Working with Azure Data Lake Storage Gen2



Reza Salehi
CLOUD CONSULTANT

@zaalion linkedin.com/in/rezasalehi2008



Overview



What is Azure Date Lake Storage (ADLS)?

- Features and use cases

Provisioning an ADLS instance

More resources for ADLS

Demo:

- Provisioning ADLS
- Uploading documents using the AzCopy tool



Azure Data Lake Storage

Is a hyper-scale repository for big data analytic workloads.



Azure Data Lake Storage



Capture data of any size, type and ingestion speed in one single place for analytics



Can be accessed from Hadoop (available with HDInsight cluster) using the WebHDFS-compatible REST APIs



Is tuned for performance for data analytics scenarios



Includes enterprise-grade capabilities: security, manageability, scalability, reliability, and availability



ADLS Generations

Gen1

Gen2

Converges the capabilities of Azure Blobs and Azure Data Lake Storage Gen1



Azure Data Lake Storage Gen2

Is a set of capabilities dedicated to big data analytics, built on top of the Azure Blob storage.



Azure Data Lake Storage Gen2

Data Lake Gen1 Features, such as file system semantics, file level security and scale

Combined with low-cost, high availability, and disaster recovery from Azure Blob storage



Microsoft recommends Azure Data Lake Storage Gen2.



Designed for Enterprise Big Data Analytics



ADLS Gen2 makes Azure Storage the foundation for building enterprise data lakes on Azure



Supporting hundreds of gigabits of throughput, ADLS Gen2 allows you to manage massive amounts of data



A fundamental part of ADLS Gen2 is the addition of a hierarchical namespace for Blob storage



Uses slashes in the name to mimic a hierarchical directory structure



Key Features of ADLS Gen2

Hadoop compatible

Manage data same as Hadoop HDFS

POSIX permissions

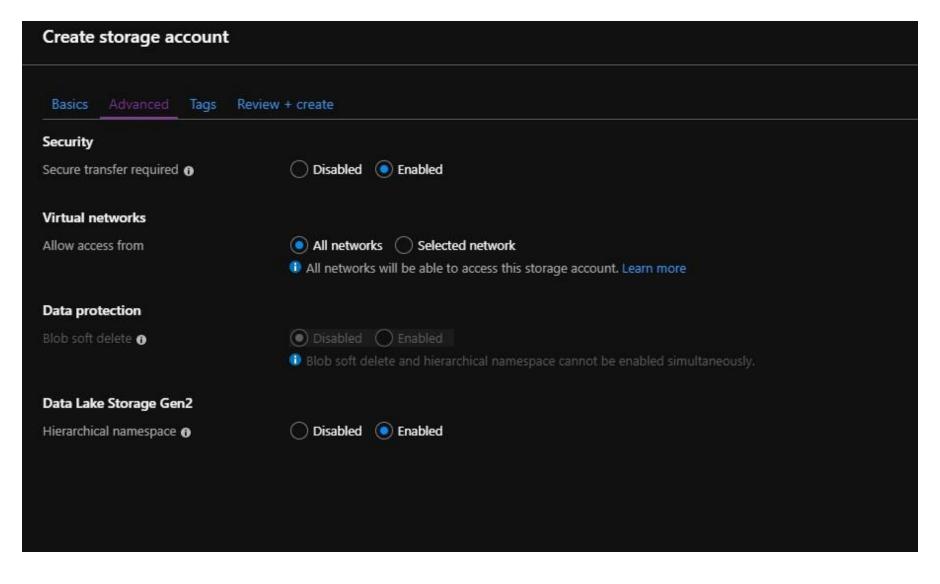
Supports ACL and POSIX file permissions

Cost effective

offers low-cost storage capacity



Provisioning ADLS Gen2





More Resources for ADLS Gen2





Demo



Provisioning Azure Date Lake Storage Gen2 using the portal

Uploading documents using AzCopy



Summary



What is Azure Date Lake Storage?

- Gen 1 vs. Gen 2

Azure Date Lake Storage features

Provisioning an Azure Date Lake Storage instance

More resources for Azure Date Lake Storage

Demo: Provisioning Azure Date Lake Storage

