|  |  |  |
| --- | --- | --- |
| **Differences** | **Azure Blob Storage** | **Azure Data Lake Storage Gen 2** |
| **Purpose** | General-purpose object storage for a variety of use cases, such as storing large amounts of unstructured data, including images, videos, documents, and backups. | Optimized for big data analytics, providing file system semantics, and hierarchical namespace for organizing large volumes of data. |
| **Data Hierarchy** | Simple flat namespace with containers, blobs (objects), and directories. | Hierarchical namespace that allows for complex directory structures, file system semantics, and directory-level permissions. |
| **Access** | Accessed via the Blob Service REST API, Azure Storage client libraries, and Azure Storage Explorer. | Accessed via the Azure Data Lake Gen2 API, Azure Storage client libraries, Azure Storage Explorer, and through Hadoop-compatible file system driver (HDFS). |
| **Performance** | Optimized for high availability and scalable throughput. | Enhanced performance for analytics workloads, such as those involving large-scale, batch, and interactive data processing. |
| **Integration** | Can be integrated with other Azure services, such as Azure Virtual Machines, Azure Functions, and Azure Kubernetes Service. | Tight integration with big data analytics tools and frameworks, such as Azure Databricks, Azure Synapse Analytics, HDInsight, and Apache Spark. |
| **Namespace** | Blob Storage uses a flat namespace | Data Lake Gen2 supports a hierarchical namespace with file system semantics |

**Local Redundancy Storage (LRS)**

* **Redundancy**: Stores multiple copies of data within a single data center.
* **Availability**: Provides high availability within the data center.
* **Use Case**: Suitable for applications that require high availability but can tolerate data loss in the event of a regional outage.
* **Cost**: Lower cost compared to ZRS and GRS due to fewer copies of data.

**Zone-Redundant Storage (ZRS)**

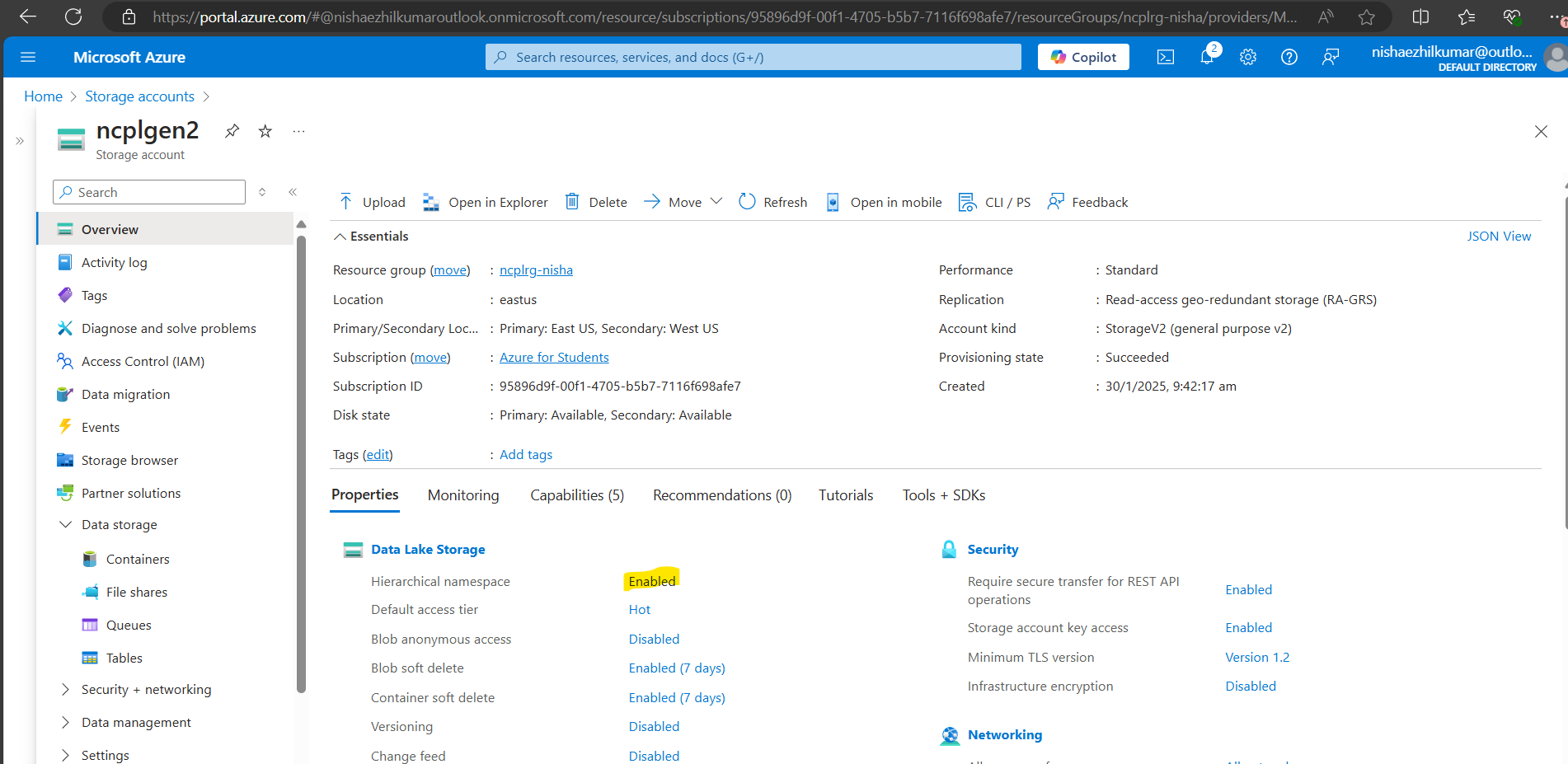
* **Redundancy**: Stores multiple copies of data across three Azure availability zones within a region.
* **Availability**: Provides higher availability than LRS by protecting against data center failures within a region.
* **Use Case**: Suitable for applications that require higher availability and can tolerate data loss in the event of a regional outage.
* **Cost**: Higher cost than LRS due to additional copies of data across multiple zones.

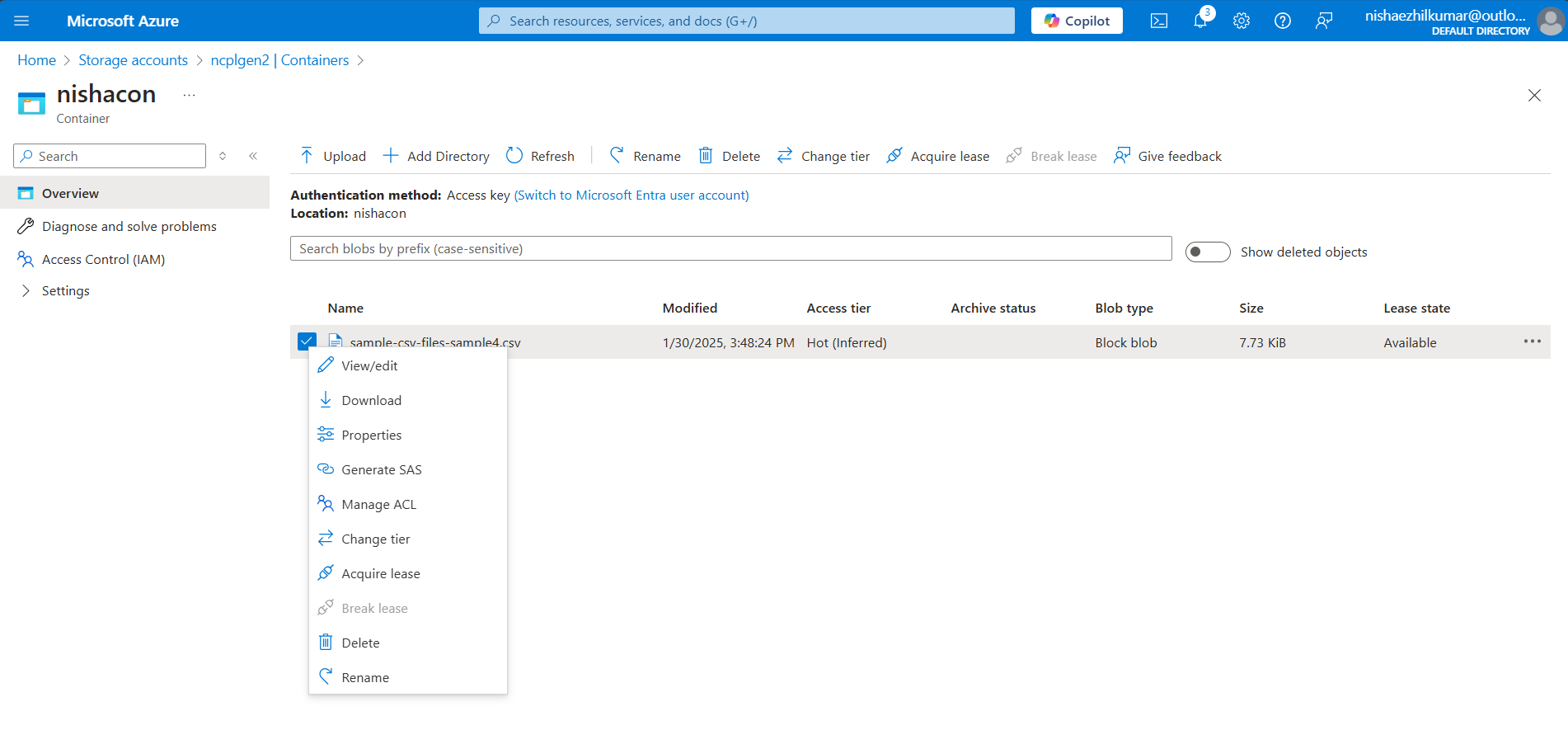
**Geo-Redundant Storage (GRS)**

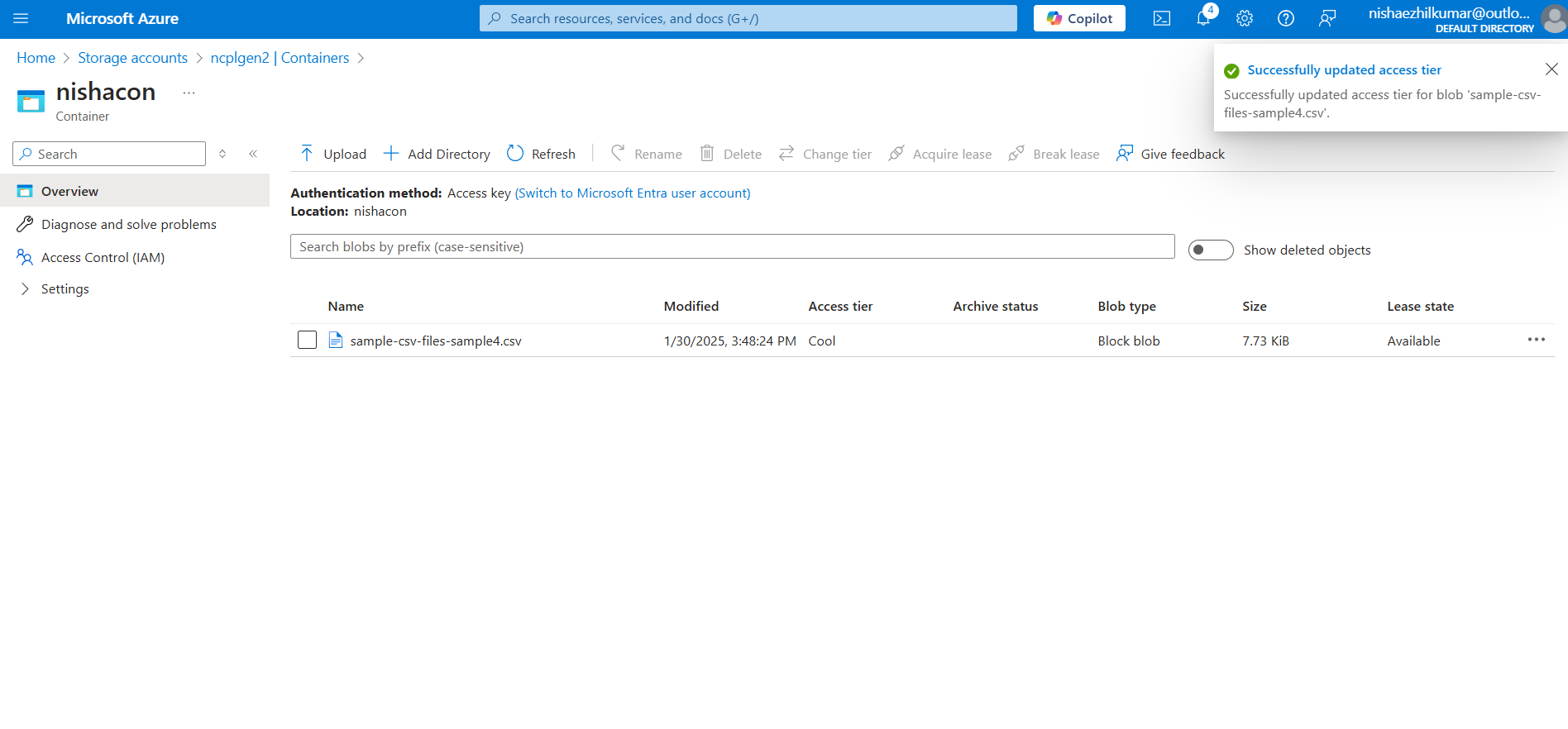
* **Redundancy**: Stores multiple copies of data in a primary region and replicates data to a secondary region (paired with the primary region).
* **Availability**: Provides the highest level of availability by protecting against data center failures and regional outages.
* **Use Case**: Suitable for applications that require the highest level of data durability and availability, even in the event of a regional outage.
* **Cost**: Highest cost among the three due to replication across regions.

2. **Access Tiers:**

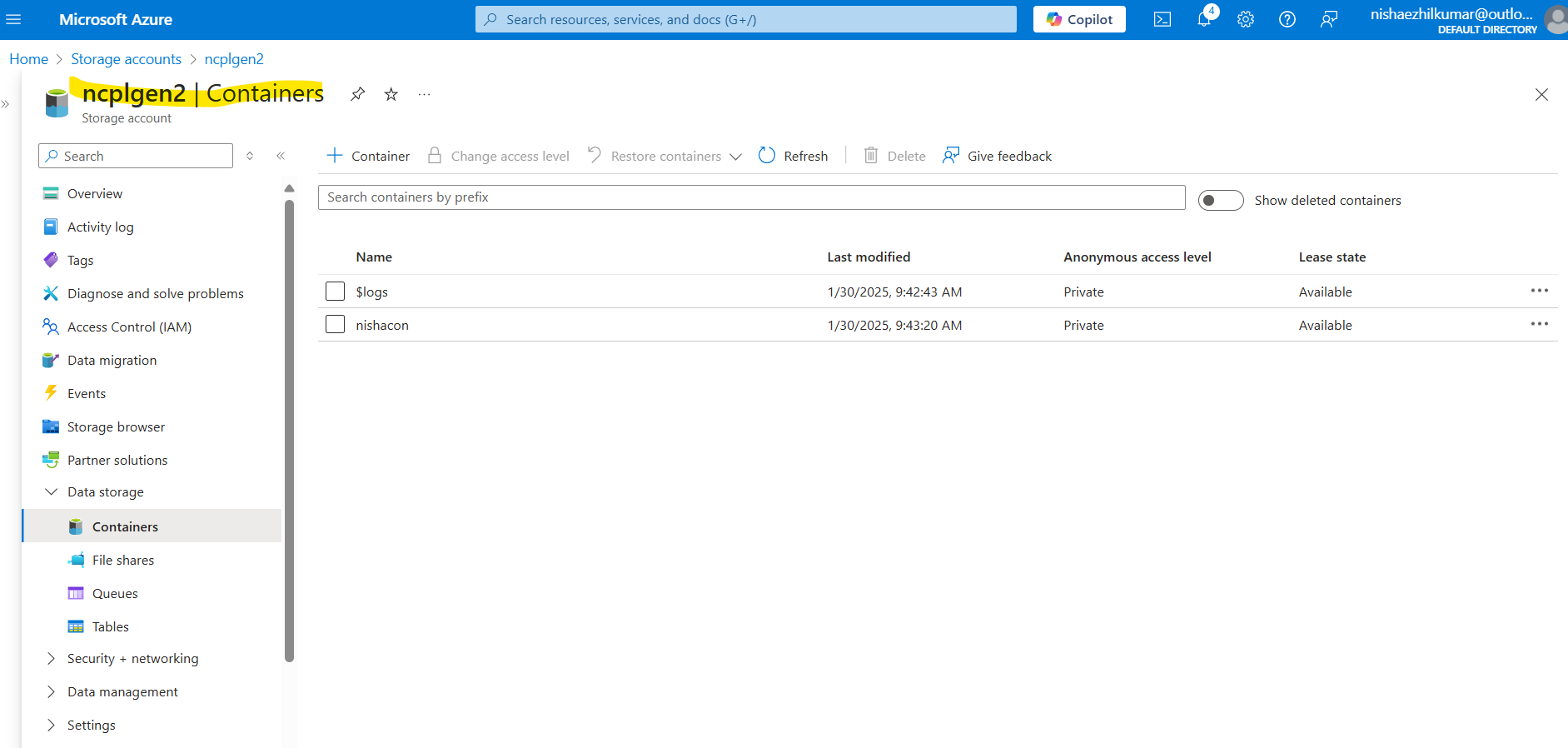
1. **Hot Tier**:
   * **Usage**: Suitable for data that is accessed frequently.
   * **Cost**: Higher storage cost but lower access (read/write) costs. Ideal for data that requires low latency and high throughput.
   * **Examples**: Active datasets, transactional data, real-time analytics.
2. **Cool Tier**:
   * **Usage**: Suitable for data that is infrequently accessed and stored for at least 30 days.
   * **Cost**: Lower storage cost compared to the Hot Tier but higher access costs. Best for data that doesn't need to be accessed regularly but still requires quick access when needed.
   * **Examples**: Backup data, disaster recovery data, long-term storage of infrequently used data.
3. **Archive Tier**:
   * **Usage**: Suitable for data that is rarely accessed and stored for at least 180 days.
   * **Cost**: Lowest storage cost but highest access and retrieval costs. Retrieval times can be in hours. Best for data that can tolerate high latency in access and is stored long-term.
   * **Examples**: Compliance data, historical data, archival data that needs to be preserved but is rarely used.

3.

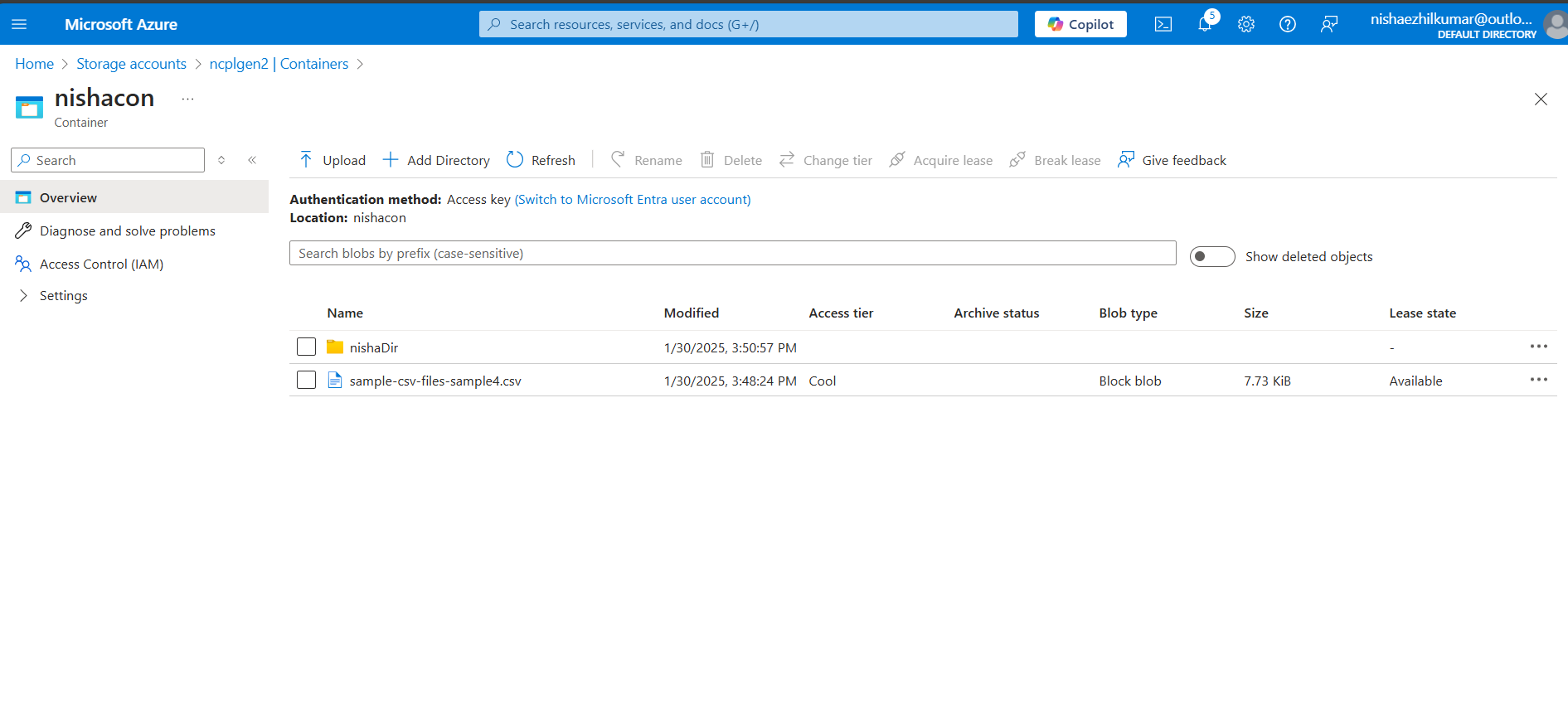




4.Containers



Directory



BLOB SAS Token: sp=r&st=2025-01-30T20:51:59Z&se=2025-01-31T04:51:59Z&spr=https&sv=2022-11-02&sr=d&sig=yTvmzHqgXS0DUYTnnYw0qB8M7m8%2BgBotE4i3JITl%2F24%3D&sdd=1

Blob SAS URL: <https://ncplgen2.blob.core.windows.net/nishacon/nishaDir?sp=r&st=2025-01-30T20:51:59Z&se=2025-01-31T04:51:59Z&spr=https&sv=2022-11-02&sr=d&sig=yTvmzHqgXS0DUYTnnYw0qB8M7m8%2BgBotE4i3JITl%2F24%3D&sdd=1>

5.  **Global Administrator** has the broadest access and control over the entire Azure subscription, including billing and user management.

 **Data Engineer** has specialized access focused on data-related tasks and resources.

 **Azure Admin** has scoped access to manage resources within a specific subscription or resource group but lacks billing and user management permissions.