**High Level Design Document**

Presented By: Manideep B

**For Pizza House Chain of Restaurant Data Management System**

**Scenario:**

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.

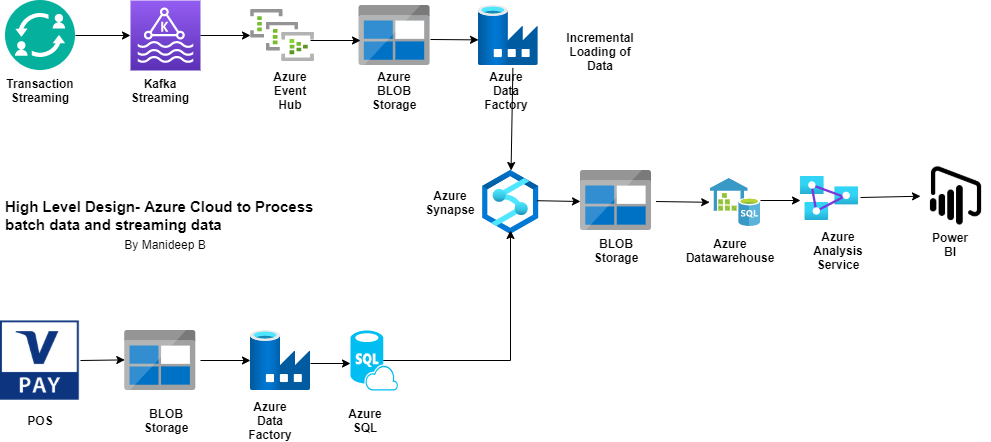
At the head office of the restaurant chain, management is concerned with the logistics of ordering, stocking and selling products while maximizing profits as well as understanding their marketing & communications. Several promotional schemes such as temporary price reductions, ads in newspapers, displays 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.

**Process Flow Diagram**



**Architecture Description:**

Two Data Flow mechanisms have been defined:

* For the batch processing of Data using Blob Storage and Azure Data Factory and Finally dumping the data to Azure SQL.
* For the real time processing, the processing has been done through Kafka stream as we have to process billions of data per day hence we need a proper queuing mechanism otherwise there would be bottleneck in the system. Finally using other components like event hub, blob storage and Azure Data Factory for incremental load of data and also making sure that same load is not loaded again and also making sure to load the data and not missing any data.

Finally both the data has been combined and consumed using Azure Analysis services and reports and dashboards are being created in Power BI to give the near real time view to the end users. Some of the components have been used as per my assumption on the requirements. Otherwise in a real scenario I would have done meetings with the stakeholder to understand some more details on the requirements.

**Why Such a System has been designed:**

* The above system can handle billions of write events per day as we are using Kafka streaming and also Azure Data Factory which are both scalable and queue-based systems.
* Again the large query read volume using time series data can be easily designed using Azure Analytics Service based on the Key so that the merchants can get the required insights.
* This system would give near to real time metrics to the customer as the data processing is being done continuously with minimal time lag.
* As this is based on cloud, so no maintenance window required as such and hence no downtime as such during the business hours.
* No issues of data reprocessing or no processing of data as the same is being monitored in the Azure Data Factory.