**Azure Time Series Insights (TSI)**

# **TSI Architecture Design**

A diagram of a software

Description automatically generated

A computer screen shot of a diagram

Description automatically generated

# **Use Case – Design Documents**

A diagram of data storage

Description automatically generated

# **Pricing**

We have considered battery data for a month for LIVE and HISTORY. We have considered 5MB as the average file size. Below screenshot is from the month May 2023 for battery INSMOMAH0106G2300E12.

A screenshot of a computer

Description automatically generated

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **No** | **Objects** | **Cost** | **Current Data Volume (in GB) for 41000 Battery data  (Assumption- 5MB/Battery/Month)** | **Current Cost per Month** | **Future Data Volume (in GB) for 200000 Battery data  (Assumption- 5MB/Battery/Month)** | **Future Cost per Month** |
| 1 | Processing up to 100GB of data per month for IOT/Event Hub | 3379.59 | 100 | 3380 | 100 | 3380 |
| 2 | Additional Processing cost per GB of data | 22.94 | 100 | 2299 | 877 | 20112 |
| 3 | Meta Data Storage per MB (Not Sure) | 4.59 |  | 5 |  | 5 |
| 4 | Warm Storage per GB per Month | 344.15 | 200 | 68896 | 977 | 336080 |
| 5 | Cold Blob Pricing (for 1 year) | 9.83 | 2402 | 23622 | 11719 | 115228 |
| 6 | Cold Blob Query charges per GB | 1.15 |  | 1 |  | 1 |
|  | **Total** | **3762.25** | **200** | **98202** | **977** | **474805** |

A screenshot of a computer

Description automatically generated

# **Business Disaster Recovery (HA and DR)**

Azure supports disaster recovery capabilities through Azure's *cross-region availability* feature.

1. Failover: Azure provides geo-replication and load balancing.
2. Data restoration and storage recovery: Azure provides several options to preserve and recover data.
3. Azure Site Recovery: Azure provides recovery features through Azure Site Recovery.
4. Azure Backup: Azure Backup supports both on-premises and in-cloud backup of Azure VMs.

# **Short Notes:**

Azure Time Series Insights is a fully managed analytics, storage, and visualization service offered by Microsoft as part of its Azure cloud platform. It is designed to help organizations process, store, and analyze large volumes of time-series data in real-time. Time series data refers to data points that are recorded with a timestamp, such as sensor readings, events, logs, and other time-stamped measurements.

Key features and capabilities of Azure Time Series Insights include:

**1. Data Ingestion:**

Time Series Insights supports the ingestion of massive amounts of time-series data from various sources, such as IoT devices, applications, and other data streams.

Data ingestion is a critical aspect of any time series data analysis platform like Azure Time Series Insights. Here are some pros and cons of data ingestion in Azure Time Series Insights:

**Pros:**

1. Scalability: Azure Time Series Insights is designed to handle large volumes of time-series data from various sources, making it suitable for IoT and industrial use cases where data volumes can be massive.

2. Real-time Ingestion: The service supports real-time data ingestion, enabling users to analyze and act on data as it arrives, leading to more timely insights and faster decision-making.

3. Ease of Integration: Time Series Insights seamlessly integrates with other Azure services like **Azure IoT Hub** and **Azure Stream Analytics**, simplifying the process of ingesting data from different sources.

4. Optimized Storage: The platform provides optimized storage mechanisms for time-series data, ensuring efficient data storage and retrieval for analysis.

5. Metadata Management: Time Series Insights allows you to define and manage metadata for ingested data, making it easier to organize and structure your data for analysis.

6. Historical Analysis: In addition to real-time data, you can also ingest and analyze historical data. The platform supports **"cold" storage for less frequently** accessed historical data.

**Cons:**

1. Complexity: Setting up and configuring data ingestion pipelines can be complex, especially when dealing with **multiple data sources and integration points**. Users need to have a good understanding of Azure services and data integration concepts.

2. Costs: Ingesting and storing large volumes of data can lead to increased costs, especially in terms of storage and data transfer fees. It's important to carefully plan and monitor your usage to manage costs effectively.

3. Data Quality and Cleanup: Ingesting raw data without proper validation and preprocessing can lead to data quality issues. It's essential to implement data validation and cleanup processes to ensure accurate analysis.

4. Latency: While Azure Time Series Insights supports real-time data ingestion, there may still be some inherent latency in the ingestion process. This can impact the timeliness of insights, especially in highly time-sensitive applications.

5. Data Source Compatibility: While Azure Time Series Insights offers integration with various Azure services, integrating with certain custom or external data sources may require additional development and configuration.

6. Learning Curve: Learning how to effectively set up data ingestion pipelines and manage the platform's features may require time and effort, particularly for users who are new to Azure or time series data concepts.

In conclusion, data ingestion in Azure Time Series Insights offers significant advantages for processing and analyzing time-series data, particularly in IoT and industrial contexts. However, there are challenges to consider, such as complexity, cost management, and data quality, which require careful planning and execution to maximize the benefits of the platform.

**Data Ingestion Support:**

1. Azure IoT Hub Integration: Azure Time Series Insights seamlessly integrate with Azure IoT Hub, allowing you to ingest data from IoT devices and sensors. This integration simplifies the process of connecting and sending data to Time Series Insights.

2. Azure Stream Analytics Integration: You can use Azure Stream Analytics to preprocess and transform data before ingesting it into Time Series Insights. This integration enables you to apply real-time analytics and filtering to your data streams.

3. Custom Ingestion: Time Series Insights supports ingesting data from custom sources using its REST API. This allows you to design your own data ingestion pipelines and integrate with various data streams.

4. Historical Data Import: In addition to real-time data, you can also import historical data into Time Series Insights. This enables you to perform historical analysis and visualize trends over time.

**Data Ingestion Limitations:**

1. Data Types: While Time Series Insights is designed for time-series data, it may not be the best fit for all types of data. It's optimized for timestamped data points, which may limit its applicability for more complex data structures.

2. Data Latency: Although Time Series Insights supports real-time data ingestion, the actual latency may vary based on factors such as data source, network conditions, and processing requirements.

3. Custom Data Formats: While you can use the REST API for custom ingestion, there might be limitations when dealing with non-standard data formats. You may need to ensure your data is formatted correctly for ingestion.

4. Scalability and Costs: Ingesting and storing large volumes of data can lead to increased costs. You should carefully consider the scalability requirements of your application and plan for potential cost implications.

5. Complexity: Setting up data ingestion pipelines and configuring integrations with Azure IoT Hub or Azure Stream Analytics can be complex, especially for users who are new to these technologies.

6. Rate Limiting: There might be rate limits on data ingestion, particularly when dealing with large bursts of data. It's important to be aware of these limits and plan your ingestion accordingly.

7. Data Validation and Cleanup: Ensuring data quality, validation, and proper cleanup processes are your responsibility. Time Series Insights doesn't handle data validation or transformation inherently.

8. Learning Curve: Like any technology, there is a learning curve associated with effectively using Time Series Insights for data ingestion and analysis.

**2. Data Storage:** The service provides a highly optimized and scalable storage backend that allows efficient storage of time-series data for quick access and analysis.

1. Provide 2 storage types, warm and cold storage. Warm Storage is optional can be enabled or disabled, and it range from 7 to 31 days range.
   1. Warm Storage – Store data for at least 7 days and max 31 days. There are no additional charges for multiple queries run on this storage.
   2. Cold Storage – History data stored in this which is aged more than 31 days.
2. Data stored in either of storage is Indexed.
3. Need to have edit and delete permission for the cold storage file. But it is not recommended to make any changes.

3. **Data Exploration:** Users can explore and visualize their time-series data using interactive tools and dashboards. This helps to gain insights, identify trends, and monitor operational metrics.

**Data Exploration Tools:**

1. Azure Time Series Insights Gen1 Explorer
2. Azure Data Catalog
3. Power BI
4. Azure Notebooks running Python script.

4. **Querying and Analytics:** Azure Time Series Insights offers powerful querying capabilities, allowing users to run ad-hoc queries on their time-series data to extract specific information and insights. The service supports functions, aggregations, and filters for data manipulation.

5. **Integration with Azure Services:** Time Series Insights is tightly integrated with other Azure services, such as Azure IoT Hub and Azure Stream Analytics, which makes it a part of end-to-end IoT and data processing solutions.

Integration with Azure Services: Azure IoT Hub and Azure Stream Analytics

6. **Scalability:** The service is designed to handle large-scale deployments and can dynamically scale resources based on demand.

7. **Security and Compliance:** Azure Time Series Insights provides features for securing data, managing access controls, and ensuring compliance with various industry regulations.

8. **Time Travel:** This feature allows users to rewind and replay historical data to analyze events and trends from the past.

9. **Alerting and Monitoring:** The service enables users to set up alerts based on specific conditions and thresholds, ensuring proactive monitoring and notifications.

10. **Integration with Power BI:** Azure Time Series Insights can be seamlessly integrated with Microsoft Power BI for advanced data visualization and reporting.

# **Conclusion**

We can use this TSI Azure service to track real time data. Our use case includes:

1. Battery Live and Battery History data
2. Station Live and History Data

Few of the application limitations have been identified for the data ingestion. Below are the lists:

1. Works well with fixed datatype mapped with timestamp (bool, datetime, double, long, string, dynamic). Not with complex datatype.
2. Custom Data format not supported.
3. Data storage volume directly impacted on the cost.
4. Ingestion rate is 1MBps or 1000 events stored per second.
5. Data Partition limit 0.5MBps or 500 events stored per second.
6. No Data Validation is Handled

The discovery made for the TSI service is on high level and would need to carry out a proper POC. POC would help us to deep dive for better understanding on the compatibility and limitation by validating it with the business requirements for further adaption.