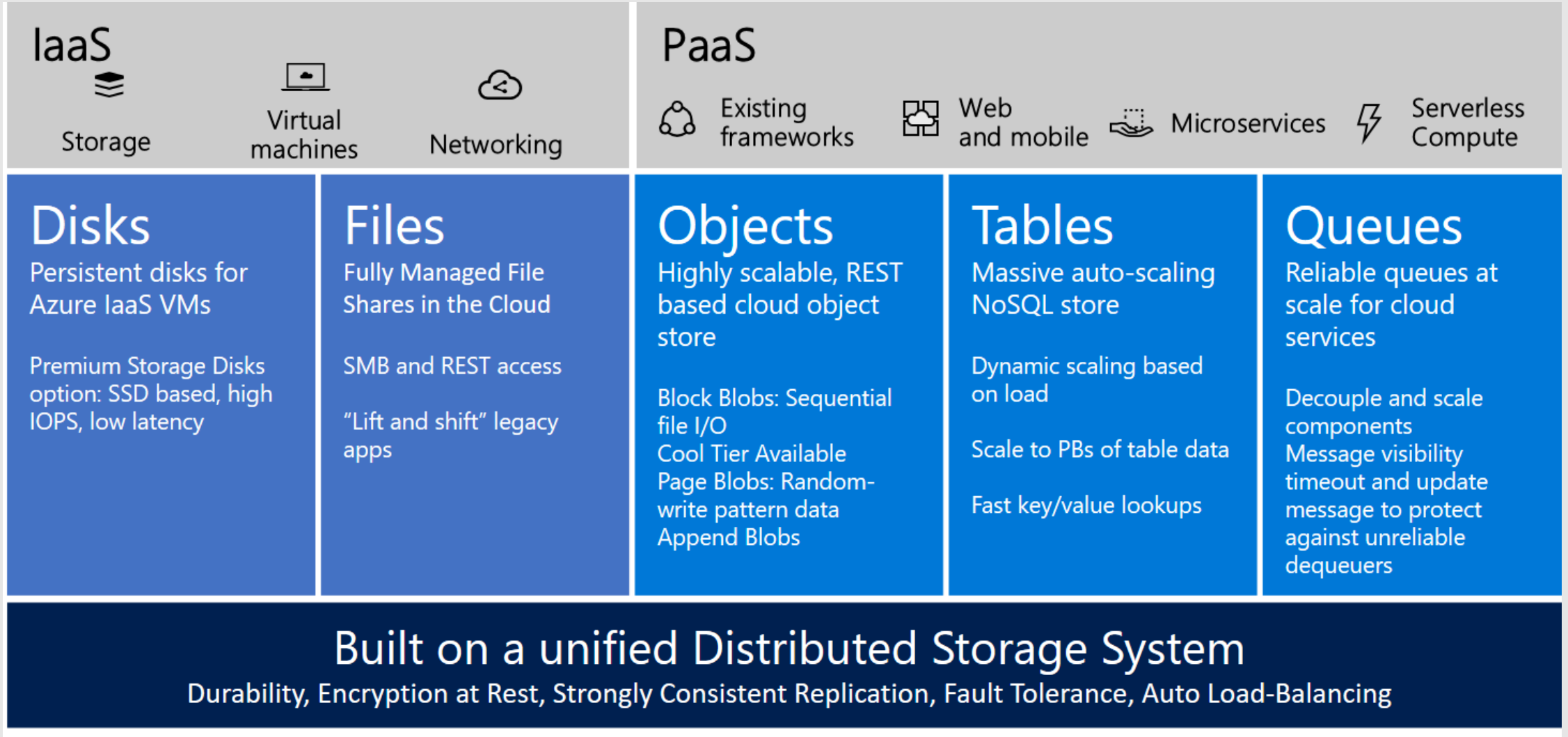
Microsoft

Azure Storage Overview

Azure Trailblazer Academy

Microsoft
Jan 27, 2021

Azure Storage Services



Secure, scalable and highly available storage options for every use case

Disk Storage

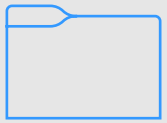
Premium
Standard

Reliable, persistent, high performing storage for Virtual Machines

Object Storage

Azure Blobs

Secure, centralized storage target for backup/disaster recovery



File storage

Azure Files
Azure NetApp Files

Lift and shift of legacy applications that require file shares to the cloud

Data Transport

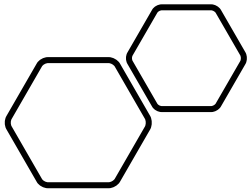
Azure Import/Export
Azure DataBox

Move or migrate data into Azure

Hybrid Storage

Azure StorSimple
Azure File Sync

Secure, intelligent data tiering between on-premises and cloud storage



Blob Storage Type

Block Blobs

Composed of blocks.

Ideal for storing Store text and binary data.

Size limit of 4.75 TB per blob (190.7 TB in preview)

Append Blobs

Made up of blocks but are optimized for append operations .

It is ideal for scenarios such as logging data from VMs.

Size limit of 195.31 GB

Page Blobs

Page blobs store virtual hard drive (VHD) files and serve as disks for Azure virtual machines.

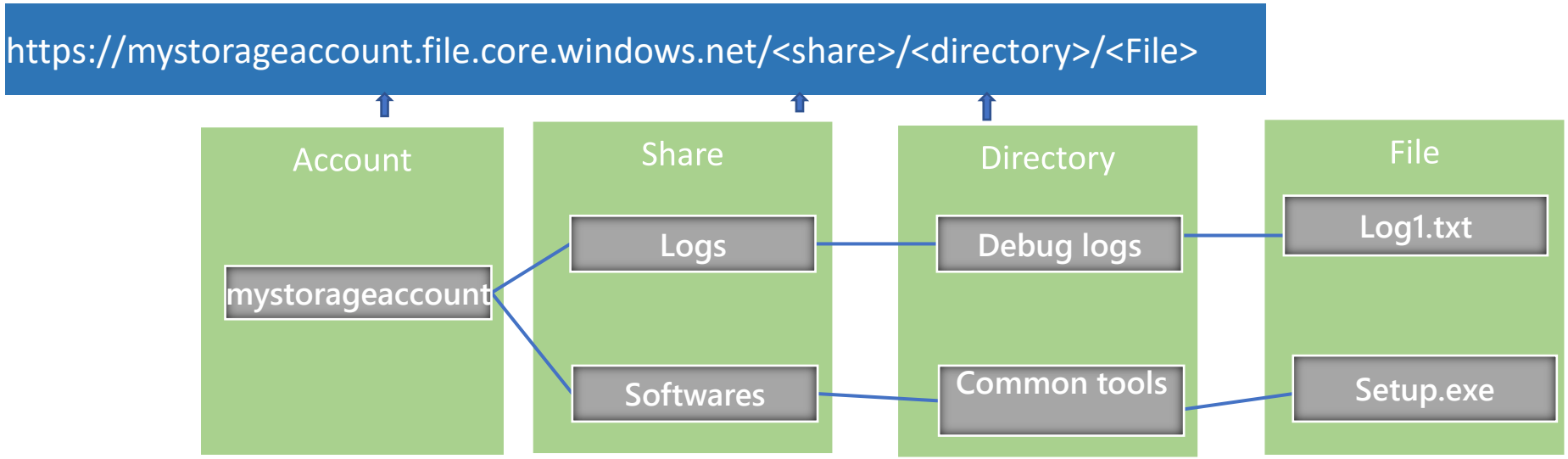
Size limit of 8 TB per blob

Azure Disks

Detail	Ultra disk	Premium SSD	Standard SSD	Standard HDD
Disk type	SSD	SSD	SSD	HDD
Scenario	IO-intensive workloads such as SAP HANA, top tier databases (for example, SQL, Oracle), and other transaction-heavy workloads.	Production and performance sensitive workloads	Web servers, lightly used enterprise applications and dev/test	Backup, non-critical, infrequent access
Max disk size	65,536 gibibyte (GiB)	32,767 GiB	32,767 GiB	32,767 GiB
Max throughput	2,000 MB/s	900 MB/s	750 MB/s	500 MB/s
Max IOPS	160,000	20,000	6,000	2,000

Azure Files

- Accessible via the industry standard Server Message Block (SMB) protocol
- Azure file shares can be mounted concurrently by cloud or on-premises deployments of Windows, Linux, and macOS
- A share can have multiple directories and directories and files must be created in a parent share.
- Primarily designed to support ‘Lift and shift’ scenario.



Azure Table Storage

- Azure Table storage is a service that stores structured NoSQL data in the cloud.
- Stores large amounts of structured data in the cloud as entities within a table.
- Table storage contains the following components:

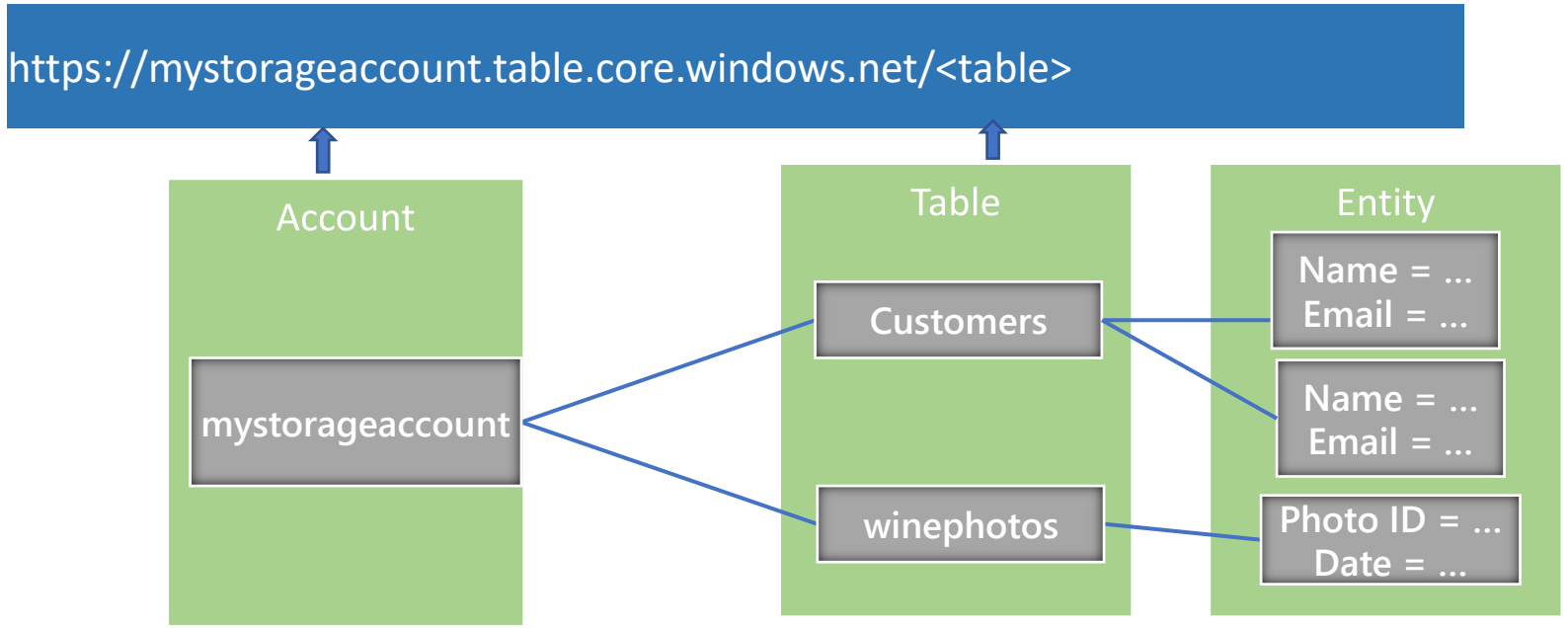
  URL format: http://<storage account>.table.core.windows.net/<table>

  Accounts: All access to Azure Storage is done through a storage account.

  Table: A table is a collection of entities

  Entity: An entity is a set of properties, similar to a database row and can be up to 1MB in size.

  Properties: A property is a name-value pair. Each entity can include up to 252 properties to store data.



Azure Queues

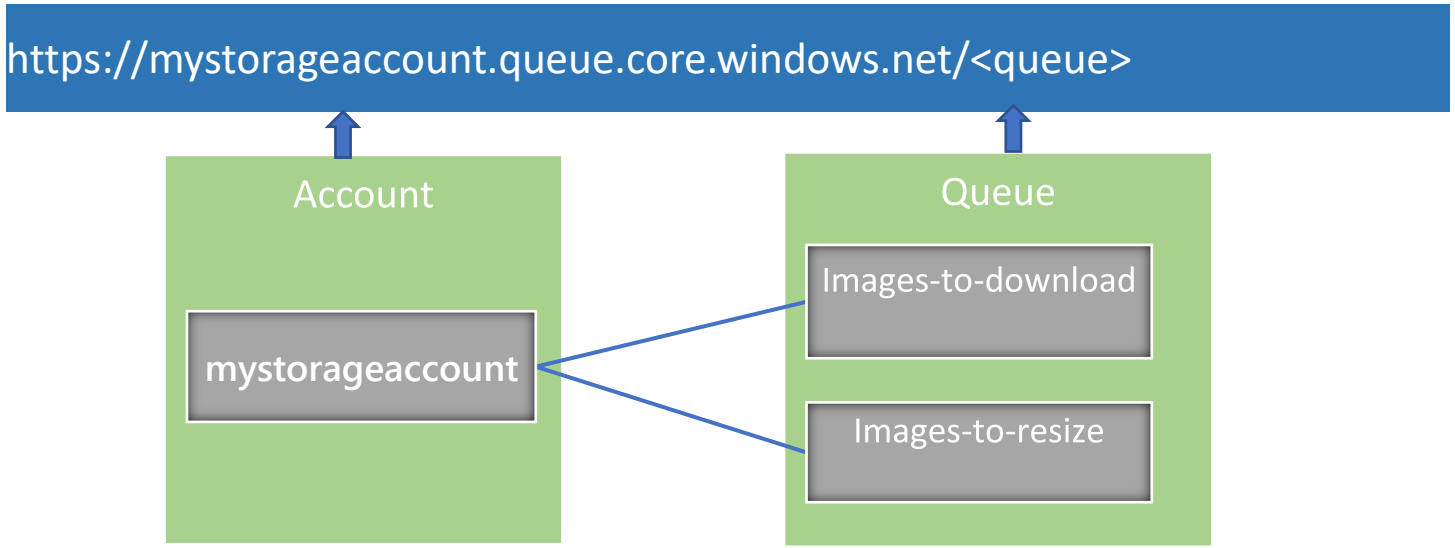
- Azure Queue Storage is a service for storing large numbers of messages. You access messages from anywhere in the world via authenticated calls using HTTP or HTTPS.
- A queue may contain millions of messages.
- The Queue service contains the following components:

  URL format: https://<storage account>.queue.core.windows.net/<queue>

  Storage account: All access to Azure Storage is done through a storage account.

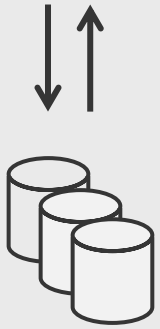
  Queue: A queue contains a set of messages. The queue name **must** be all lowercase.

  Message: A message, in any format, the maximum time that a message can remain in the queue is 7 days.



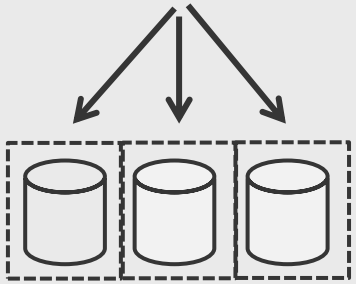
Azure Storage Durability

LRS



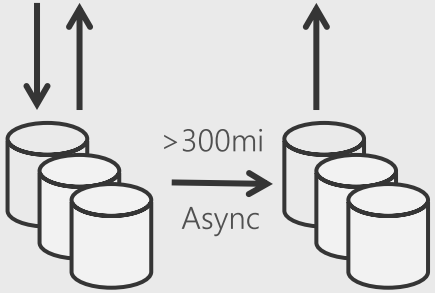
3 replicas
1 region

ZRS



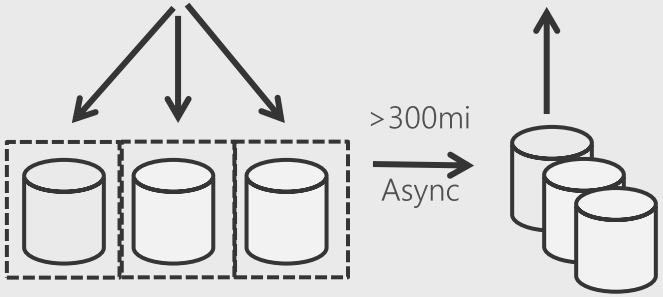
Multiple availability zones
3 replicas
1 region

(RA-)GRS



6 replicas
2 regions

(RA-)GZRS



Multiple availability zones in primary,
single DC in secondary
6 replicas
2 regions

Lower ← Cost → Higher

Storing Data Cost Effectively

	Premium	Hot	Cool	Archive
	Low and consistent latency data	Frequently accessed data	Less frequently accessed data	Rarely accessed data
RETRIEVAL TIMES	Immediate (SSD)	Immediate (HDD)	Immediate (HDD)	Hours
USE CASE EXAMPLES	Interactive Transactions Telemetry	Cloud native application data	Server backups	Medical records archive

Choose between online Cool tier and offline Archive tier

No more management or migrations of storage hardware!

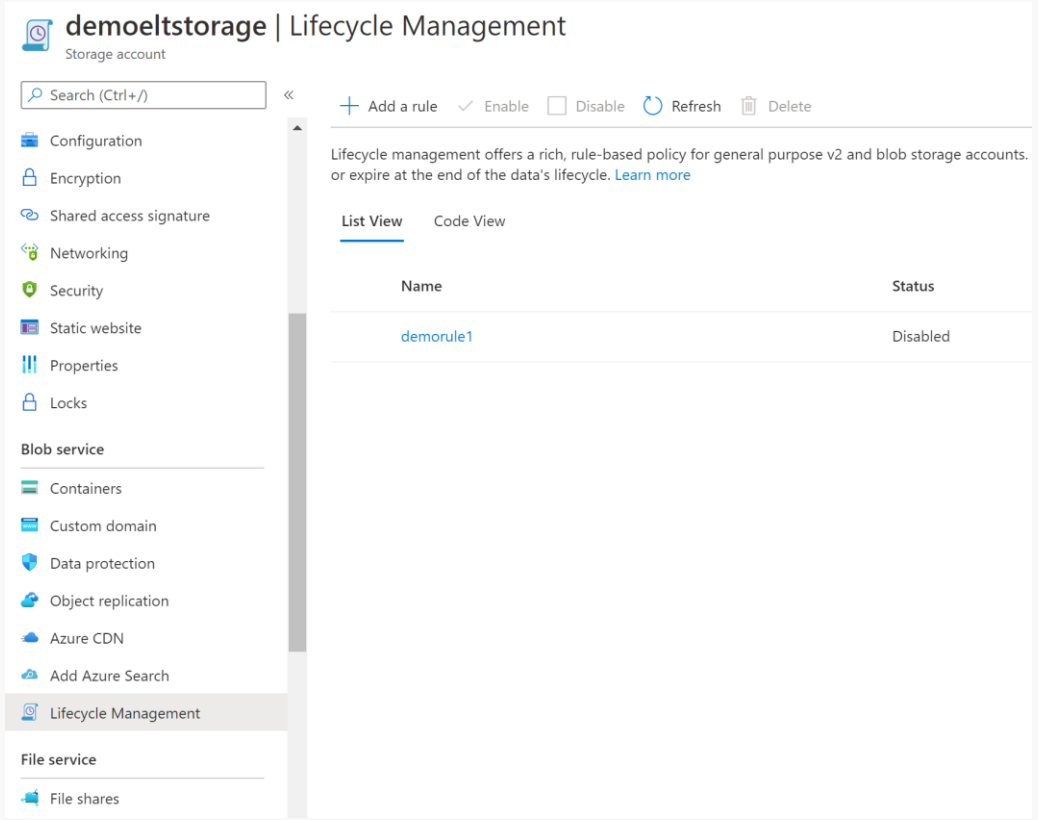
Aug-2019 – **50% price drop on the Archive tier** in major regions

New – Save even more with **Reserved capacity** –
Up to 38% savings 100TB and 1PB pre-purchases for 1-3 years

Life Cycle Management

Rule based automation for data tiering and retention management providing transitions:

- Hot to Cool
- Hot to Archive
- Cool to Archive
- Deletion



Rules are executed daily at the storage account

Support GPv2, Blob storage, Premium block blob accounts

Rules can be applied to containers or a subset of blobs (using prefixes as filters)

Comprehensive data protection capabilities

Object Storage

GA

Point-in-time restore

Provides the ability to restore a subset of containers or blobs within a storage account to a previous state

Versioning

When enabled, you can restore an earlier version of a blob to recover your data if it is modified or deleted

Block Storage

GA

Incremental snapshots

Point-in-time backup of disks, consisting only of changes since last snapshot was taken

Direct upload

Restore VHD directly to an empty disk from on-premises or cloud , providing a simplified workflow

File Storage

PUBLIC PREVIEW

Azure NetApp Files - Volume snapshot policy

Snapshot policies to automatically create volume snapshots

GENERALLY AVAILABLE

Azure Files – Azure Backup support

Azure Files supports snapshots natively and snapshot creation and lifecycle management are seamlessly integrated into Azure Backup

Modern data protection capabilities for all storage on Azure

Azure Private Link

Securely access your data stored on Azure Storage with a private endpoint

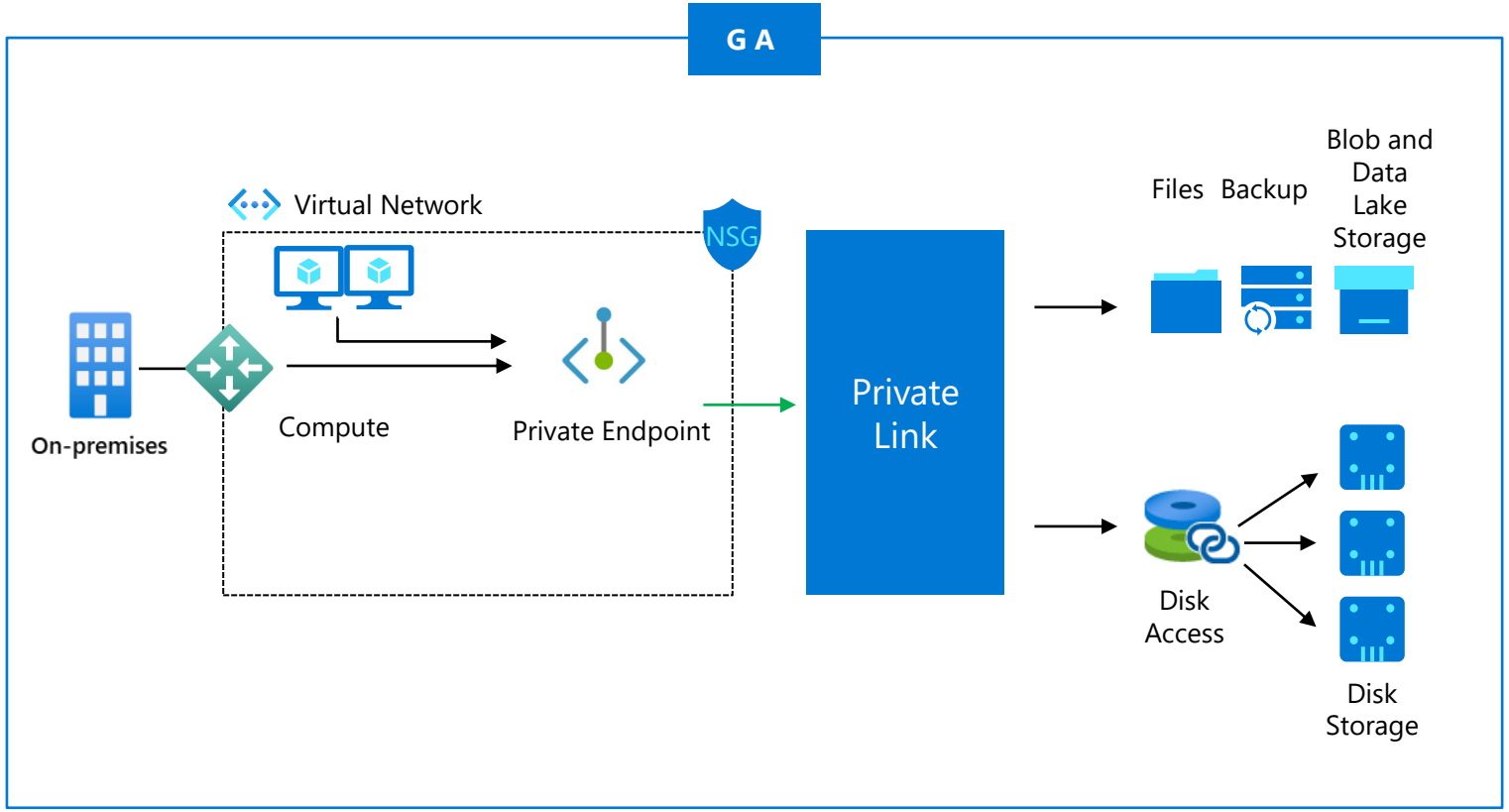
Private connectivity to storage on Azure - traffic remains on the Microsoft network, with no public internet access

Integration with on-premises and peered networks

Block exfiltration of data from the virtual network for increased security

Restrict export and import of disks to your private Azure virtual network

Supported on Azure Disk Storage, Files, Blob and Data Lake Storage and Backup



ADLS Gen2 key audiences & benefits

Data Engineer

- Supports familiar REST API and HDFS file system
- Integration with leading Hadoop and Spark analytic engines like Databricks
- Low Cost Storage for structured and unstructured data
- Strong analytics ISV support
- Granular file and folder level security
- Supports Azure Active Directory Integration and other advanced security features available in Blob Storage

Data Scientists

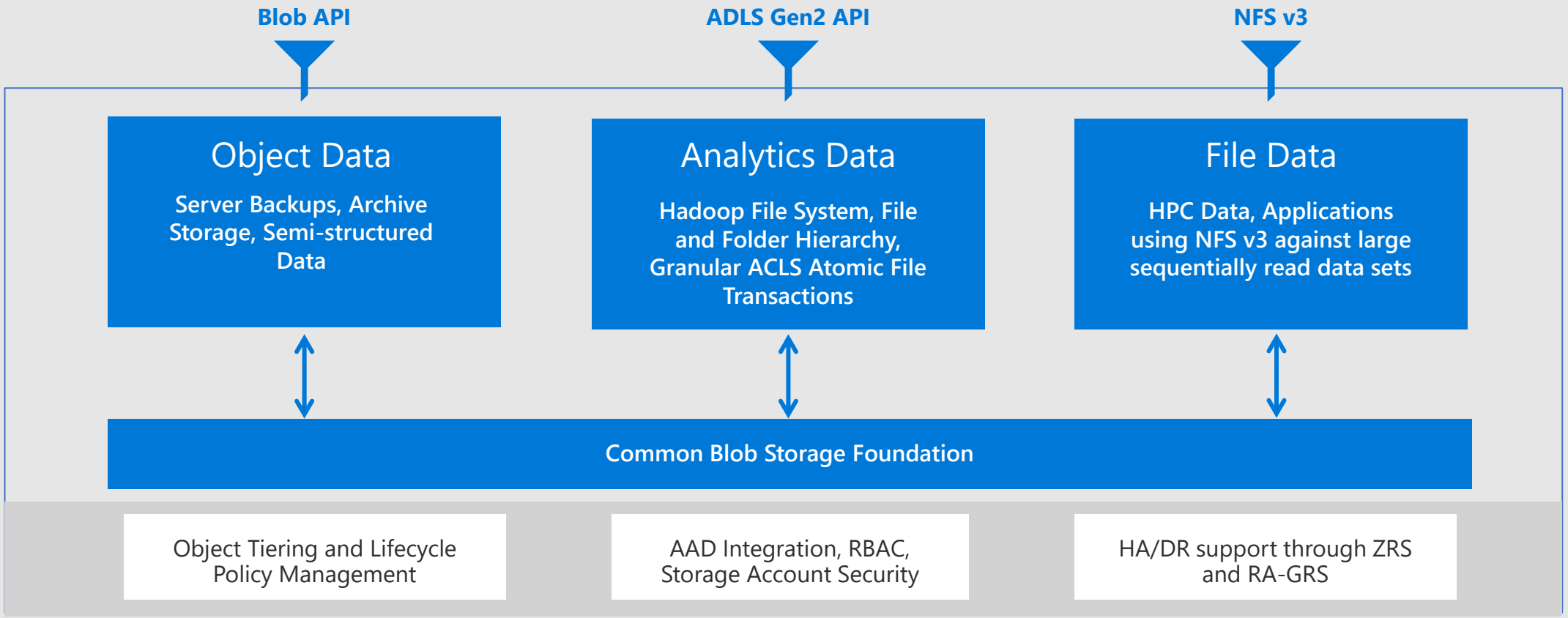
- Fast performance leads to lower end to end analytic job run times
- Integration with leading Hadoop and Spark analytic engines like Databricks
- Supports integration with visualization tools like Power BI and Tableau
- Granular file and folder level security

CDO, VP of analytics

- Low Cost Storage for structured and unstructured data
- Meets data sovereignty requirements in multiple geographies
- Granular file and folder level security
- Supports Azure Active Directory Integration and other advanced security features available in Blob Storage
- Inherits object level tiering and advanced HA/DR capabilities of Blob Storage

Azure Data Lake Storage Gen2

ADLS Gen2 adds a high performance HDFS Endpoint to Azure Blob Storage and inherits the rich feature set of Azure Blob Storage *



Public Cloud Object storage access through NFS v3 is an industry first

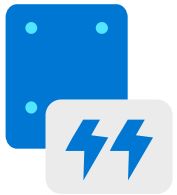
Accelerate cloud migration with Azure Storage

Azure Data Lake Storage

Native analytics workloads support using Azure Databricks, Synapse Analytics, or HDInsight

Scale storage and compute independently without limits

Automated lifecycle management policies for optimizing storage costs.



Azure Ultra Disks

For data-intensive and transaction-heavy workloads like SAP HANA

High IOPS, throughput, and sub-millisecond latency

Dynamic scaling of IOPS and throughput without disruption

Azure NetApp Files

Lift-and-shift any enterprise file workloads seamlessly into Azure

NetApp's trusted technology delivered as a cloud native service

Bare-metal performance, sub-millisecond latency

Unique storage capabilities for enterprise-class workloads

Azure Data Box family

OFFLINE DATA TRANSFER

Data Box

- Capacity: 100 TB
- Weight: ~50 lbs.
- Secure, ruggedized appliance

Data Box enables bulk migration to Azure when network isn't an option

Data Box Disk

- Capacity: 8TB ea.; 40TB/order
- Secure, ruggedized USB drives orderable in packs of 5 (up to 40TB)

Perfect for projects that require a smaller form factor e.g., autonomous vehicles

Data Box Heavy

- Capacity: 1 PB
- Weight 500+ lbs.
- Secure, ruggedized appliance

Same service as Data Box, but targeted to petabyte-sized datasets

ONLINE DATA TRANSFER

Data Box Gateway

- Virtual device provisioned in your hypervisor
- Supports storage gateway, SMB, NFS, Azure blob, files

Virtual network transfer appliance (VM), runs on your choice of hardware

Data Box Edge

- Local Cache Capacity: ~12 TB
- Includes Data Box Gateway and Azure IoT Edge

Data Box Edge manages uploads to Azure and can pre-process data prior to upload

Network Data Transfer

Edge Compute

Data Box offline device selection

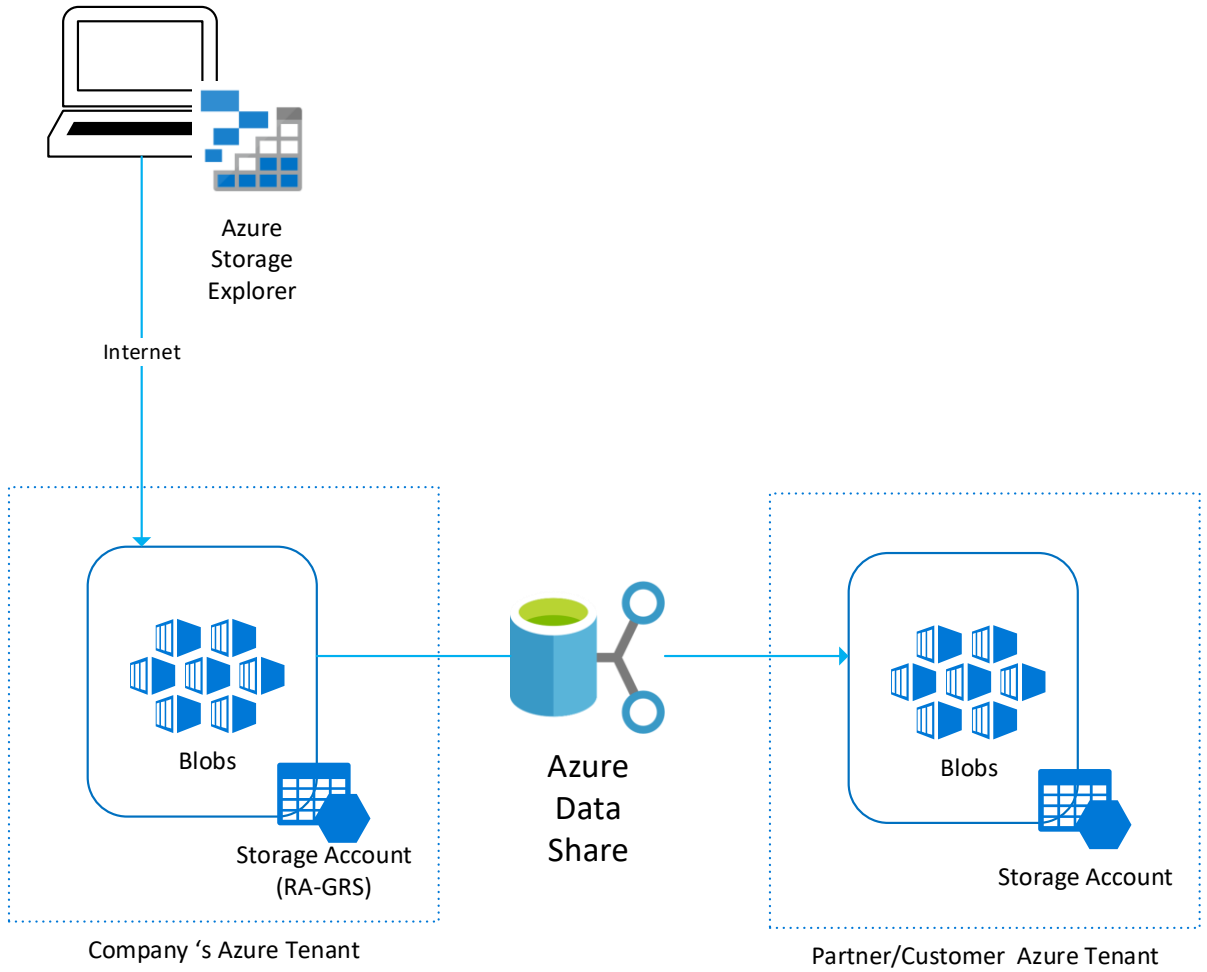
Data qty	45 Mbps (T3)	100 Mbps	1 Gbps	10 Gbps
1 TB	2 days	1 day	2 hours	14 minutes
10 TB	22 days	10 days	1 day	2 hours
35 TB	76 days	34 days	3 days	8 hours
80 TB	173 days	78 days	8 days	19 hours
100 TB	216 days	97 days	10 days	1 day
200 TB	1 years	194 days	19 days	2 days
500 TB	3 years	1 years	49 days	5 days
1 PB	6 years	3 years	97 days	10 days
2 PB	12 years	5 years	194 days	19 days

Key
Use the network
Use Data Box Disk
Use Data Box
Use Data Box Heavy

Use a Data Box when data volume exceeds network capacity

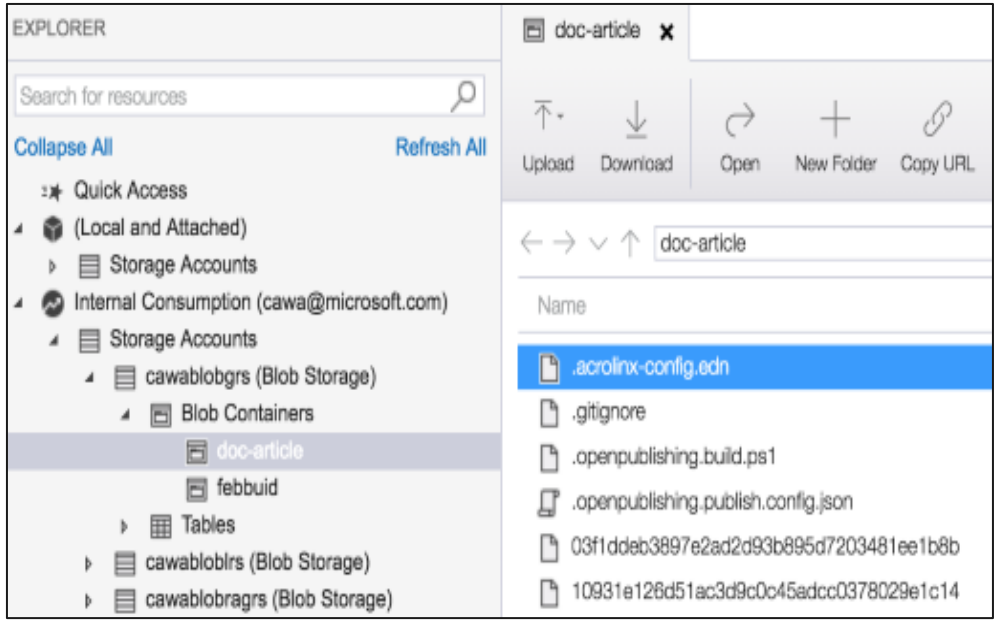
Note: the table is based on 100% available network bandwidth

Lab Architecture



Azure Storage Explorer

- Access multiple accounts and subscriptions
- Create, delete, view, edit storage resource
- View and edit Blob, Queue, Table, File, Cosmos DB storage and Data Lake Storage
- Obtain shared access signature (SAS) keys
- Available for Windows, Mac, and Linux



How customers are using Azure Data Share

1. Cross organization big data analytics
2. Share data collected on behalf of customer
3. Analytics outsourcing
4. Industry-specific data consortium
5. Data monetization and marketplace

How data share works

Cross tenant data sharing

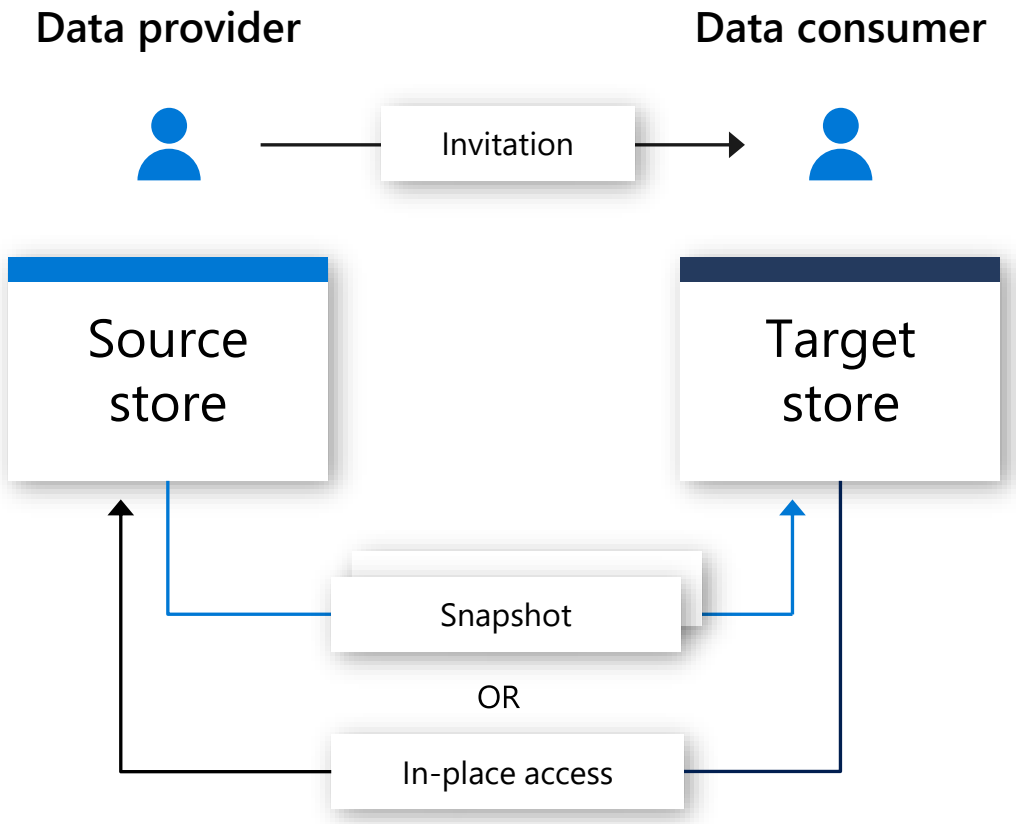
Data provider initiates sharing

- What to share
- Who to shared with
- Terms of use
- Snapshot or In-place

Data consumer accepts share

- Where to receive

Starting with Blob, ADLS, Azure SQL DB, Azure Synapse Analytics, and Azure Data Explorer



Supported Azure data stores

Heterogenous source and target so that data provider and consumer can use different storage resources

Source	Target					
	Blob Storage	ADLS Gen1	ADLS Gen2	Azure SQL DB	Azure Synapse Analytics	Azure Data Explorer
Blob Storage	Snapshot		Snapshot			
ADLS Gen1	Snapshot		Snapshot			
ADLS Gen2	Snapshot		Snapshot			
Azure SQL DB	Snapshot		Snapshot	Snapshot	Snapshot	
Azure Synapse Analytics	Snapshot		Snapshot	Snapshot	Snapshot	
Azure Data Explorer						In-place

Q & A