# An Introduction about TDengine

Guangqi Jiang 2020141450013 LaTeX

July 2022

**Abstract**

TDengine is a domestic time-series database, which is developed, managed and maintained by Taos Data, a high-tech company based in Peking. It is designed mainly for the big data derived from the Internet of Things(IoT) . To be precise, it is good at processing huge amount of data with time tags. In this article, we will discuss the database system in detail. We are going to spell out its characteristics and study some industrial cases based on TDengine.

**Key words:** TDengine, time-series database, big data, industrial monitoring system

## 1 Introduction

TDengine is a database system with fully independent property rights.It supports SQL statements and has high performance which is written in C language. In addition, it is highly reliable, distributed and thus scalable. The team of TDengine developers lays great emphasis on the user-friendly features. As a consequence, TDengine is easy to learn with zero management and convenient installation.

To conclude, TDengine is a very successful and outstanding big data tool in China. It provides caching, data subscription, streaming computing and other functions in order to minimize the complexity of Research and Development(R&D) as well as Operating and Management(O&M).

We learned in the class that, time-series data has its own characteristics, including but not limited to that data needs to be written more than read, and that data is rarely deleted or updated, etc. TDengine is intentionally designed to fit these characteristics. Moreover, it adopts an innovative storage engine design to significantly improve performance. As can be seen from its official website, the experimental test results are much better than the existing similar products such as influxDB.TDengine adopts a distributed development model, with partitioning, sharding, and virtual nodes, makes scaling very easy.

Many other general-purpose time-series databases have their own languages, such as influxDB and Prometheus. As a comparision, TDengine supports SQL, which greatly reduces the learning costs. Also, TDengine designs a table for each data point. As Mr. Tao said in our course, such a modeling approach leads to the generation of a huge number of tables. Therefore, TDengine carries another super table, which makes the management and operation of individual tables much easier. The commands for creating a super table and creating a normal table are as follows:

```
CREATE STABLE meters (
ts timestamp,
current float,
voltage int,
phase float) TAGS (location binary(64), groupId int);

CREATE TABLE d1001 USING meters TAGS ("California.SanFrancisco", 2);
```

# 2 Outstanding features of TDengine

TDengine has many excellent performance characteristics. We will select some of the most typical and impressive ones to explain in detail.

## 2.1 In terms of development technology and software functionality

### 2.1.1 All in one

TDengine has consistently been developed to reduce user complexity. Accordingly, with built-in caching, stream calculation, data subscription and other powerful functions, system complexity as well as operation and maintenance difficulties are lowered to a minimal level. It is no exagerration to say that our TDengine is almost equivalent to an integrated big data platform with full functions. Besides, TDengine is easy to install, which means it does not require any dependencies, and can be used through the command line familiar to users.

### 2.1.2 Good statistical analysis of data visualization

We have just mentioned that TDengine can be accessed through cmmand line. Nevertheless, it does more than that. In addition to querying through the Shell terminal, TDengine can also be used to query and analyze using tools such as Grafana, Matlab and Python. These visualization tools can be used to quickly build visual monitoring platforms. As was also introduced in class, our instructor, Mr. Tao, highlighted and sang high praise fot Grafana, an analysis tool. All the tools above are integrated seamlessly and can even be use with just one line of code. This shows that TDengine has very powerful support for data processing.

## 2.2 From the point of view of the user and personal use

### 2.2.1 Detailed official documentation

The official website and GitHub have complete descriptions, from building to installing and developing. There is easy access to documentation from version 1.6 to version 2.6 on the official website. From this semester's class: Introduction to Software Engineering, we learned about the lifecycle of a large software project. During the whole life of a software project, building and coding are only part and parcel of it. The remaining biggest cost of software development lies in the maintenance to follow. For that matter, it can be safely concluded that TDengine does a great job with detailed documentation. It is always in the process of being updated and iterated. Since the official commercial release in August 2018, TDengine has made the kernel and community version fully open-source in July 2019. Afterwards, TAOS released version 2.0 in August 2020, and again in 2021, version 2.2 came up. As we can see, the development of TDengine is a model of an evolving and iterative process.

### 2.2.2 Desirable community ecology

Mr. Tao introduced us to the concept of open source during our very first class. It has been a popular trend for the past years. TDengine fully follows the concept of open source. There are 169 repositories on GitHub when we search for the keyword "TDengine", including the open source code of TDengine itself, some extension applications of TDengine developed by other contributors and projects made by some developers using TDengine. There are also posts on Zhihu, a domestic platform for exchanging knowledge, discussing TDengine. One of the issues heated debated is titled with "Talking about the time-series database TDengine" which was published in 2019. Under this question, a respondent tested TDengine and gave very detailed test results and personal suggestions. To our surprise, there are also replies and interactions from netizens including Mr. Jeff Tao himself in the comment section. In a nutshell, considering the enthusiam and passion of Mr. Tao, we deeply see the growth of TDengine in our market.

# 3 Industrial cases based on TDengine

One of the major differences between technology and science is that the former one needs to be put into practical use. TDengine has been used extensively in industrial area. We have obtained examples of projects and studies using TDengine from the Internet. The main applications are in the industrial Internet of Things, monitoring and inspection systems, big data analytics, and other research areas that are currently popular in industry.

### 3.1 Case 1. Research on data storage method of smart gridmonitoring system based on TDengine

NARI Group has done some research on the data storage method of smart grid monitoring system based on TDengine. They came up with a solution to the problem that the general time series database cannot model the grid model. Their analysis provides the interface and performance of high-density data storage and query. What's more, they provide accurate historical data for various applications. As a result, their new technology meets the requirements of various applications for real-time data storage, large volume and security as well as providing powerful data support for smart grid safety operation monitoring and accident analysis.

### 3.2 Case2.Design of CPS Real-Time Monitoring System Based on Time Series Database

A National Natural Science Foundation of China (61873068) led by some scholars from Guangdong University of Technology also chose us, TDengine, to build a database system. They designed a real-time monitoring system and built a monitoring cloud platform using Grafana, an open source visualization tool. They meet the needs of industrial enterprises for equipment monitoring data visualization and alarm pushing. As a result, thhey improve operation and maintenance efficiency. Meanwhile, the experiment also proved the convenience of TDengine deployment and superior performance.

## 4 Conclusion

To conclude, TDengine is an excellent time-series database system. It bears a lot of innovative features and attained hororable achievements. We are glad to see that TDengine receives many stars in the GitHub and are becoming more and more popular among industry employemnt as well as scientific reserach. After the classes taught by Mr. Tao, we gain more insights into time-series database and his successful work, TDengine.

## 5 References

[1].DONG Xue, GAO Yuan, AO Bing. Research on data storage method of smart grid monitoring system based on TDengine [J]. Electrotechnical Application,2021,40(08):68-74.
[2].TDengine Documentation and its White Paper TAOS DATA https://docs.taosdata.com/
[3]. Zhong Zihan, Zhang Lichen. Design of CPS Real-Time Monitoring System Based on Time Series Database [J]. Modern Computer,2021,27(25):102-106.