**Why do we need a time-series database?**

Neal A. Maxwell once said, "Time is clearly not our natural dimension. Thus it is that we are never really at home in time. Alternately, we find ourselves wishing to hasten the passage of time or to hold back the dawn. We can do neither, of course, but whereas the fish is at home in water, we are clearly not at home in time--because we belong to eternity." For me, time is a very ambiguous thing. Because we know that we can hardly feel the existence of time as opposed to something as concrete and palpable as color, length or even speed. I think it is due to such a characteristic of time that many literary works since ancient times have been related to time, or even have time as a theme. The invisibility of time, the effect of its passage on our emotions is a natural vehicle for literary works. In addition, time has become closely related to each of us, from the most intuitive alternation of day and night, to the clocks and watches around us that are used for timekeeping, large and small. I think no one denies that time is a very important thing in this world, and because of its importance, we all include time in our research process when we conduct scientific studies.

## What is a time series database?

Time series database is a database which is completely optimized for timestamp or time series data. And time series is a series of data points ordered by time. These data points may be a sequence of events from an application user, the number of sunspots, changes in weather over a period of time, or data from sensors in the car that emit data about the acceleration and speed of the vehicle. Unlike traditional databases that typically update old data entries with the latest data points to show the latest status, time series databases accumulate data points over time. In this way, one can derive insights from the evolution of metrics over time and thus derive meaningful insights from the data. Even better is to predict future events through machine learning models based on historical time series analysis.

Time-series databases have been around for a short time compared to relational databases (RDBMS) or NoSQL, but they can be more effective in solving many of today's popular research problems. For example, in recent years, the very popular auto self-driving, Internet of Things(IoT), smart home, these application scenarios are very emphasis on time or "real-time", in these applications we urgently need a form of data that can measure the change of transactions over time, where time is not just a metric, but a coordinate of the primary axis. This is time series data, which is gradually taking on a larger role in our world and could revolutionize the way we live our daily lives in the near future.

There are at least two reasons why TSDBs are the fastest growing database category today: scalability and usability.

Scalability: Time-series data accumulates very quickly, and ordinary databases are designed to handle that scale (at least not in an automated way). Relational databases perform poorly on very large datasets, while NoSQL databases are better at scale. In contrast, temporal databases (whether relational, or nosql-based) introduce benefits that are only possible if you consider time as the first consideration. These benefits allow them to provide massive performance improvements, including higher throughput and faster large-scale queries, as well as better data compression.

Usability: TSDBs also typically include built-in functions and operations commonly used for time-series data analysis, such as data retention policies, continuous queries, flexible temporal aggregation, and more. Even if you're just starting to collect this type of data and don't need to think about scale just yet, these functions can still provide a better user experience and make data analysis tasks easier. Using the built-in functions and features to analyze readily available trends in the data layer often reveals unexpected value, whether your data set is large or small.

## When to use time series database now?

Today, anything that can be a component can be further broken down into smaller scale components. To further improve human life, we make the objects around us more "intelligent" to make them more convenient for us to use. So we are amazed at all the things around us that could be intelligent - cars, factories, power grids, satellites, phones, refrigerators, streets. All of these things already have or will have sensors, and with the data they sense, they can make smarter feedback. Thus, these "non-intelligent" everyday objects that used to be all around us will continue to release streams of event or time-series data in the future. How to deal with such a huge amount of data becomes a challenge.

This means that the underlying infrastructure needs to be further upgraded to support these new massive workloads - more data points, more data sources, more monitoring, more control. What we are about to see, and what is needed for these new areas, is a paradigm shift in how we handle data infrastructure and how we handle building, monitoring, controlling and managing systems. In other words, what we need is a high-performance, scalable, dedicated time series database.

Time series databases are not a very new term, but the first generation of time series databases in the past focused primarily on the study of financial data, the volatility of stock trades, and systems built to solve trades. That means the application of this technology was still very limited at that time. Today, however, financial data is no longer the only application for time series data - in fact, it is just one of many applications across a wide range of industries. The basic conditions for computing have changed dramatically in the last decade. Everything has become compartmentalized. The single mainframe no longer exists, replaced by serverless servers, containers.

In our daily life, time series data also has typical application scenarios. For example, some features on Zhihu, a Q&A community in China, are characterized by time series data. The typical interaction pattern of this website is that a questioner posts a question, and other users can post their thoughts or answers under such a relatively independent question space. In such an application scenario, the question posted by the questioner automatically generates a log of the questions asked or modified, and the bottom left corner of each answer indicates when the current answer was created and when the answer was edited. Such an unobtrusive feature actually contributes a lot to providing a good online community atmosphere. It provides a dynamic approach that allows users to have a more comprehensive understanding of how a topic is perceived at different times. Moreover, with the addition of time as a metric, we can easily distinguish which views are highly forward-looking. So time-series data is not something completely new that suddenly appeared, but has been used in our lives for a long time.

## Advantages of Time-Series Database

Based on the characteristics of time-series data, relational database cannot meet the effective storage and processing of time-series data, so there is an urgent need for a database system that is specifically optimized for time-series data, i.e., time-series database.

The storage and processing of time-series big data are often handled by relational databases, but the inherent disadvantages of relational databases make it impossible to store and query data efficiently. By using a special storage method, the temporal big data solution makes it possible to efficiently store and quickly process massive temporal big data, which is an important technology for solving massive data processing. The technology uses a special data storage method, which greatly improves the processing capability of time-related data, halves the storage space and greatly increases the query speed compared to relational databases. The superior query performance of time series functions far exceeds that of relational databases, and Informix Time Series is well suited for IoT analytics applications.

In addition to this, the time series database has many advantages. For example, more accurate as well as more meaningful time series measurements. As mentioned earlier, we know that time series databases make it easy to measure how datasets change over time. This means that users can look at past, present and future datasets at the same time in order to get more accurate and meaningful reports from them.

Resource efficient data storage is also one of its advantages. Due to the nature of the data type, processing it can require a lot of storage, which can be difficult to manage. And we also find it very expensive. Time series databases have tools to aggregate data into predetermined time periods and eliminate any data streams as needed. There are also compression algorithms to optimize data storage.

Finally, it is very fast to query data. TSDB also makes it easy to query and retrieve data based on a specific time period. Imagine you don't remember information about a particular product you recently browsed, but you know it was seen in a certain month. Then the Time Series Database can help users find that information quickly and without the need for extensive wildcard searches. In other words, with a time series database, users are able to quickly find information based on a time range. This is a very important feature for browsers and shopping sites to improve the user experience.