Azure Storage Part I

CSCI E-94

Fundamentals of Cloud Computing - Azure

Joseph Ficara

Portions © 2013-2025

Agenda

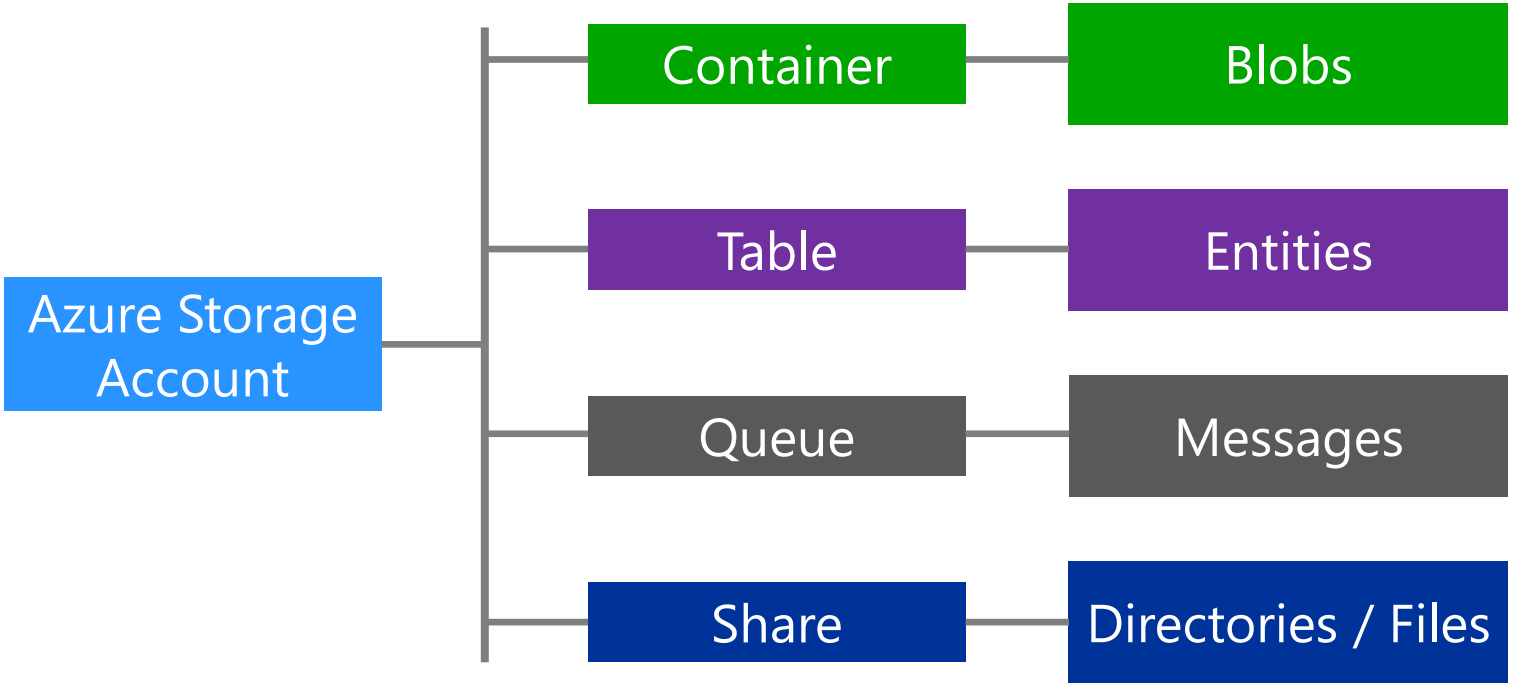
- Overview
- Azure Storage Provisioning
- Azure Block Blob Storage
  - Implementation

Overview

- "*Azure Storage*" has 4 storage facilities
  - Blobs
  - Queues
  - Tables
  - Files

Azure Storage

- Azure Storage resources relationships



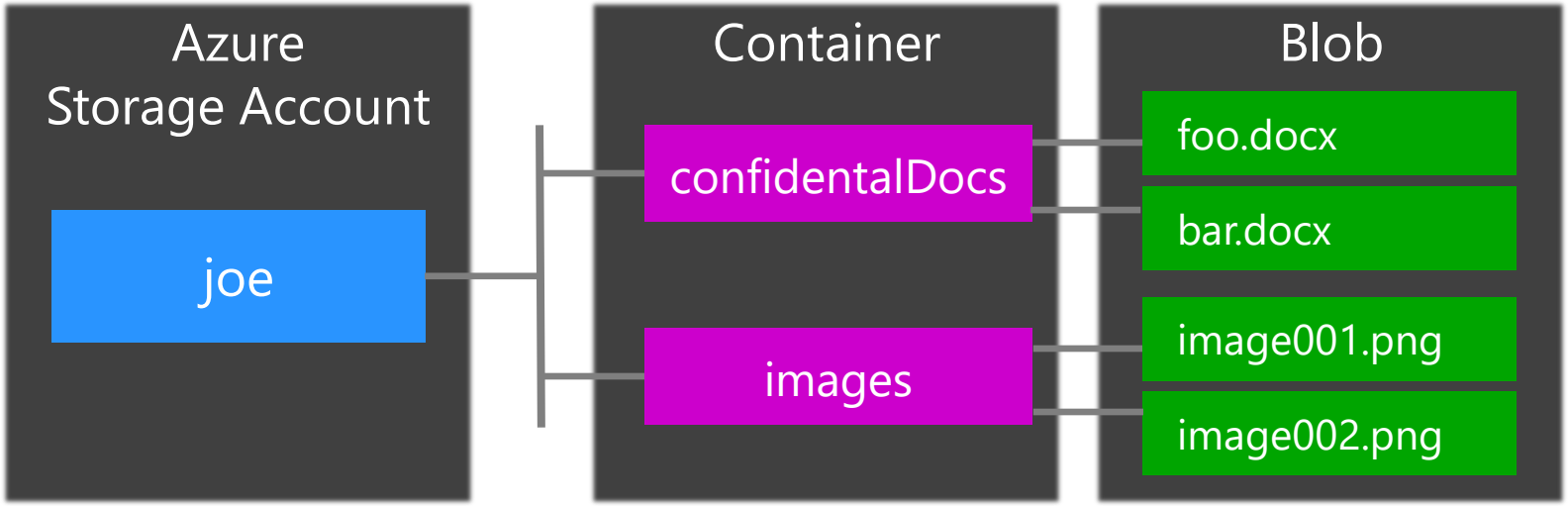
See: Azure Storage Concepts: http://bit.ly/3twKleu

Overview
Azure Blob Storage

- What is it?
  - Highly scalable storage of unstructured data
- Common uses
  - Serving images/documents direct to browser
  - Storing files for distributed access
  - Streaming video & audio
  - Secure backup and recovery
  - Offloading data
    - For large scale processing in Azure

Overview
Azure Blob Storage

- Blob Service Concepts



URL to access blob:
http://<storage account>.blob.core.windows.net/<container>/<blob>

Adapted from: Link

Overview
Azure Blob Storage

- General Features
  - Low storage cost
    - East US as of 2/15/2025
      - 1000 GB Locally Redundant $20.80 USD
        - 10 Million Write Storage Operations: $5.00
      - 1000 GB Geo Redundant $45.80 USD
        - 10 Million Write Storage Operations: $10.00
    - See Azure Calculator: http://bit.ly/1cMjBOY

Overview
Azure Blob Storage

- Options
  - Locally-redundant storage LRS
    - 99.999999999% (11 9's)
      - durability of objects over a give year.
  - Zone-redundant storage (ZRS)
    - 99.9999999999% (12 9's)
      - durability of objects over a given year
  - Geo-redundant storage (GRS)
    - 99.99999999999999% (16 9's)
      - durability of objects over a given year

Overview
Azure Blob Storage

- Options …
  - Hot storage (default)
    - Frequent read/write data
      - Staged for processing and migrated to cool tier
  - Cool tier
    - Infrequently accessed and stored for 30+ days
      - Short-term backup and disaster recovery
      - Logs for periodic analysis
      - Minimum retention of 30 days.
      - Available with low latency (milliseconds)

Overview
Azure Blob Storage

- Options …
  - Cold tier
    - Infrequently accessed and stored for 90+ days
      - Rarely access data where some latency is tolerable
      - Medium term compliance & regulatory storage
      - Older analytics data sets
      - Minimum retention of 90 days.
      - Available with low latency (milliseconds)

Overview
Azure Blob Storage

- Options …
  - Archive storage
    - Rarely accessed and stored for 180+ days
      - Regulatory or legal hold files
        - Must be kept for years
      - Compliance and archival data
        - Security camera footage, X-Rays/MRIs,
        - Audio Recordings
        - Transcripts of calls for financial service
      - Any data you do not anticipate needing to retrieve quickly
      - Latency in the order of hours

Overview
Azure Blob Storage

- Options …
  - Premium storage
    - High performance hardware
  - SSD vs Disk
  - Encryption at rest
  - Highly scalable
    - As of February 15, 2025
      - 5 PiB for US and Europe, can request increases
        - See Azure Limits: http://bit.ly/2HcRdpt

Overview
Azure Blob Storage

- Azure Block Blobs
  - Optimized for streaming & storing objects
    - Maximum size of a block in a blob
      - 4000 MiB = $\mathbf{4000\ x\ (20^{20})}$
    - Maximum number of blocks in a blob
      - 50,000 blocks

Overview
Azure Blob Storage

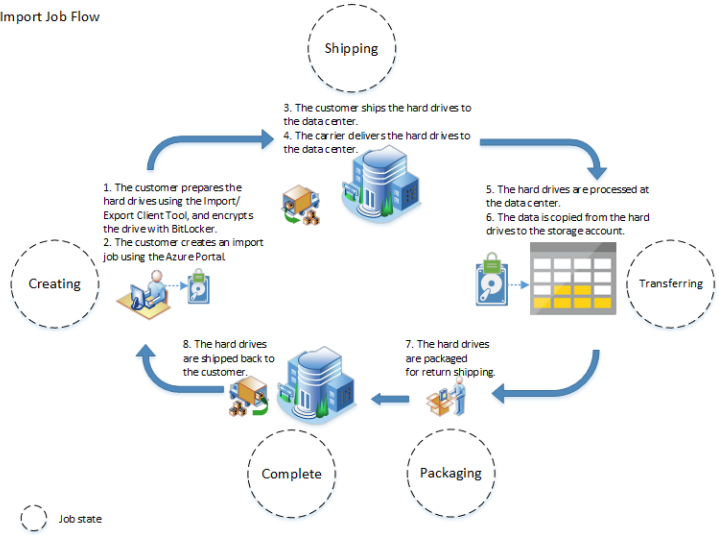
- Azure Block Blobs
  - Storage of unstructured data
    - Documents, Backups
    - Media, Images, Audio, Streaming
    - Big data
      - Logs & Large Datasets

Overview
Azure Blob Storage

- Azure Page Blobs
  - Optimized for representing IaaS disks
    - Up to 8 TiB = $\mathbf{8\ x\ (20^{40})}$ in size per page blob
    - Azure virtual machine network attached disk
      - VHD stored as a page blob
  - Supports random writes
- Azure Append Blobs
  - Optimized for append operations
    - Update and delete are not supported
      - 195 GiB = $\mathbf{195\ x\ (20^{30})}$ in size per appended block blob

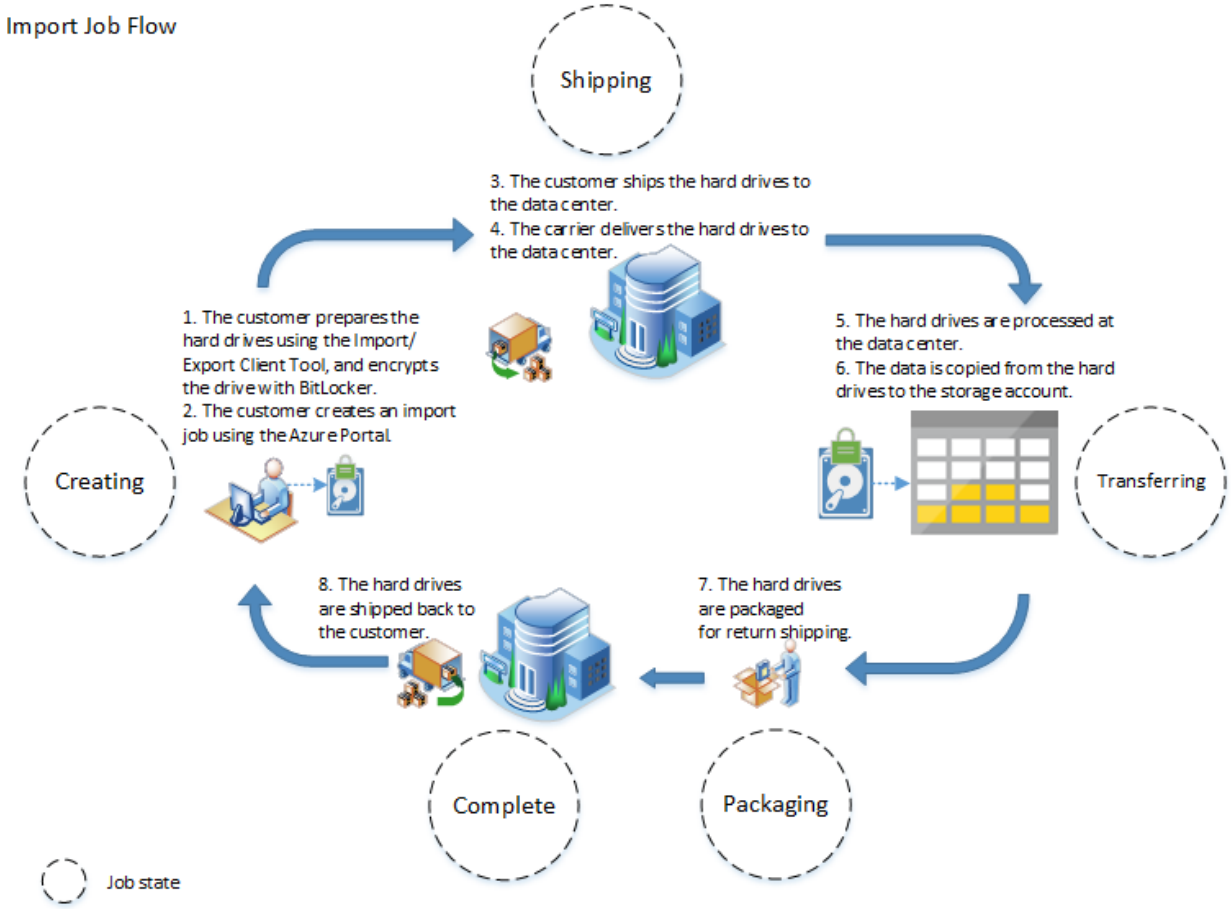
Overview

- Azure Import/Export Service
  - Import or Export data directly
    - 3.5-inch SATA II/III drives
      - Multiple drive import/export supported



Courtesy Microsoft documentation:[Using Azure Import/Export to transfer data to and from Azure Storage]

Overview



Courtesy Microsoft documentation: http://bit.ly/2BYtLgc

Overview
Azure Blob Storage

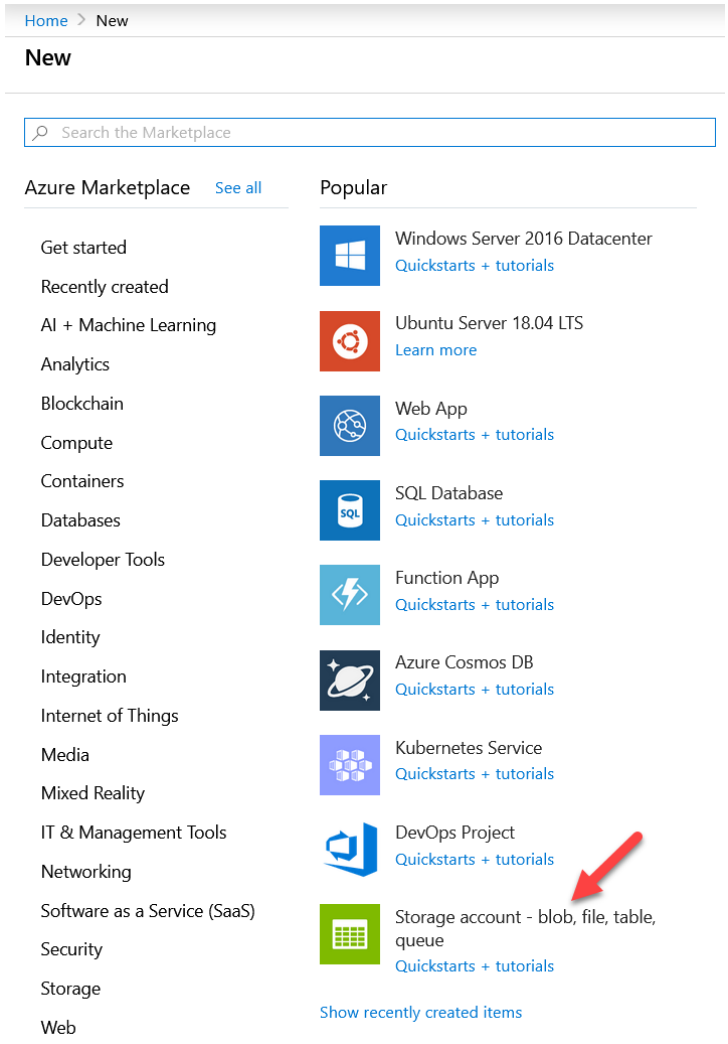
- Rich tooling support
  - PowerShell
    - Automate tasks
      - Create Containers, Upload, Download, Etc…
  - AzCopy
    - Bulk load & download
  - Cross-Platform CLI
    - Command line interface
  - Node.js
- Accessible via REST API

Overview
Azure Blob Storage

- Rich client libraries
  - .NET
  - Java
  - C++
  - Go
  - Node.js
- Accessible via REST for everyone else
  - PHP
  - Python

Overview
Azure Blob Storage

Azure Storage Provisioning



Azure Storage Provisioning

Basics | Advanced | Networking | Data protection | Encryption | Tags | Review + create

Azure Storage is a Microsoft-managed service providing cloud storage that is highly available, secure, durable, scalable, and redundant. Azure Storage includes Azure Blobs (objects), Azure Data Lake Storage Gen2, Azure Files, Azure Queues, and Azure Tables. The cost of your storage account depends on the usage and the options you choose below. Learn more about Azure storage accounts

Project details

Select the subscription in which to create the new storage account. Choose a new or existing resource group to organize and manage your storage account together with other resources.

Subscription * : Staff Labs for jficara.harvard

Resource group * : rg-classdemo

Create new

Instance details

If you need to create a legacy storage account type, please click here.

Storage account name * : stcscie94demo

Region * : (US) East US

Performance * :

- **Standard:** Recommended for most scenarios (general-purpose v2 account)
- **Premium:** Recommended for scenarios that require low latency.

Redundancy * : Locally-redundant storage (LRS)

Review + create | < Previous | Next : Advanced >

Locally-redundant storage (LRS):
Lowest-cost option with basic protection against server rack and drive failures. Recommended for non-critical scenarios.

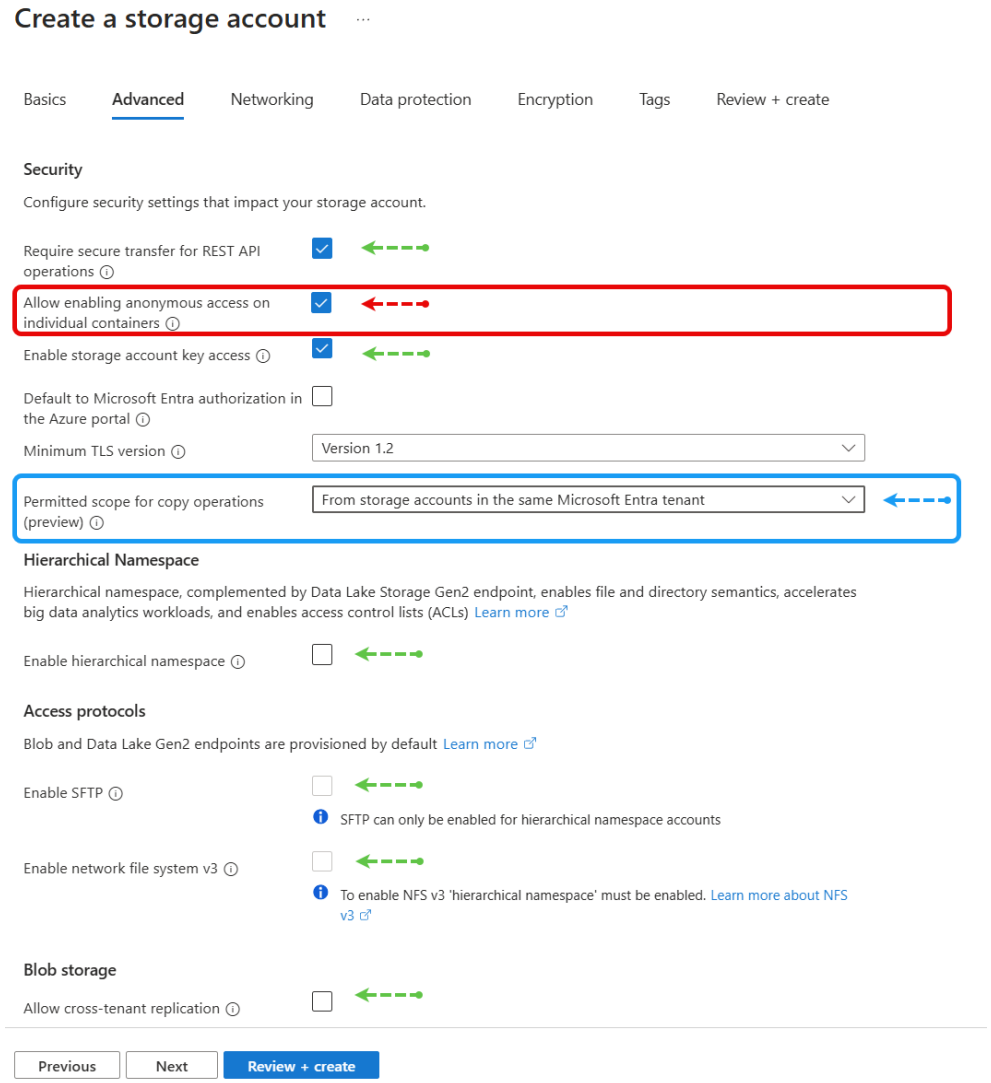
Geo-redundant storage (GRS):
Intermediate option with failover capabilities in a secondary region. Recommended for backup scenarios.

Zone-redundant storage (ZRS):
Intermediate option with protection against datacenter-level failures. Recommended for high availability scenarios.

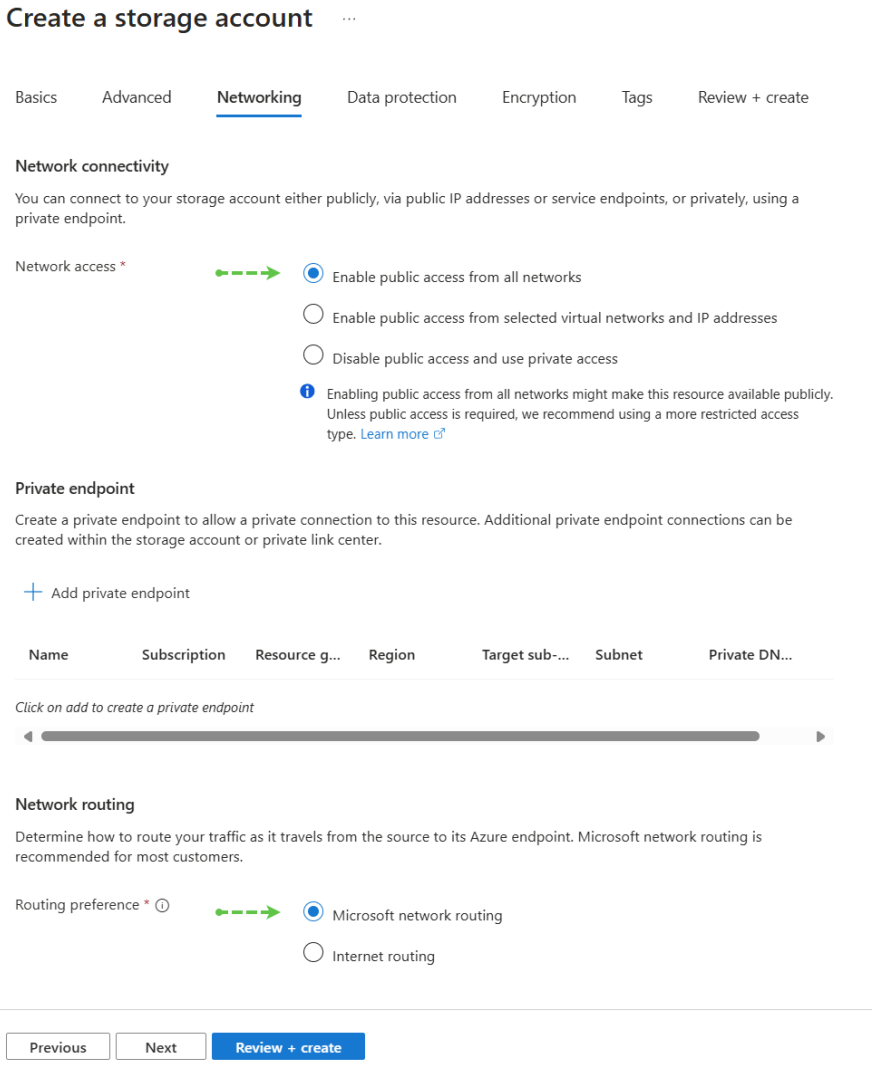
Geo-zone-redundant storage (GZRS):
Optimal data protection solution that includes the offerings of both GRS and ZRS. Recommended for critical data scenarios.

Locally-redundant storage (LRS)

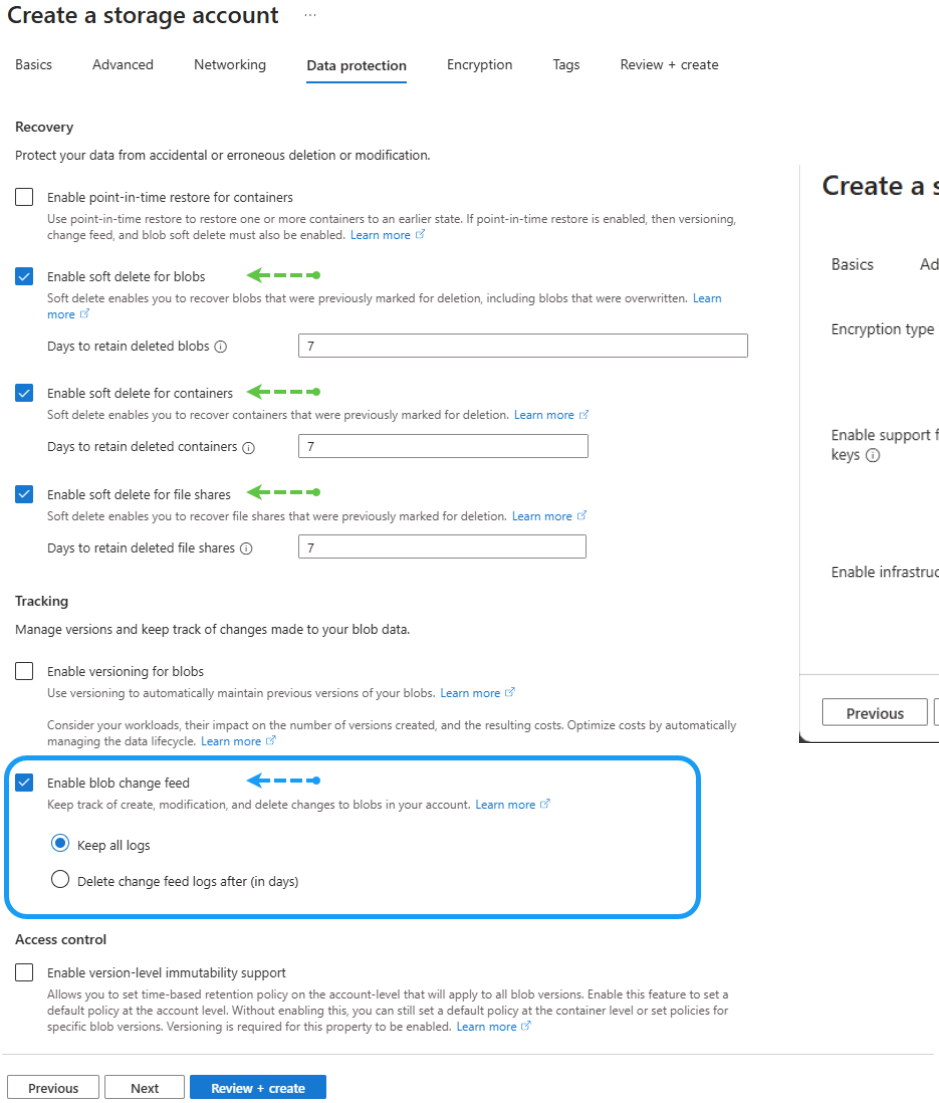
Azure Storage Provisioning

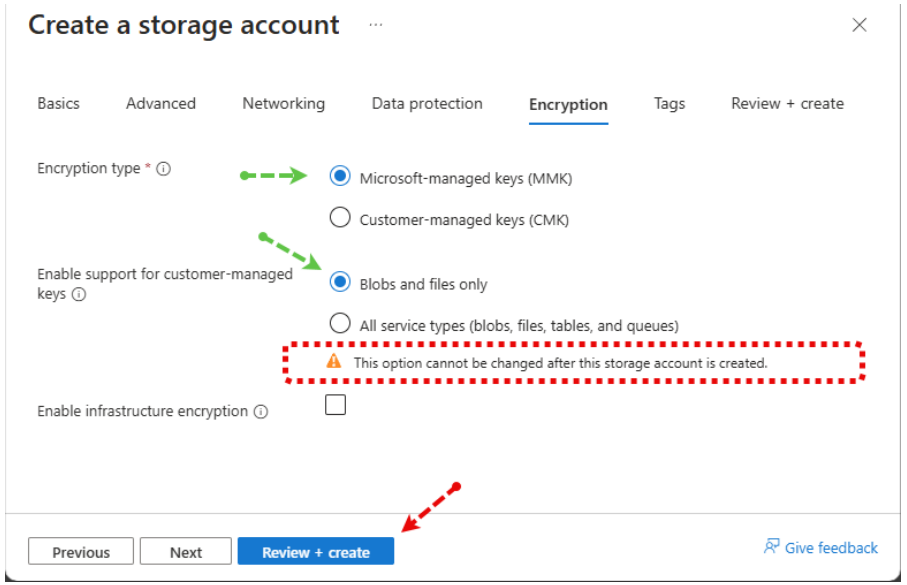


Azure Storage Provisioning

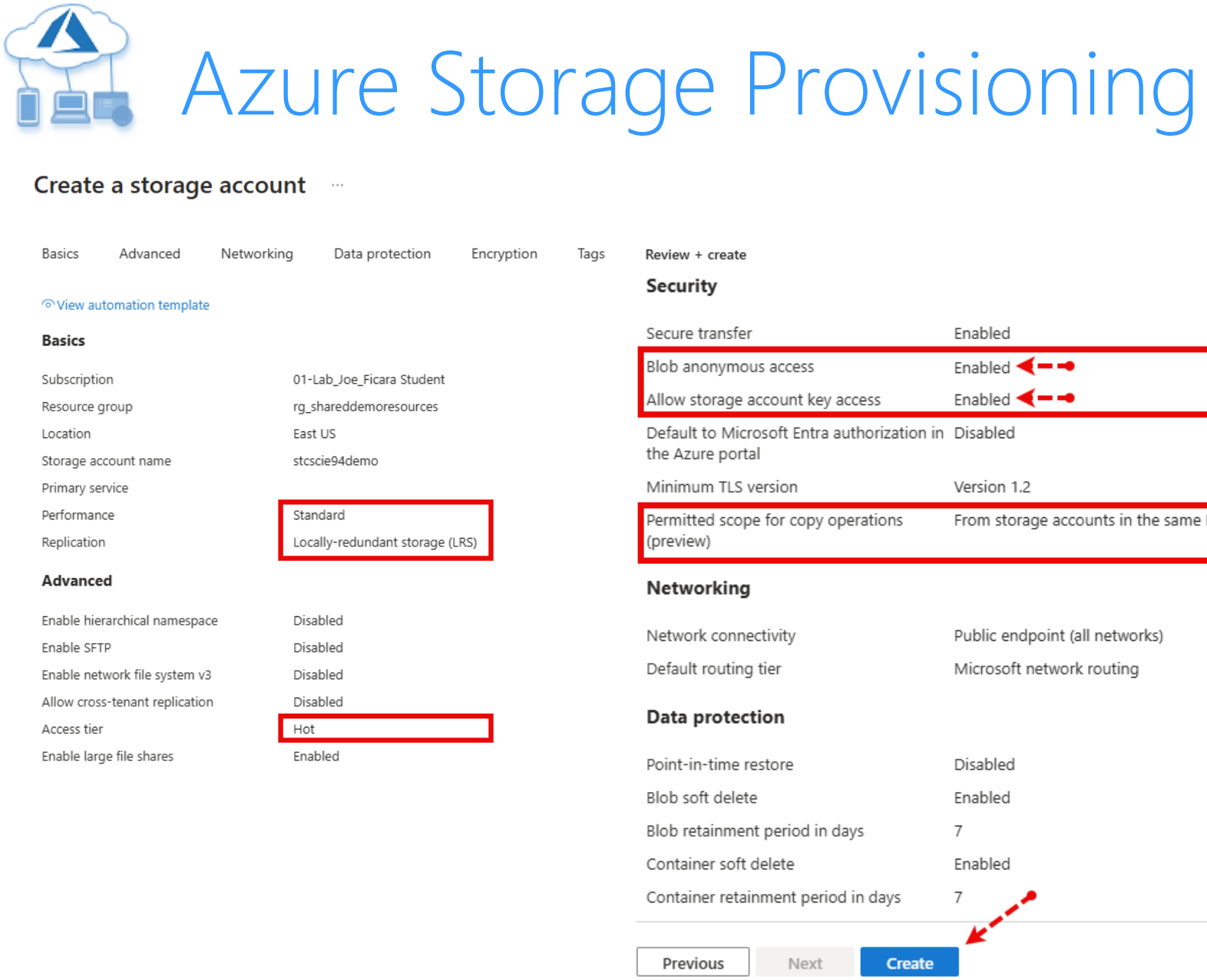


Azure Storage Provisioning





Azure Storage Provisioning



Azure Storage Provisioning

- There are several redundancy options
  - See: Link & for a summary see: Link
- When developing locally:
  - LRS: Use local redundancy -> Lowest cost
    - 3 Copies of data
    - Same facility
    - Does not protect against facility failure
    - Protects against limited hardware failures

Azure Storage Provisioning

- Production-> Meet business requirements
  - ZRS: Zone Redundancy
    - Same region or two regions
      - Only available for block blobs
      - Isolated infrastructure
        - 3 Copies of data
  - GRS: Geo Redundancy
    - Protects again catastrophic failure
      - Within a geographic location
    - Two locations
      - 6 copies of data

Azure Storage Provisioning

- Production-> Meet business requirements
  - RA-GRS: Read Access Geo Redundancy
    - Protects again catastrophic failure
      - Within a geographic location
    - Two locations
      - 6 copies of data
      - Read access from both locations

Azure Storage Provisioning

- Production-> Meet business requirements
  - Geo Zone Redundant (GZRS)
    V2 Storage Accounts Only
    - RA-GZRS, RA-GRZ, & GZRS
      - 99.99999999999999% (16 9's)
        - Durability of objects over a given year
    - Included
      - Block blobs
      - Page Blobs (Except VHD disks)
      - Files
      - Tables
      - Queues

Azure Storage Provisioning

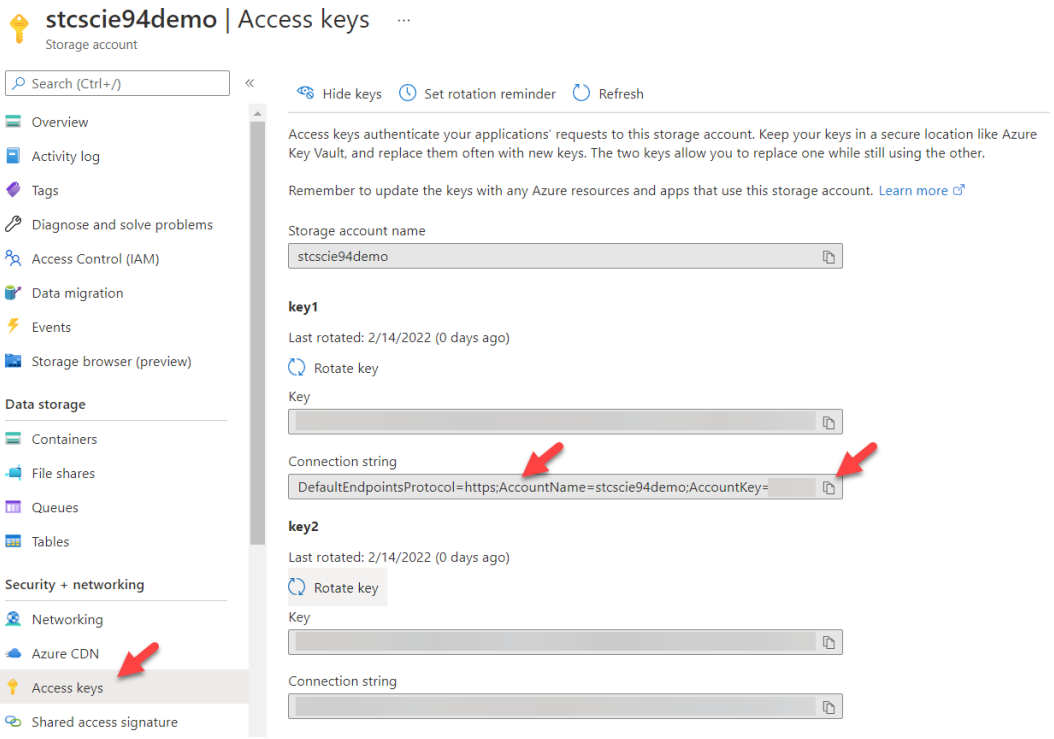
- How do you access the storage account?
  - Storage Explorers
    - Recommended:
      - Microsoft's Storage Explorer:
      - Azure Storage Explorer
    - A list of others:
      - Microsoft client tools for working with Azure Storage
  - Note:
    - Azure Portal includes a full featured GUI
      - Block, Page Blob, Append Blob, Table, Queue, Files

Azure Storage Provisioning

- How do you access the storage account?
  - Azure.Storage.Blobs Nuget Package
    - The SDK
      - Using the storage connection string
        - DefaultEndpointsProtocol=https;
          AccountName=[accountname];
          AccountKey=[accesskey]

Azure Storage Provisioning

- Where do you get the connection string?
  - In the new dashboard -> Keys



Azure Storage Provisioning

- Do store the connection string
  - Dashboard config
- Don't put the connection string in code
  - Or in your client/mobile app
  - Or in your JSON config file
- The SDK is available via Nuget
  - Azure.Storage.Blobs
- Can also access via REST API
  - Use the SDK when you can though

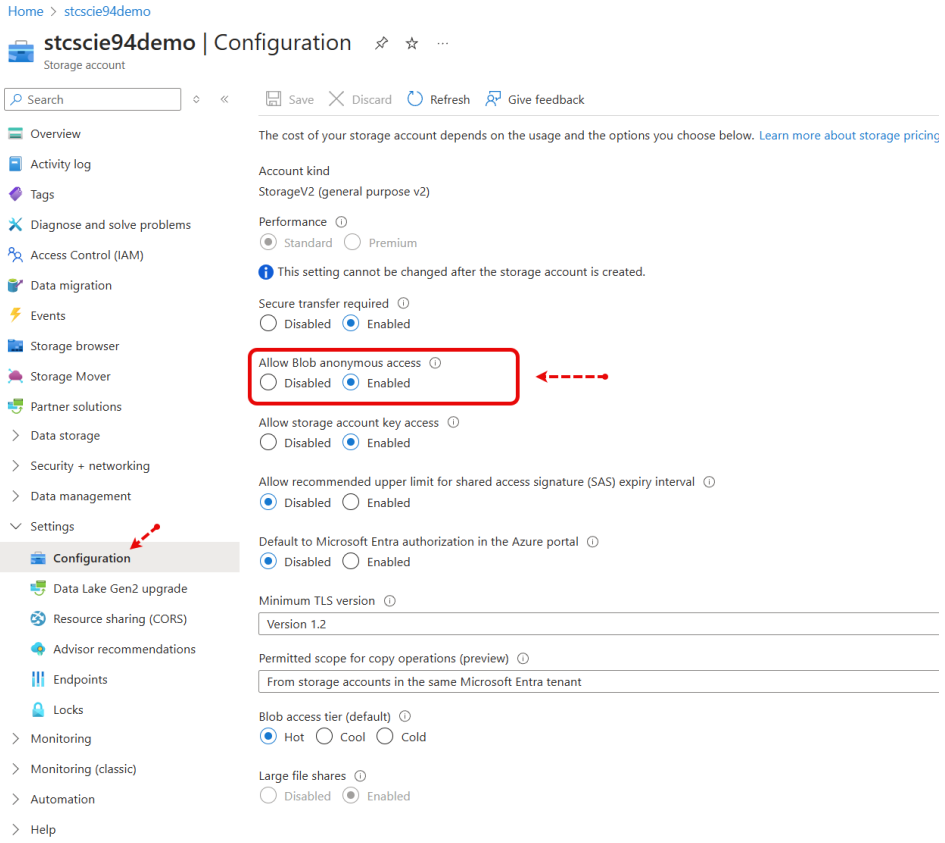
Demo

Azure Storage Provisioning

Azure Dashboard Demo

Azure Storage Provisioning

- If you forget to allow anonymous access



Azure Block Blob Storage
Implementation

- Adding a Block Blob requires 4 steps
  - 1. Create a BlobContainerClient
    - Using the cloud storage connection string
  - 2. Create a BlobClient
    - Using passing in the blob Name
  - 2a. Optionally set the access policy
    - On the container
    - Using the BlobContainerClient
      - SetAccessPolicyAsync

Azure Block Blob Storage
Implementation

- Adding a Block Blob requires 4 steps …
  - 3. Open a read stream
    - OpenReadStream
  - 4. Upload data to the blob
    - Using the BlobClient
  - 4a. Optionally set metadata on the blob
    - Using SetMetadataAsync
  - Done!

Azure Block Blob Storage
Implementation

- Notes:
  - Blobs support
    - Leases
    - Metadata
    - Etags
    - Updating and more …
  - Nested containers, not supported by default
    [*unless hierarchy is enabled on the storage account*]
    - Do support pseudo directories
      - Using a forward slash in the blob name
        - directory/blobname

Demo

Azure Block Blob Storage

Implementation

`AzureBlobSolution`

Azure Block Blob Storage

Azure Block Blob Storage
Implementation 2

- Two changes:
  - Let's use a business layer
    - Via dependency injection
    - Makes the code cleaner
    - Easier to unit test
  - Managed Identities to access storage
    - Removes secrets from callers
      - Developers in Visual Studio
      - Azure App Services
      - Etc…

Azure Storage
Managed Identities

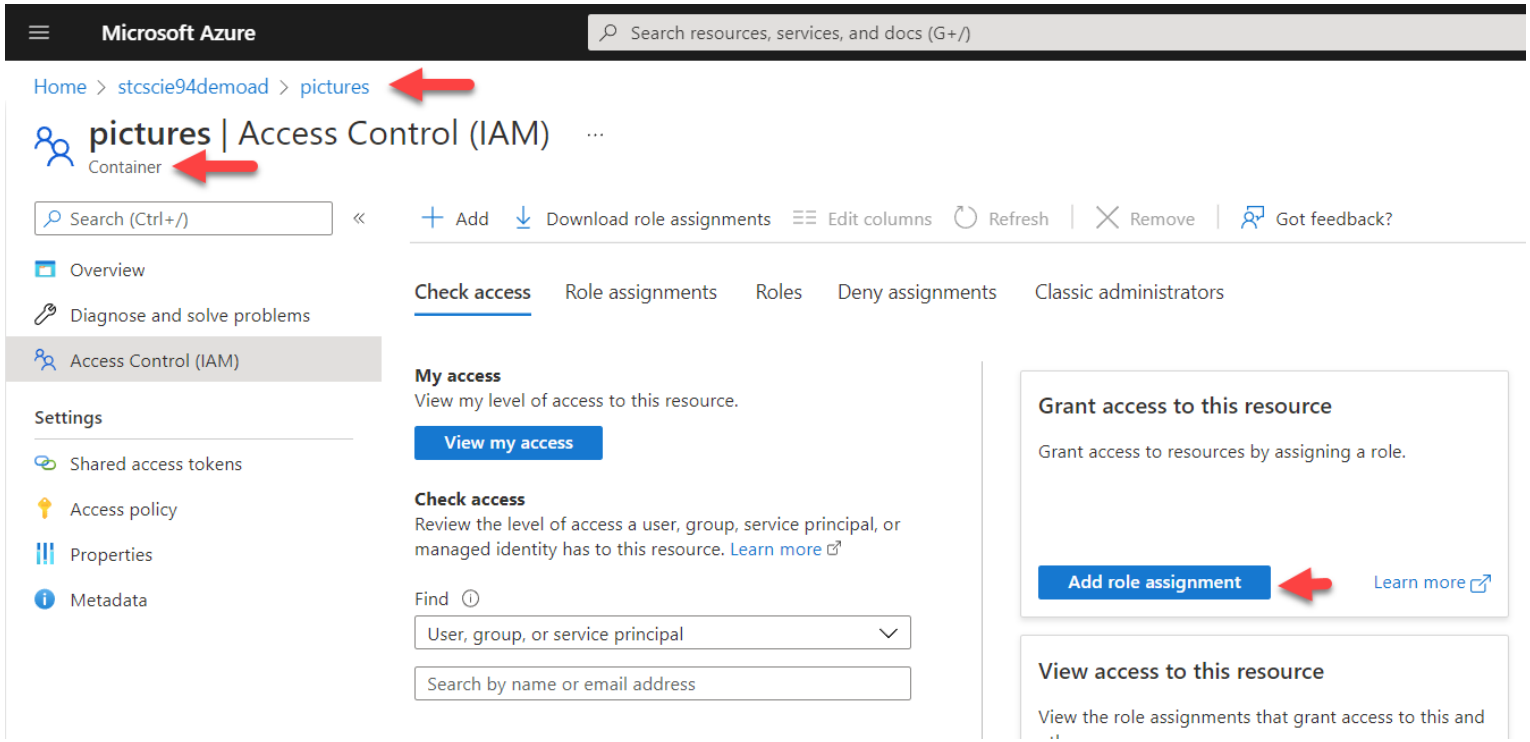
- Azure Storage Supports Managed Identity
- Requires caller to have
  - A Microsoft Entra ID Account
  - Role(s) assigned to grant access
    - Some of the common roles
      - Storage Blob Data Contributor
      - Storage Blob Data Owner
      - Storage Blob Data Reader
      - See: Built-in-roles

Azure Storage
Managed Identities

- Requires assignment
  - Account level
    - Storage Blob Data Owner
  - At the container
    - Storage Blob Data Contributor, Reader

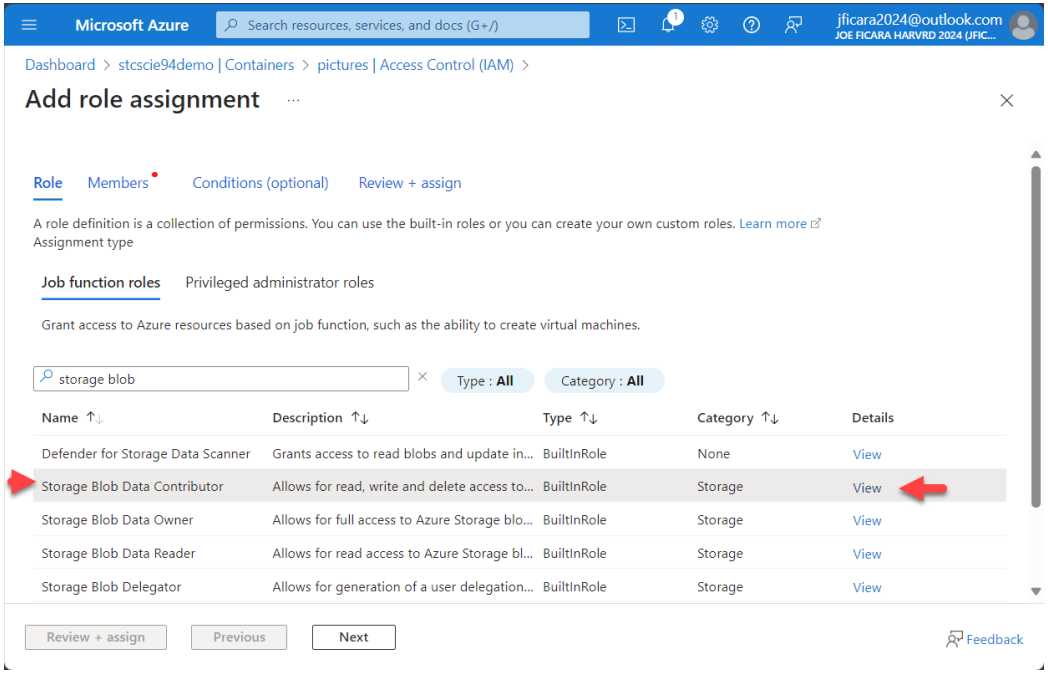
Azure Storage
Managed Identities

- Add a role assignment at the container
  - Here the pictures container



Azure Storage
Managed Identities

- Add a role assignment at the container
  - Add the "Storage Blob Data Contributor" role



Azure Storage
Managed Identities

- Storage Blob Data Contributor
  - Permissions – Actions (control plane)

Storage Blob Data Contributor
BuiltInRole

Permissions | JSON | Assignments

Description: Allows for read, write and delete access to Azure Storage blob containers and data

Search permissions | Type : **All**

(•) Actions ( ) DataActions

Showing 6 of 6 permissions

Type	Permissions	Description
Microsoft.Storage		
Other	Generate a user delegation key	Returns a user delegation key for the blob service
Write	Patch blob container	Returns the result of patch blob container
Delete	Delete blob container	Returns the result of deleting a container
Read	Get blob container	Returns a container
Read	List of blob containers	Returns list of containers
Write	Put blob container	Returns the result of put blob container

Azure Storage
Managed Identities

- Storage Blob Data Contributor
  - Permissions – Data Actions (data plane)

Storage Blob Data Contributor
BuiltInRole

Permissions | JSON | Assignments

Description: Allows for read, write and delete access to Azure Storage blob containers and data

Search permissions | Type : **All**

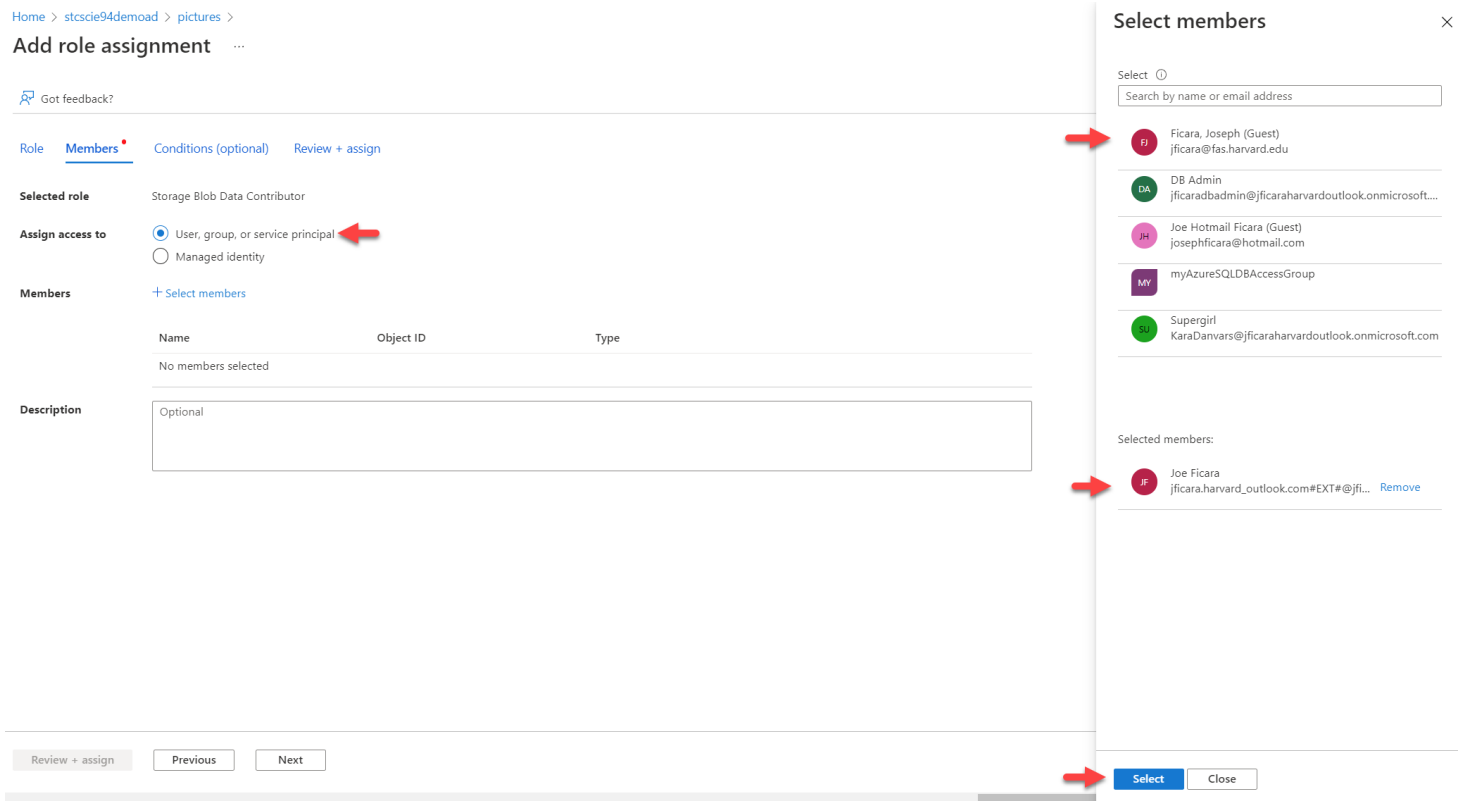
Actions | DataActions

Showing 5 of 5 permissions

Type	Permissions	Description
Microsoft.Storage		
Read	Read Blob	Returns a blob or a list of blobs
Write	Write Blob	Returns the result of writing a blob
Delete	Delete blob	Returns the result of deleting a blob
Other	Add blob content	Returns the result of adding blob content
Other	Move blobs	Moves the blob from one path to another

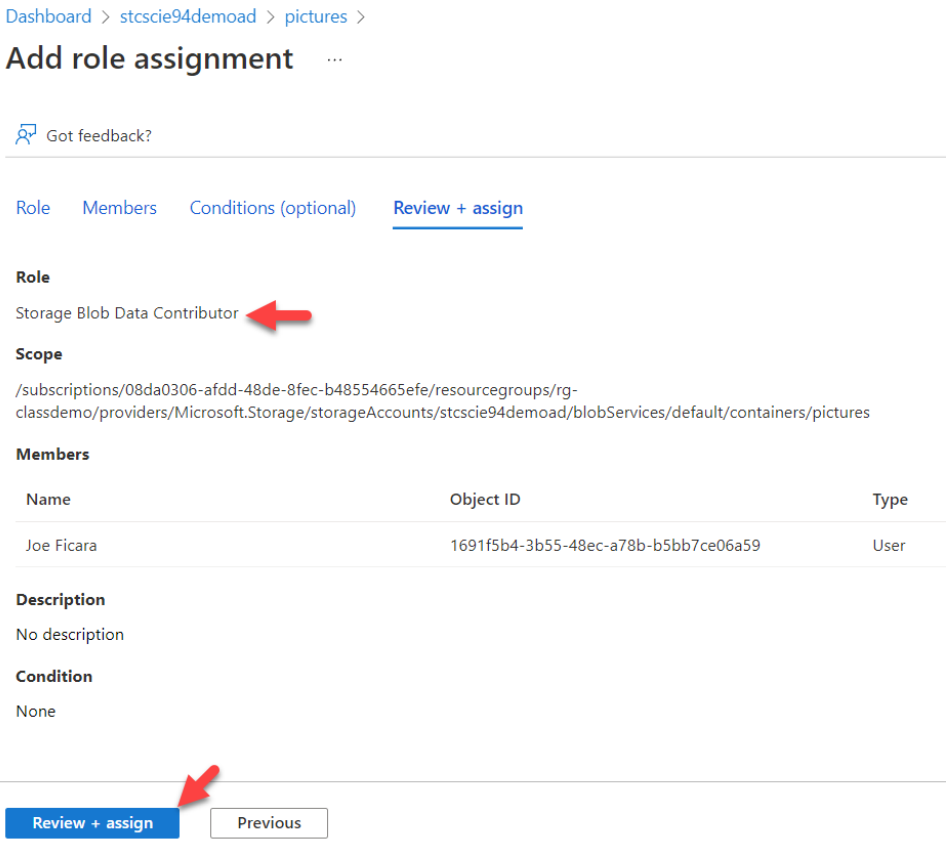
Azure Storage
Managed Identities

- Pick the user group or service principal



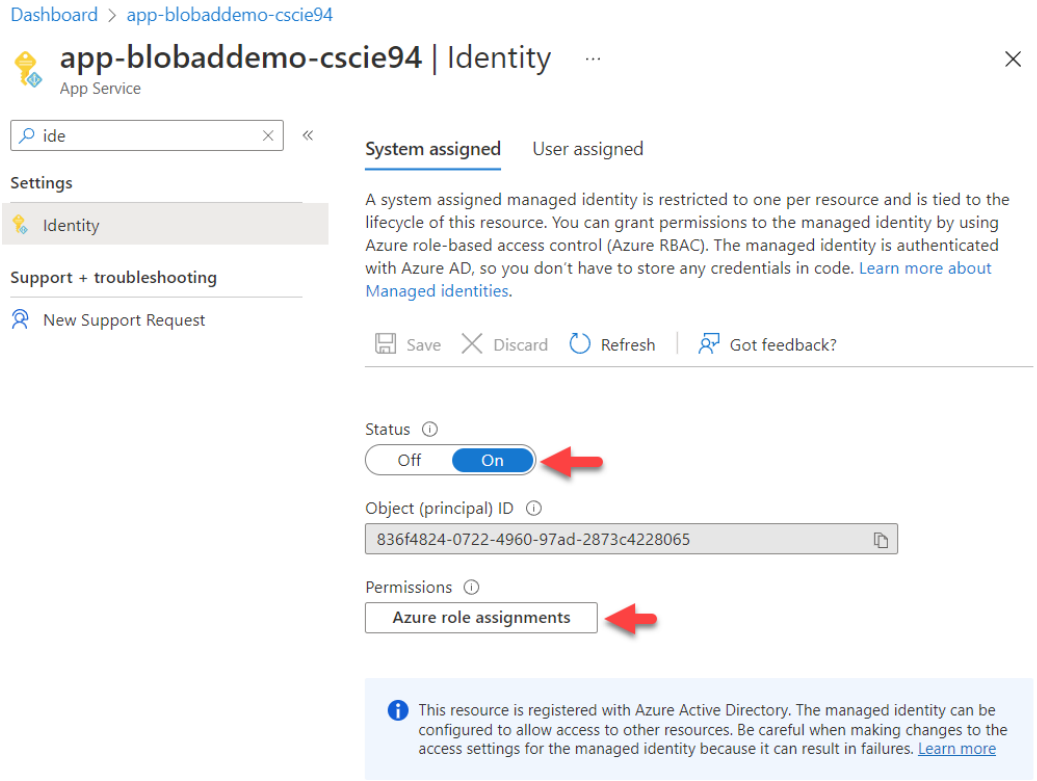
Azure Storage
Managed Identities

- Pick the user group or service principal



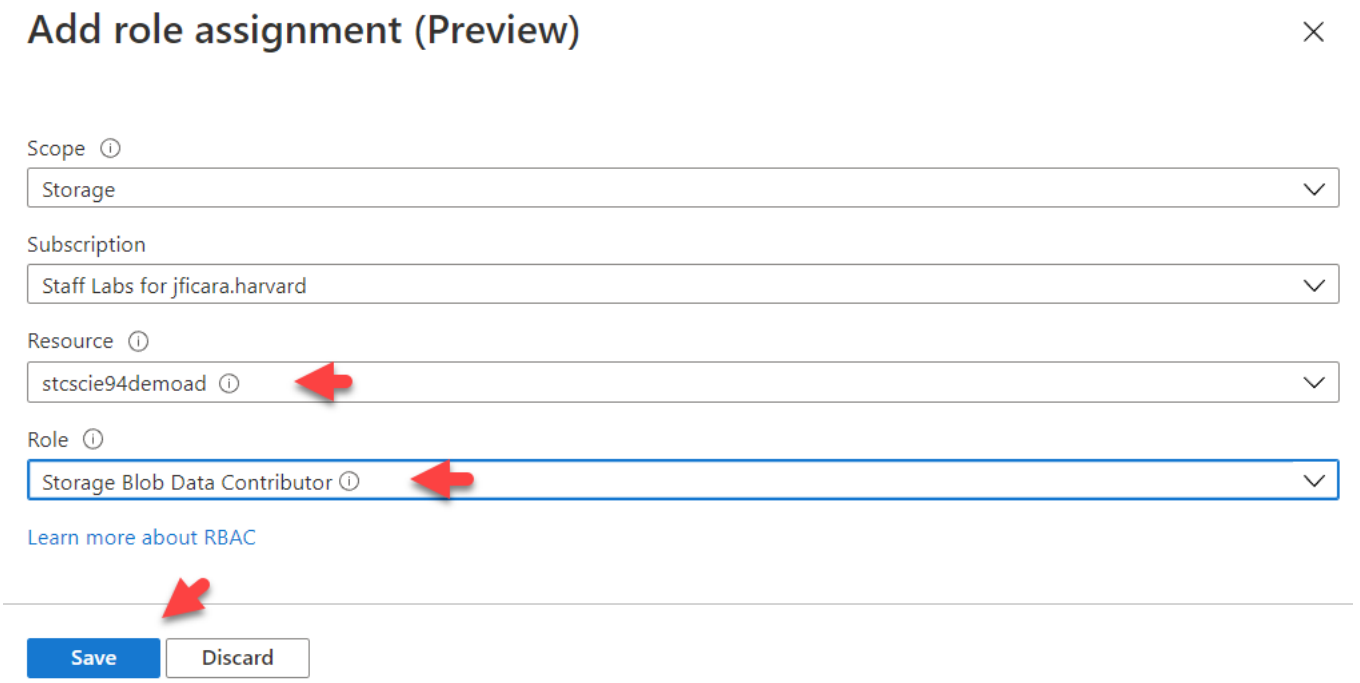
Azure Storage
Managed Identities

- Can also assign roles at the App Service



Azure Storage
Managed Identities

- Can also assign roles at the App Service



Azure Storage
Managed Identities

- Some code changes are needed
  - Use DefaultAzureCredential
    - To gain access to the host assigned AD credentials
  - Use DefaultAzureCredentialOptions
    - To specify the TenantId
      - Obtained from the Azure AD Overview Page
  - Create a BlobContainerClient
    - Singleton using the DefaultAzureCredential

Demo

Azure Block Blob Storage

Managed Identity

- AzureBlobManagedIdentitySolution.sln
- Managed Identity Configuration: Lines 49 – 60 Program.cs
- Implementation of Azure Storage Access: StorageRepository.cs

Azure Storage

Managed Identities

Links & Resources

- Introduction to Azure Storage
  - Introduction to Azure Storage - Cloud storage on Azure
- Azure Import/Export service
  - Using Azure Import/Export to transfer data to and from Azure Storage
- Azure Storage Scalability & Performance
  - Scalability and performance targets for standard storage accounts - Azure Storage
- Azure subscription & service limits, quotas, & constraints
  - Azure subscription and service limits, quotas, and constraints - Azure Resource Manager

Links & Resources

- How to use:
  - Azure Blob Storage
    - Azure Blob Storage | Microsoft Azure
  - Azure Table Storage
    - Table storage | Microsoft Azure
  - Azure Queues
    - Queue Storage | Microsoft Azure
  - Azure File Storage
    - Azure Files | Microsoft Azure

Links & Resources

- How to use:
  - Storage Service REST APIs
    - Azure Storage REST API Reference | Microsoft Docs
  - Azure Storage PowerShell Cmdlets
    - Install Azure PowerShell with PowerShellGet | Microsoft Docs

Links & Resources

- SDKs
  - .NET
  - Java
  - Go
  - Python
  - Node.JS
  - REST APIs
    - Azure documentation | Microsoft Docs
    - Azure SDK Latest Releases | Azure SDKs

Links & Resources

- Azure Storage & Managed Identities
  - Control Plan vs Data Plane
  - Authorize access to blob or queue data from a native or web application
  - Authorize access to blob data with managed identities for Azure Resources
  - Blob service error codes