Why Do We Need a Time-Series Database?

Databases are used for storing, maintaining and accessing any sort of data.Databases are useful in many different scenarios for storing data. They collect information on people, places or things. That information is gathered in one place so that it can be observed and analyzed. Databases can be thought of as an organized collection of information.In the era of big data, databases and database management systems are becoming more and more important.

However, in some cases, traditional databases do not perform so well.As 5G technology matures, the Internet of Things will make everything connected. Before the Internet of Things era, only mobile phones and computers can be connected to the Internet, but in the future, all devices will be connected to the Internet. These devices will spit out a large amount of time-organized data every moment.This is time-series data, a series of data points indexed/listed in time order.Time series data is often a continuous flow of data like measurements from sensors and intraday stock prices. Time series data is very different from normal data in many ways.These data should not only be generated in real time, written and stored, but also support fast query, achieve visual display, help managers to analyze and make decisions, and be used for big data analysis,which is a big challenge for traditional databases.

There are huge differences between sequential business and ordinary business in many aspects, which can be summarized as follows:

1. Consistently generating huge amounts of data without peaks and troughs. To take a simple example, a sentry monitoring system that monitors 100 metrics per second on 1W servers would generate 100 WATTS of TPS per second. For example, if 100 million people are wearing a popular sports wristband and each wristband only collects 3 metrcis (heartbeat, pulse and step count) per second, it will also generate 300W TPS per second.

2. Data is inserted, and there is no update or deletion. Based on the fact that data generated by sequential services is rarely updated or deleted, there can be considerable simplification in the design of sequential database architectures.

3. More attention will be paid to recent data, and streaming processing will be paid more attention in the future. Time-old data are rarely accessed or even discarded. It's easy to understand, sentry systems we tend to care most about the last hour, most about the last three days, very little about the last three days. With the advent of streaming computing, sequential data will inevitably pay more attention to the value of real-time data in the future development, and this part of data is undoubtedly the most valuable. It is a very common and important scenario that the data can be generated and the alarm can be generated according to certain rules. The more timely the alarm is, the better for the business.

4. Labels of multiple dimensions exist in data, which often requires multi-dimensional joint query and statistical query. Another very important function of time series data is multi-dimensional aggregated statistical query. For example, the business needs to calculate the click-through rate and total revenue of advertisements published by Google in USA in the last hour, which is a typical multi-dimensional aggregated statistical query demand. This requirement usually has a low requirement for effectiveness, but a high requirement for query aggregation performance.

These are challenges for traditional database and that is why we need a time-series database.The invention of time-series database product is to solve the deficiency and defect of traditional relational database in timing data storage and analysis. They optimize write, store, and query processes for the characteristics of sequential data, which are closely related to the characteristics of sequential data.A time-series database lets you store large volumes of timestamped data in a format that allows fast insertion and fast retrieval to support complex analysis on that data.In recent years, the popularity of time-series database continues to rise.Unlike traditional databases in which older data entries are typically updated with the most recent data point to show the latest state, time-series databases (“TSDB”) continuously accumulate data points over time. This way, one can draw insights from the evolution of metrics over time to conclude meaningful insights from the data. Better yet is to predict future events through machine learning models based on historical time-series analysis.

This is how time-series database works:A time-series database captures a set of fixed values and dynamic values. For example, in a web application, a set of data points might be actions performed by a group of visitors. The dynamic values are the number of desired actions, such as eCommerce purchases, performed by the visitors over time. By analyzing purchases over time, the organization can understand the value of each group of users, and prioritize marketing activity for different [customer segments](https://formation.ai/blog/the-problems-with-customer-segmentation/).  Ideally, time series records should be written into a repository in a format that enables quick time-based writes and reads. Because the records are time-stamped, the order of data points becomes a native aspect of the data. You can then use this order to deliver the data to a stream processing engine, which can then treat the ordered data like a data stream. By leveraging a fast stream processing engine you can ensure your time-series database is fast.

The time-series database has the following major advantages:

1.More accurate and meaningful time series measurement:A time-series database makes it easy to measure how datasets change over time. You can concurrently view past, present, and future datasets for reporting that is more accurate and meaningful.

2. Resource-efficient data storage:By the very nature of the data type, processing it can require massive amounts of storage, which can be difficult to manage. It's also very expensive. Time-series databases have tooling to aggregate data into predetermined time periods and to eliminate any data streams as needed. There are also compression algorithms that optimize data storage.

3. Lightning-fast data queries:A time-series database can also make it easy to query and retrieve data based on specific periods. Imagine that you can't remember the title of a book you recently read, but you know it was three months ago. Time series databases can help you figure out what the book was without having to use a bunch of wildcard searches.

Using a time series database, we can quickly find information based on timeframe, which is important in some cases.There are some examples:

1.Monitoring sensor data :One of the use cases is the monitoring of sensor data for safety measurements, predictive maintenance, or assistance functions. E.g. a car stores and uses all kinds of sensor data like tyre pressure, surrounding temperature and humidity for driver assistance and maintenance support. [An aircraft](https://everydaymatters.rpi.edu/too-many-airplane-systems-rely-on-too-few-sensors/)monitors gravity and aerodynamic principles to reassure pilots that everything is alright – or to alert them that something has gone wrong. In fact, a Boeing creates on average half a terabyte of data per flight, most of which is time-series data.

### 2.[Tracking assets](https://objectbox.io/time-series-and-objects-using-data-on-the-edge/):Tracking assets is ideal for a time-series database as you constantly want to monitor where assets are, e.g. the cars of a fleet or any goods you might be stocking or shipping. These applications typically include unique vehicle or asset IDs, GPS coordinates, and additional metadata per timestamp. Apart from keeping track of the assets in realtime, you also can use the data for logistics and optimize e.g. your stocking and delivery processes.

### 3.[Tracking assets](https://objectbox.io/time-series-and-objects-using-data-on-the-edge/):Tracking assets is ideal for a time-series database as you constantly want to monitor where assets are, e.g. the cars of a fleet or any goods you might be stocking or shipping. These applications typically include unique vehicle or asset IDs, GPS coordinates, and additional metadata per timestamp. Apart from keeping track of the assets in realtime, you also can use the data for logistics and optimize e.g. your stocking and delivery processes.

### With the development of the Internet of Things and many emerging industries, the processing of temporal data is becoming more and more important. As a result, we need better, faster, and more accurate time-series database systems to meet higher business requirements. Business needs will lead to technological advances. Therefore, I believe that the time-series database space will become more popular in the future, and more high-quality time-series databases will be created.