2019 IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services (SmartCNS)

Survey of Time Series Data Processing in Industrial Internet

Miao-qiong WANG
China Academy of Information and Communication Technology
Cloud Computing and Big Data Research Institute
Beijing, China
wangmiaoqiong@caict.ac.cn

Kai WEI
China Academy of Information and Communication Technology
Cloud Computing and Big Data Research Institute
Beijing, China
weikai@caict.ac.cn

Chun-yu JIANG
China Academy of Information and Communication Technology
Cloud Computing and Big Data Research Institute
Beijing, China
jiangchunyu1@caict.ac.cn

***Abstract*—This paper focuses on the processing requirements of time series data in industrial and IoT fields. The study of time series data processing in industry is continued for a long time, and a mature solution of using real-time/ historian database has been formed. However, with the new demand of Industrial Internet, old architectures are unable to fully support requirements (i.e., large amount and real-time analysis of industrial data). Meanwhile, a new architecture for processing real-time data in mobile Internet started to mature, this forms a solution called time-series database, which provides lots of new advantages. When we try to use a new technology to replace an old one, many aspects should be considered. This paper focuses on these challenges, starting with the demands of the industry, to analyze how to solve traditional problems with new technologies. This paper also analyzes the development trend of time series data processing and puts forward some general requirements for the application of new technology in the field of Industrial Internet, which lays a theoretical foundation for the application and development of basic technologies of Industrial Internet.**

Keywords*—*Time Series Data; Time Series Database; Real-time/ historian Database; Industrial Internet

I. INTRODUCTION

The rapid development of the Internet has led to a trend of technological innovation. A lot of new technologies has been applied to the industry to solve practical needs. In industrial fields, concepts like “Industrial Internet”, “Intelligent Manufacturing” begin to emerge. People are actively exploring the application of new technologies in industrial fields. Whether all the new technologies are suitable for promoting? When a new technology is used well in one industry, it may not be suitable for other industries, especially in the industrial field, which has more complex environment. Industrial software system has extremely high requirements, especially for the abilities like real-time, stability and security. Industrial systems were relatively close for many years, so they were formed to be mature and independent[1-3].

In industrial fields, more than 80 percent of the monitoring data are real-time data, and all of them are time-series data with timestamps[4]. These data from sensors or monitoring systems are collected in real time and used for quick feedback of system status. In traditional industry, real-time/ historian databases are often used as the core solution to collect, store, query and analyze these data.

However, outside the field of industry, with the emerging of new concepts such as mobile Internet, Internet of things, Internet of vehicles, smart grid and so on, more requirements for real-time data processing are putting forward. Another solution with a new architecture is gradually formed. It is called time series database, and it is designed to meet the need of monitoring and analyzing massive real-time data from the Internet. This new database is kind of similar with the real-time/ historian database. Compared with the traditional industrial solution, the Internet solution has better scalability. In addition, it’s natural integration with big data ecosystem will undoubtedly challenge the original technical framework[5-10].

This paper focuses on what new challenges will occur when processing real-time data in industrial systems in the age of Industrial Internet, and what are the differences between the mature technologies and the emerging technologies. Besides, this paper also researches how to seize the trend of technology to meet the requirements of Industrial Internet.

II. REAL-TIME DATA PROCESSING IN TRADITIONAL INDUSTRY

In traditional industrial control field, there are a lot of real-time data processing requirements, especially in the Process Industry. The monitoring requirements are stringent in production. Real-time monitoring data will reflect the status of system, therefore, the processing of real-time data is very important. After a long time of accumulation, a unique and mature solution has been formed. The application of real-time/ historian database is an important part of it, which has been used for many years. In the field of industrial control, real-time/ historian database is mainly used to collect, storage, query and analyze industrial process data, and realize the real-time monitoring of process status[11]. Data in industry have these characteristics: 1) most of industrial data have timestamps and are generated in sequence; 2) most of industrial data are structured data; 3) the frequency of data collection is high, and the amount is large; 4) the characteristics of a time period are more important than of a single time point.

The requirements of software in industry are always very rigorous, so the real-time/ historian database is polished as practical, precise, stable, close, and with high-performance. Take a medium-sized industrial enterprise as an example. When monitoring process, it may have 50,000 to 100,000 measuring points. The amount of data produced per day can reach hundreds of GB. Data in industrial enterprises are required to be stored over a long time, so that historical trends can be queried at any time. These simple requirements have demonstrated some of the capabilities that traditional real-time/ historian databases needs to have, such as:

Project supported by National Key R&D Program of China (2016YFB1000601)

978-1-7281-5209-7/19/$31.00 ©2019 IEEE
DOI 10.1109/IUCC/DSCI/SmartCNS.2019.00151

1) High writing performance (because we have great amount of measurement points and high frequency of data collection)

2) Efficient data compression capacity (because it requires long-time storage)

3) Quick response of query (because people are not willing to wait minutes for a single search)

4) Real-time analysis ability (because the status need to be quickly reflected)

Table I shows the requirements of a real-time/ historian database in a normal industrial project:

TABLE I. REQUIREMENTS OF TRADITIONAL REAL-TIME/ HISTORIAN DATABASE

Requirements	Value
Capacity of monitoring points in a single machine	1 million
Scalability	Support
Throughput	1 million TPS
Precision	1 millisecond
Minimum data refresh cycle	<100 millisecond
Average compression ratio	30:1
Data collection rate	100%
Backup and recovery	On-line backup and recovery
User-defined online computing	Support
Scripts	VBScript, C#, Lua
Computation library	Support

Characteristics of traditional industrial real-time /historian database can be summarized as follows:

1) High writing performance: Industrial real-time/ historian databases often require high writing speed. Take the process industry as an example, sensors will be placed in each step with pretty high collection frequency, so the concurrent volume of writing will be extremely high, sometimes will reach millions of measurement points per second. Therefore, except for the wellness of software, it also requires expansive and high-performance hardware to ensure the performance.

2) Quick response of query: On the one hand, quick response to query is required to ensure that the system status can be monitored in real time; on the other hand, historical data stored also need to be quickly searched. The amount of historical data is always very large. When drawing a trend for a very long period, it takes time to aggregate these data first. No one is willing to wait for a plot for a very long time, so the delay of the query must be very small, even if one-year period data is being queried, the result should be quickly reflected.

3) Extreme data compression ability: The data compression requirements are particularly high, because in industry data will be stored for a very long time, like 5 years or even 10 years. In the case of limited storage, data need to be compressed in lossless compression or lossy compression mode. Compression ratio of lossy compression will be higher than lossless compression, sometimes up to 30:1. Lossy compression will use additional algorithms to retain the details of data after compression.

4) Amounts of tools accumulated: Traditional industrial solutions often have rich toolkit for various scenarios, such as hundreds of protocols and different data models, which are important and competitive in industrial fields.

5) Perfect stability: Stability requirements of industrial software are particularly high, in addition to use Active-Standby solution to ensure high availability, the extremely high quality of software is another way to ensure the continuous operation of the program. A program running for ten years without errors is normal in industrial fields.

III. NEW PROBLEMS IN THE AGE OF INDUSTRIAL INTERNET

With the gradual maturity of Internet, Internet of things, cloud computing, big data, artificial intelligence and other new generation of technologies, a new round of technological revolution is emerging. A series of new modes of production, organization, and business models began to emerge, as well as the concept of Industrial Internet. Industrial Internet is a deep integration of the new generation of information technology and traditional industrial systems. It is the key infrastructure to the development of industrial intelligence. People started to use new technologies into design, production, management, services and other aspects of industrial production[12]. Whether to explore individually or learn from other industries, it is inevitable that new and old technologies need to be connected.

Take real-time data analysis as an example, new applications are emerging with the growth of Internet of things technology[13]. Increasing sensors, soaring amount of data, as well as higher big data analysis demand start to challenge the original technical architecture.

The old architecture faces a series of challenges:

Scalability bottlenecks: Although the traditional technical architecture can ensure that a single machine has a very high performance, it cannot realize dynamic and flexible scaling like distributed systems, the scaling process needs to be planned with in advance. When the system needs to be expanded for business upgrade, the lack of scalability of the old architecture will be the bottleneck.

Connection with big data ecosystem: The ultimate purpose of data collection is making data understood and used. Big data industry has already have a very mature solution for mass data storage and analysis. The connection requirements are imperative whether it is Hadoop or Spark ecosystem. Many industrial enterprises have to upgrade or replace their existing database systems since they want to use the new big data analysis technology[14].

High cost: Traditional industrial solutions are expensive. The cost usually includes machines, software, and related operation and maintenance services. In the process of upgrading, due to its long history and outdated structure, you may spend extra large amount of money to reconstruct the software system. On the one hand, the support service will be insufficient for many years has passed. On the other hand, it will cost a lot of manpower and money to reconstruct the system. Companies will naturally start looking for new, cheaper and more effective solutions to replace them. In addition, with the development of the Industrial Internet, more and more small and medium enterprises are aware of the importance of data, but due to limited funds, they will also tend to look for cheaper solutions.

IV. TIME SERIES DATABASE IN INTERNET INDUSTRY

Development History

Due to the rapid growth of data in the Internet industry, a database with new architecture called Time Series Database has been formed. This database solution and the traditional real-time/ historian database are like twins in different times.

After entering the era of Internet, with the innovation of communications technology and the decline of communication costs, another trend of Internet of everything starts. Not only the computer system needs to collect data, mobile phones, smart devices, shared bikes and cars that people use every day are constantly sending real-time data to the cloud. These data will be analyzed with big data technologies to monitor and forecast the business, and help enterprises reduce costs, as well as serve the public[15-16].

These data share some of the same characteristics as most of the real-time data in industrial fields:

1) The length of a single data is not very large, but the amount of data is very large;

2) They are all time-stamped, and they're generated in sequence;

3) Most of the data are structured and are used to describe the characteristics of a parameter at a certain time point;

4) The writing frequency is much higher than the query frequency;

5) There are very few requirements to update data;

6) Users are more interested about the characteristics of data over a time-period than of a single time point;

7) Most of the queries are based on a certain time-period or a certain numerical range;

8) Need calculation and visualization.

Data from smart meters, environmental monitoring equipment, and industrial production lines also have these characteristics.

However, due to the difference of application scenarios, industrial solutions may differ from Internet solutions to a certain extent, which can be seen in Table II.

TABLE II. DIFFERENCES BETWEEN TWO DATABASE SOLUTIONS

Items	Real-time/ historian database	Time series database
Growth environment	Industrial enterprises, mostly enterprise or group-level applications	Internet enterprises, mainly based on cloud platforms
Deployment	Active-Standby mode	Distributed
Functional requirements	Read and write data, aggregate query, data compression	Read and write data, aggregate query, data compression
Performance requirements	Extreme processing speed with single machine	High throughput and performance scaling with cluster
Charge mode	One-time License fee, high unit price	Pay as much as use
Software ecology	Integrated toolkit	Combine with other independent services
Main advantages	Single machine performance and compatibility with industrial systems	Cloud platform and architecture advantages
Development tendency	Distributed and cloud platform	Gradually infiltrate into industrial field

The time series database with new architecture is the same, to a certain extent, as the traditional real-time/ historian database when processing these Internet data. They have the same functional requirements realized in different areas. When the new Internet technology permeates into industry, it will reflect certain strengths and weaknesses. The development of Industrial Internet technology requires the mutual penetration and integration of technologies from two sides. So that they can absorb advantages and compensate disadvantages from each other.

Technology Trends

With the development of Industrial Internet, demands are becoming more and more clear. When these two database technologies bump into each other, we can observe some trends of technology development. We conclude them as the following 6 points:

1) Transition to distributed architecture: Traditional real-time databases mostly use Active-standby architectures, usually requiring expensive machines with higher hardware configuration to achieve extreme performance of a single machine; at the same time, it requires extreme stability of the running software. The quality of the software ensures error-free running for many years; It will also require ultra-high data compression ratio because of limited storage. But with the development of distributed technology, the system can be easily expanded, so that the database is no longer dependent on expensive hardware and storage devices. It can achieve high availability with the natural advantages of clusters, and single point failure will never occur. It can be run on a normal x86 server or even on a virtual machine. Distributed architecture will greatly reduce the cost of use[17].

2) Diversified data structure: In industry, the traditional real-time/ historian database often uses single-value model. A parameter under monitoring is called a measuring point. A model will be built for each measuring point when writing data to a database. For example, an index like temperature of a wind turbine can be calculated as a measuring point, ten indexes of ten wind turbines are 100 measuring points. Each measurement point has some information (like name, precision, data type, switching/ analog value and so on). The writing efficiency of single-value model is very high. Fig 1 shows the structure of a single-value model.

Timestamp	Value	Tagname	Description
1537948697000	23.4	qd.1.speed	Speed/No.1 Wind Turbine in Tsingdao
1537948697000	38.1	qd.1.temp	Temp/No.1 Wind Turbine in Tsingdao
1537948697000	23.4	bj.1.speed	Speed/No.1 Wind Turbine in Beijing
1537948697000	38.12	bj.1.temp	Temp/No.1 Wind Turbine in Beijing

TagList

Fig. 1. Structure of a single-value model

Time series database in the Internet fields usually use multi-value model, which is similar with the object-oriented model. For example, we create a model called wind turbine. Its parameters include temperature, pressure, as well as latitude and longitude, ID and other tag information. This will make it more appropriate for analysis when providing services. Technically, the single-value model and the multi-value model can be converted to each other. Many databases provide services with multi-value model, but the underlying storage is still single-value model. Fig 2 shows the structure of a multi-value model.

	Field		Tag	
Timestamp	Speed	Temp	No	Organization
1537948697000	23.4	38.1	1	Tsingdao
1467627245001	23.4	38.2	1	Tsingdao
1467627245000	23.4	38.3	2	Beijing
1467627245001	23.4	38.3	2	Beijing

Metric

Fig. 2. Structure of a multi-value model

3) SQL support: Most of the time series databases choose NoSQL storage which has better scalability[18]. Compared with the relational database, the data model of NoSQL is more flexible, which is very suitable for the multi-value model, because it is easier to be extended. It is easy to scale out the cluster when resources are limited or when performance needs to be improved. The query efficiency is high, and the cost of open source software is quite low. Most time series databases use various types of NoSQL models, and Table III shows some examples:

TABLE III. NoSQL MODEL USED IN TIME SERIES DATABASES

Products	Database model
InfluxDB	Key-value
OpenTSDB	Wide column store
Graphite	Key-value/ wide column store
KairosDB	Wide column store
Prometheus	Key-value

However, the use of NoSQL models can result in the loss of some original features, such as transactions. Therefore, we need to include other methods to ensure data consistency. Support for SQL is also missing. SQL has been used as a standard query method in industry for many years, and its learning cost is relatively low. Time series database companies are trying to integrate SQL engine in NoSQL models, so that their products can be accessed with SQL operations to reduce the barriers.

4) Diversified query model: In the Internet age, requirements for queries are not only with criteria or interpolations. With the development of Internet of things and people's needs to fully control the information, more and more map-based applications occur. Queries will gradually expand from dimension of time to dimension of space. Higher requirements for visualization will also take place. In addition to the guarantee of real-time analysis, it is necessary to present the whole status of a system dynamically in a more vivid way.

5) Transition to cloud service: In traditional industry, solutions to process real-time data are deployed in private environment. The cost includes machines, software, as well as service fee, which is a very high expense. Sometimes it also requires professional technicians to maintain the system. As services move to cloud, there is no need to purchase machines, no need to hire engineers to maintain machines and software systems. One just need to think about how to develop and maintain the business. In addition, you can only buy resources that you need. Traditional One-time purchase of services will cause waste of resources or lack of resources for secondary construction. Cloud service can reduce large amounts of expenditure. With the maturity of network and cloud computing technology, the related performance and security are also constantly upgrading. Cloud service will eventually be the same as localized service, and this will become an irresistible trend[19-20].

6) Edge computing: The industrial field is an important experimental area of Internet of things, the development of Industrial Internet will bring more sensors and more data collection. When data is too large, the centralized processing method will be difficult to respond in time, which brings the data computing to the edge side. Data with real-time monitoring requirements will be processed and generate feedback through the edge devices in time. The data needed for large-scale analysis are transmitted to centralized storage. This hierarchical processing can effectively enhance the value of time-sensitive data, while reducing the burden of storage systems. So many time series databases have edge computing versions, combined with the ability of stream computing to make the functionality more diverse[21-23].

Fig 3 shows the Industrial Internet real-time processing solution with edge computing.

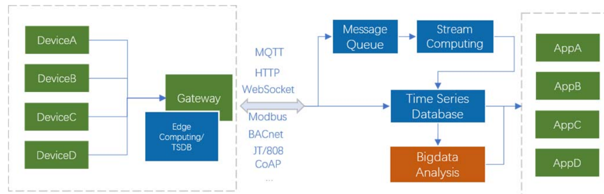

Fig. 3. Real-time processing solution with edge computing

V. SURVEY ON PRODUCTS

In 2018, China Academy of Information and Communications Technology (CAICT) organized several companies who have time-series database (TSDB) products to investigate the development status of TSDB markets in China. This work group set an industry standard about technical requirements of TSDB including 33 items which are needed in common business scenarios.

Based on this standard, Data Center Alliance (DCA) tested 7 TSDB products of different companies. Here we replaced their names with letters.

Specific contents of the test results are listed in Table IV. Optional requirements are marked with *.

Among them, Company A, C, D, F are large cloud service providers, providing cloud version of TSDB products; Company B, E, G mainly provides local version of products. Company B is a startup company, Company E is a senior company providing solutions in industrial fields, while Company G provides common big data solutions for enterprise users.

Technical Architecture: Only 2 of the 7 products use architecture that is 100% self-developed. The rest of them are mostly developed based on OpenTSDB (an opensource TSDB product), with one based on Elasticsearch (an opensource search engine). Although their architectures are different, all products could cover more than 75% of the technical requirements.

Functional Requirements: Although these products could reach most functional requirements of the standard, their ways of realization are different. For example, when considering the function "Dynamically add time series", mature products would encapsulate a direct interface for users to add new time series. When adding new time series, products using NoSQL architecture will have better scalability and better user experience than relational architecture since the latter one needs to change its schema every time. Carefully designed products would have functions to monitor the status of time series， while others need users to do extra aggregate operations to realize the function. The same as "Query latest data", more professional products have direct interface to acquire the latest data, beginners ask users to write conditional queries to get data.

Company E is special among these companies due to its industrial background. Its product has very obvious industry features and most mature functions. 95% of its operations can be done with visual interface. Other products need IDEs to set the operations. Company E is also the only one that has label management function, while others need to check the data structure to get the meta data.

From overall situation, 7 companies have their own emphasis on the design of product architecture and functions. Although their functions have different degrees of maturity and ease of use, all products can cover most of the technical requirements.

TABLE IV. FUNCTIONAL SATISFACTION OF DIFFERENT DATABASES

Requirements	A	B	C	D	E	F	G
Data type	●	●	●	●	●	●	●
Data accuracy	●	●	●	●	●	●	●
Write time-series data	●	●	●	●	●	●	●
Dynamically add time series	●	●	●	●	●	●	●
Query time series	●	●	●	●	●	●	●
Query tags	●	●	●	●	●	●	●
Interpolation queries	●	●	●	●	●	●	●
Query value*	●	●	●	●	●	●	○
Query latest data	●	●	●	●	●	●	●
Aggregation queries	●	●	●	●	●	●	●
User-defined functions*	●	○	○	●	●	○	○
Geographic position queries*	●	○	●	●	○	○	○
Data lifetime management	●	●	●	●	●	●	●
Compatibility with mainstream hardware and operating systems	●	●	●	●	●	●	●
Deploy with containers*	○	○	●	●	○	○	○
Connection with big data ecosystem*	●	○	●	●	●	○	●
Easy deployment	●	●	●	●	●	●	●
Configuration management	●	●	●	●	●	●	●
Real-time monitoring	●	●	●	●	●	●	●
User management	●	●	●	●	●	●	●
Online update*	●	○	●	○	●	○	●
Meta data management	●	●	●	●	●	●	●
Import and export data	●	●	●	●	●	●	●
Hardware fault tolerance	●	●	●	●	●	●	●
Operating system fault tolerance	●	●	●	●	●	●	●
Database service fault tolerance	●	●	●	●	●	●	●
Overload protection	●	●	●	●	●	●	●
Multiple replicas	●	●	●	●	●	●	●
Online scale up or scale out	●	●	●	●	●	●	●
Online scale down*	●	○	●	○	●	○	●
Authentication	●	●	●	●	●	●	●
Operation audit	●	●	●	●	●	●	●
Encryption communication	●	●	●	●	●	○	●

VI. CONCLUSION

In this paper, we analyze the challenges of processing time series data in industry from several aspects, describe the traditional processing methods in industrial field, and the new processing architecture after the development of the Internet. In order to solve the problems emerging in the industry, old method needs to take advantages form new one and they need to complement with each other, which is also the essence of Industrial Internet. The following part of the paper describes the technology trend in industry. Through the research on demand and supply side, we list some technical requirements for processing industrial real-time data in the new era. This work is of great significance for further development of time series database and real-time processing in industry. Our team will continue study the performance requirements of time series database to evaluate products which are suitable for industrial scenarios.

ACKNOWLEDGEMENT

This paper's relevant project is supported by National Key R&D Program of China (2016YFB1000601). In addition, Special Thanks to Jeff TAO, Jie-ying HU, Miao HUANG, Yu ZHONG, Le-qiang AI.

REFERENCES

[1] Industrial internet platform white paper. Alliance of Industrial Internet, 2019.

[2] Technical architecture of industrial big data white paper. Alliance of Industrial Internet, 2018.

[3] Industial big data technology and application white paper. Alliance of Industrial Internet, 2017.

[4] Survey of industrial enterprise data asset management status. Alliance of Industrial Internet, 2018.

[5] S. K. Jensen, T. B. Pedersen and C. Thomsen, "Time Series Management Systems: A Survey," in IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 11, pp. 2581-2600, 1 Nov. 2017.

[6] Garima and S. Rani, "Review on time series databases and recent research trends in Time Series Mining," 2014 5th International Conference - Confluence The Next Generation Information Technology Summit (Confluence), Noida, 2014, pp. 109-115.

[7] L. Xu, et al. "Telecom big data based user offloading self-optimisation in heterogeneous relay cellular systems," International Journal of Distributed Systems and Technologies, 8(2), pp. 27-46, April 2017.

[8] L. Xu, et al. "Data mining and evaluation for Base Station deployment," in Proc. ICSINC, 13-15 Sept 2017, Chongqing China, pp. 356-364.

[9] D. Han and E. Stroulia, "A three-dimensional data model in HBase for large time-series dataset analysis," 2012 IEEE 6th International Workshop on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems (MESOCA), Trnto, 2012, pp. 47-56.

[10] D. Ramesh, A. Sinha and S. Singh, "Data modelling for discrete time series data using Cassandra and MongoDB," 2016 3rd International Conference on Recent Advances in Information Technology (RAIT), Dhanbad, 2016, pp. 598-601.

[11] Mi Yeong Hwang, Cheng Hao Jin, Yang Koo Lee, Kwang Deuk Kim, Jung Hoon Shin and Keun Ho Ryu, "Prediction of wind power generation and power ramp rate with time series analysis," 2011 3rd International Conference on Awareness Science and Technology (iCAST), Dalian, 2011, pp. 512-515.

[12] D. Ramesh and A. Kumar, "Query Driven implementation of Twitter base using Cassandra," 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, 2018, pp. 1-4.

[13] H. Talei, M. Essaaidi and D. Benhaddou, "An End to End Real Time Architecture for Analyzing and Clustering Time Series Data: Case of an Energy Management System," 2018 6th International Renewable and Sustainable Energy Conference (IRSEC), Rabat, Morocco, 2018, pp. 1-7.

[14] M. Tahmassebpour, "A New Method for Time-Series Big Data Effective Storage," in IEEE Access, vol. 5, pp. 10694-10699, 2017.

[15] Lei Yan, Xin Ai, Yao Wang and Huiying Zhang, "Impacts of electric vehicles on power grid considering time series of TOU," 2014 IEEE Conference and Expo Transportation Electrification Asia-Pacific (ITEC Asia-Pacific), Beijing, 2014, pp. 1-5.

[16] S. Aljawarneh, V. Radhakrishna, P. V. Kumar and V. Janaki, "A similarity measure for temporal pattern discovery in time series data

generated by IoT," 2016 International Conference on Engineering & MIS (ICEMIS), Agadir, 2016, pp. 1-4.

[17] Pelkonen T , Franklin S , Teller J , et al., "Gorilla: A Fast, Scalable, In-Memory Time Series Database," Proceedings of the Vldb Endowment, 2015, 8(12):1816-1827.

[18] A. K. Kalakanti, V. Sudhakaran, V. Raveendran and N. Menon, "A comprehensive evaluation of NoSQL datastores in the context of historians and sensor data analysis," 2015 IEEE International Conference on Big Data (Big Data), Santa Clara, CA, 2015, pp. 1797-1806.

[19] H. S. Raju and S. Shenoy, "Real-time remote monitoring and operation of industrial devices using IoT and cloud," 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I), Noida, 2016, pp. 324-329.

[20] E. Elazab, T. Awad, H. Elgamal and B. Elsouhily, "A cloud based condition monitoring system for industrial machinery with application to power plants," 2017 Nineteenth International Middle East Power Systems Conference (MEPCON), Cairo, 2017, pp. 1400-1405.

[21] E. Oyekanlu and K. Scoles, "Real-Time Distributed Computing at Network Edges for Large Scale Industrial IoT Networks," 2018 IEEE World Congress on Services (SERVICES), San Francisco, CA, 2018, pp. 63-64.

[22] P. Porambage, J. Okwuibe, M. Liyanage, M. Ylianttila and T. Taleb, "Survey on Multi-Access Edge Computing for Internet of Things Realization," in IEEE Communications Surveys & Tutorials, vol. 20, no. 4, pp. 2961-2991, Fourthquarter 2018.

[23] E. Oyekanlu, "Predictive edge computing for time series of industrial IoT and large scale critical infrastructure based on open-source software analytic of big data," 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, 2017, pp. 1663-1669.