

Custom Learnings

Day 6

To communicate between data storage account and data factory:

1. Shared Access Signature: Time bounded
2. Access Key: Life time access

Creation of storage account and shared access signature(sas):

The screenshot shows the Microsoft Azure portal interface. The top navigation bar includes the Microsoft Azure logo, a search bar, and the user profile 'Shellunext_1693422208... UNEXT'. The main content area displays the 'idashellsa03' storage account overview. The left sidebar contains a navigation menu with options like Overview, Activity log, Tags, Diagnose and solve problems, Access Control (IAM), Data migration, Events, Storage browser, Data storage (Containers, File shares, Queues, Tables), and Security + networking (Networking). The main content area is divided into sections: Essentials (Resource group: RG_IDA_SHELL_03, Location: East US, Subscription: npunext-1680261698581, Subscription ID: 7179a1e4-9824-4360-9eda-2930d7e2dc08, Disk state: Available, Tags: Add tags), Properties (Data Lake Storage settings: Hierarchical namespace: Enabled, Default access tier: Hot, Blob anonymous access: Disabled, Blob soft delete: Enabled (7 days), Container soft delete: Enabled (7 days), Versioning: Disabled, Change feed: Disabled), Security (Require secure transfer for REST API operations: Enabled, Storage account key access: Enabled, Minimum TLS version: Version 1.2, Infrastructure encryption: Disabled), and Networking (Allow access from: All networks).

The screenshot shows the 'Shared access signature' page for the storage account 'idashellsa03'. The top navigation bar is the same as the previous screenshot. The main content area is titled 'idashellsa03 | Shared access signature'. The left sidebar is the same as the previous screenshot. The main content area displays a 'Generate SAS and connection string' button. Below this, there are sections for 'Connection string', 'SAS token', 'Blob service SAS URL', 'File service SAS URL', 'Queue service SAS URL', and 'Table service SAS URL'. Each section contains a text box with the generated SAS token and a copy icon. The SAS token is: ?sv=2022-11-02&ss=bfqt&srt=c&sp=rwdlacupyx&se=2023-09-05T04:50:10Z&st=2023-09-05T03:50:10Z&spr=https&sig=uaRnscB0EzbWZxeQoqjMo...

In case of data, we can't give hard-coded passwords to give access to database. Hence we use Azure Key Vault to give a time bounded name that is mapped to a key password that can be used as a login credential.

Creation of key vault and secret:

The screenshot shows the Microsoft Azure portal interface for a Key Vault named 'idashellKey003'. The left sidebar contains navigation links for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Access policies, Events, Objects (Keys, Secrets, Certificates), and Settings (Access configuration, Networking, Microsoft Defender for Cloud). The main content area displays the 'Essentials' section with the following details:

- Resource group: [RG_IDA_SHELL_03](#)
- Location: East US
- Subscription: [npunext-1680261698581](#)
- Subscription ID: 7179a1e4-9824-4360-9eda-2930d7e2dc08
- Vault URI: <https://idashellkey003.vault.azure.net/>
- Skus (Pricing tier): Standard
- Directory ID: dce87315-8ffa-4a01-ab40-0de5a7214b2f
- Directory Name: Unnext
- Soft-delete: [Enabled](#)
- Purge protection: [Disabled](#)

Below the Essentials section, there are tabs for 'Get started', 'Properties', 'Monitoring', 'Tools + SDKs', and 'Tutorials'. A message at the bottom states: 'Manage keys and secrets used by apps and services'.

The screenshot shows the Microsoft Azure portal interface for the 'idashellKey003' Key Vault, specifically the 'Secrets' page. A notification banner at the top right states: 'Creating the secret 'idashellkeyname'. The secret 'idashellkeyname' has been successfully created.' Below the notification, a table lists the created secret:

Name	Type	Status	Expiration date
idashellkeyname	secret key	✓ Enabled	

The left sidebar is identical to the previous screenshot. The main content area also includes links for 'Generate/Import', 'Refresh', 'Restore Backup', 'View sample code', and 'Manage deleted secrets'.

Access Control (IAM): Check your access level and assign role to other users.

Azure Data Factory:

Transformation: ADB

E&L: ADF, using for data flow activity

ADF Components:

1. Pipelines: Combination of one or more activities
2. Activities: Action to be carried out
3. Datasets : To access the particular file or table we create datasets.
4. Linked services : To access the particular path or database we create linked service.
5. Integration runtime: To communicate to any source or destination, we create a connection.
Gives access to a higher level.
6. Triggers
7. Data flow

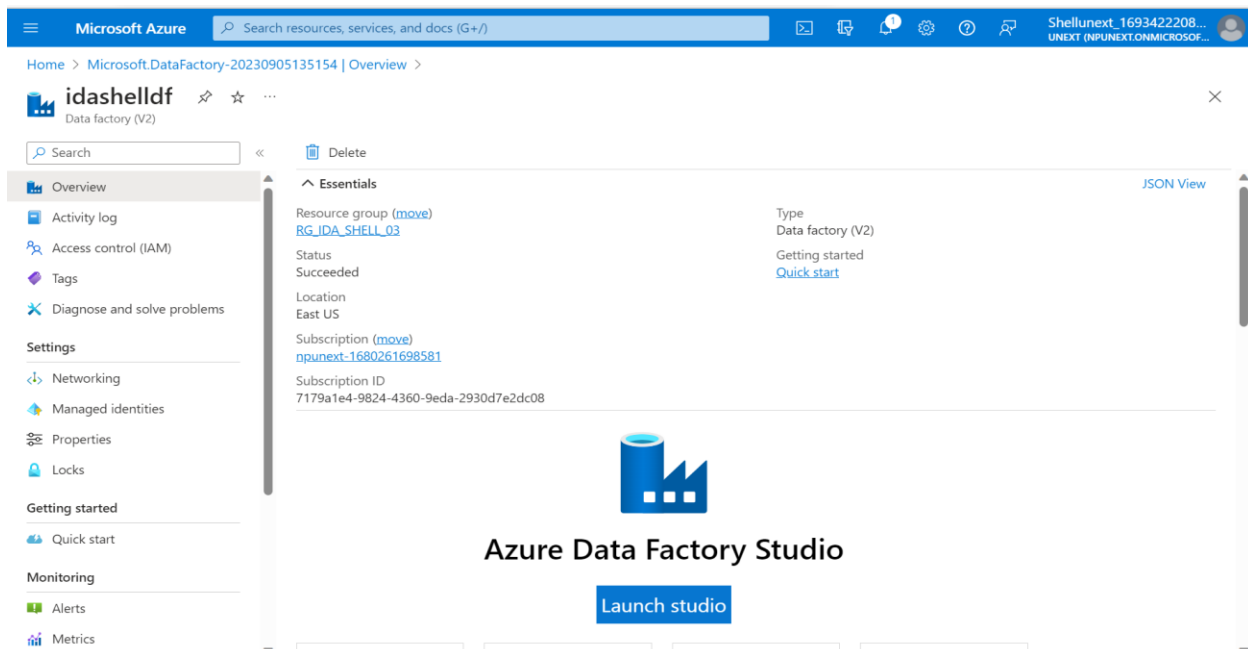
ELT: Advantage is we need not extract data from the source each and every time.

ETL: We transform all the data that is extracted.

IR:

1. Azure Auto Resolved: If we have any cloud based storage.
2. Self Hosted: If we have any on-premises or local storage and want to connect to it.
3. Azure SSIS: If any existing SSIS package and want to run that in Azure SSIS.

Creation of Data Factory:



Creation of integration runtime, linked service and dataset:

Microsoft Azure | Data Factory | idashell | Search factory and documentation | Shellunext_1693422208756@npunext.onmicrosoft.com | UNEXT

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

» Data Factory | Validate all | Publish all | Preview experience | Off

Integration runtimes

The integration runtime (IR) is the compute infrastructure to provide the following data integration capabilities across different network environment. [Learn more](#)

+ New Refresh

Filter by name

Showing 1 - 2 of 2 items

Name	Type	Sub-type	Status	Related	Region	Version
AutoResolveIntegrationR...	Azure	Public	Running	0	Auto Resolve	---
IRidashell	Self-Hosted	---	Unav... More	0	---	---

Microsoft Azure | Data Factory | idashell | Search factory and documentation | Shellunext_1693422208756@npunext.onmicrosoft.com | UNEXT

Microsoft recently announced the public preview of Microsoft Fabric, a brand new and exciting way to build cloud-first data analytics. Click [here](#) to get started with Fabric Data Factory!

» Data Factory | Validate all | Publish all | Preview experience | Off

Linked services

Linked service defines the connection information to a data store or compute. [Learn more](#)

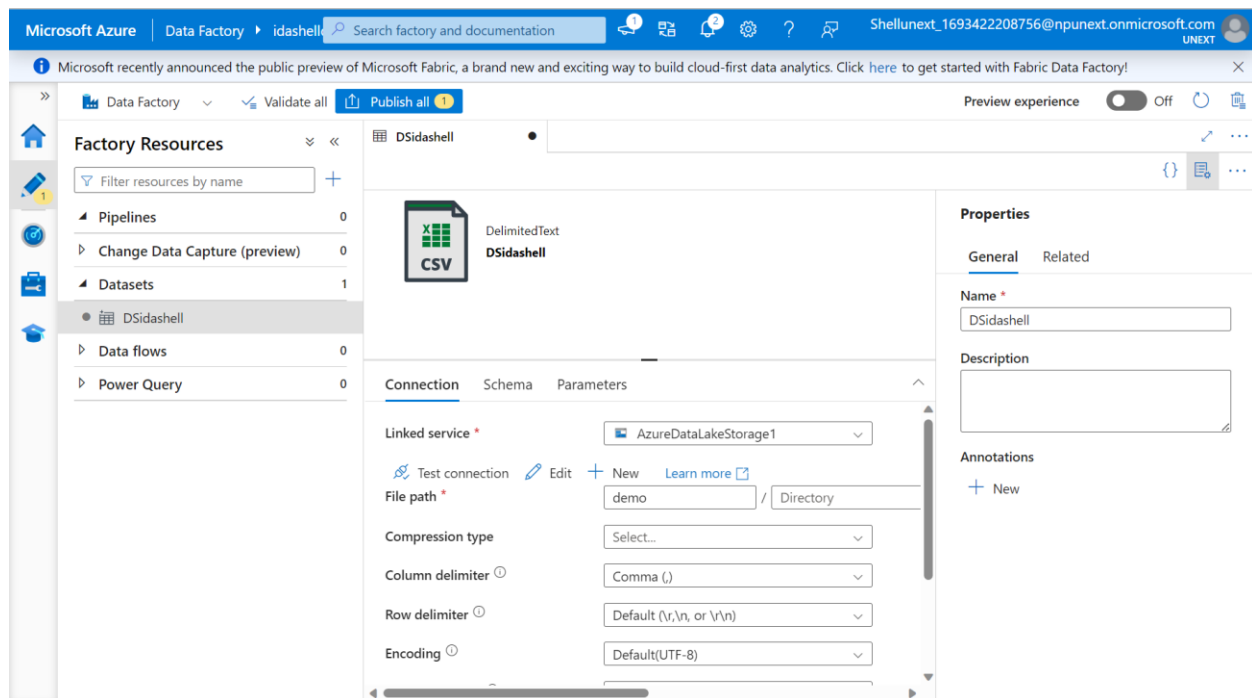
+ New

Filter by name Annotations: Any

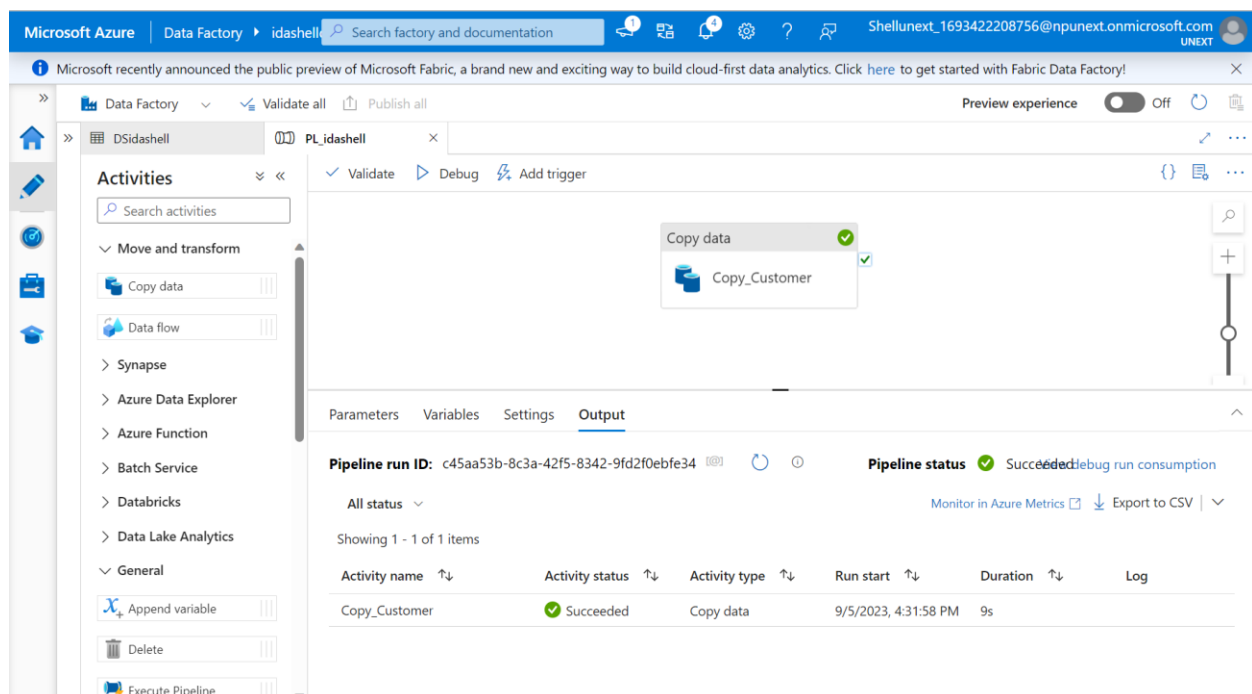
Showing 1 - 1 of 1 items

Name	Type	Related	Annotations
AzureDataLakeStorage1	Azure Data Lake Storage Gen2	0	

Successfully created
Successfully created AzureDataLakeStorage1 (Linked service).



Creation of pipeline:



Parameterized dataset and linked service: Each time it will run, we can give the name of dataset and linked service so that the same pipeline can be used for different parameters.

Pipeline: Overview of pipeline and simple copying of data from one source to another with and without parameterization.