



1. **Source systems** : These are the different source systems from where the data is ingested to warehouse as part of existing solutions.
2. **Data ingestion from Sources**: The data from different source systems will be transferred on daily basis to ADLS (Azure Data Lake Storage) Gen2 in the underlying Bronze Layer via Azure Data Factory.
3. **Data Transformation through Databricks**: We can use Azure Databricks for transforming the data. Azure Databricks shall fetch the ingested data from the ADLS Bronze layer and transform the data.
4. **Storing transformed data in Gold Layer**: Transformed data shall be placed in the ADLS Gen 2 Gold Layer.
5. **Copying data to Databricks SQL Warehouse**: The data from the Gold layer shall be copied to Databricks SQL Warehouse where the dimensions and facts can be created and data be modelled for reporting purposes.
6. **Visualization through Power BI**: The data residing in the Azure SQL Warehouse can be exposed as a dataset in the Power BI Workspace, which can be consumed by the business stakeholders or Power BI Developers for generating reports. This will be controlled by a respective Active Directory group. Read-only access is expected to be of interest for users.
7. **Key Vault** : The credentials to connect to respective Storage Azure files and the access keys are stored in Azure Key vault as secrets.