening/closing/deleting) against a SMB Premium File share will receive the biggest benefit

  - Web \ App Services
  - Indexing \ Batch Jobs
  - Virtual Desktop Infrastructure
  - Business Application
  - CI \ CD DevOps pipeline

Advanced Caching for Metadata



Advanced Caching for Metadata

- **Metadata Operations [Create/Open/Close/Delete] –**
  - Up to 55% lower P50 Metadata Latency

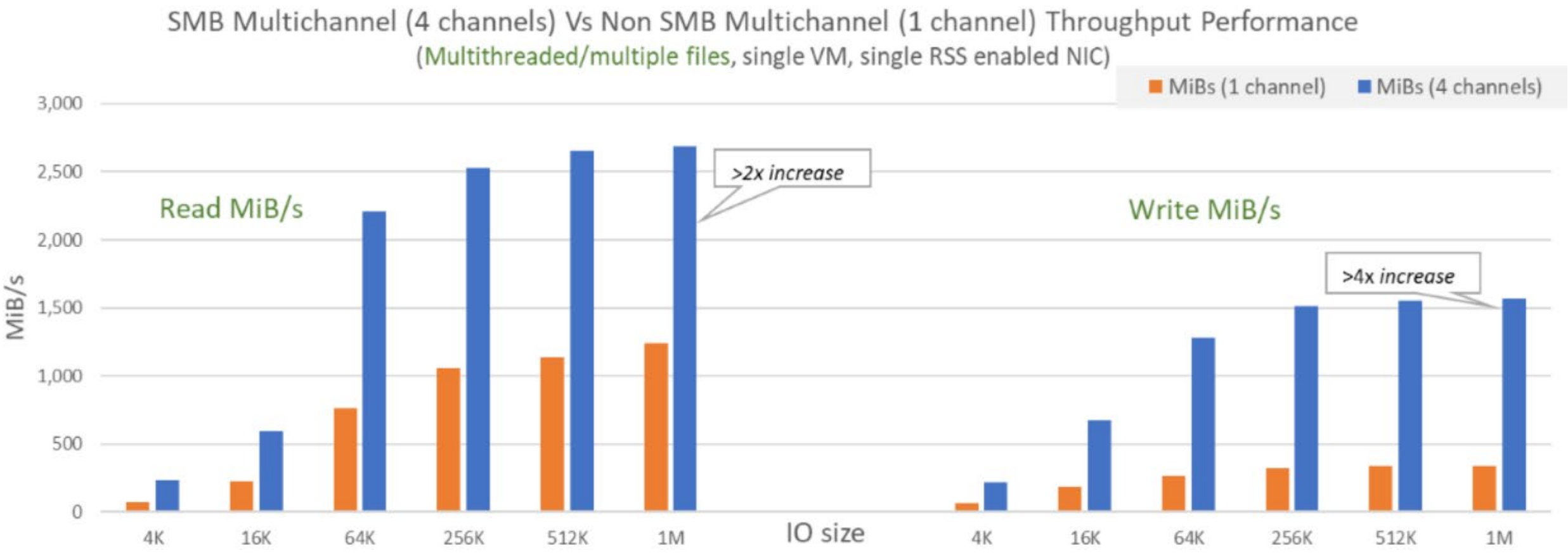
- **Metadata Scale Increase –**
  - Perform Up to 3x more metadata operations at scale (high queue depth).

Advances Caching Documentation

Azure Files – Multichannel (SMB) / NConnect (NFS)

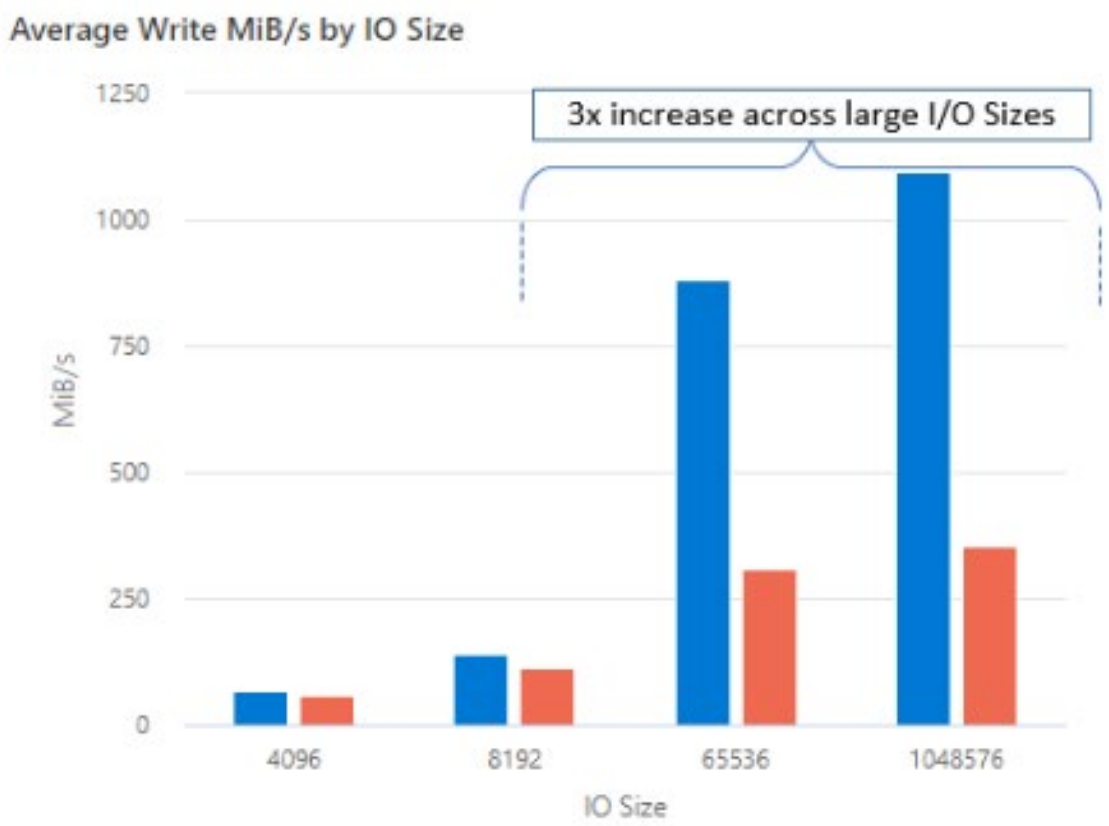
- Benefits :
  - Higher throughput
  - Increased IOPS
  - Network fault tolerance
  - Cost optimization

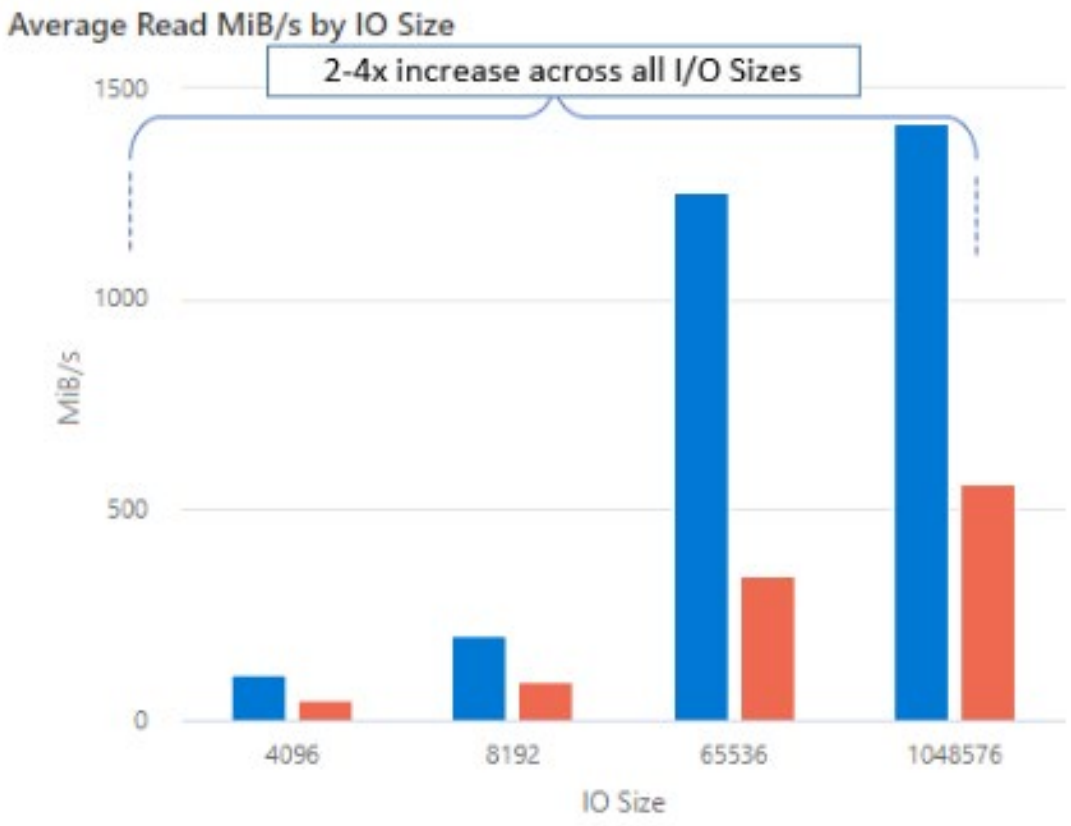
SMB Multichannel – Throughput gains



[Multichannel documentation]

NFS NConnect – Throughput gains





Lease Break Optimization

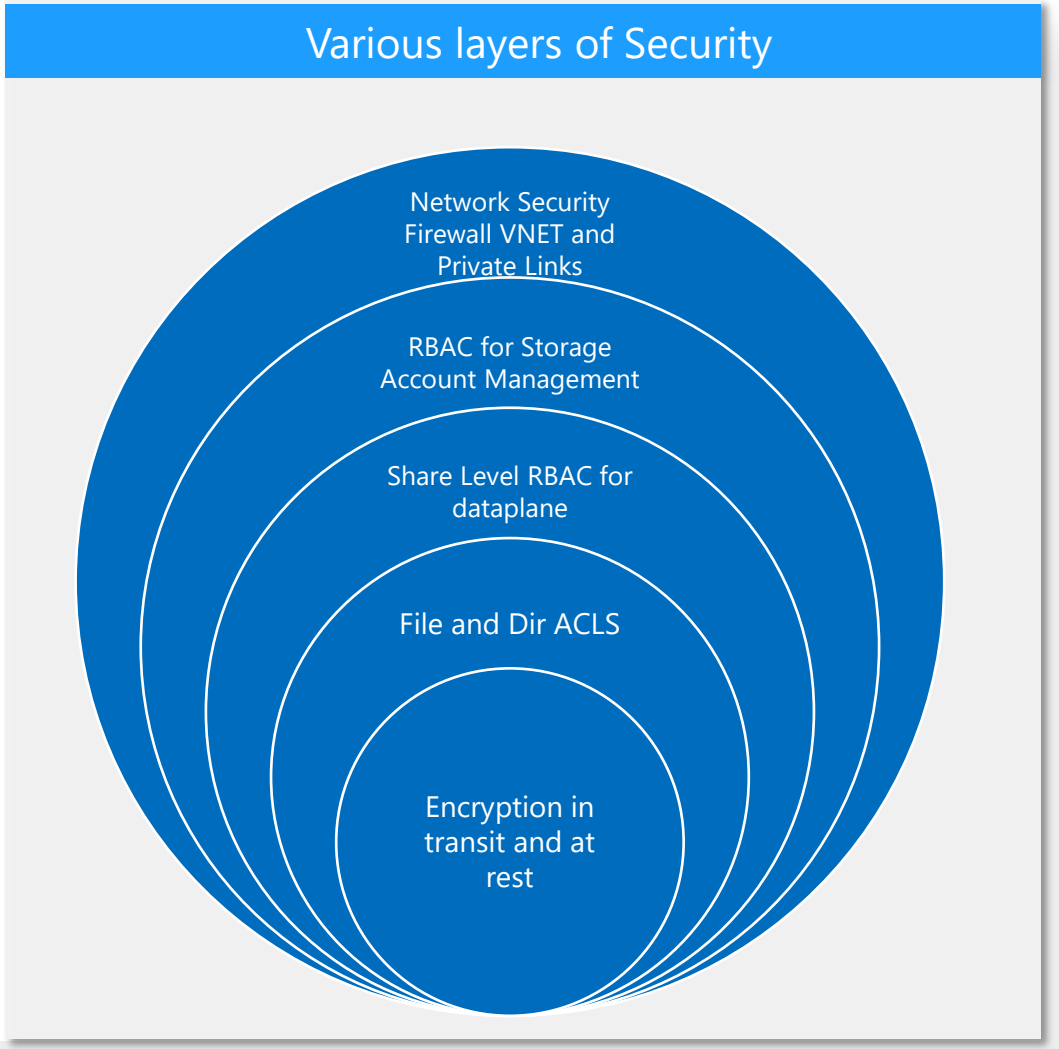
- Scenario :
  - Simultaneous opens to the same file, requesting exclusive (RWH) lease
  - Old linux clients fail to acknowledge the lease break.
- Issue :
  - These handles are short-lived, so we end up doing a lot of unnecessary lease-break work
  - Causes increase in latency for file open calls from other clients while waiting for lease break ack
- Solution :
  - In-memory cache of the file's last lease state, do not issue exclusive lease if the lease is going to be broken soon.
- Benefit
  - Reduce the # of lease break requests sent and need for acknowledgement from clients
  - Improving latencies for file open calls

Security

Supporting hybrid Identities

Security Landscape

- Security is even more critical in cloud
- Security needs to come in multiple layers
- Customers need traditional auth presented in modern way

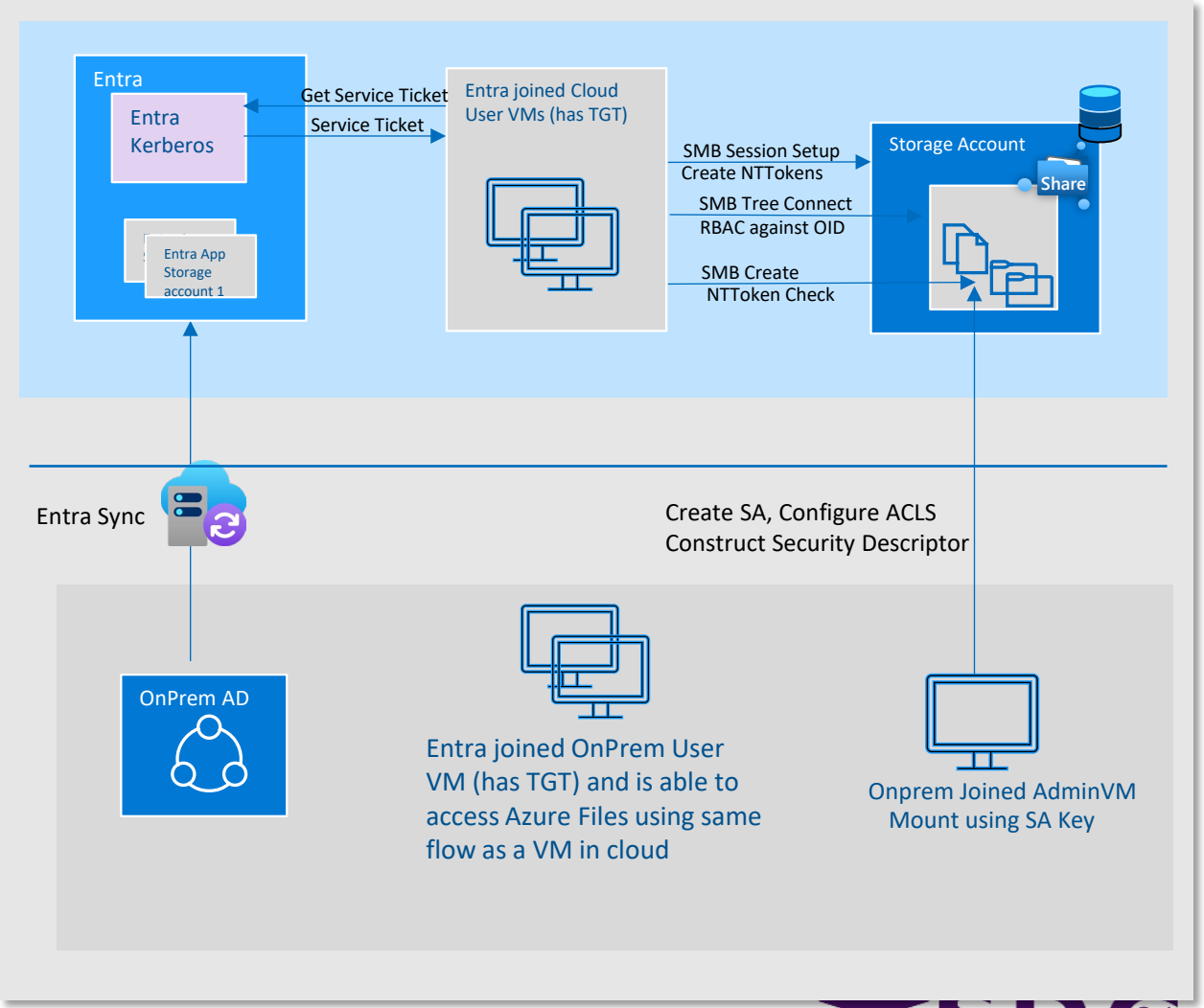


Typical Customer Journey and enablement

Identity Fully on-premises	Hybrid Identity	No on-premises - Cloud DS
Self hosted Active Directory (AD) Domain Controller	**Entra Kerberos (Azure Files only)**	**Entra Domain Services (Azure AD DS)**
Considered when: • Starting with on premises AD DS • Your company requires your AD DS to stay on premises • AD on VM owned • Self Managed	**Use Entra for Kerberos authentication** • No additional domain control setup/management • Requires hybrid identities (Identities created on AD DS and synced to Azure AD)	**Considered as:** • Managed service alternative to AD DS • Planned – Cloud-native identities
Notes		
Requires client to have line-of-sight to domain controller	Once ACLs are configured, users can access file shares over the internet, without line-of-sight requirement to domain controller	Independent domain service to manage

Hybrid setup: Request flow for the successful mount

- Customers stay on hybrid setup
- Bridge “legacy” on-premises security (client Kerberos support, file-level ACLs) with “modern” cloud-based security (Azure RBAC, Entra ID, etc.)
- Entra Kerberos provides “KDC” in the cloud. Windows Client enabled a feature to check for tickets from cloud KDC

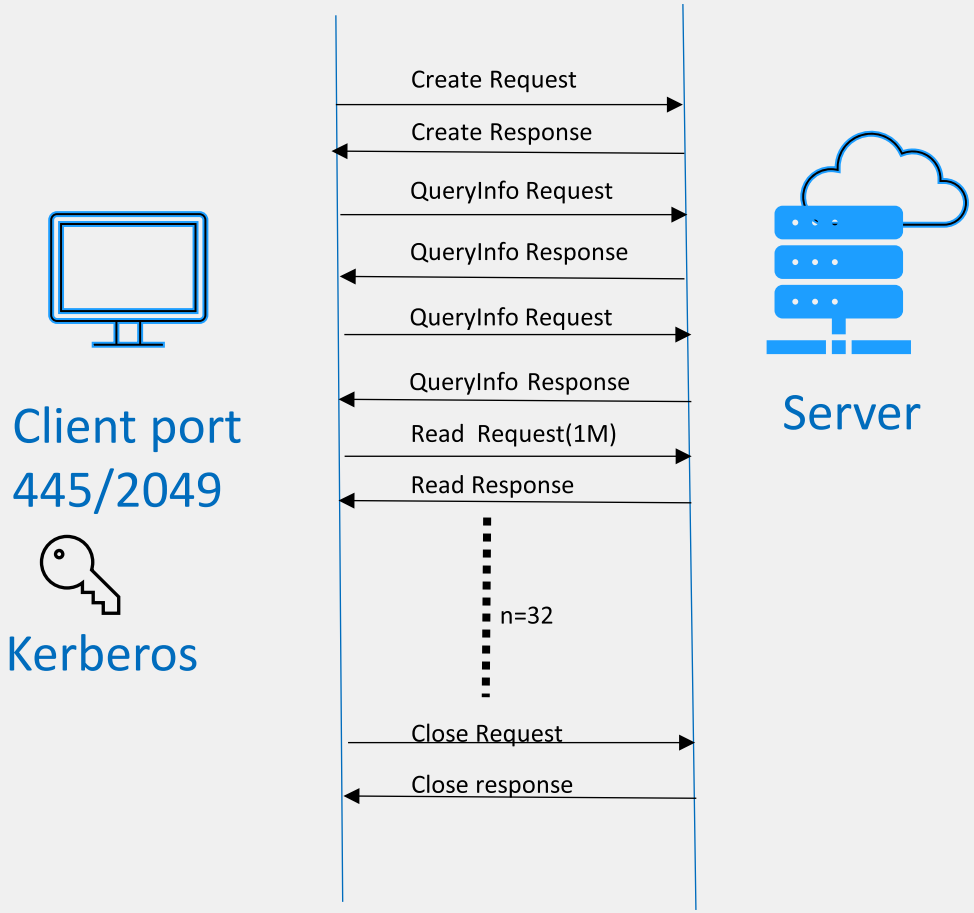


REST interop with SMB/NFS

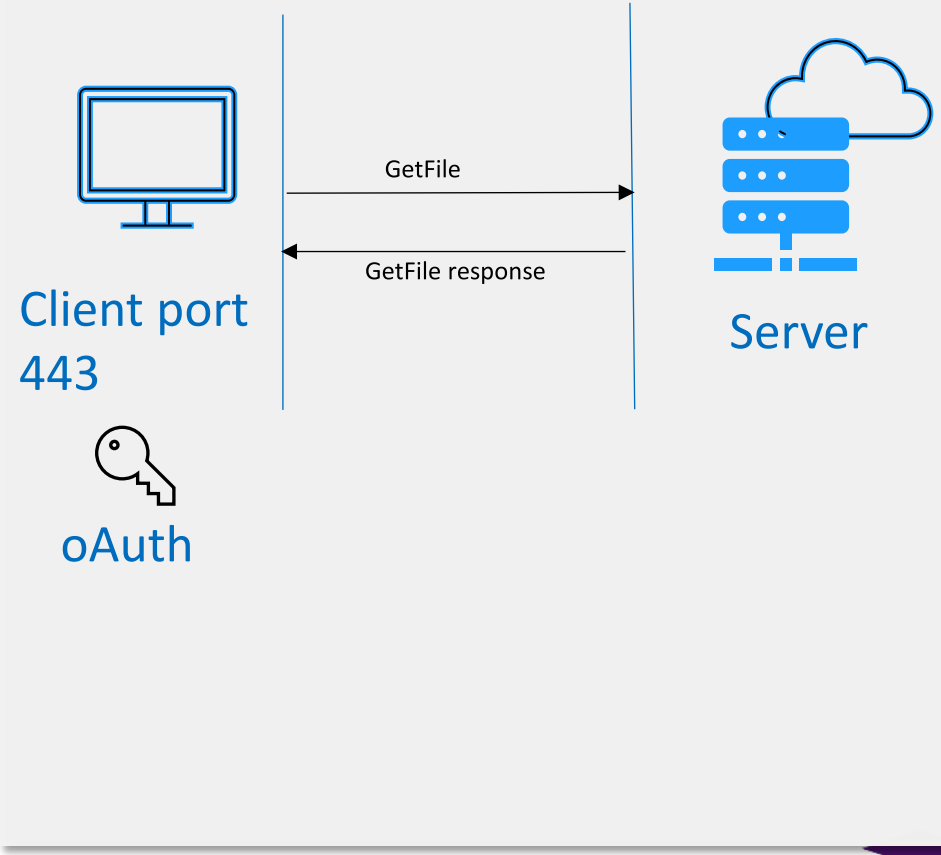
Enabling ISV scenarios

SMB/NFS vs REST – 32 MB ReadFile Example

Stateful Protocol



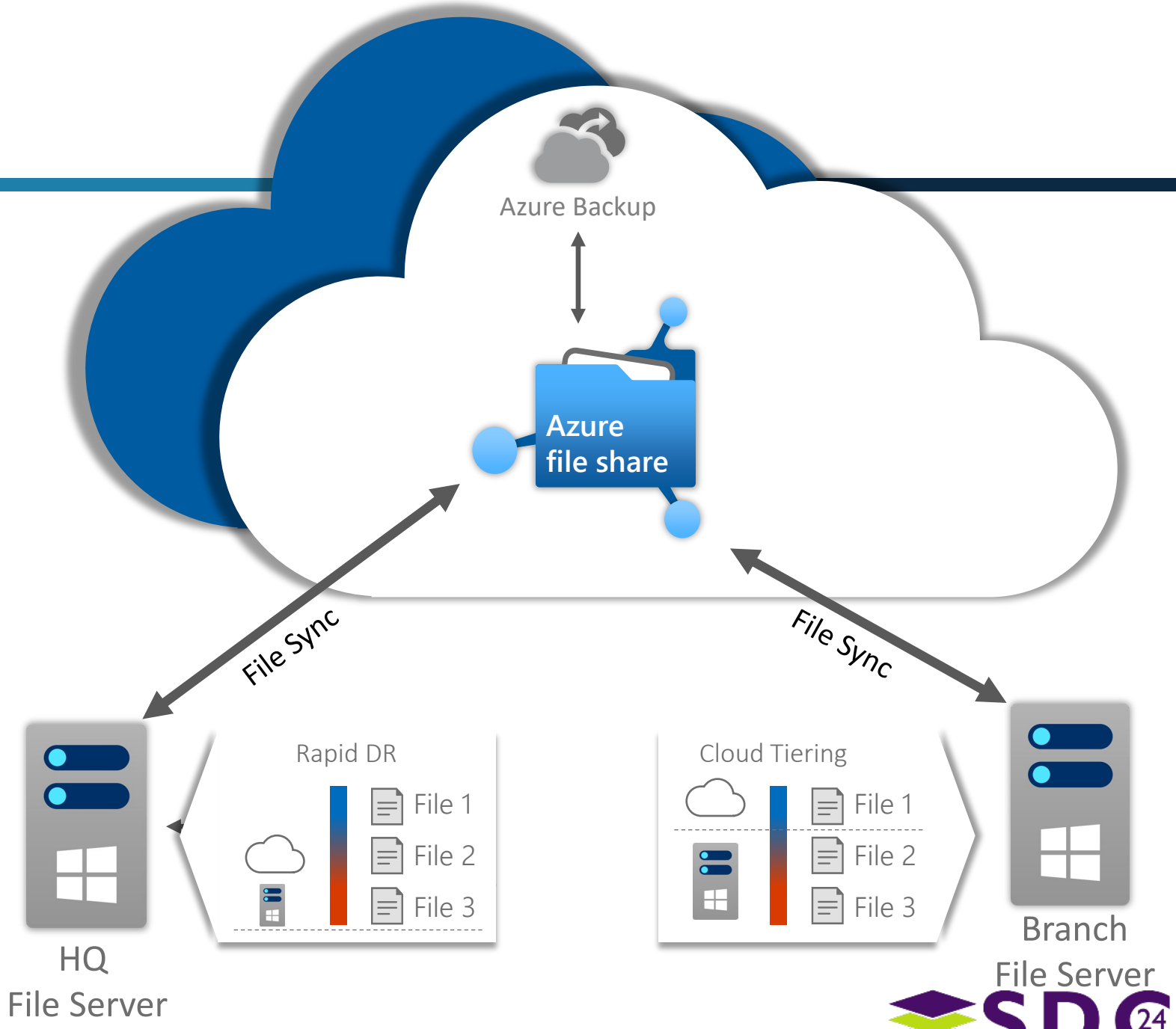
REST (Stateless)



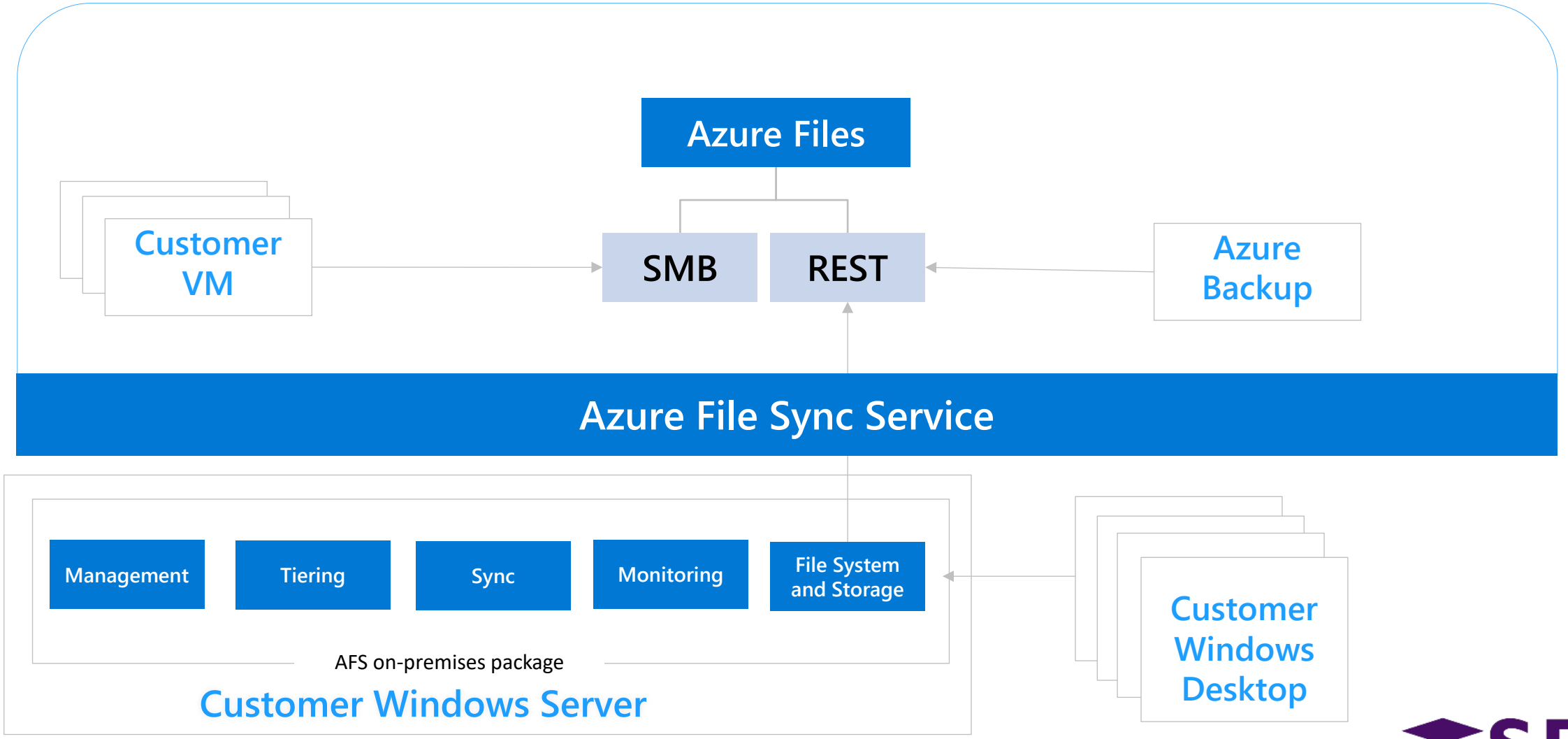
Azure File Sync

Hybrid Azure file shares

- Multi-site Sync
- Cloud Tiering
- Cloud Backup
- Disaster Recovery

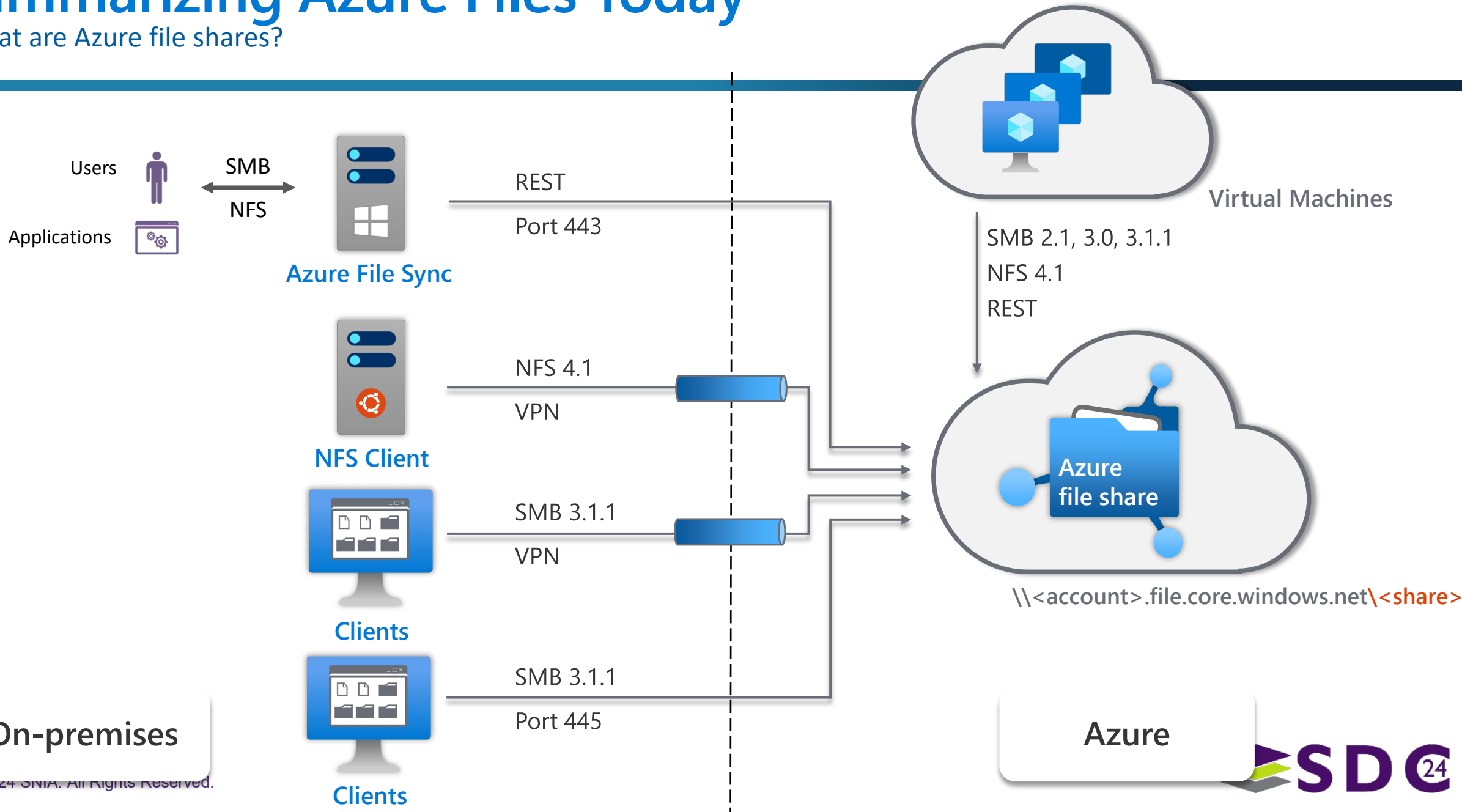


REST Example - Inside Azure File Sync



Summarizing Azure Files Today

What are Azure file shares?



Act now

1. **Lift and Shift** file workloads to cloud. Cloud file system market doubling every three years. Most companies are shifting toward cloud.

2. ISVs **partner with us** and build SaaS services and data ecosystem services with Azure Files leveraging REST, SMB and NFS.

3. **Learn more/provide feedback** about Azure Files (links in references). Attend/watch additional Azure sessions at SNIA.

Please take a moment to rate this session.

Your feedback is important to us.

Thank You

Please reach out to us at azurefiles @ microsoft.com

: umohata

: renashah

Reference

- Azure Files docs
- Azure Files REST reference
- Azure File Sync Doc
- SambaXP Talk
- Past Azure Storage SNIA talks
  - 2015
  - 2017

