

Development of AI-equipped Machine Automation Controller (2)

Data collection synchronized with machine control and realization of Time-Series database

OTA Masanori and NISHIYAMA Yoshihide

In recent years, in the manufacturing industry, the need to utilize AI (Artificial Intelligence) and IoT (Internet of Things) at the manufacturing site continues to rise in order to improve availability of facilities and stably produce high quality products.

Conventionally, in order to utilize the data of the manufacturing site, data of each device such as a sensor and a motor was accumulated in a dedicated database server via a network and analyzed. However, as controllers controlling the apparatus become faster and more accurate, it is becoming necessary to gather data necessary for more accurate data analysis, but in the conventional system, it is impossible to store all the data in the server It was.

Therefore, in developing a machine automation controller equipped with the AI function, we developed a data collection mechanism synchronized with the control cycle and a time series database that accumulates the collected data in the controller. As a result, we made it possible to create highly accurate AI machine learning model, realized controller based on model based AI function and abnormality detection of facility.

1. Introduction

The use of information technology (IT) on the production line is an increasing trend these days, and an approach utilizing the data collected from the production line is accelerated. For example, the manufacturing and quality data stored in the database and analyzed can be used for management of the capacity utilization and for preventive maintenance^{1,2)}. Consideration is also made to introduce AI technology, which is remarkably advanced these days, into factory automation, especially into production management and facility maintenance, and demonstration of the failure detection technique of the production equipment utilizing the data obtained from the control system of the production facility was reported^{3,4)}. The use of AI technology will make analysis of a large volume of data possible, which is impossible by human power, but in order to improve the quality of the databased on an equipment maintenance service, data acquisition at a frequency sufficient to detect the behavior of the operating production equipment is necessary. Considering the above, the authors have established a data collection mechanism synchronized to the control cycle of the controller controlling the equipment and the data collection technique in the development of the AI-equipped machine automation controller.

Contact : *OTA Masanori* masanori.ota@omron.com

2. Challenges

2.1 Conventional data acquisition method

The machine automation controller used on the production line is the control device that provides sequence control according to the prescribed sequence and motion control controlling position of the motor. The controller collects input and output data (manufacturing data) from diversified control devices composing the production line, such as sensors and motors, at an interval of 0.125 to a few milliseconds via the industrial network, typically EtherCAT (field network). These manufacturing data are designated as variables, and the results of the operation of these variables by the control program embedded in the controller can be substituted into another variable. The controller also provides an interfacing function between the production line and the production management host system using the advanced information technology of the TCP/IP based network.

In manufacturing industries, accumulation of production information and quality data in a database has been made using the controller as an interface for the Manufacturing Execution System (MES) that performs the functions of monitoring the manufacturing process and providing work instructions to the operator for the purpose of a cause analysis of the product defect or equipment failure and for recording of the production history of the process. The method frequently used is to install

the data collection software connected via network between the controller and the server and to store the data collected by the software in the database (Fig. 1a). The controller that allows direct storage of the data in a database connected by the network is recently available (Fig. 1b)¹⁾.

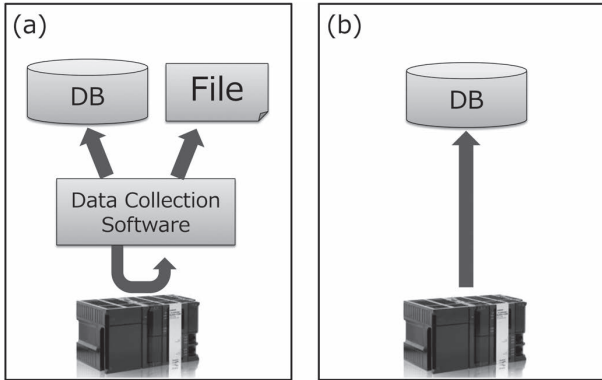


Fig. 1 Conventional Data Collection Method

2.2 Issues involved in high-speed and high-accuracy data acquisition

A data collection cycle in any conventional method of data collection is on the order of a few milliseconds to several hundreds of milliseconds because the data are acquired using the network. Fluctuation of the data collection cycle is unavoidable due to the load on the network, data acquisition software, and the database server.

The controller, on the other hand, can operate at high speed and at a high accuracy control cycle at the minimum 0.125 ms. Fig. 2 shows the graphs of the analog values representing the sine waveform input to the controller (Fig. 2a) and the input data of the analog values acquired at 10 ms intervals (Fig. 2b).

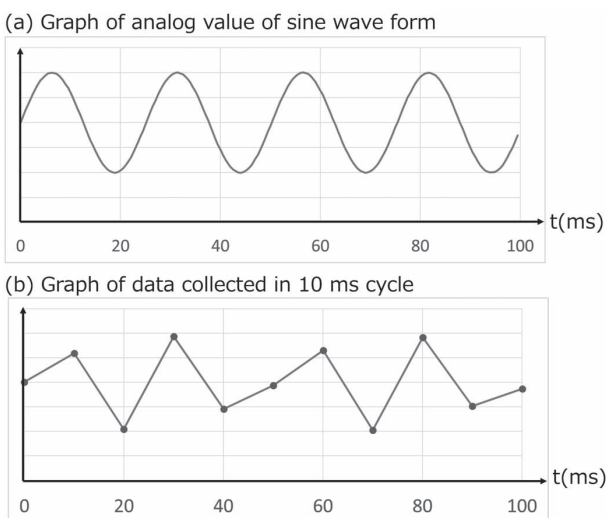


Fig. 2 Graph of Data Collected by Conventional Method in Sine Waveform

The authors consider that the high-speed and high-accuracy data collection synchronized with the control cycle of the controller is necessary to realize high-accuracy data analysis utilizing AI techniques, but the conventional data acquisition method can provide an acquisition cycle of about 10 ms, and it is impossible to analyze when the original data are in sine waveform. Accordingly, the authors developed the Time Series Database where the data are collected, stored, and synchronized with the control cycle of the machine controller in the development of the AI-equipped machine automation controller that has the machine learning and feature extraction functions using AI.

3. Details of the Technique

3.1 Time-series database

In order to resolve the issue in Section 2, the time-series database (DB) was constructed in the machine controller in order to acquire the manufacturing data produced in time-series synchronized with the machine controller. Generally, the time-series DB is a database system specialized in the handling of the data produced at a constant cycle (time-series data) and suitable for handling of large volumes of data with a simple data structure using a timestamp as an index.

Fig. 3 shows the system configuration of the time-series DB of the AI-equipped machine automation controller. The machine controller is connected to the sensors and motors via the IO units or the field network. The control program is a program executing sequence control and motion control of the equipment. The data collection process collects data from the sensors and motors and stores the data collected in the time-series DB. The time-series DB exports the stored data as a CSV file. Data in the CSV file are analyzed by the host PC.

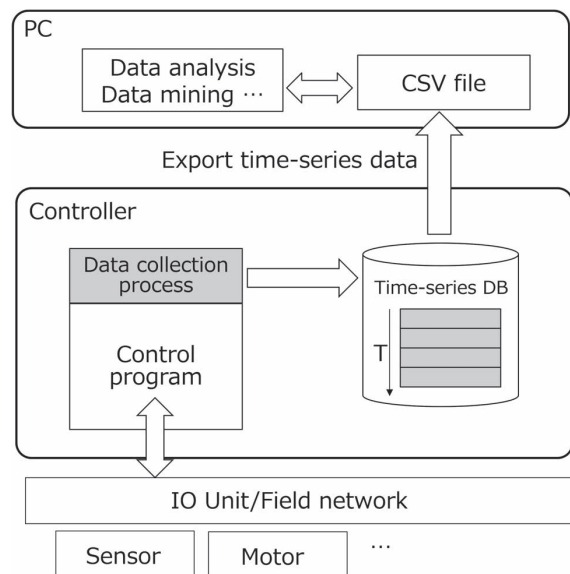


Fig. 3 System Configuration of the Time-Series DB

The authors established the following design guidelines in the development of the database incorporated in the AI-equipped automation controller.

<Design guideline>

- 1) The variable values must be acquired and synchronized with the control task for data acquisition and processing
- 2) The time when the data are acquired and the variable values must be precisely linked.
- 3) The data must be recorded without affecting the execution time of the control program.
Data acquisition and processing must not be blocked by recording and processing of the data into the time-series DB.
- 4) The mechanism to acquire the data recorded in the time series DB must be provided.

3.2 Data structure of the time-series DB

The relational database has long been the mainstream of the database system. The relational database is generally the database having the data structure with the following features:

- The data are stored in 2-D table format.
- A number of tables can be combined.
- Select, add, update, and delete are possible by SQL, the generic query language.

The data volume processed by the web-based system increased, and high-speed processing becomes necessary. For such a system, a strict definition of the table as in a relational database and a query combining complex conditions are not required. Accordingly, the database having a data structure other than that of a relational database is proposed. Such databases other than the relational database are collectively called NoSQL (Not Only SQL)⁵⁾.

The NoSQL type database can be classified into a few types based on the difference in the data structure; for example, the key-value pair type where the key and the value are stored as a pair, the document type specializing in storage of the document data, and the column type where the data are processed by grouping in a column, and the graph type where the data are stored in a graphic structure. In the developed system, the key-value pair type suitable for high-speed data storage and acquisition is used that are enabled by a simple data structure consisting of the key and the value in order to store large volumes of data produced in a time-series linked to time.

The data structure of the time-series DB is a pair of the key data unique in the database and the value storing the manufacturing data, and the key value pair constitutes the record. The data are stored in the database by the unit of record in the time-series DB. The key has the index information

containing the time. The value is a set of the variable values to be acquired. The variable value consists of the information showing the data type of the variable and the binary column showing the variable value (see Fig. 4).

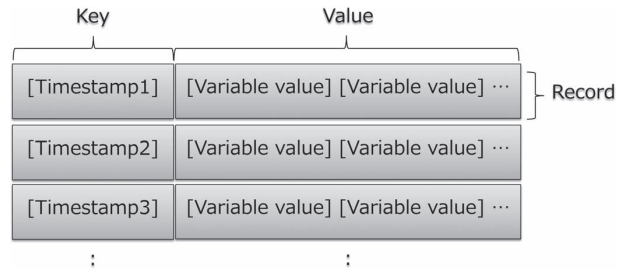


Fig. 4 Data Structure of the Time-Series DB

By the time-series DB, as shown in Fig. 4, the time and the manufacturing data can be precisely linked by time-series management of the time stamp and the variable value.

3.3 Data acquisition synchronized with the machine control

The OMRON machine automation controller executes the control program at a constant cycle of 0.125 ms. The control cycle is executed by the minimum cycle of 0.125 ms. Data communication with the sensors and motors connected to the field network are executed and synchronized with the control cycle. The task-scheduling model used by the OMRON machine automation controller is shown in Fig. 5.

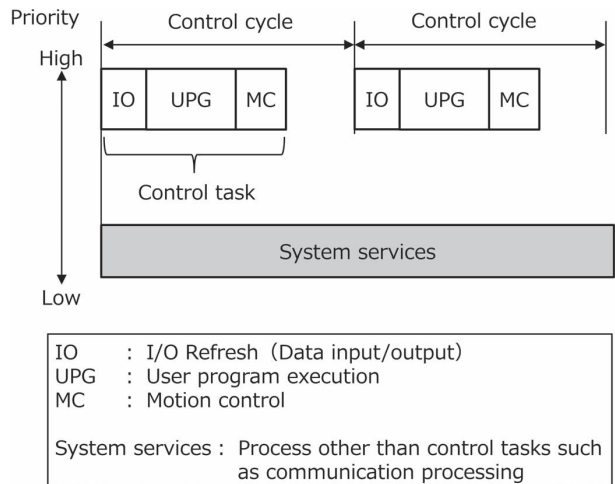


Fig. 5 Task Scheduling Model

The control tasks execute processing of I/O refresh (data input/output), user program execution, and motion control according to the strict control cycle. By the system service, communication processing with the programming devices, programmable display devices, and host PC, access to the SD memory card other than the control tasks is executed. The

OMRON machine automation controller has the task priority management feature and task scheduling feature and executes the control task with higher priority than the system service to realize high accuracy control cycle without fluctuation. In this system, data acquisition fully synchronized to the control cycle is realized by execution of data acquisition in the control task (see Fig. 6).

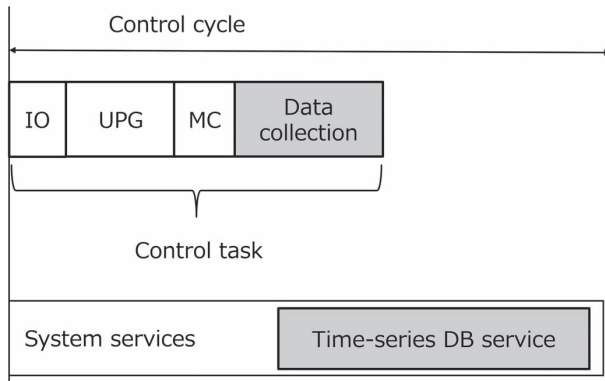


Fig. 6 Task Scheduling Model of the Time-series DB

The data collection and processing create a record stored in the time-series DB of the variable values that are acquired at every control cycle with the acquisition time according to the list of registered variables to be acquired. As the control tasks must strictly follow the control cycle, and the execution time of the control task must not exceed the control cycle, the processing whose processing time cannot be guaranteed such as writing in the storage media cannot be executed in the control task. So the processing to write the record created by the data collection and processing in the time-series DB is separated to the time-series DB service that is executed in the system service to enable the data acquisition and storage in the time-series DB strictly synchronized with the control cycle.

3.4 Writing of the data in the time-series DB

The SD memory card, SSD (solid state drive), and the RAM disk can be used as the storage media of the controller for the time-series DB. Writing data to the storage media may take a very long time depending on the type of media used. For example, latency in writing data to the SD memory card varies widely, and it may take from hundreds of milliseconds to a few seconds. The time required for writing data of the unit volume to the storage media is generally quicker when a large volume of data is handled in one execution. Accordingly, the non-blocking queue (NBQ) that absorbs the difference and the fluctuation of the processing time for writing the data from the control program to the storage media is realized in the time-series DB by writing a number of records cached in the built-in

buffer memory in one execution to the storage media (see Fig. 7).

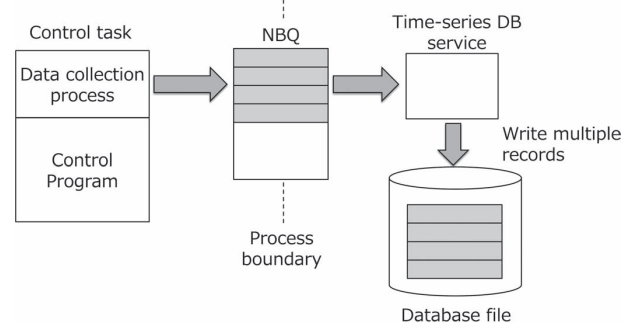


Fig. 7 Writing of the data in the Storage

The following features in writing to the time-series DB are realized based on the techniques explained earlier.

1) To store the record in the NBQ by each control cycle

The record is stored in the NBQ by each control cycle for data collection and processing. It is configured to minimize extension of the execution time of the control task by the operation to store the record without blocking.

2) To place the NBQ in the shared memory

It is configured to enable delivery of the data to the time-series DB that works at lower priority from the control task that works at higher priority and to execute data exchange across the process boundary in one operation by placing the NBQ in shared memory.

3) Writing of number of records in one execution

The time-series DB service fetches the records cached in the NBQ. The records fetched are bundled in the writing buffer and are written in the database file in one execution. Because of this scheme, the time required to save unit data volume is reduced, and data overflow of the NBQ is unlikely to occur. While writing is in progress, the record is retained in the NBQ to absorb fluctuations in the latency in writing to the storage media.

4) Parallel writing to multiple time-series DBs.

It is configured to have multiple NBQs to enable parallel data writing to multiple time-series DBs with different performance characteristics. It is also configured to allow an adjustable number of records and record sizes stored in the NBQ according to the latency of the storage media and the number of records stored in the unit time. Such a configuration makes writing in a short cycle to a small volume and high-speed storage media like a RAM disk possible while writing is processed to the large volume and low speed media like an SD memory card using two NBQs.

3.5 Exporting the time-series data

The time series DB has a function to export the information of the record to utilize the manufacturing data stored in the time-series DB. Various formats used in exporting are available depending on the users. For the AI-equipped machine automation controller, a method to output the text file in CSV format that is highly readable and versatile is used considering easy analysis of the collected data. Fig. 8 is an example of the CSV file exported from the time-series data shown in Microsoft Excel.

	A	B	C	D	E	F	G	H	I	J	K	L
1	indx	tmstmp	tmstmp2	schcont	schcont2	aid	aidx	MC_Axis00	MC_Axis00	MC_Axis00	MC_Axis00	MC_Axis00
2	12735891	12.4516.745	1.53742E+18	6.98E+12	0	0	0	21435	21435	21435	21435	21435
3	12735892	12.4516.745	1.53742E+18	6.98E+12	0	0	0	21408	21408	21408	21408	21408
4	12735893	12.4516.746	1.53742E+18	6.98E+12	0	0	0	21381	21381	21381	21381	21381
5	12735894	12.4516.746	1.53742E+18	6.98E+12	0	0	0	21355	21355	21355	21355	21355
6	12735895	12.4516.747	1.53742E+18	6.98E+12	0	0	0	21327	21327	21327	21327	21327
7	12735896	12.4516.747	1.53742E+18	6.98E+12	0	0	0	21300	21300	21300	21300	21300
8	12735897	12.4516.748	1.53742E+18	6.98E+12	0	0	0	21273	21273	21273	21273	21273
9	12735898	12.4516.748	1.53742E+18	6.98E+12	0	0	0	21246	21246	21246	21246	21246
10	12735899	12.4516.749	1.53742E+18	6.98E+12	0	0	0	21218	21218	21218	21218	21218
11	12735900	12.4516.749	1.53742E+18	6.98E+12	0	0	0	21191	21191	21191	21191	21191
12	12735901	12.4516.750	1.53742E+18	6.98E+12	0	0	0	21163	21163	21163	21163	21163
13	12735902	12.4516.750	1.53742E+18	6.98E+12	0	0	0	21135	21135	21135	21135	21135
14	12735903	12.4516.751	1.53742E+18	6.98E+12	0	0	0	21107	21107	21107	21107	21107
15	12735904	12.4516.751	1.53742E+18	6.98E+12	0	0	0	21080	21080	21080	21080	21080
16	12735905	12.4516.752	1.53742E+18	6.98E+12	0	0	0	21051	21051	21051	21051	21051
17	12735906	12.4516.752	1.53742E+18	6.98E+12	0	0	0	21023	21023	21023	21023	21023
18	12735907	12.4516.753	1.53742E+18	6.98E+12	0	0	0	20995	20995	20995	20995	20995
19	12735908	12.4516.753	1.53742E+18	6.98E+12	0	0	0	20967	20967	20967	20967	20967
20	12735909	12.4516.754	1.53742E+18	6.98E+12	0	0	0	20939	20939	20939	20939	20939

Fig. 8 Example of the CSV File

The export function is extensible in order to handle formats other than CSV files. For example, it is possible to have an export method to output the record to the data collection system operating on the host computer by socket communication.

4. Utilization of the Time-series DB

4.1 Utilization in AI-equipped Automation Controller

There are principally two activities required to realize predictive maintenance by the AI-equipped automation controller.

- 1) Data collection and analysis

To collect the manufacturing data of the production equipment in question and to construct the AI machine learning model consisting of the feature values and the learned data reflecting a failure. In this paper, this activity is called the analysis phase.

- 2) Detection of outlier

The outlier is detected in real time from the manufacturing databased on the AI machine-learning model obtained in 1) above. In this paper, this activity is called the utilization phase.

Utilization of the time-series DB in each phase is explained in the subsequent subsections.

4.2 Utilization of the time-series DB in analysis phase

The analysis phase is an activity to collect the data of the production equipment in question and to construct the AI machine-learning model based on the collected data for analysis. In the AI-equipped automation controller, the data for analysis are collected from the time-series DB and based on

such data, the analysis tool AI Easy Modeler selects the feature values used for anomaly detection, learns the outlier detection model, and determines the threshold value used to discriminate normal and abnormal conditions, and then constructs the AI machine-learning model. The flow diagram of the processing in the analysis phase is shown in Fig. 9.

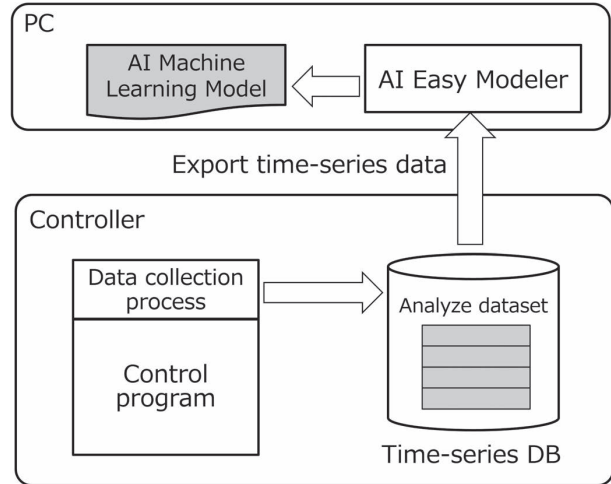


Fig. 9 Processing Flow Diagram of the Analysis Phase

So that the high accuracy AI machine-learning model can be constructed by the AI Easy Modeler, collection of the highly accurate analysis data is critical. Fig. 10 is a graph of the analysis data of the sensor data in sine waveform collected at the 0.5 ms cycle synchronized with the control cycle. The measurement condition is as follows:

- Control cycle: 0.5 ms
- Variables to be collected: LREAL data (double precision floating point number type)
- Storage media: SD memory card

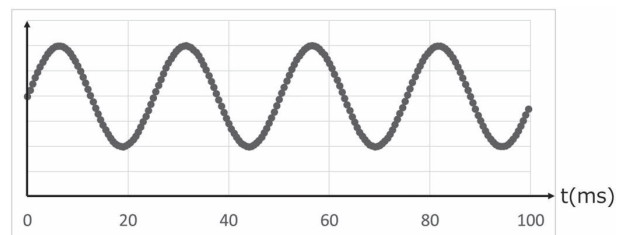


Fig. 10 Graph of Data collected at 0.5 ms Cycle

As explained earlier in Subsection 2.2 “Issues involved in high-speed and high-accuracy data acquisition,” the data collection cycle by the conventional method expands from a few milliseconds to hundreds of milliseconds, and it is impossible to analyze the original waveform as a sine wave, but the data that are collected, synchronized with the control task, and stored in the time-series DB can be correctly analyzed when the sensor

data are in sine waveform. As explained above, the AI-equipped automation controller can construct the AI machine-learning model with high accuracy by collecting data synchronized with the control task using the time-series DB.

4.3 Utilization of the time-series DB in utilization phase

The utilization phase is an activity to determine anomalies based on the AI machine-learning model constructed by the AI Easy Modeler while the equipment is operating. The processing flow of the utilization phase is as follows (see Fig. 11):

- 1) Feature extraction function: To operate the feature value that is used in anomaly determination selected in the analysis phase.
- 2) Outlier detection function: To determine anomalies based on the feature value operated in 1) and the AI machine-learning model.
- 3) AI Viewer visualization tool: To visualize the feature value in 1) and determination the result in 2).

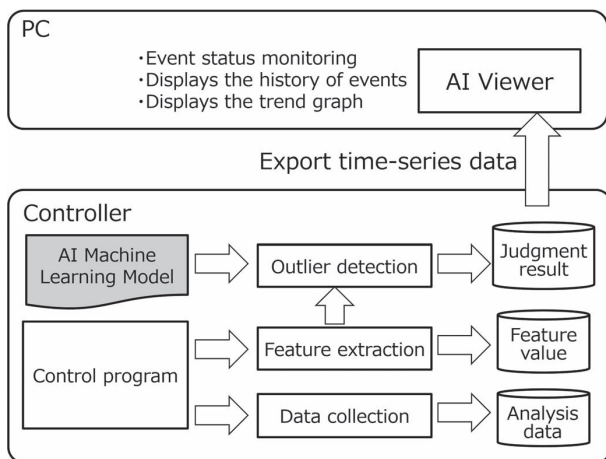


Fig. 11 Processing Flow Diagram of the Utilization Phase

In the utilization phase, three types of data—analysis data, operation results of feature values, and determination results of outlier detection—are produced at different times and frequencies. The analysis data are created at a minimum of a 0.125 ms cycle synchronized with the control, and the feature values and the results of determination are produced at about hundreds of milliseconds to a few seconds cycle synchronized with the production tact. The AI-equipped automation controller stores three types of data—analysis data, feature values, and results of the determination—in different time-series DBs. The user can know that an anomaly occurred from the result of the determination visually indicated on the AI viewer screen and can analyze the feature values and the analysis data to identify the factors affecting the anomaly.

4.4 Utilization in applications other than AI-equipped Automation Controller

The manufacturing data from the production line contain the temperature, vibration, motor speed, and position that are constantly generated. The time-series DB technique that can collect such data without interruption synchronized with the control cycle can be used for various applications. One of such applications is edge computing.

Huge costs and large communication zones are required to store large volumes of data that are constantly generated on the production line into the cloud data server. The data from the production line are generally not significant while production is normal and are not actually utilized. For such applications, a concept of edge computing is becoming popular⁶⁾.

Fig. 12 shows an example where the time-series DB is utilized in the edge-computing layer. The cloud-computing layer is a system operating on the internet, and the fog computing layer is a system operating on the factory local network that provides the services like MES and data analysis. The controller in which the time-series DB is operating is placed in the edge-computing layer. The manufacturing data are stored into the time-series DB in every control cycle. The time-series DB stores the data during a certain period of time, and the old data are destroyed. An application will be possible, where storage of the data before and after the failure of the equipment on the server in the fog computing layer, only when there is a failure of the equipment.

CLOUD Computing Layer

- Big-data
- Business logic
- Internet

FOG Computing Layer

- MES
- Data analysis
- Local network

EDGE Computing Layer

- Machine Controller
- Sensor/Motor device
- Field network

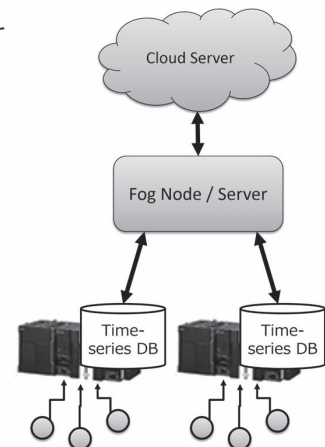


Fig. 12 Utilization of Time-series DB in the Edge Computing

Table 1 Estimation of Data Volume Collected

Condition of Estimation	Data Volume
Data volume when entire data are stored	About 82 GB per day
Data volume when the data before and after occurrence of a failure are stored	About 59 MB per one minute

Table 1 shows the estimation of the data volume when 128 items of the manufacturing data of double precision floating point number type are collected during a 1 ms cycle. When the

entire data are stored on the host server, the data volumes generated are approximately 82 GB per day. The data volumes can be reduced to approximately 59 MB by storing the data in the time-series DB restricted to before and after a failure, which can be handled by a common network.

5. Conclusion

5.1 Discussion

The time-series DB that collects the manufacturing data generated in a time series synchronized with the control cycle is realized as the technique to collect and utilize the data from the production equipment at high speed with high accuracy, which is incorporated into the AI-equipped automation controller. The use of the time-series DB makes data collection synchronized with the control cycle possible, which is difficult by the conventional method via the network, and construction of the highly accurate AI machine-learning model becomes possible. It also becomes possible to use the data for maintenance of the production equipment by collecting and visualizing the results of feature value operations and outlier determination based on the AI machine-learning model.

5.2 Development towards the future

The technique of the time-series DB that collects and accumulates the data from the production line synchronized with the control cycle can be utilized for various applications as represented in edge computing. The authors intend to expand application of the technique using the time-series DB as the high-speed database engine that can be incorporated into the device like the controller.

References

- 1) OMRON Corporation. "Equipment data for big-data analysis" (in Japanese). <https://www.fa.omron.co.jp/solution/sysmac/innovation/bigdata.html>, (accessed 2018-10-11).
- 2) OMRON Corporation. "Improvement in manufacturing floor utilizing big data (in Japanese)". <https://www.fa.omron.co.jp/solution/sysmac/technology/bigdata/index.html>, (accessed 2018-10-11).
- 3) Tsuruta, K.; Minemoto, T.; Hirohashi, Y. Development of AI Technology for Machine Automation Controller (1), Anomaly Detection Method for Manufacturing Equipment by Utilizing Machine Data. OMRON TECHNICS. 2018, Vol.50, No.1, p.6-11 (in Japanese).
- 4) Abe, Y.; Ueyama, Y.; Sakatani, N.; Fujii, T. Development of AI Technology for Machine Automation Controller (2), Demonstration Example in the Development of Anomaly Detection Controller Prototype. OMRON TECHNICS, 2018, Vol.50, No.1 p.12-17 (in Japanese).
- 5) Tsuyuzaki, K.; Onizuka, M. NoSQL Database Characteristics and

Benchmark System. NTT Technical Review. 2012, Vol.24, No.10, p.22-25. <http://www.ntt.co.jp/journal/1210/files/jn201210022.pdf>, (accessed 2018-10-26).

- 6) Kaifu, M. Status of Edge Computing, Next Big Thing after Cloud. KDDI Research, R&A, 2018 April. <https://rp.kddi-research.jp/article/RA2018003>, (accessed 2018-11-01).

About the Author

OTA Masanori

Controller Development Dept. 1, Controller Div.
Product Business Division H.Q.
Industrial Automation Company
Speciality: Software Engineering

NISHIYAMA Yoshihide

Controller Development Dept. 1, Controller Div.
Product Business Division H.Q.
Industrial Automation Company
Speciality: Software Engineering

Microsoft and Excel are registered trademarks and product names of Microsoft Corporation in the United States and other countries. The names of products in the text may be trademarks of each company.