**Company Overview**

CloudPortalHub owns and operates 300 convenience stores across the US. The company sells a variety of packaged foods and drinks, as well as a variety of prepared foods, such as sandwiches and pizzas.

CloudPortalHub has a loyalty club whereby members can get daily discounts on specific items by providing their membership number at checkout. CloudPortalHub employs business analysts who prefer to analyze data by using Microsoft Power BI, and data scientists who prefer analyzing data in Azure Databricks notebooks.

**Requirements**

**Business Goals**

CloudPortalHub wants to create a new analytics environment in Azure to meet the following requirements:

* See inventory levels across the stores.
* Data must be updated as close to real time as possible.
* Execute ad hoc analytical queries on historical data to identify whether the loyalty club discounts increase sales of the discounted products.
* Every four hours, notify store employees about how many prepared food items to produce based on historical demand from the sales data.

**Technical Requirements**

CloudPortalHub identifies the following technical requirements

* Minimize the number of different Azure services needed to achieve the business goals.
* Use platform as a service (PaaS) offerings whenever possible and avoid having to provision virtual machines that must be managed by CloudPortalHub.
* Ensure that the analytical data store is accessible only to the company's on-premises network and Azure services.
* Use Azure Active Directory (Azure AD) authentication whenever possible.
* Use the principle of least privilege when designing security.
* Stage Inventory data in Azure Data Lake Storage Gen2 before loading the data into the analytical data store.
* CloudPortalHub wants to remove transient data from Data Lake Storage once the data is no longer in use.
* Files that have a modified date that is older than 14 days must be removed.
* Limit the business analysts' access to customer contact information, such as phone numbers, because this type of data is not analytically relevant.
* Ensure that you can quickly restore a copy of the analytical data store within one hour in the event of corruption or accidental deletion.

**Planned changes**

The following changes are planned

The application development team will create an Azure event hub to receive real-time sales data, including store number, date, time, product ID, customer loyalty number, price, and discount amount, from the point of sale (POS) system and output the data to a storage account in Azure.

Customer data, including name, contact information, and loyalty number, comes from Salesforce, a SaaS application, and can be imported into Azure once every eight hours. Row modified dates are not trusted in the source table.

Product data, including product ID, name, and category, comes from Salesforce and can be imported into Azure once every eight hours. Row modified dates are not trusted in the source table. Daily inventory data comes from a Microsoft SQL server located on a private network.

CloudPortalHub currently has 5 TB of historical sales data and 100 GB of customer data. The company expects approximately 100 GB of new data per month for the next year. CloudPortalHub will build a custom application named FoodPrep to provide store employees with the calculation results of how many prepared food items to produce every four hours. CloudPortalHub does not plan to implement Azure ExpressRoute or a VPN between the on-premises network and Azure.