

# Module 3 Azure storage

## Module Overview

Microsoft Azure storage  
Azure Storage replication  
Azure storage types  
Azure services and tools  
Premium storage

# Microsoft Azure Storage

Azure Storage or Storage Account is a service from Azure, which provides storage service for various use cases.

Azure Storage is the cloud storage solution for modern applications that rely on durability, availability, and scalability to meet the needs of their customers.

- Cloud storage - anywhere and anytime access
  - Blobs, Tables, Queues and Files
- Highly durable, available and massively scalable
  - Easily build “Internet scale” applications
  - More than 25 trillion stored objects
  - 2.5+ Million requests/sec on average
- Pay for what you use
- Exposed via easy and open REST APIs, cross-platform client libraries and tools

## Security & Management

- Security Center
- Portal
- Azure Active Directory
- Azure AD B2C
- Multi-Factor Authentication
- Automation
- Scheduler
- Key Vault
- Store/ Marketplace
- VM Image Gallery & VM Depot

## Platform Services

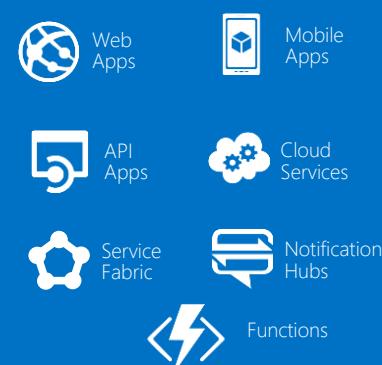
### Media & CDN



### Integration



### Application Platform



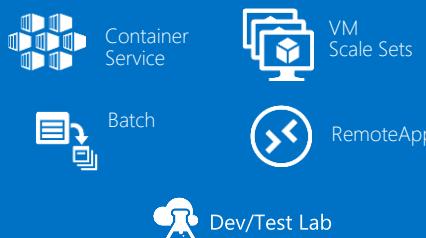
### Data



### Intelligence



### Compute Services



### Developer Services



### Analytics & IoT



## Hybrid Cloud

- Azure AD Health Monitoring
- AD Privileged Identity Management
- Domain Services
- Backup
- Operational Analytics
- Import/Export
- Azure Site Recovery
- StorSimple

### Compute



### Storage



### Infrastructure Services



Datacenter Infrastructure (42 Regions Announced, 36 Online)

# Azure Storage

## Foundational Building Block of Azure

Azure Services: SQL Data Warehouse, HDInsight, Data Lake Store, Event Hubs, IoT Hubs...

Microsoft Services: Office 365, OneDrive, XBox, Skype...

Hyper Scale

>30 million transactions per second, trillions of objects

Durable

Never lose your data. Multiple redundancy options. Automatic data checks

Secure

Encryption at Rest. Client side Encryption. Integration with KeyVault

Highly Available

Fault tolerance to hardware/software issues. Automatic load balancing

Open

REST API, Open sourced Client Libraries – .NET, Java, C++, Python, Node.js, iOS, Android, Xamarin...

Hybrid

Extensive partner ecosystem. Azure Stack for private/hosted clouds.

# Azure Storage Services

IaaS



Storage



Virtual  
machines



Networking

PaaS



Existing  
frameworks



Web  
and mobile



Microservices



Serverless  
Compute

## Disks

Persistent disks for Azure IaaS VMs

Standard Storage Disks: Magnetic disk based, low IOPS, moderate latency

Premium Storage Disks: SSD based, high IOPS, low latency

Managed Disks

## Files

Fully Managed File Shares in the Cloud

SMB and REST access

"Lift and shift" legacy apps

## Blobs

Highly scalable, REST based cloud object store

Block Blobs: Sequential I/O, Hot, Cool and Archive Tiers

Page Blobs: Random-write pattern data

Append Blobs

## Tables

Massive auto-scaling NoSQL store

Dynamic scaling based on load

Scale to PBs of table data

Fast key/value lookups

## Queues

Reliable queues at scale for cloud services

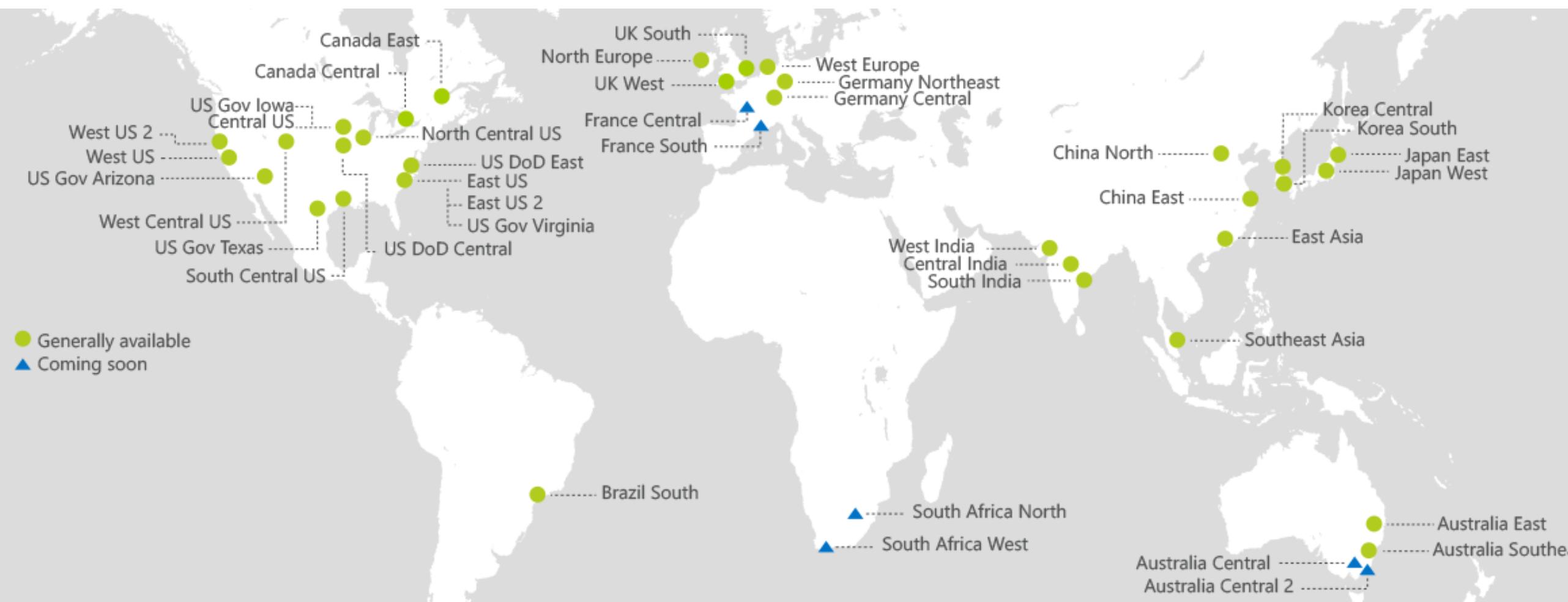
Decouple and scale components

Message visibility, timeout and update message to protect against unreliable dequeuers

Built on a unified Distributed Storage System

Durability, Encryption at Rest, Strongly Consistent Replication, Fault Tolerance, Auto Load-Balancing

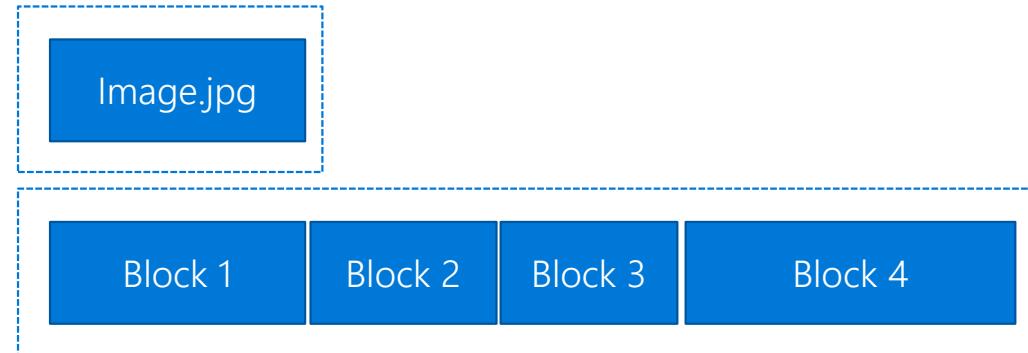
# Azure Scale



36 GA, 6 coming soon – Storage is available in every region

# What is the Blob Storage Service?

Azure Blob Storage is a service that stores unstructured data in the cloud as Objects/blobs. Blob storage can store any type of text or binary data, such as a document, media file, or application installer. It is also referred to as Object Storage.

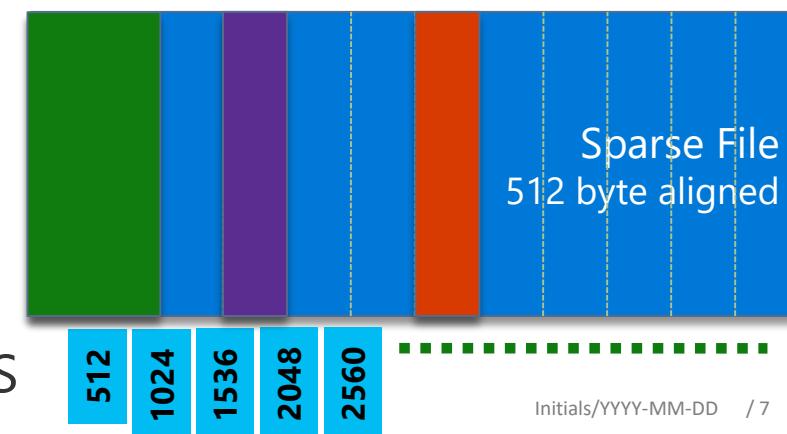


## Types of Blobs

Block Blobs - Most object storage scenarios

Append Blobs - Multi-writer append only scenarios

Page Blobs - Page aligned random reads and writes



# Blob Storage Service



## Block Blobs

Most object storage scenarios

Image.jpg

Block 1 | Block 2 | Block 3 | Block 4

## Append Blobs

Multi-writer append only scenarios

Writer 1

Writer 2

Writer 3

Block 1

Block 2

Block 3

## Page Blobs

Page aligned random reads & writes

Sparse File

512 byte aligned

512

1024

1536

2048

2560

1. Block Blobs : Ideal for working with large files. Suppose we have taken a backup of a VM which is a backup operation of huge large binary size files and we can upload that storage account as block blob.

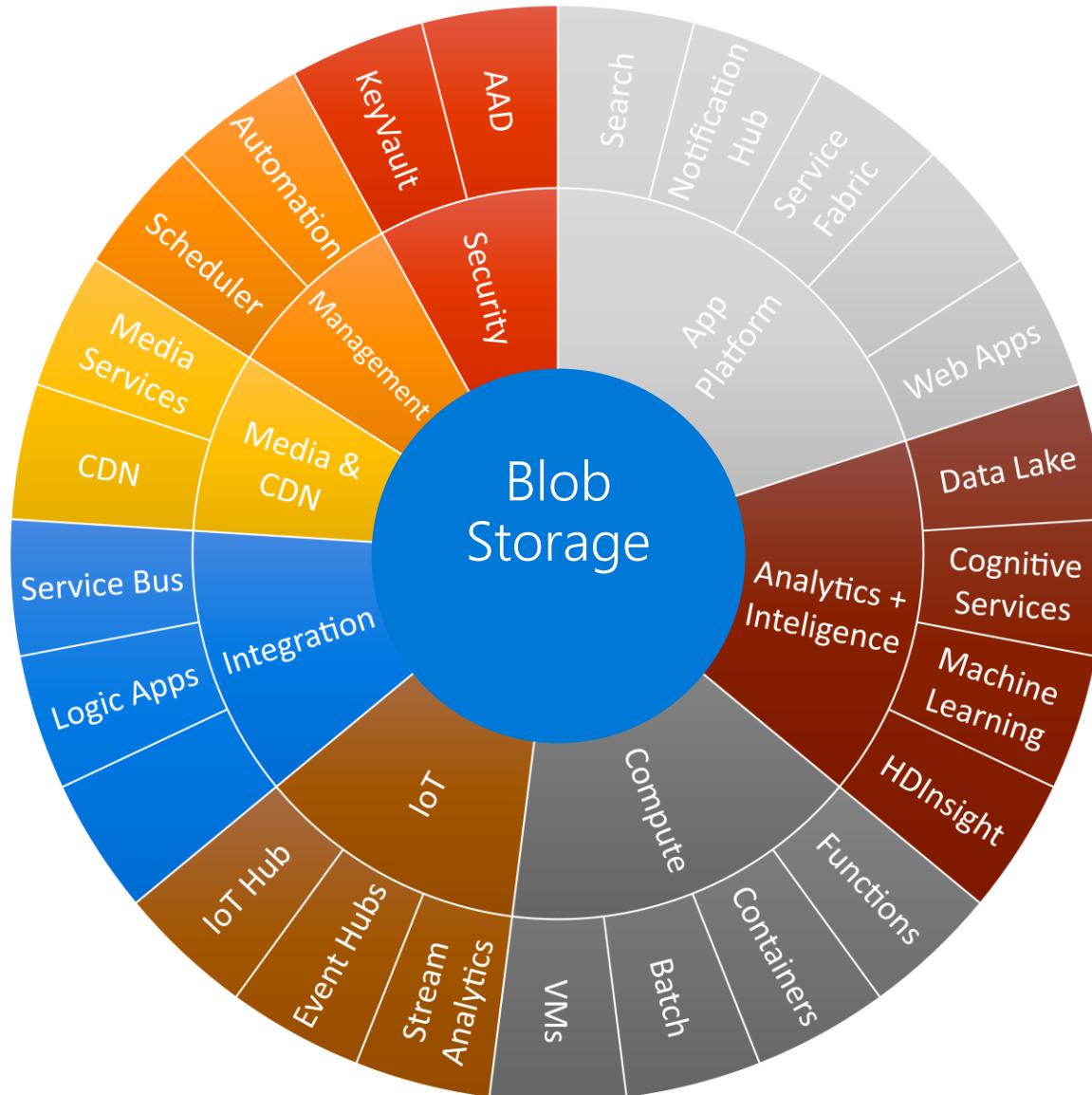
Block sizes are dynamic and it can be upto a maximum of 100MB for single block and then a block blob can contain upto 50000 block which is maximum upto 4.75 TB. This is optimised for large blocks.

2. Page Blobs : There is smaller limits on this and so page blobs can have maximum upto 1 TB. This is optimised for randomised Read and Write Operations.Thats why VHD files are ideal for page blobs.

It is going to write anything we save in a VM as VHD. it will write on individual page on blobs with maximum of 1 TB.

3. Append Blobs: This is optimised for Append operations.That means when we append anything that we add to the end of the blob. It is similar to like logging or something like systemware adding at the end of operations.

# Azure Ecosystem and Blob Storage



Broad integration for Blobs  
across Azure services

Enables many scenarios

# What is the Table Storage Service?

It provides a NoSQL Key-Value store for massive scale structured data. You can use the Table service API to create tables for structured storage, and to insert, update, delete, and query data.

Azure Table storage is a service that stores structured NoSQL data in the cloud, providing a key/attribute store with a schemaless design. Because Table storage is schemaless,

it's easy to adapt your data as the needs of your application evolve. Access to Table storage data is fast and cost-effective for many types of applications,

and is typically lower in cost than traditional SQL for similar volumes of data.

You can use Table storage to store flexible datasets like user data for web applications, address books, device information, or other types of metadata your service requires.

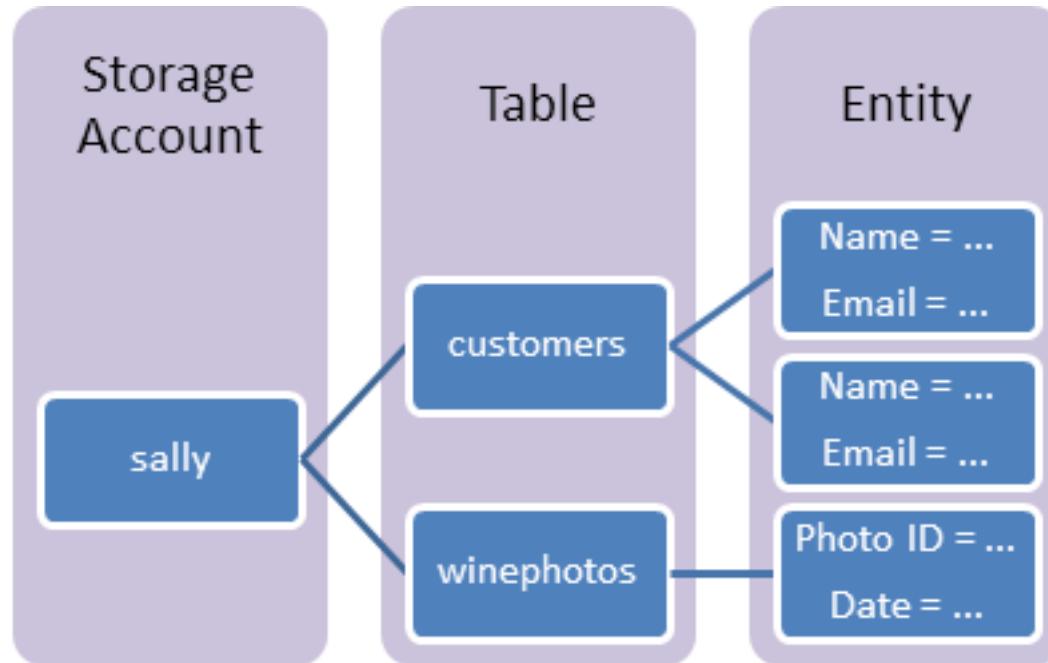
You can store any number of entities in a table, and a storage account may contain any number of tables, up to the capacity limit of the storage account.

Azure Table storage stores large amounts of structured data. The service is a NoSQL datastore which accepts authenticated calls from inside and outside the Azure cloud.

Azure tables are ideal for storing structured, non-relational data.

You can use Table storage to store and query huge sets of structured, non-relational data, and your tables will scale as demand increases.

Table storage contains the following components:

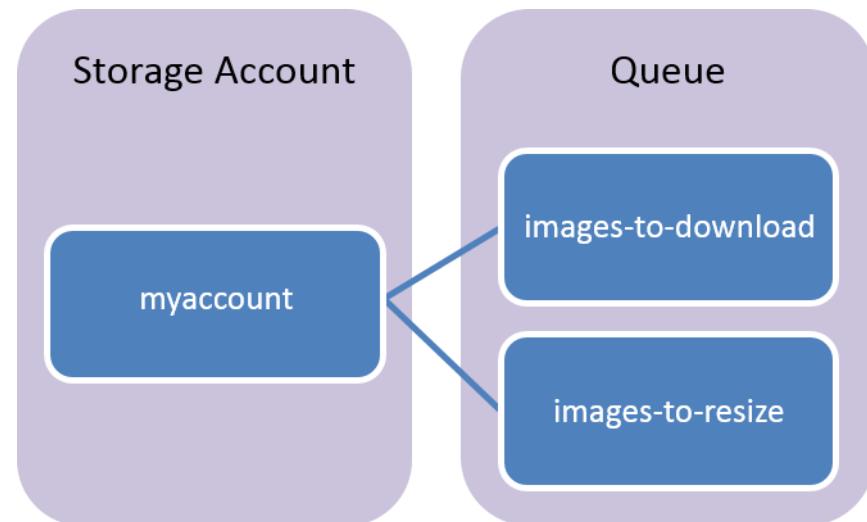


URL format: Azure Table Storage accounts use this format:  
`http://<storage account>.table.core.windows.net/<table>`

# What is the Queue Storage Service?

Azure Queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS.

A single queue message can be up to 64 KB in size, and a queue can contain millions of messages, up to the total capacity limit of a storage account.



Common uses of Queue storage include:

1. Creating a backlog of work to process asynchronously
2. Passing messages from an Azure web role to an Azure worker role

Concept : Storage Account - myaccount

Queue - images-to-download and images-to-resize

Queue: A queue contains a set of messages. All messages must be in a queue.

Note that the queue name must be all lowercase. For information on naming queues.

Message: A message, in any format, of up to 64 KB. The maximum time that a message can remain in the queue is seven days.

Azure Queue storage is a service for storing large numbers of messages that can be accessed from anywhere in the world via authenticated calls using HTTP or HTTPS.

A Single queue message can be up to 64 KB in size, and a queue can contain millions of messages, up to total capacity limit of a storage account.

URL format: Queues are addressable using the following URL format:

`http://<storage account>.queue.core.windows.net/<queue>`

# What is the File Storage Service?

Fully managed file shares in the cloud, accessible via standard Server Message Block (SMB) protocol. Enables sharing files between applications using Windows APIs or REST API.

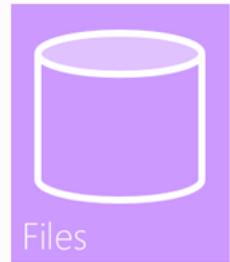
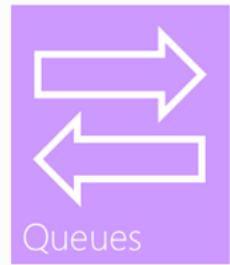
Simple, secure and fully managed cloud file shares

- 1.Extend your servers to Azure with SyncPreview for on-premises performance and capability
- 2.Secure data at rest and in-transit using SMB 3.0 and HTTPS
- 3.Simplify cloud file share management using familiar tools
- 4.Azure File storage offers network file shares in the cloud
- 5.Azure Files can be used to completely replace or supplement traditional on-premises file servers or NAS devices.
- 6.Azure Files makes it easy to "lift and shift" applications to the cloud that expect a file share to store file application or user data.



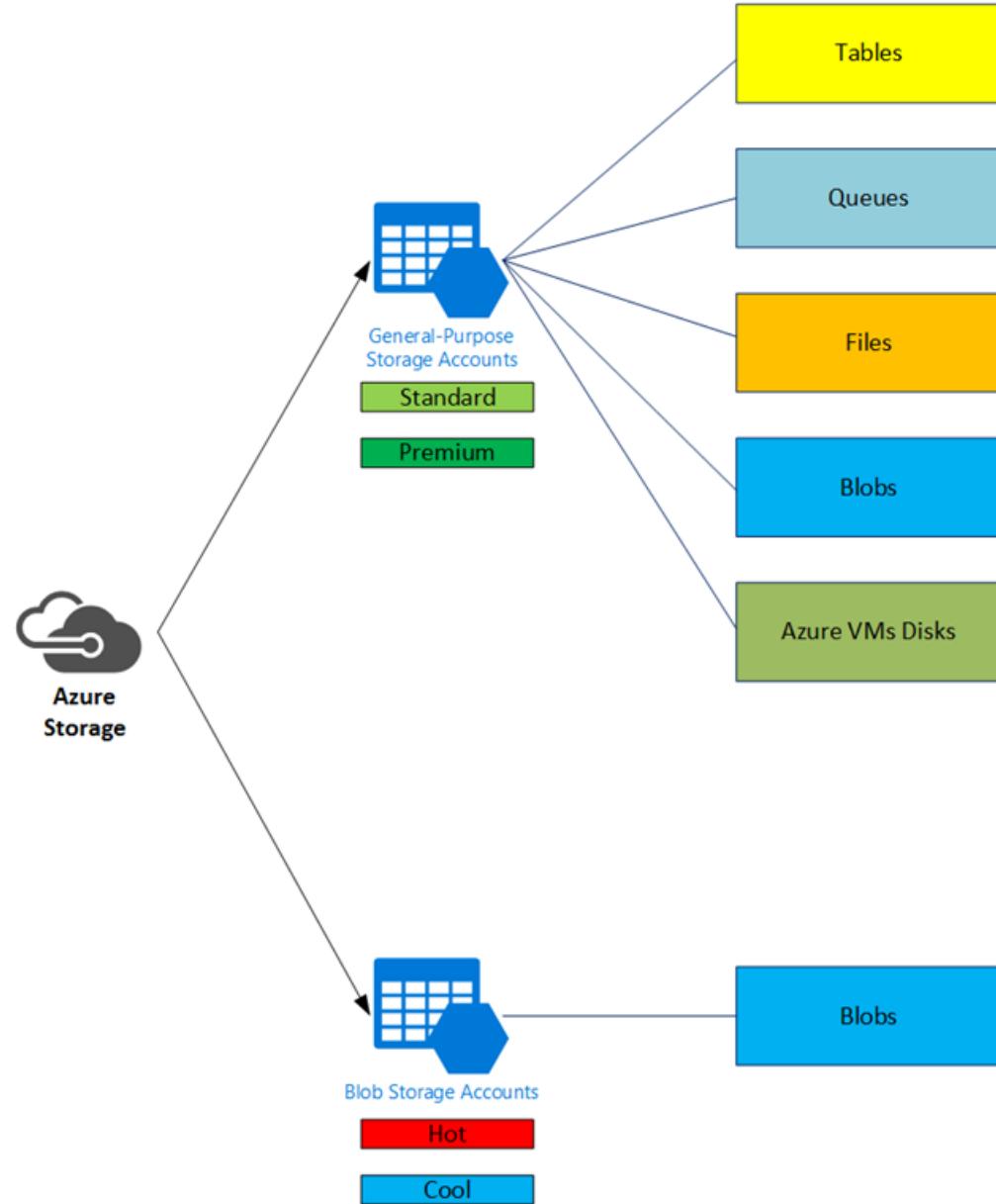
Storage for any type of data, analogous to files in a file system, with individual blobs storing up to 1 TB of data

NoSQL storage of semi-structured data for rapid development and fast access to large quantities of data

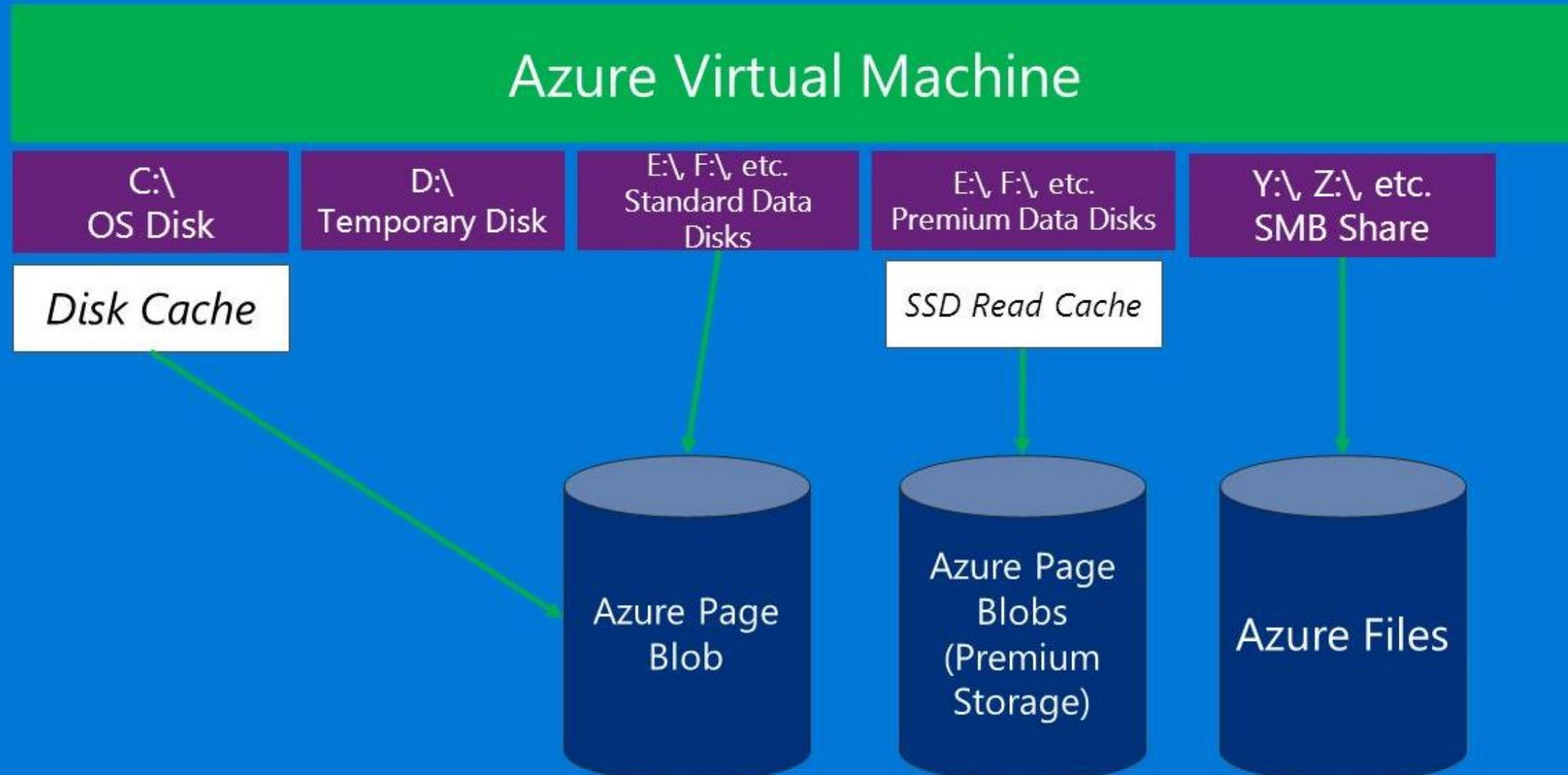


Reliable messaging for workflow processing and for communication between applications or application components

Shared storage for VMs and cloud services using Server Message Block (SMB) protocol



# Summary - Virtual Machine Storage Architecture



An Azure virtual machine supports attaching a number of data disks. This article describes scalability and performance targets for a VM's data disks. Use these targets to help decide the number and type of disk that you need to meet your performance and capacity requirements

## Azure Managed Disks:

The disk limit for managed disks is per region and per disk type. The maximum limit, and also the default limit, is 10,000 managed disks per region and per disk type for a subscription. For example, you can create up to 10,000 standard managed disks and also 10,000 premium managed disks in a region, per subscription. Managed snapshots and images count against the managed disks limit.

## Standard storage accounts:

A standard storage account has a maximum total request rate of 20,000 IOPS. The total IOPS across all of your virtual machine disks in a standard storage account should not exceed this limit.

You can roughly calculate the number of highly utilized disks supported by a single standard storage account based on the request rate limit. For example, for a Basic Tier VM, the maximum number of highly utilized disks is about 66 ( $20,000/300$  IOPS per disk), and for a Standard Tier VM, it is about 40 ( $20,000/500$ IOPS per disk).

## Premium storage accounts:

A premium storage account has a maximum total throughput rate of 50 Gbps. The total throughput across all of your VM disks should not exceed this limit.

# Managed virtual machine disks

## Standard managed virtual machine disks

STANDARD DISK TYPE	S4	S6	S10	S20	S30	S40	S50
Disk size	32 GB	64 GB	128 GB	512 GB	1024 GB (1 TB)	2048 GB (2TB)	4095 GB (4 TB)
IOPS per disk	500	500	500	500	500	500	500
Throughput per disk	60 MB/sec	60 MB/sec	60 MB/sec				

### Premium managed virtual machine disks: per disk limits

PREMIUM DISKS TYPE	P4	P6	P10	P20	P30	P40	P50
Disk size	32 GB	64 GB	128 GB	512 GB	1024 GB (1 TB)	2048 GB (2 TB)	4095 GB (4 TB)
IOPS per disk	120	240	500	2300	5000	7500	7500
Throughput per disk	25 MB/sec	50 MB/sec	100 MB/sec	150 MB/sec	200 MB/sec	250 MB/sec	250 MB/sec

### Premium managed virtual machine disks: per VM limits

RESOURCE	DEFAULT LIMIT
Max IOPS Per VM	80,000 IOPS with GS5 VM
Max throughput per VM	2,000 MB/s with GS5 VM

# Unmanaged virtual machine disks

## Standard unmanaged virtual machine disks: per disk limits

VM TIER	BASIC TIER VM	STANDARD TIER VM
Disk size	4095 GB	4095 GB
Max 8 KB IOPS per persistent disk	300	500
Max number of disks performing max IOPS	66	40

## Premium unmanaged virtual machine disks: per account limits

RESOURCE	DEFAULT LIMIT
Total disk capacity per account	35 TB
Total snapshot capacity per account	10 TB
Max bandwidth per account (ingress + egress <sup>1</sup> )	<=50 Gbps

<sup>1</sup>Ingress refers to all data (requests) being sent to a storage account. Egress refers to all data (responses) being received from a storage account.

## Premium unmanaged virtual machine disks: per disk limits

PREMIUM STORAGE DISK TYPE	P10	P20	P30	P40	P50
Disk size	128 GiB	512 GiB	1024 GiB (1 TB)	2048 GiB (2 TB)	4095 GiB (4 TB)
Max IOPS per disk	500	2300	5000	7500	7500
Max throughput per disk	100 MB/s	150 MB/s	200 MB/s	250 MB/s	250 MB/s
Max number of disks per storage account	280	70	35	17	8

## Premium unmanaged virtual machine disks: per VM limits

RESOURCE	DEFAULT LIMIT
Max IOPS Per VM	80,000 IOPS with GS5 VM
Max throughput per VM	2,000 MB/s with GS5 VM

# Premium Storage



High Bandwidth with Low Latency

Up to **64 TB** of storage per VM

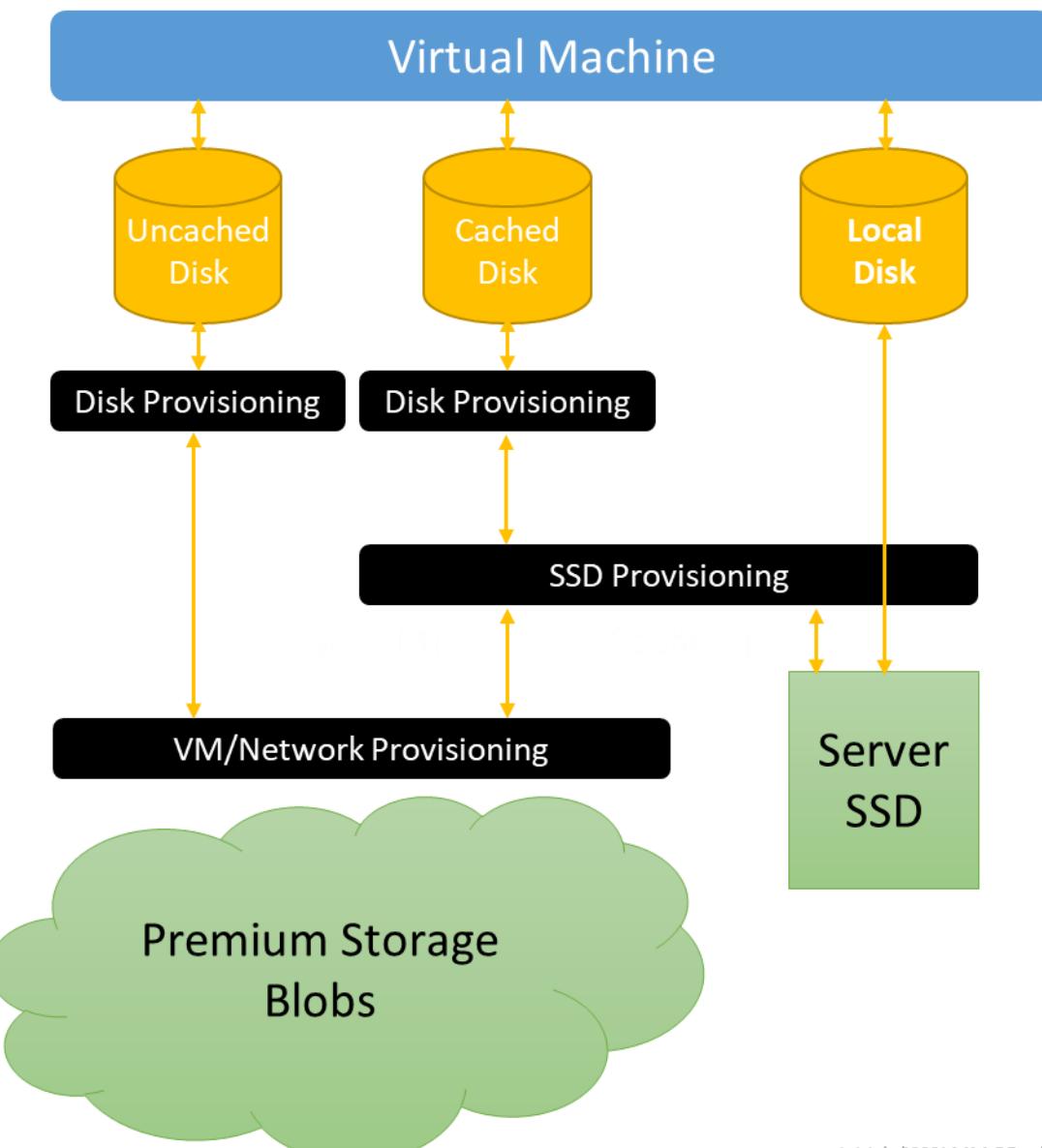
80,000 IOPS per VM

5,000 IOPS per disk

~5 ms read/write (no cache)

Less than 1ms read latency (cache)

*\*Supports only Azure VM*



# Azure Storage

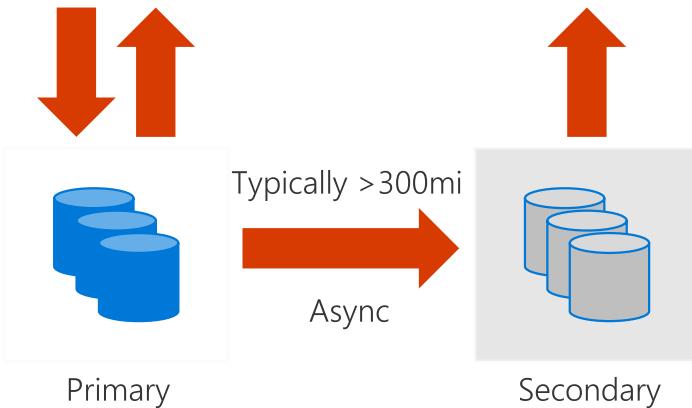
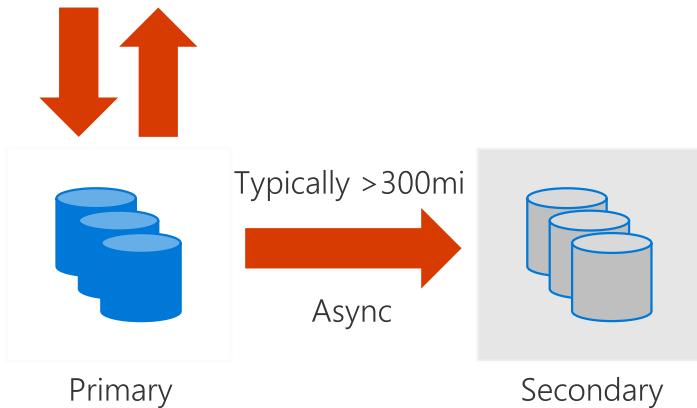
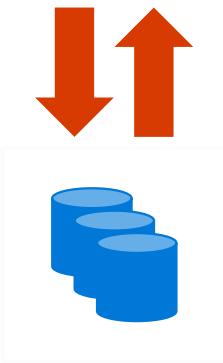
Azure Premium Storage delivers high-performance, low-latency disk support for virtual machines running I/O-intensive workloads. Virtual machine (VM) disks that use Premium Storage store data on solid state drives (SSDs). You can migrate your application's VM disks to Azure Premium Storage to take advantage of the speed and performance of these disks.

## Premium storage accounts

Virtual machine disks: per account limits

Resource	Default Limit		
Total disk capacity per account	35 TB		
Total snapshot capacity per account	10 TB		
Max bandwidth per account (ingress + egress <sup>1</sup> )	<=50 Gbps		
<small><sup>1</sup>Ingress refers to all data (requests) being sent to a storage account. Egress refers to all data (responses) being received from a storage account.</small>			
Virtual machine disks: per disk limits			
Premium Storage Disk Type	P10	P20	P30
Disk size	128 GiB	512 GiB	1024 GiB (1 TB)
Max IOPS per disk	500	2300	5000
Max throughput per disk	100 MB per second	150 MB per second	200 MB per second
Max number of disks per storage account	280	70	35

# Azure Storage Durability



## LRS

3 replicas, 1 region  
Protect against disk, node, rack failures  
Write is ack'd when all replicas are committed  
Superior to dual-parity RAID

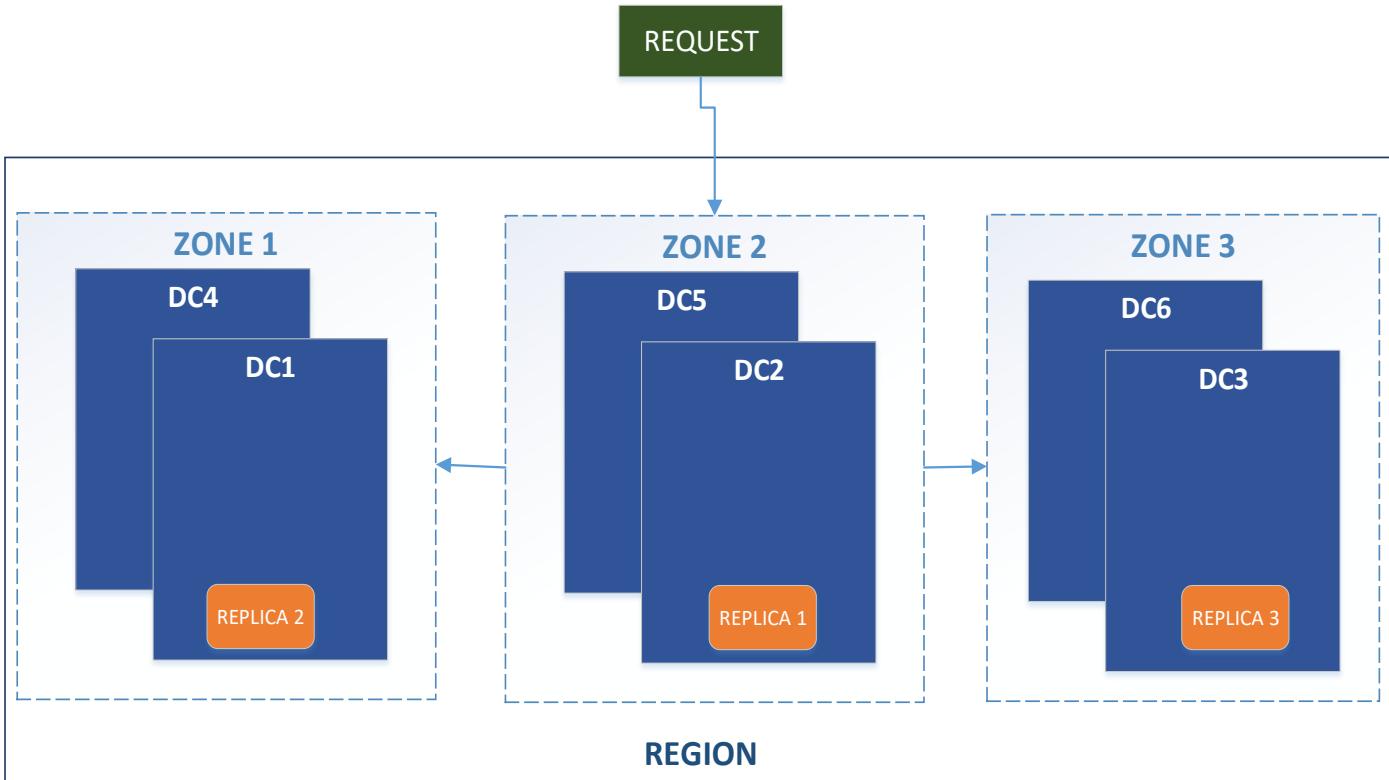
## GRS

6 replicas, 2 regions (3/region)  
Protects against major regional disasters  
Asynchronous to secondary

## RA-GRS

GRS + Read access to secondary  
Separate secondary endpoint  
RPO delay to secondary can be queried

# NEW - Azure Zone Redundant Storage (ZRS)



Read / write resilience against single cluster / DC unavailability

Support for Blob, Table, File, Queue Storage

Public Preview in Q4 CY 2017 in multiple regions.  
GA in H1 CY 2018

Synchronous data replication across [Azure Availability Zones](#) within region

## LRS

Resilient to disk/node/rack failures

## ZRS

Resilient to single cluster / datacenter outage

## GRS

Resilient to regional outage

## RA-GRS

Resilient to regional outage  
Read access to second region

# Microsoft Azure Storage: Redundancy

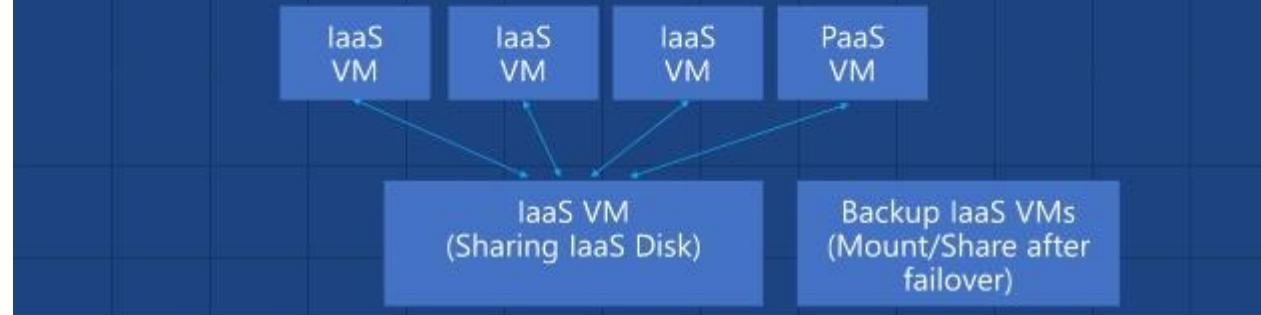
	LOCALLY REDUNDANT STORAGE (LRS)	ZONE REDUNDANT STORAGE (ZRS)	GEOGRAPHICALLY REDUNDANT STORAGE (GRS)	READ-ACCESS GEOGRAPHICALLY REDUNDANT STORAGE (RA-GRS)
How it works	Makes multiple synchronous copies of your data within a single datacenter	Stores three copies of data across multiple datacenters within or across regions. For block blobs only.	Same as LRS, plus multiple asynchronous copies to a second datacenter hundreds of miles away	Same as GRS, plus read access to the secondary datacenter
Total copies	3	3	6	6
Why use it	For economical local storage or data governance compliance	An economical, higher durability option for block blob storage	For protection against a major datacenter outage or disaster	Provides read access to data during an outage, for maximum data availability and durability
Availability SLA	99.9% read/write	99.9% read/write	99.9% read/write	99.9% write 99.99% read

- Locally redundant storage (LRS) replicates your data three times within a storage scale unit, which is hosted in a datacenter in the region in which you created your storage account. A write request returns successfully only once it has been written to all three replicas. These three replicas each reside in separate fault domains and upgrade domains within one storage scale unit.
- Zone-redundant storage (ZRS) replicates your data asynchronously across datacenters within one or two regions in addition to storing three replicas similar to LRS, thus providing higher durability than LRS. Data stored in ZRS is durable even if the primary datacenter is unavailable or unrecoverable.
- Geo-redundant storage (GRS) replicates your data to a secondary region that is hundreds of miles away from the primary region. If your storage account has GRS enabled, then your data is durable even in the case of a complete regional outage or a disaster in which the primary region is not recoverable.
- Read-access geo-redundant storage (RA-GRS) maximizes availability for your storage account, by providing read-only access to the data in the secondary location, in addition to the replication across two regions provided by GRS

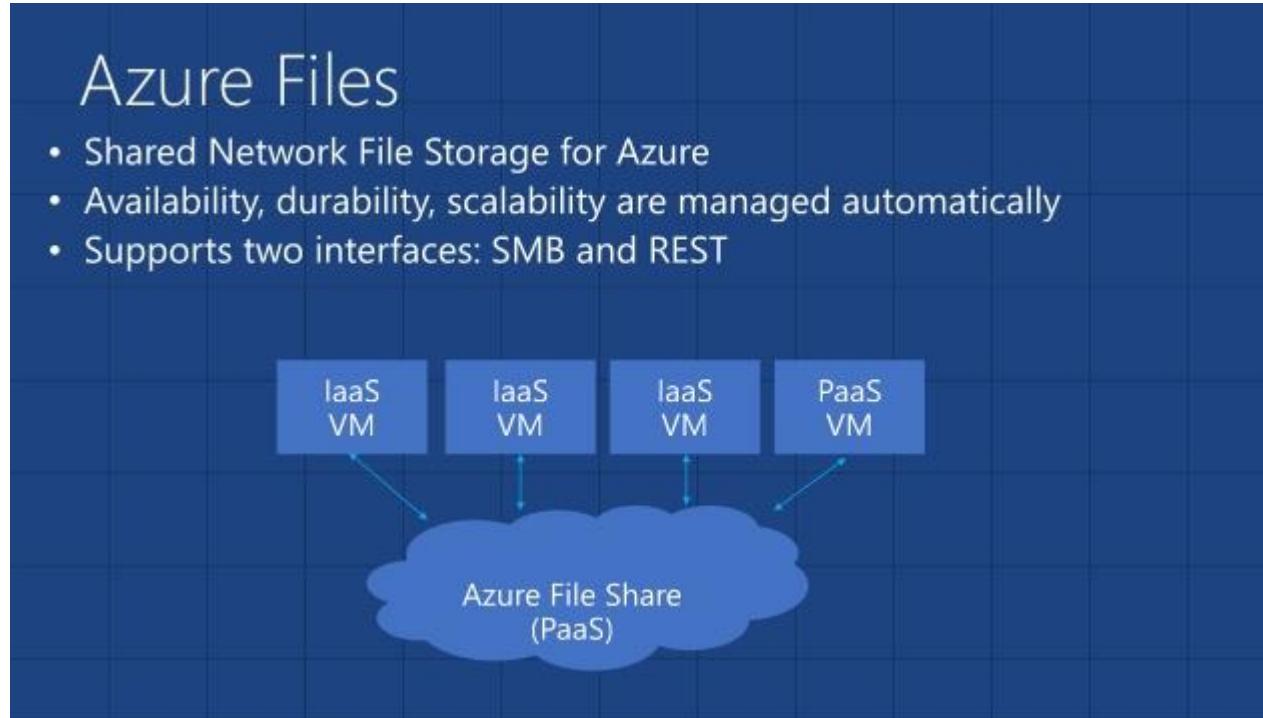
# File sharing the old way in Azure

## Sharing Files – The old way

- Setup an IaaS VM to host a File Share backed by an IaaS Disk
- Write code to find the IaaS File Share from the rest of the VMs in your service.
- Write some code to provide high availability
  - Handle host upgrades, node failures
- You can only access the File Share from other VMs

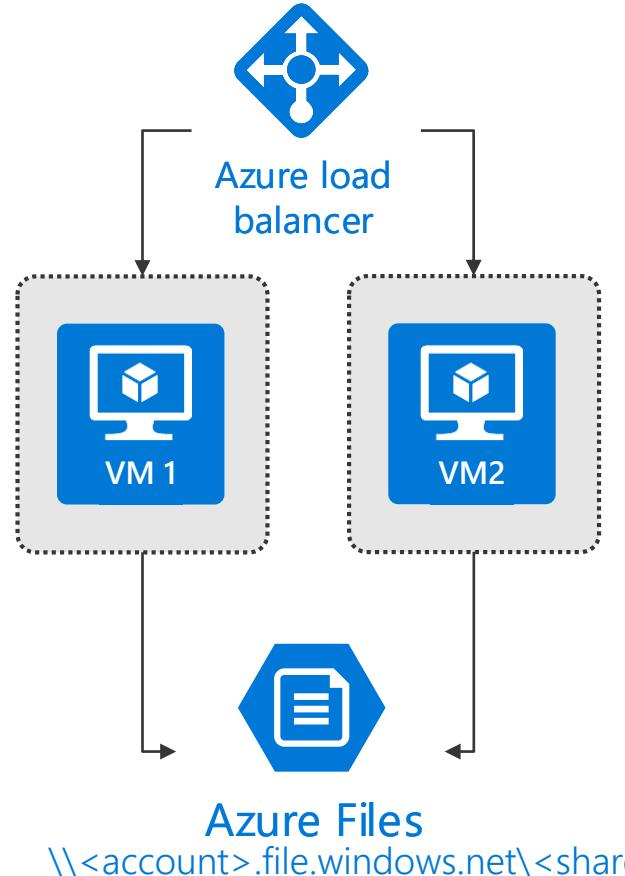


# File sharing nowadays in Azure



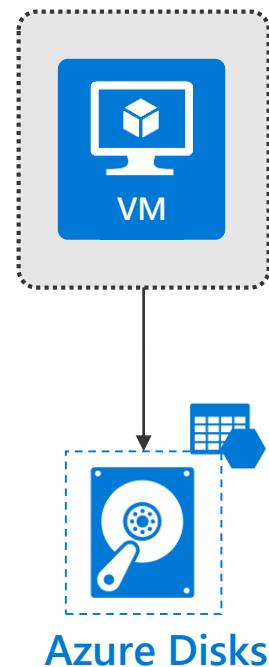
# Top Use Cases: Highly Available FTP Server

FTP server using Files



VS

FTP server using Disks



Increased availability through multiple VM instances

Single VM instance is prone to availability loss

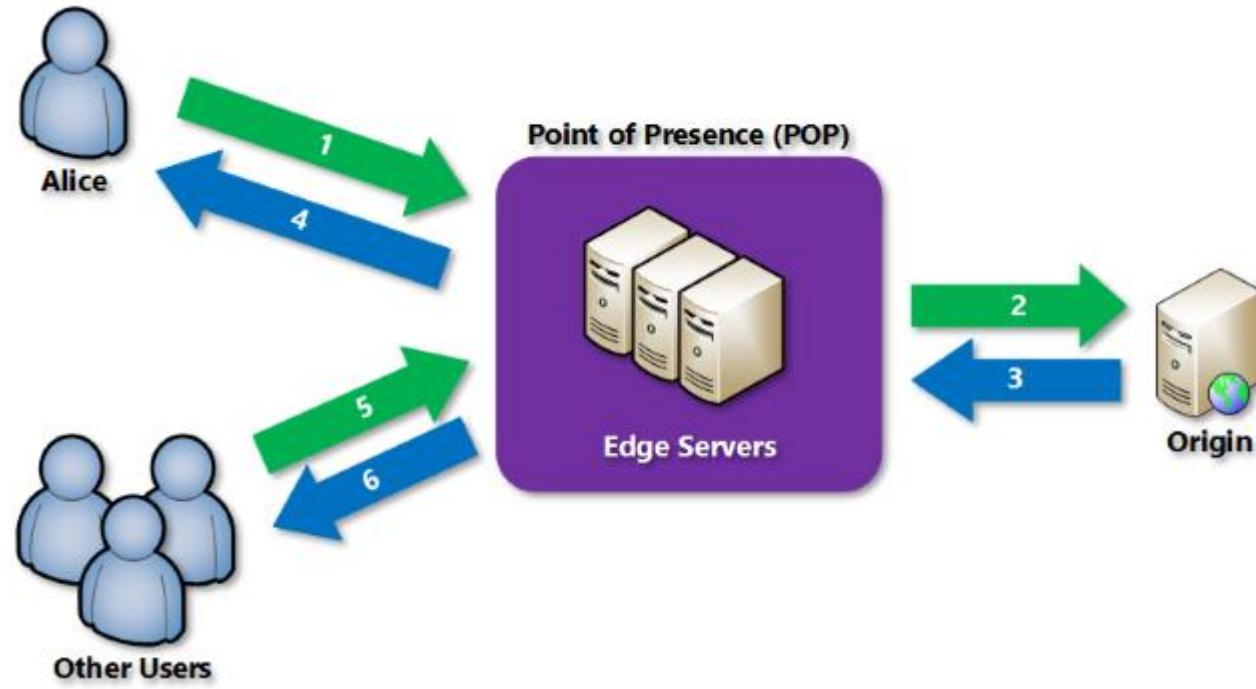
# Content Delivery Network(CDN)

The Azure Content Delivery Network (CDN) caches static web content at strategically placed locations to provide maximum throughput for delivering content to users. The CDN offers developers a global solution for delivering high-bandwidth content by caching the content at physical nodes across the world.

It is designed to send audio, video, images, and other files faster and more reliably to customers using servers that are closest to the users. This dramatically increases speed and availability, resulting in significant user experience improvements.

## Benefits

- Better performance and user experience for end users,
- Large scaling
- By distributing user requests and serving content from edge servers, less traffic is sent to the origin.



There are three Azure CDN products:

- Azure CDN Standard from Akamai,
- Azure CDN Standard from Verizon,
- Azure CDN Premium from Verizon

You can Integrate an Azure storage account with Azure CDN

# Azure CDN POP Locations

Region	Verizon	Akamai
North America	Atlanta, GA, Boston, MA , Chicago, IL , Dallas, TX, Denver, CO ,Los Angeles, CA Miami, FL ,New York, NY ,Philadelphia, PA, San Jose, CA ,Seattle, WA , Washington DC	Canada Mexico United States
South America	Buenos Aires, Argentina , Rio de Janeiro, Brazil ,São Paulo, Brazil ,Valparaíso, Chile Barranquilla, Colombia ,Medellin, Colombia , Quito, Ecuador ,Lima, Peru	Argentina ,Brazil ,Chile ,Colombia ,Ecuador ,Peru ,Uruguay
Northern and Eastern Europe	Copenhagen, Denmark, Helsinki, Finland ,Warsaw, Poland Stockholm, Sweden	Bulgaria , Denmark ,Finland ,Norway ,Poland ,Sweden
Western Europe	Vienna, Austria ,Paris, France ,Frankfurt, Germany , Amsterdam, Netherlands, London, UK	Austria , Belgium ,France ,Germany ,Ireland ,Netherlands ,Switzerland United Kingdom
Southern Europe	Milan, Italy Madrid, Spain	Greece ,Italy , Portugal ,Spain
East Asia	Hong Kong ,Batam, Indonesia ,Jakarta, Indonesia, Osaka, Japan, Tokyo, Japan ,Singapore Seoul, South Korea, Kaohsiung, Taiwan	Hong Kong, Indonesia , Japan, Macau, Malaysia, Philippines, Singapore, ,South Korea ,Taiwan ,Thailand
South and Central Asia	Bangalore, India , Chennai, India , Delhi, India ,Mumbai, India ,New Delhi, India	India Sri Lanka
Middle East/West Asia	Muscat, Oman Fujirah, United Arab Emirates	Israel, Kuwait, Qatar ,Turkey, United Arab Emirates
Africa		Egypt, South Africa
Australia and New Zealand	Melbourne, Australia, Sydney, Australia, Auckland, New Zealand	Australia, New Zealand

# Cross-Origin Resource Sharing(CORS)

CORS is an HTTP feature that enables a web application running under one domain to access resources in another domain. Web browsers implement a security restriction known as same-origin policy that prevents a web page from calling APIs in a different domain. CORS provides a secure way to allow one domain (the origin domain) to call APIs in another domain.

We can set CORS rules individually for each of the storage services(i.e. blob, file, queue, table).Once we set the CORS rues for the service then a properly authenticated request made against the service from different domain will be evaluated to determine whether it is allowed according to the rules we have specified.

CORS rules allow clients to access blobs from any web domains you authorize.

## CORS configuration

Enable API based CORS Configuration	<input checked="" type="checkbox"/>
Access Control Allow Origins	<input checked="" type="checkbox"/> All Allow Origins
Access Control Allow Headers	authorization <span style="border: 1px solid #ccc; padding: 2px;">x</span> Access-Control-Allow-Origin <span style="border: 1px solid #ccc; padding: 2px;">x</span> Content-Type <span style="border: 1px solid #ccc; padding: 2px;">x</span> SOAPAction <span style="border: 1px solid #ccc; padding: 2px;">x</span>
Access Control Allow Methods	GET <span style="border: 1px solid #ccc; padding: 2px;">x</span> PUT <span style="border: 1px solid #ccc; padding: 2px;">x</span> POST <span style="border: 1px solid #ccc; padding: 2px;">x</span> DELETE <span style="border: 1px solid #ccc; padding: 2px;">x</span> PATCH <span style="border: 1px solid #ccc; padding: 2px;">x</span> OPTIONS <span style="border: 1px solid #ccc; padding: 2px;">x</span>
Access Control Allow Credentials	<input checked="" type="checkbox"/>

Save Next : Manage >

Cross origin resource sharing is required when you are dealing with multiple domains and all of them need to be able to make calls to specific sub-domain or the API layer. Many times we even need to allow the Partner networks to have access to such API sub-domains. One can do this on backend servers but it gets complicated quickly and every change needs to be replicated on multiple backend servers in the setup. Doing the same through the load balancing setup is a much simpler way to get there.

## Process flow for CORS requests

