Design of Big Data Control System for Electrical Automation

Lin-ze Gao(✉)

Guilin University of Electronic Technology, Guilin 536000, China
gaolz20@126.com

Abstract. The traditional big data control system is limited by hardware and software resources, which leads to low efficiency of data access and calculation. Therefore, a big data control system for electrical automation is designed. The system is mainly designed from two aspects of hardware and software. The hardware mainly designs a digital acquisition circuit, selects the model of each part in the overall working module of the acquisition board, and designs a pseudo dual-port RAM to store data, which is convenient for the processing of massive data and Control; in software design, according to the needs of the system, light MySQL and distributed HBase are used to jointly design the ER diagram and database table of the database user database to make the database performance more optimized. In the control algorithm of the system, the butterfly operation is used to optimized calculate the speed, obtain the operation error through the operation and correct it to ensure the efficiency of the system operation. The experimental results show that the performance of the designed system is better than the traditional system in data insertion speed, access speed and calculation speed, which fully verifies the application value of the system.

Keywords: Electrical automation · Big data · Database · Butterfly algorithm · Digital acquisition circuit

1 Introduction

The invention of electric energy is known as one of the greatest achievements since the 18th century. Its wide application has set off the second climax of industrialization, and made the human society march into the electrical age with its head held high. Since its development, electric energy has entered every aspect of people's life, which is not only related to the quality of people's life, but also of great significance to national economic development, national defense security and social stability [1]. Many countries in the world have used their degree of application as an important indicator for judging the national development level. Industrial production is inseparable from electricity, and many fields in the industry have put forward the requirements of high-speed real-time computing. Although traditional industrial field data collection systems have good reliability, accuracy and compatibility, they are large and costly high. Although the cost of the system based on single chip microcomputer is low, there are many devices, the system is complex, the development time is relatively long, and the

S. Liu and L. Xia (Eds.): ADHIP 2020, LNICST 347, pp. 66–76, 2021.
https://doi.org/10.1007/978-3-030-67871-5_7

reliability of the system is poor [2, 3], which has a certain impact on the processing efficiency of massive data.

In view of the above problems, a big data control system for electrical automation is designed. The hardware of the system mainly includes digital acquisition circuit and database, which is convenient for processing massive data. On the basis of hardware design, the system software is mainly designed. The user block E-R diagram and database table of database are designed by using lightweight MySQL and distributed HBase. Butterfly operation is used to improve the calculation speed of the system, and the calculation error is corrected. The experimental results show that the designed system has the advantages of data insertion speed, access speed and calculation speed, and has high utilization value.

2 Hardware Design of Electric Automation Big Data Control System

2.1 Design Digital Acquisition Circuit

The data acquisition board adopts PC/104 bus standard, 16-bit data transmission controlled by I/O commands, and has program trigger and external trigger functions. The external analog signal is converted into a digital signal by an AD converter. The conversion accuracy is 120 digital signals, and then the number of digits is expanded to 16 bits, which is sent to the FPGA for FFT conversion [4, 5]. The control signals required for the entire collection and subsequent data transmission are completed by the FPGA. The FPGA interface control module is mainly composed of address decoding, address generation, and FFT processing unit. The address decoding part translates the corresponding port through the 10 bit address from the bus, and generates the corresponding control signal; the address generation part is mainly responsible for generating the address of on-chip dual port RAM and off chip SRAM. The designed dual port RAM adopts ping-pong time sharing mode and has two sets of independent data, address and control line. Two data ports are read-only and write only (also known as pseudo dual port RAM).

The A/D conversion model is AD 1672. A belongs to a monolithic analog-to-digital converter (ADC). The chip contains 4 high-performance sample-and-hold amplifiers (SHA), 4 flashing ADCs and a voltage reference [6]. The output is equipped with an error correction logic circuit to ensure 12-bit accuracy at a sampling speed of 3MSPS and no missing codes over the entire operating temperature range. The sampling voltage is −2.5 V to +2.5 V. The data latch 74HC374 is an 8-bit tristate buffer. At the rising edge of the clock, the data is sent from the d end to the Q end to complete the cache. The schematic diagram of the overall working module of the acquisition board is shown in Fig. 1:

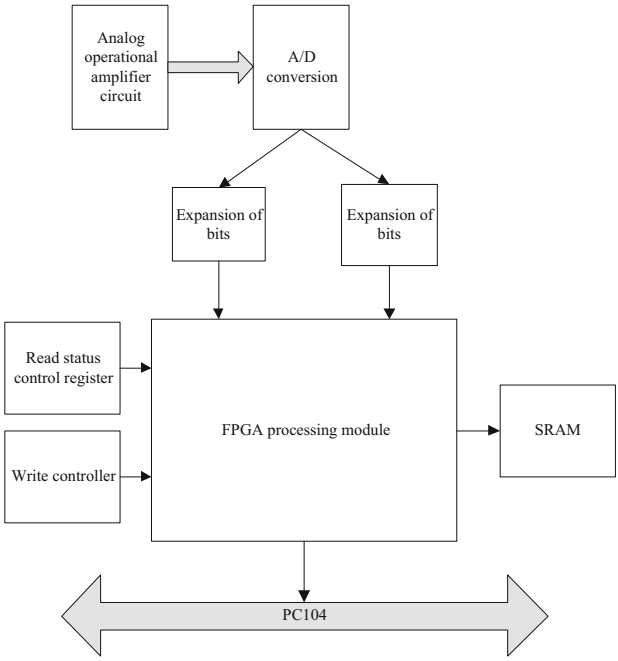

Fig. 1. Schematic diagram of the overall working module of the acquisition board

In the figure above, the rising edge of a clock AD 1672 completes the transfer of a data. In this design, AD 1672 and 74HC374 share a clock INCLK to maintain synchronization, INCLK is provided by the address generation unit inside the FPGA. In the design, two 74HC374s are used for bit number expansion, and the 12-bit signed number after A/D sampling is expanded to 16 bits. The control register model is 74HC373. Two circuit control registers are used in the circuit design. One is used to record the status signal on the acquisition board for query, and one is used to write the control word to control the working state of the entire acquisition board [7, 8].

The data memory model is IS61C1024, which is an 8-bit 128 KB high-speed CMOS static RAM with a minimum storage time of 12 ns. It can meet the requirements of high-speed data reading and writing, and is also suitable for the temporary storage of large capacity data. This design uses four pieces of IS61C1024 to cache FFT processing results. The bidirectional data bus transceiver model is HC245HC245, which is an eight bidirectional bus transceiver used for bidirectional asynchronous communication between data buses. The output allows the control terminal/G to be effective when the level is low, and when the level is high, the two ends are high impedance. In the circuit, MegaWizard designed a 2048×32 bits dual-port RAM to store the intermediate data of FFT operation.

The internal storage unit is divided into two parts. When the first part is used to provide operation data to the current level, the second part is used to store the operation results of the previous level. When the 1024 point operation of the current level is completed, the functions of the two parts are exchanged.

2.2 Design Database

Because the system database not only needs to store massive unstructured power data, but also involves the association operation of structured data such as user information, job information, system block information, etc. Therefore, plans to use lightweight MySQL database to process structured data and relational computing, and distributed HBase database to access massive unstructured data. MySQL is a small relational database management system, which has the characteristics of small size, high speed and low cost. User information, system block information and Hadoop job information in the system will involve query methods such as condition and association [9], so MySQL database is selected for storage. Among them, the E-R diagram of user board is shown in the Fig. 2, and there is a many to many relationship between them.

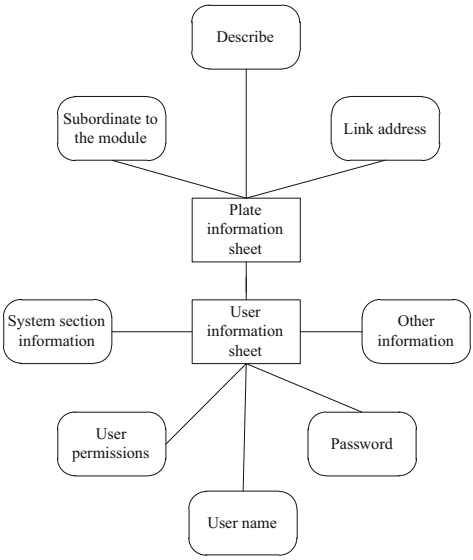

Fig. 2. E-R diagram of user sector

In the system designed in this paper, the data table structure of MySQL database is shown in Table 1:

Table 1. Database table

Field name	Field type	Field size	Allow to be empty
Uid	Int	4	Not null
User name	Varchar	30	Not null
Pass word	Varchar	20	Not null
Authority	Int	3	Not null
Other	Varchar	50	Null
Mid	Int	4	Not null
Module	Varchar	10	Null
Depict	Varchar	40	Null
Address	Varchar	100	Not null
Authority	Int	3	Not null
Jid	Int	4	Not null
Job name	Varchar	10	Null
Start time	Varchar	20	Null
Percentage	Varchar	10	Null
State	Int	2	Null
Job maker	Varchar	30	Null
Fid	Int	4	Not null
File name	Varchar	30	Null
Start time	Varchar	20	Null
Job maker	Varchar	30	Null

HBase is suitable for massive unstructured data storage. The structure of the table is determined by a row key (Row Key), time stamp (Time Stamp) and column (Column) three dimensions to determine a cell (Cell). The data type in the cell is not limited, and it is converted to binary in practice. Open the database through external commands of revit, so as to realize the control connection between the database and the user. However, if want to further realize the problem of data processing efficiency on the system, the large database automatically feeds back preventive measures, resolution measures, occurrence probability, etc., and further programming is needed to achieve it. At this point, the design of the system database is completed.

3 Software Design of Electric Automation Big Data Control System

In this paper, the control algorithm of the system is butterfly operation, the speed of operation directly affects the speed of the whole design, so how to speed up the processing speed of butterfly operation unit is the key to improve the calculation speed of the system operation unit. Butterfly unit is composed of real part and virtual part, and then passes through complex multiplication unit, buffer unit and complex addition unit. Butterfly operation includes complex multiplication and complex addition. Only by increasing the speed of complex multiplication can the processing speed of the butterfly unit be accelerated. The realization of complex multiplication can be realized directly by multiplier operation, but the multiplication operation is difficult to achieve in hardware and the calculation speed is slow [10]. To solve this problem, the CORDIC (Coordinate Rotation Digital Computer) algorithm is used to implement complex multiplication. The CORDIC algorithm can not only convert the complex multiplication into hardware addition and subtraction and shift operations, but according to its iteration principle, the CORDIC unit It can be represented by a pipeline structure, which can make vector rotation parallel processing, which greatly speeds up the speed of butterfly operation [11]. The coordinate rotation digital algorithm includes three rotation systems: circular system, linear system, and hyperbolic system. In this paper, only the circular system is used for optimization, and the plane coordinate rotation is shown in the Fig. 3:

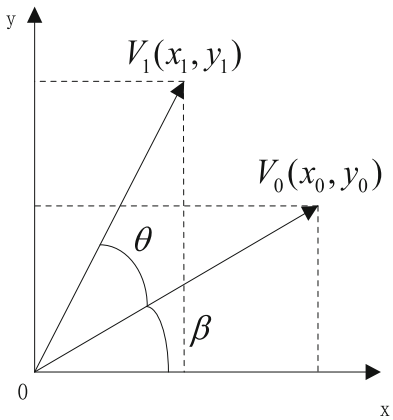

Fig. 3. Plane coordinate rotation

The initial vector V_0 rotates the angle θ to get the vector V_1. The coordinates can be expressed as follows:

$$\begin{cases} x_1 = R\cos(\beta+\theta) = x_0\cos\theta - y_0\sin\theta \\ y_1 = R\sin(\beta+\theta) = y_0\cos\theta + x_0\sin\theta \end{cases} \tag{1}$$

In the above formula, R is the radius of the circle and θ is the rotation angle. Iterate the above formula and divide the rotation process into multiple steps. Then the rotation formula in step i is:

$$\begin{cases} x_{i+1} = (x_i - y_i \cdot \tan\theta_i)\cdot\cos\theta_i \\ y_{i+1} = (y_i + x_i \cdot \tan\theta_i)\cdot\cos\theta_i \end{cases} \tag{2}$$

In order to achieve convenience in hardware, the following convention is made: each rotation angle θ tangent value is a multiple of 2, namely: $\theta_i = \arctan 2^i$, that is, the sum of all iteration angles is equal to the rotation angle. S_i represents the direction of rotation. Set the angle that differs from θ after n rotations to Z_n, then:

$$Z_n = \theta - \sum_{i=1}^{n} S_i\theta_i \tag{3}$$

Therefore, the i time single-step rotation formula can be changed to:

$$\begin{cases} x_{i+1} = x_i - S_i \cdot y_i \cdot 2^i \\ y_{i+1} = y_i + S_i \cdot x_i \cdot 2^i \end{cases} \tag{4}$$

In this way, CORDIC algorithm can basically realize the function operation needed in the mathematical operation. Because of the generality of CORDIC algorithm in calculation and the simplicity of its implementation, it is suitable to be used as a module unit of data processing [12].

The output error caused by the actual rotation angle and the ideal rotation angle is called the approximate error, which is caused by the limited rotation series and the limited binary digits used to express the angle [13]. If the ideal selection result of vector V_0 is V_1, the ideal rotation angle is Z_0, and the actual angle is Z_0', then the angle error can be calculated:

$$Z_0 - Z_0' = \delta \le \arctan\left(2^{n+1}\right) \tag{5}$$

The approximate error can be expressed as:

$$\left|V_1 - V_1'\right| = \sqrt{(\cos\delta - 1)^2 + \sin^2\delta}\left|V_1'\right| \tag{6}$$

In addition, the limited calculation accuracy in each stage will also cause rounding error, which is caused by the limited operand width. The accuracy of CORDIC

algorithm is related to N (rotation Series) and B (operation digit width). The number of iterations is fixed, and the value of approximate error is constant. Therefore, the rounding error can be reduced by increasing the number of processing bits of CORDIC operation unit, so as to reduce the overall error of CORDIC unit. But at the same time, the processing speed will be reduced by increasing the processing bits of CORDIC. The best point must be found between the number of processing bits and the processing speed [14, 15]. By extending the adder and subtracter of CORDIC operation unit to 19 bits, the operation speed of CORDIC operation unit is not greatly affected, but the calculation accuracy is improved. Here, only the adder and subtracter in CORDIC operation unit is extended to 19 bits, ram is still 16 bits wide, 16 bits of data will be sent to the high position of 19 bits adder and subtracter for operation, the calculated 19 bits result will be truncated at the low position, and finally the 16 bits wide result will be output.

4 Experiment

4.1 Experiment Preparation

In order to verify whether the mass data analysis system meets the requirements specification definition, this chapter will conduct system tests to find out areas that are inconsistent or inconsistent with the demand stage, and propose improvements and perfect plans. According to the system function, the system test is divided as shown in Fig. 4:

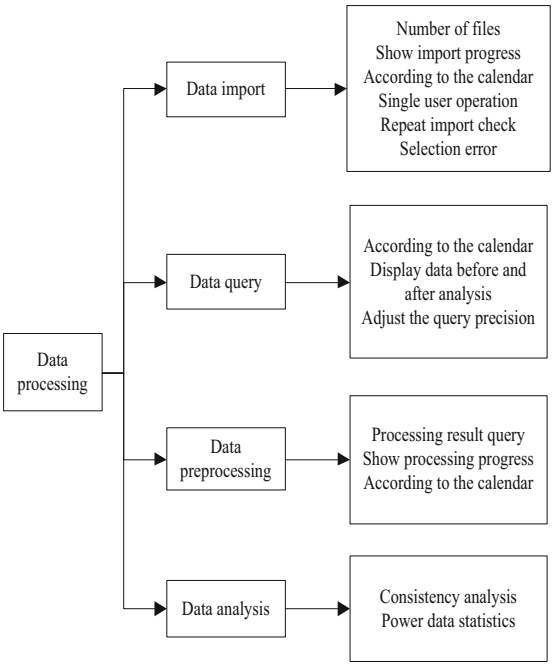

Fig. 4. System function test

In order to verify the effectiveness of the big data control system designed in this paper, the control system based on single-chip microcomputer is used as a comparison method, and the two systems are tested in the same test environment by script. The test environment is shown in Table 2:

Table 2. Test environment

Environmental classification	Project	Parameter
Hardware environment	Blade server	5 units, each with 24 cores
	Memory	65 GB
	Hard disk	2T
Software Environment	Operating system platform	Linux CentOS
	Server platform	Tomcat
	Data base	One MySQL installation, five HBase clusters
	Software platform	Java Web project, Hadoop distributed system

In order to test the performance of the system in this paper in all aspects, we choose to insert performance, query performance and computing performance for testing.

4.2 Statistics and Analysis of Experimental Results

In the insertion performance test, in order to ensure database consistency, data insertion is set to single-threaded execution. Imported 50,000, 500,000, and 5 million pieces of battery data into the two systems to calculate the efficiency. The insertion performance test results are shown in Table 3.

Table 3. Comparison of data insertion test results

Data volume/M	Insert time/s		
	Control system based on MCU	Text system	Efficiency ratio/%
300	10.74	1.08	9.9
600	92.21	8.58	10.7
900	1033.41	88.24	11.7

From the data insertion test results, it can be intuitively compared that the insertion time of this system is much less than the control system based on the single-chip microcomputer with the same data volume, and as the data volume increases, the efficiency gap becomes more obvious. So for the import of massive data, the performance of this system is better.

During the query performance test, the multi-threaded concurrent access to the MySQL database of the simulation system and the distributed HBase database of the system To simulate multi-users reading data at the same time and counting the query time. The query performance test results are shown in Table 4.

Table 4. Comparison of data query

Thread concurrency/one	Query time/s	
	Control system based on MCU	System of this paper
Single-threaded query 100	1.37	0.67
10 threads query 100	2.14	0.68
Single thread query 1000	3.44	0.65
10 threads query 1000	3.76	0.74
100 threads query 1000	24.35	0.79

From the data query results, it can be clearly seen that the query performance of the control system based on single chip microcomputer is still far behind that of the system in this paper, and with the increase of query data and thread concurrency, the gap is more obvious. However, the data volume of tens of millions level is far from the processing limit for the system database of this cluster. For the change of data volume and concurrent thread access volume, the query time tends to be stable.

In the calculation performance test process, the power data variance calculation test is performed on the specified level of data volume by using the microcontroller-based control system and the system of this article, and the data of the two systems are compared calculate ability. Make final analysis and statistics on the test results. The calculation performance test results are shown in Table 5.

Table 5. Comparison of data calculation

Data volume/M	Data calculation time/s		
	Control system based on single chip microcomputer	System of this paper	Efficiency ratio/%
300	873	247	3.5
600	4375	1337	3.3
900	8864	2977	3.0

From the data calculation results, the calculation time of the system in this paper is far less than the calculation time of the control system based on the single chip microcomputer. Although the processing efficiency of the cluster cannot reach 5 times that of a single server due to Hadoop task startup, file reading, and communication delay between clusters, this has greatly improved processing performance. In summary, the performance of the system designed in this paper is significantly better than the traditional data control system.

5 Conclusion

Based on the analysis of some shortcomings in the existing big data control system, this paper optimizes the design from both hardware and software, and designs a big data control system for electrical automation. Starting with data processing, calculation and other methods, the hardware and software configuration has been re-optimized. It is expected that the designed system can have high data processing efficiency. Compared with the existing system, the system in this paper has significant advantages, but there are still some areas that need to be further improved, such as a large number of redundant data in big data, if the redundant data is not processed, the control effect will be affected. The next step will focus on redundant data processing, and further optimize the system to make the system really suitable for the needs of power production.

References

1. Linlin, S.: Design and implementation of big data cloud storage system for monitoring system: an example of monitoring system optimization for port service center in Yangpu Port. Value Eng. **037**(026), 215–217 (2018)
2. Ming, S., Baofeng, W., Yujuan, C., et al.: Design and verification of big data processing system oriented spacecraft control mission. Spacecr. Eng. **27**(6), 69–76 (2018)
3. Xuemei, Z.: Fault diagnosis for electro-hydraulic control system of hydraulic support based on big data. Ind. Mine Autom. **44**(12), 38–42 (2018)
4. Baohua, S., Lei, C., Dong, X., et al.: Design and application of an intelligent operation and maintenance control platform of the power distribution network based on the big data platform. Electr. Autom. **40**(6), 85–88 (2018)
5. Fengjie, S., Chengmin, W., Ning, X.: Frequent pattern network model for association rule mining of big data in smart grid. Electric Power Autom. Equip. **38**(5), 110–116 (2018)
6. Qing, L., Song, L., Xiaojie, L., et al.: Intelligent control system based on big data technology for whole production line of sintering quality. Iron Steel **53**(7), 1–9 (2018)
7. Liu, S., Li, Z., Zhang, Y., Cheng, X.: Introduction of key problems in long-distance learning and training. Mobile Netw. Appl. **24**(1), 1–4 (2018). https://doi.org/10.1007/s11036-018-1136-6
8. Ruiguo, B., Lishan, X., Kuo, B., et al.: Application of big data process quality control system in iron and steel production. China Metal. **28**(8), 76–80 (2018)
9. Haiba, C., Quan, G., Yangcheng, D., et al.: Power grid dispatching intelligent alarm system based on big data. Electron. Des. Eng. **27**(11), 91–95 (2019)
10. Jianguo, X., Haifeng, X.: Realization of a wisdom agricultural internet of things system based on big data. Softw. Guide **190**(8), 133–136+144 (2018)
11. Sifah, E.B., et al.: Chain-based big data access control infrastructure. J. Supercomput. **74**(10), 4945–4964 (2018). https://doi.org/10.1007/s11227-018-2308-7
12. Ndikumana, A., Tran, N.H., Ho, T.M., et al.: Joint communication, computation, caching, and control in big data multi-access edge computing. IEEE Trans. Mob. Comput. **19**(6), 1359–1374 (2019)
13. Gosine, A.: Industrial control systems are converging with Big Data. Control Eng. **65**(3), 40–42 (2018)
14. Shuai, L., Gelan, Y.: Advanced Hybrid Information Processing, pp. 1–594. Springer, Heidelberghttps://doi.org/10.1007/978-3-319-73317-3
15. Fu, W., Liu, S., Srivastava, G.: Optimization of big data scheduling in social networks. Entropy **21**(9), 902 (2019)