**Context:**

A Pizza Restaurant chain “Pizza House” has more than 2000 stores across the country. Each store manages its own inventory of raw materials. Each store prepares pizzas, side dishes, etc. and sells them along with ready to eat products such as cookies, drinks, etc. The sale can happen by Point of Sale (POS) or Online. The online transactions would be flowing in real time whereas the transactions made by POS can be synced every 15 minutes in batches. They offer pick-up and deliveries by 3rd party providers.

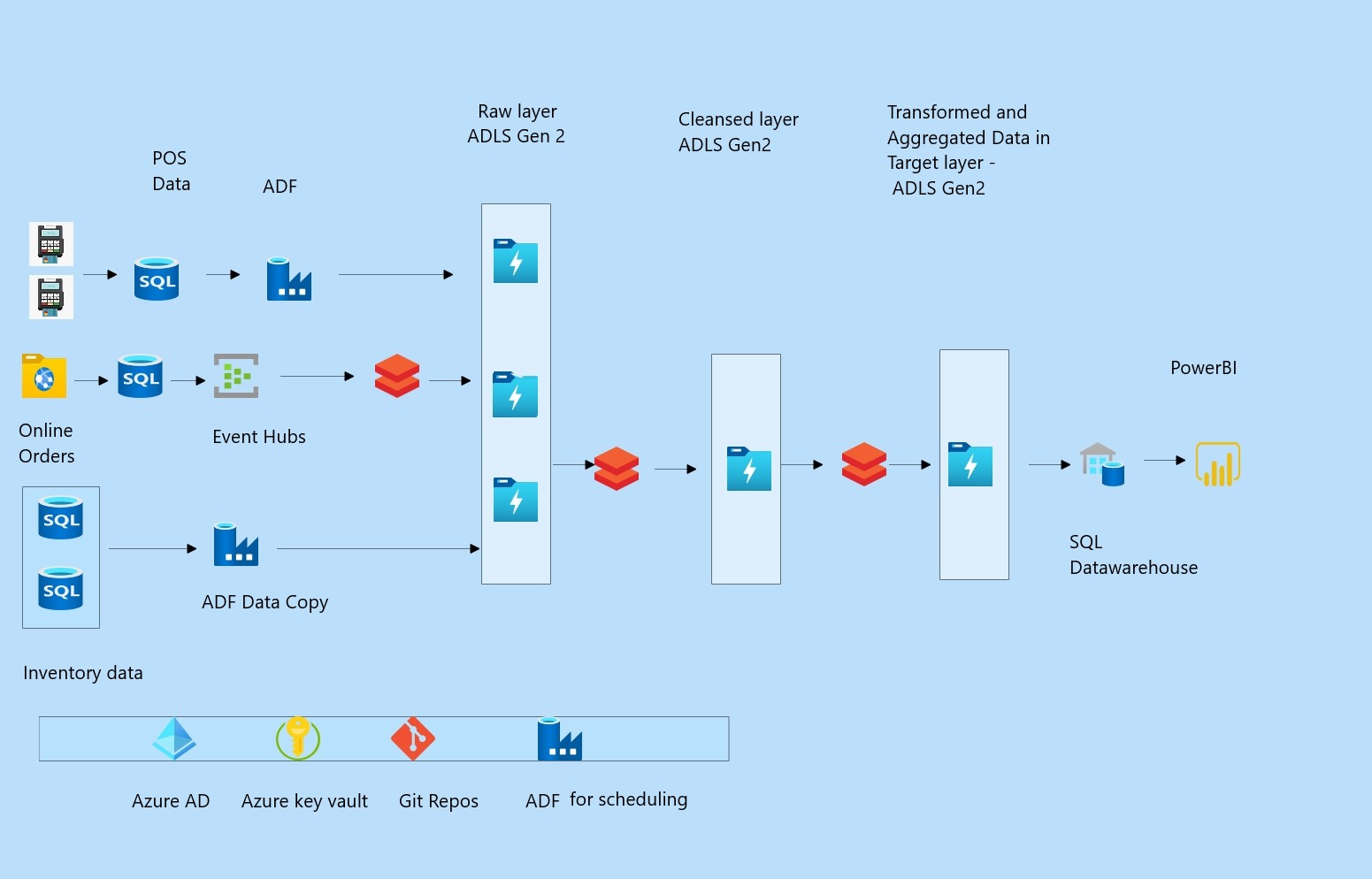
At the head office of the restaurant chain, management is concerned with the logistics of ordering, stocking and selling products while maximizing profits as well as understanding their marketing & communications. Several promotional schemes such as temporary price reductions, ads in newspapers, displays etc., also keep rising. Considering the huge data volumes (hundreds of GB per month) and the variety of the data they have; management wants the architecture to be robust enough to handle the varying data loads.

Design a cloud data platform to process and deliver insights based on the above. Please provide a high level solution design for the architecture. Feel free to choose any cloud provider you want.

**Requirements**

1. Handle large write volume: Billions of write events per day.
2. Handle large read/query volume: Millions of merchants wish to gain insight into their business. Read/Query patterns are time-series related metrics.
3. Provide metrics to customers with at most one hour delay.
4. Run with minimum downtime.
5. Have the ability to reprocess historical data in case of bugs in the processing logic.

**Solution Design:**



Data Sourcing:

1. POS Data is sent to every 15 minutes in Azure SQL Database.

Using ADF Copy, we can the data 15 minutes batch interval level to ADLS Gen2 storage.

1. Online orders data gets updated at Azure SQL DB in real time.

Using Azure event hubs and Spark streaming, we can load the data into ADLS.

1. Stores inventory data is stored in Azure SQL DB.

Using ADF copy we can load data from SQL DB to ADLS.

Data Cleaning:

1. Check for null values, data types, duplicates, rows with values outside business range while moving data from raw layer to cleansed layer
2. If there are bad records, please move those records into bad records table

Data Transformations:

1. Identify and understand the transformations and aggregations required to generate the business KPIs
2. Implement all the transformations in Data pipeline and insert the data in ADLS Gen2.

Data warehousing & Analytics:

1. Create the External table in SQL Data warehouse and pass ADLS path in DDL.
2. Connect to SQL DW from Power BI and create required reports and dashboards.