

Assignment-19

Name-Ajay Chaudhary
Batch-Data Engineering

Handwritten notes-

Database characteristics

- Security features to ensure the data only be accessed by authorized users.
- ACID transactions to ensure data integrity.
- Query languages & APIs to easily interact with the data in the database.
- Indexes to optimize query performance.
- Full-text search
- Optimizations for mobile devices
- Flexible deployment topologies to isolate workloads to a specific set of resources.

Why use a database?

- Patient medical records
- Items in an online store
- Financial records
- Articles & blog entries
- Sports scores & statistics
- Online gaming information
- Student grades & scores
- IoT device readings
- Mobile application information

Database Examples

Relational database : Oracle , MySQL

OLAP + data warehouses & data lakes

Both data warehouses & data lakes are meant to support Online Analytical Processing (OLAP).

OLAP systems are typically used to collect data from a variety of sources.

What is a data warehouse?

A data warehouse is a system that stores highly structured information from various sources.

Data warehouse typically store current & historical data from one or more systems.

goal of datawarehouse -

analyze the data

look for insights

create business intelligence (BI) in the form of reports & dashboards.

Data Warehouse Characteristics

- They store large amounts of current & historical data from various sources.
- Extract, transform, load (ETL) processes move data from its original source to the data warehouse.

ETL vs ~~ELT~~ ELT

→ Data warehouses have pre-defined & fixed relational schema.

AWS SE
~~MongoDB~~
 AWS Athena

	Database	Data Lake	Data Warehouse
Workloads	Operational & transactional	Analytical	Analytical
Data type	Structured or semi-structured	Structured, semi-structured & unstructured	Structured or semi-structured
Schema flexibility	Rigid or flexible schema depending on database type	No schema definition required for ingest (schema on read)	Pre-defined & fixed schema definition for ingest (schema on read)
Data freshness	Real time	May not be up-to-date based on frequency of ETL processes	May not be up-to-date based on frequency of ETL
Users	Application developers	Business analysts, app dev, & data scientists	Business analysts & data scientists
Pros	Fast querying for storing & updating data	Easy data storage	fixed schema
Cons	Limited analytic capabilities	Requires efforts to organize & prepare data for use.	Difficult to design & evolve schema

How to create a storage account

The screenshot shows the 'Create a storage account' wizard in the Microsoft Azure portal. The 'Basics' tab is selected. In the 'Project details' section, the subscription is set to 'Azure subscription 1' and the resource group is 'rg-azuser1071_mml.local-84y1N'. In the 'Instance details' section, the storage account name is 'hexadeb1071'. At the bottom, there are 'Review' and 'Next : Advanced >' buttons.

The screenshot shows the 'Create a storage account' wizard in the Microsoft Azure portal. The 'Advanced' tab is selected. In the 'Instance details' section, the storage account name is 'hexadeb1071', the region is '(Asia Pacific) Central India', and the performance tier is 'Standard'. Under 'Redundancy', 'Geo-redundant storage (GRS)' is selected, and the 'Make read access to data available in the event of regional unavailability' checkbox is checked. At the bottom, there are 'Review' and 'Next : Advanced >' buttons.

Chrome File Edit View History Bookmarks Profiles Tab Window Help 0KB/s 28% Thu 15 Feb 3:43 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | Create a storage account - M... +

portal.azure.com/#create/Microsoft.StorageAccount-ARM

Microsoft Azure Search resources, services, and docs (G+) azuser1071_mml.local@... IIHT (IITHONMICROSOFT.COM)

Home > Storage accounts > Create a storage account ...

Basics Advanced Networking Data protection Encryption Tags Review

Access tier Hot
Enable SFTP Disabled
Large file shares Disabled

Networking

Network connectivity Public endpoint (all networks)
Default routing tier Microsoft network routing
Endpoint type Standard

Security

Secure transfer Enabled
Allow storage account key access Enabled
Default to Microsoft Entra authorization in the Azure portal Disabled
Blob anonymous access Disabled
Minimum TLS version Version 1.2
Banned from running operations From any storage account

Create < Previous Next > Download a template for automation Give feedback

Apple Chrome File Edit View History Bookmarks Profiles Tab Window Help 0KB/s 28% Thu 15 Feb 3:43 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | Create a storage account - M... +

portal.azure.com/#create/Microsoft.StorageAccount-ARM

Microsoft Azure Search resources, services, and docs (G+) azuser1071_mml.local@... IIHT (IITHONMICROSOFT.COM)

Home > Storage accounts > Create a storage account ...

Basics Advanced Networking Data protection Encryption Tags Review

Blob retention period in days 7
Container soft delete Enabled
Container retention period in days 7
File share soft delete Enabled
File share retention period in days 7
Versioning Disabled
Blob change feed Disabled
Version-level immutability support Disabled

Encryption

Encryption type Microsoft-managed keys (MMK)
Enable support for customer-managed keys Blobs and files only
Enable infrastructure encryption Disabled

Create < Previous Next > Download a template for automation

Apple Chrome File Edit View History Bookmarks Profiles Tab Window Help 0KB/s 28% Thu 15 Feb 3:43 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | Create a storage account - M... +

portal.azure.com/#create/Microsoft.StorageAccount-ARM

Microsoft Azure Search resources, services, and docs (G+) azuser1071_mml.local@... IIHT (IITHONMICROSOFT.COM)

Chrome File Edit View History Bookmarks Profiles Tab Window Help ↓ 31KB/s 28% Thu 16 Feb 3:43 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | hexadeb1071_1707992018183 | +

portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/-/overview/id/%2Fsubscriptions%2F984f097c-963c-4eb6-a20d-839457ae9f0... azuser1071_mml.local@... IIHT(HTMLONMICROSOFT.COM)

Microsoft Azure Search resources, services, and docs (G+)

Home > hexadeb1071_1707992018183 | Overview

Deployment

Search Delete Cancel Redeploy Download Refresh

Overview Inputs Outputs Template

Deployment is in progress

Deployment name: hexadeb1071_1707992018183
Subscription: Azure subscription 1
Resource group: rg-azuser1071_mml.local-84y1N

Start time: 15/02/2024, 15:43:41
Correlation ID: a02b1804-9231-47eb-a110-edc0b50a2077

Deployment details

Resource	Type	Status	Operation details
No results.			

Give feedback Tell us about your experience with deployment

Microsoft Defender for Cloud
Secure your apps and infrastructure
[Go to Microsoft Defender for Cloud >](#)

Free Microsoft tutorials
[Start learning today >](#)

Work with an expert
Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support.
[Find an Azure expert >](#)

Chrome File Edit View History Bookmarks Profiles Tab Window Help 0KB/s 28% Thu 15 Feb 3:45 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | hexadeb1071_1707992018183 | +

portal.azure.com/#view/HubsExtension/DeploymentDetailsBlade/-/overview/id/%2Fsubscriptions%2F984f097c-963c-4eb6-a20d-839457ae9f0... azuser1071_mml.local@... IIHT (IITHONMICROSOFT.COM)

Microsoft Azure Search resources, services, and docs (G+) Home > hexadeb1071_1707992018183 | Overview

Deployment

Search Delete Cancel Redeploy Download Refresh

Overview Inputs Outputs Template

Your deployment is complete

Deployment name: hexadeb1071_1707992018183
Subscription: Azure subscription 1
Resource group: rg-azuser1071_mml.local-84y1N

Start time: 15/02/2024, 15:43:41 Correlation ID: a02b1804-9231-47eb-a110-edc0b50a2077

Deployment details Next steps Go to resource

Give feedback Tell us about your experience with deployment

Cost Management Get notified to stay within your budget and prevent unexpected charges on your bill. Set up cost alerts >

Microsoft Defender for Cloud Secure your apps and infrastructure Go to Microsoft Defender for Cloud >

Free Microsoft tutorials Start learning today >

Work with an expert Azure experts are service provider partners who can help manage your assets on Azure and be your first line of support. Find an Azure expert >

Chrome File Edit View History Bookmarks Profiles Tab Window Help 175KB/s 36% Thu 15 Feb 3:52 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | hexadeb1071 - Microsoft Azure | +

portal.azure.com/#@iith.onmicrosoft.com/resource/subscriptions/984f097c-963c-4eb6-a20d-839457ae9f08/resourcegroups/rg-azuser1071... azuser1071_mml.local@... IIHT (IITHONMICROSOFT.COM)

Microsoft Azure Search resources, services, and docs (G+) Home > hexadeb1071

Storage account

hexadeb1071 | Containers

Container Change access level Restore containers Refresh Delete Give feedback

Overview Activity log Tags Diagnose and solve problems Access Control (IAM) Data migration Events Storage browser Storage Mover

Containers

Name	Last modified	Anonymous access level	Lease state
\$logs	15/02/2024, 15:44:12	Private	Available
hexa1071	15/02/2024, 15:51:21	Private	Available

Search containers by prefix Show deleted containers

Data storage

Containers File shares Queues Tables

Security + networking

Networking Front Door and CDN

The screenshot shows the Microsoft Azure Storage Explorer interface. On the left, the navigation pane lists 'hexa1071 | Containers' under 'Containers'. The 'Overview' section is selected. It displays basic information: 'Authentication method: Access key (Switch to Microsoft Entra user account)', 'Location: hexa1071', and a search bar 'Search blobs by prefix (case-sensitive)'. Below this is a table with columns 'Name', 'Modified', 'Access tier', and 'Arc'. A message 'No results' is shown. On the right, a large modal window titled 'Upload blob' is open. It features a central area with a cloud icon and the text 'Drag and drop files here or Browse for files'. Below this are checkboxes for 'Overwrite if files already exist' and 'Advanced'. At the bottom are 'Upload' and 'Give feedback' buttons.

The screenshot shows the Microsoft Azure Storage Explorer interface. At the top, there's a toolbar with various icons like Home, History, Bookmarks, Profiles, Tab, Window, Help, and network status (136KB/s, 38% battery). Below the toolbar is a tab bar with several tabs: Data Engineering training, Azure Storage Explorer - cloud, MakeMyLabs, hexa1071 - Microsoft Azure, and a new tab indicator. The main title bar says "Microsoft Azure". The left sidebar shows the navigation path: Home > hexa1071 | Containers > hexa1071. It also lists sections: Overview, Diagnose and solve problems, Access Control (IAM), Settings (Shared access tokens, Access policy, Properties, Metadata), and a search bar. The "Overview" section is currently selected. The main content area displays blob details for "hexa1071". It shows the authentication method as "Access key" and the location as "hexa1071". A search bar allows filtering by prefix. A table lists the blobs: Name (test.csv), Modified (15/02/2024, 15:55:23), Access tier (Hot (Inferred)), Archive status (Not yet archived), Blob type (Block blob), Size (247 B), and Lease state (Available). There are also "Add filter" and "Show deleted blobs" buttons.

Name	Modified	Access tier	Archive status	Blob type	Size	Lease state
test.csv	15/02/2024, 15:55:23	Hot (Inferred)	Not yet archived	Block blob	247 B	Available

Chrome File Edit View History Bookmarks Profiles Tab Window Help 11KB/s 46% Thu 15 Feb 4:02 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | hexadeb1071 - Microsoft Azure

portal.azure.com/#@iiht.onmicrosoft.com/resource/subscriptions/984f097c-963c-4eb6-a20d-839457ae9f08/resourcegroups/rg-azuser1071_...

Microsoft Azure Search resources, services, and docs (G+) azuser1071_mml.local@... IIHT (IIHTONMICROSOFT.COM)

Home > hexadeb1071

hexadeb1071 | Shared access signature

Storage account

Search Give feedback

A shared access signature (SAS) is a URI that grants restricted access rights to Azure Storage resources. You can provide a shared access signature to clients who should not be trusted with your storage account key but whom you wish to delegate access to certain storage account resources. By distributing a shared access signature URI to these clients, you grant them access to a resource for a specified period of time.

An account-level SAS can delegate access to multiple storage services (i.e. blob, file, queue, table). Note that stored access policies are currently not supported for an account-level SAS.

Learn more about creating an account SAS

Allowed services: Blob, File, Queue, Table

Allowed resource types: Service, Container, Object

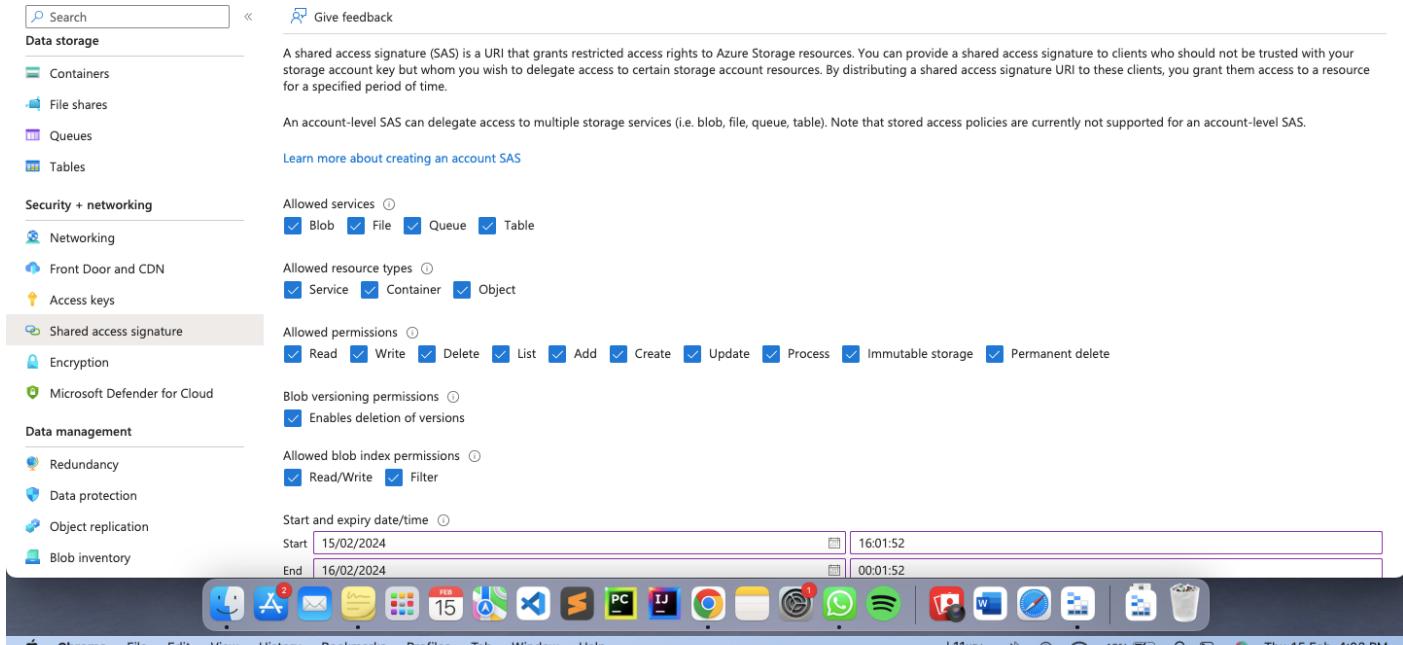
Allowed permissions: Read, Write, Delete, List, Add, Create, Update, Process, Immutable storage, Permanent delete

Blob versioning permissions: Enables deletion of versions

Allowed blob index permissions: Read/Write, Filter

Start and expiry date/time:

Start: 15/02/2024 End: 16/02/2024



Chrome File Edit View History Bookmarks Profiles Tab Window Help 11KB/s 46% Thu 15 Feb 4:02 PM

Data Engineering training | Azure Storage Explorer - clo... | MakeMyLabs | hexadeb1071 - Microsoft Azure

portal.azure.com/#@iiht.onmicrosoft.com/resource/subscriptions/984f097c-963c-4eb6-a20d-839457ae9f08/resourcegroups/rg-azuser1071_...

Microsoft Azure Search resources, services, and docs (G+) azuser1071_mml.local@... IIHT (IIHTONMICROSOFT.COM)

Home > hexadeb1071

hexadeb1071 | Shared access signature

Storage account

Search Give feedback

Enables deletion of versions

Allowed blob index permissions: Read/Write, Filter

Start and expiry date/time:

Start: 15/02/2024 End: 16/02/2024

(UTC+05:30) Chennai, Kolkata, Mumbai, New Delhi

Allowed IP addresses: For example, 168.1.5.65 or 168.1.5.65-168.1.5.70

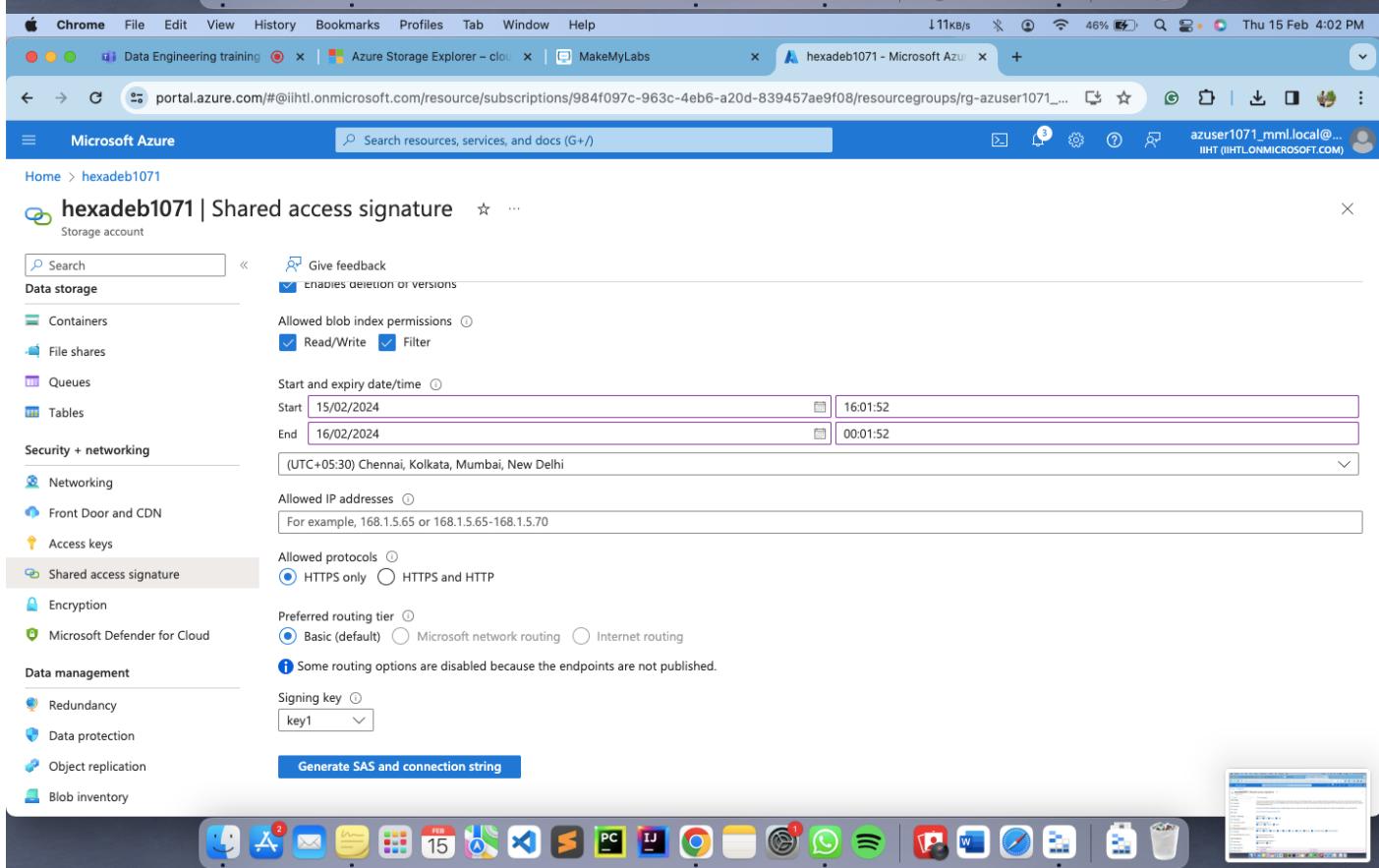
Allowed protocols: HTTPS only

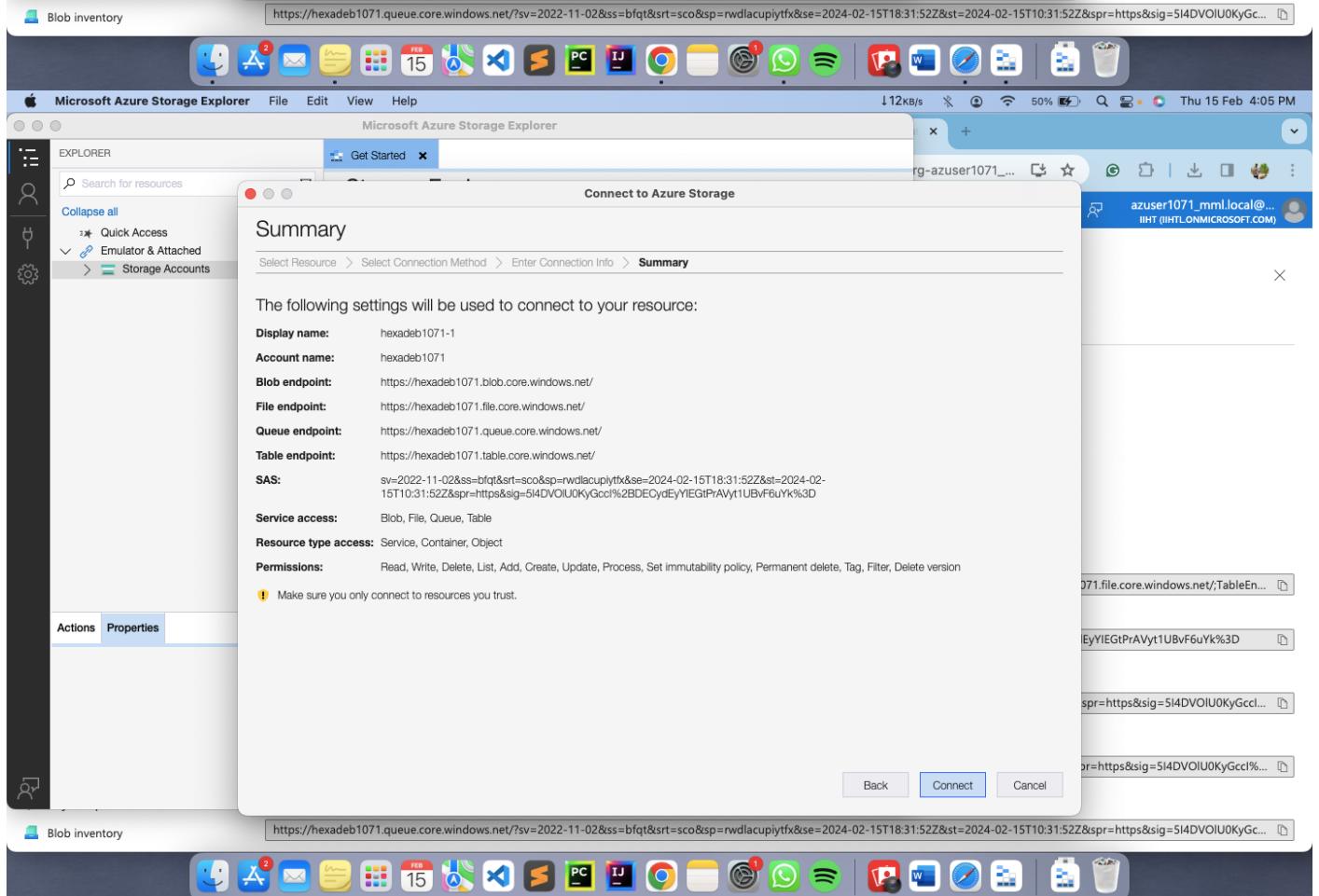
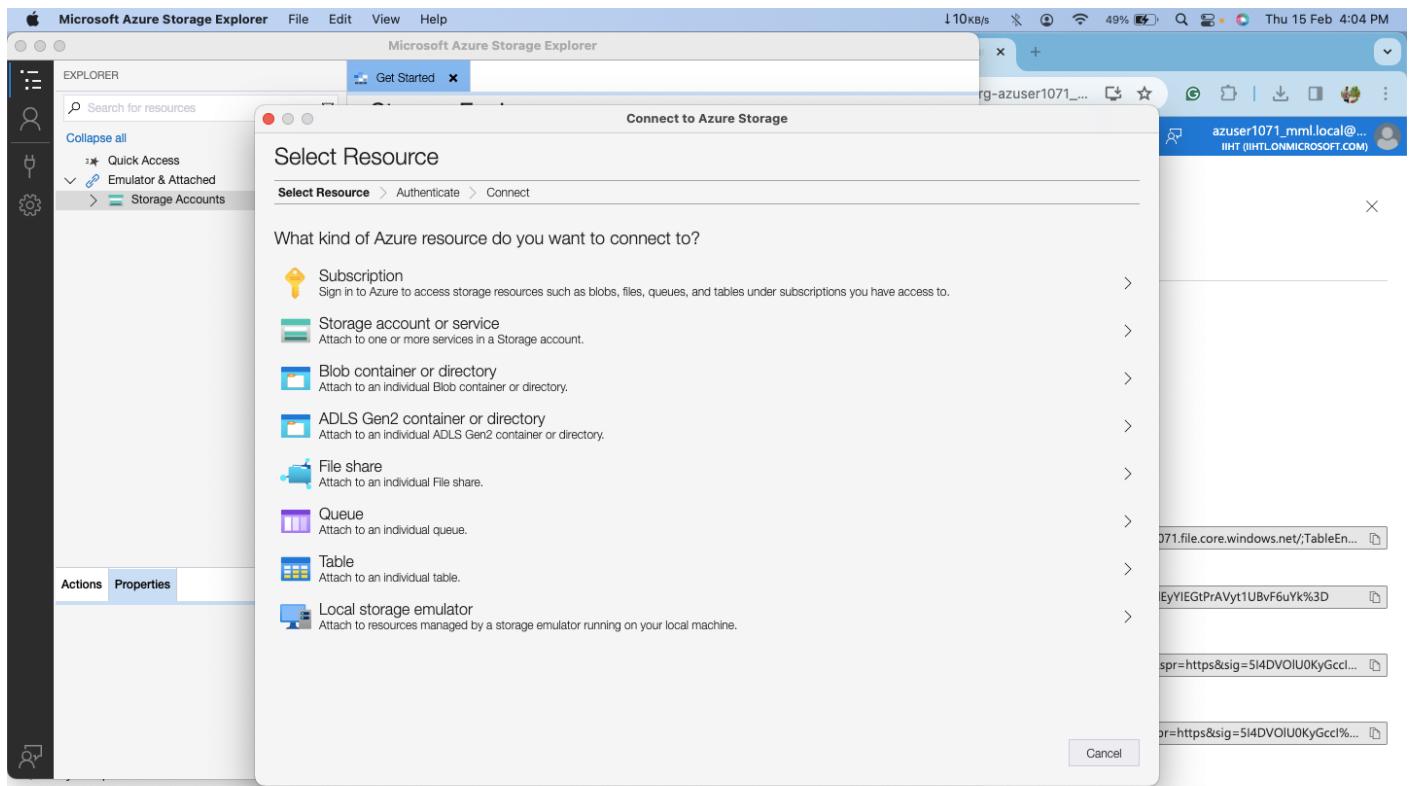
Preferred routing tier: Basic (default)

Some routing options are disabled because the endpoints are not published.

Signing key: key1

Generate SAS and connection string





Microsoft Azure Storage Explorer

File Edit View Help

0KB/s

Thu 15 Feb 4:05 PM

Microsoft Azure Storage Explorer

Get Started

Storage Explorer

Get Started

Sign in with Azure
Sign in with your Azure account to access all of your Azure resources.

Attach to a resource
Don't have permissions to access Azure subscriptions? You can connect to individual storage resources using various authentication methods.

Tasks

View Azure resources
See your subscriptions and Azure resources all in one place.

Manage accounts and subscriptions
Manage your Azure accounts and choose which subscriptions you want to see.

Connect to Azure resources
Sign into Azure or attach to individual storage resources.

Show this page on startup

Activities

Clear completed Clear successful

Successfully added new connection.

File Endpoint: https://hexadeb1071.file.core.windows.net/;TableEn...

sig=5i4DV0lU0KyGcl%2BDECyDyYIEGTPrAVyt1UBvF6uYk%3D

Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Queue Service SAS URL: https://hexadeb1071.queue.core.windows.net/?sv=2022-11-02&ss=bfqt&srt=sco&sp=rw&lac=up&tfx=&se=2024-02-15T18:31:52Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Blob inventory

Actions Properties

Custom Domain:

Soft Delete: Enabled

Retention Days: 7

Microsoft Azure Storage Explorer

File Edit View Help

0KB/s

Thu 15 Feb 4:05 PM

Microsoft Azure Storage Explorer

Get Started

Storage Explorer

Get Started

Sign in with Azure
Sign in with your Azure account to access all of your Azure resources.

Attach to a resource
Don't have permissions to access Azure subscriptions? You can connect to individual storage resources using various authentication methods.

Tasks

View Azure resources
See your subscriptions and Azure resources all in one place.

Manage accounts and subscriptions
Manage your Azure accounts and choose which subscriptions you want to see.

Connect to Azure resources
Sign into Azure or attach to individual storage resources.

Show this page on startup

Activities

Clear completed Clear successful

Successfully added new connection.

File Endpoint: https://hexadeb1071.file.core.windows.net/;TableEn...

sig=5i4DV0lU0KyGcl%2BDECyDyYIEGTPrAVyt1UBvF6uYk%3D

Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Queue Service SAS URL: https://hexadeb1071.queue.core.windows.net/?sv=2022-11-02&ss=bfqt&srt=sco&sp=rw&lac=up&tfx=&se=2024-02-15T18:31:52Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Blob inventory

Actions Properties

Custom Domain:

Soft Delete: Enabled

Retention Days: 7

Microsoft Azure Storage Explorer

File Edit View Help

0KB/s

Thu 15 Feb 4:21 PM

Microsoft Azure Storage Explorer

hexa1071

Storage Explorer

hexa1071

EXPLORER

Search for resources

Collapse all Refresh all

Quick Access

Emulator & Attached

Storage Accounts

(Attached Containers)

(Emulator - Default Ports) (Key)

hexadeb1071-1 (SAS)

Blob Containers

Slogs

hexa1071

File Shares

Queues

Tables

Actions Properties

URL: https://hexadeb1071...

Custom Domain:

Type: Blob Container

HNS Enabled: false

Lease State: available

Lease Status: unlocked

Public Read Access: off

Last Modified: 15 February 2024 at 3...

Object replication

Start: 15/02/2024

End: 16/02/2024

Activities

Clear completed Clear successful

Successfully added new connection.

File Endpoint: https://hexadeb1071.file.core.windows.net/;TableEn...

sig=5i4DV0lU0KyGcl%2BDECyDyYIEGTPrAVyt1UBvF6uYk%3D

Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Queue Service SAS URL: https://hexadeb1071.queue.core.windows.net/?sv=2022-11-02&ss=bfqt&srt=sco&sp=rw&lac=up&tfx=&se=2024-02-15T18:31:52Z&st=2024-02-15T10:31:52Z&spr=https&sig=5i4DV0lU0KyGcl%...

Blob inventory

Actions Properties

Custom Domain:

Soft Delete: Enabled

Retention Days: 7

Microsoft Azure Storage Explorer

File Edit View Help

0KB/s

Thu 15 Feb 4:21 PM

Microsoft Azure Storage Explorer

hexa1071

Storage Explorer

hexa1071

EXPLORER

Search for resources

Collapse all Refresh all

Quick Access

Emulator & Attached

Storage Accounts

(Attached Containers)

(Emulator - Default Ports) (Key)

hexadeb1071-1 (SAS)

Blob Containers

Slogs

hexa1071

File Shares

Queues

Tables

Actions Properties

URL: https://hexadeb1071...

Custom Domain:

Type: Blob Container

HNS Enabled: false

Lease State: available

Lease Status: unlocked

Public Read Access: off

Last Modified: 15 February 2024 at 3...

Object replication

Start: 15/02/2024

End: 16/02/2024

Activities

Clear completed Clear successful

Successfully added new connection.

Name

Access Tier

Access Tier Last Mod

test.csv

Hot (inferred)

Showing 1 to 1 of 1 cached items

<< < > >> Load more

Permanent delete

Microsoft Azure Storage Explorer

File Edit View Help

0KB/s

Thu 15 Feb 4:21 PM