

## Design and Implementation of Data Cloud Platform for CNC System

Sheng-Yang SHI<sup>1,a</sup>, Wen-Zhong TANG<sup>1,b</sup>, Ji-Qiang XIA<sup>2,c</sup>, Lian-Sheng GAO<sup>2,d</sup>,  
Yan-Yang WANG<sup>3,e,\*</sup>

<sup>1</sup>School of Computer Science and Engineering, Beihang University, XueYuan Road No.37, Haidian District, Beijing, China

<sup>2</sup>School of Mechanical Engineering and Automation, Beihang University, XueYuan Road No.37, Haidian District, Beijing, China

<sup>3</sup>School of Aeronautic Science and Engineering, Beihang University, XueYuan Road No.37, Haidian District, Beijing, China

<sup>a</sup>shishengyang@buaa.edu.cn, <sup>b</sup>tangwenzhong@buaa.edu.cn, <sup>c</sup>xiajiqiang@buaa.edu.cn,  
<sup>d</sup>gaoliansheng@buaa.edu.cn, <sup>e</sup>wangyanyang@buaa.edu.cn

\*Corresponding author

**Keywords:** CNC Machine Tools, Cloud Platform, Data Acquisition, Data Service.

**Abstract.** Reliable CNC machine tools can make the industrial production efficient and stable. The reliability of CNC system is a hot topic in the field of mechanical engineering. It is very meaningful to make suggestions to improve the reliability of CNC machine tools by collecting and analyzing the operation data of CNC machine tools. The data cloud platform of CNC system is designed and implemented in this paper. It is used to collect the data of CNC system and provide data service to the public. This data cloud platform, based on server cluster and big data processing framework, can effectively receive, forward, store and publish data of CNC system. The experiment proved that the system has high availability, high concurrency, high throughput and reliable disaster recovery ability.

### Introduction

Reliability is the ability of the product to complete the specified function within the prescribed time and within the specified conditions. It has become an important index to determine the competitive power of the product market [1]. The system of CNC machine tool is based on technology of microelectronics, computer, automatic control, information processing. It is a special computer system consisting of NC program, input device, output device, computer numerical control device (CNC device), programmable logic controller (PLC), spindle drive device and feed (servo) driving device (including detection special device) [2]. The reliability of CNC machine tools is influenced by the reliability of CNC system. Therefore, the research on the reliability of CNC system is of great significance.

The use of CNC machine tools is more and more common. It is meaningful to collect and analyze the data of CNC system, which can effectively help the on-site processing personnel and management staff to understand the whole process, improve the digital monitoring and control of manufacturing process, provide data support for the optimization of production scheduling and the analysis of machine fault [3]. The reliability test and analysis of CNC system need a large number of operation data of CNC system. These data mainly include machine tool running data, environment data and alarm data. By analyzing the changing rules of these data, the factors that affecting the reliability of CNC system can be obtained.

The data needed in the research process have the characteristics of real time, high frequency and large amount. It is impractical to collect data manually. It is a feasible way to collect these data by developing special data acquisition system based on the CNC interface [4,5,6]. The system needs to be extensible because machine tools may join or exit the system during the process of data acquisition.

Besides, the reliability of the data acquisition system needs to be guaranteed because it needs a long time to collect enough data.

Cloud platform is based on the architecture of "Terminal + Cloud" [7]. In the cloud platform system of SaaS (Software as a Service) [8] service model, the "Terminal" is composed of distributed system modules for data collection, transmitting and receiving; the "Cloud" is based on the cluster servers, which provides the function of data receiving, storage, calculation, mining. Communication between the "Terminal" and the "Cloud" is achieved by means of Socket and WebService. Cloud platform system of SaaS service model can meet the requirements of data collection of CNC system. "Terminal" is the distributed CNC system, which is responsible for collecting real-time data and sent to the "Cloud"; "Cloud" is the cluster servers and big data processing system, which provide services of data receiving, storage and computing for data acquisition, services of condition monitoring and data publishing to users.

## Data Acquisition Platform

### System Architecture

Cloud platform system consists of cloud and terminals [9]. The terminals consists of two kinds, one is the distributed data acquisition system, which collects the data from CNC interface and sends them to the cloud; the other one is the terminals which request the condition monitoring and data publishing services from the cloud platform. The cloud provides two kinds of services: data acquisition and data service.

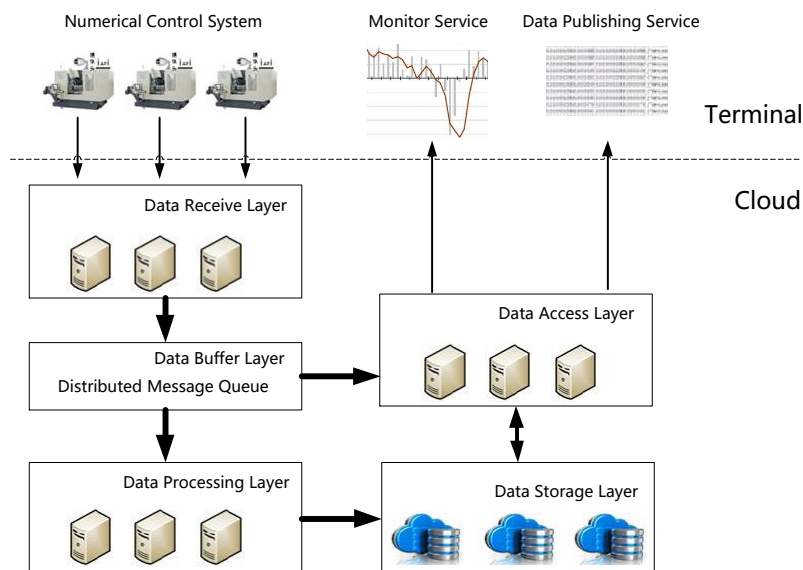


Fig. 1. System Architecture

Fig. 1 shows the architecture of the data cloud platform for CNC systems. Data acquisition module is composed of data receiving layer, data buffer layer, data processing layer and data storage layer. The data receiving layer provides an interface for data access, and the interface is WebService. It sends the received data to the distributed message queue; A system of distributed message queue is deployed in the data buffer layer. All data received by cloud will be cached here, which can be used by other modules; the data processing layer gets data from the message queue, it processes the data and stores the analysis result in the data storage layer; all the original data and the processed data will be stored in the data storage layer.

The data services are primarily provided by the data access layer. Data access layer gets data from the data buffer layer and data storage layer. It provides condition monitoring services and data

publishing services in the form of web, which can be easily accessed by the Web browser in the terminal.

Cloud platform has user access control mechanism. For all terminals which send data or request service, they will access to the interface of cloud after getting the security certification. User access control mechanism can resist malicious attacks from the network, but also to ensure the security isolation of user data in the cloud.

### **Functional requirements of cloud platforms**

The cloud platform has three functions, including data acquisition, data service and user access control mechanism.

#### **Data Acquisition**

In order to facilitate the development, reuse and maintenance of the system, the data acquisition function is divided into the following four sub modules.

##### **(1) Data receiving module**

The data receiving module provides the data interface of the cloud platform. Cloud platform needs to receive different types of data, including strings, pictures, and videos. At present, the mainstream data transmission mode includes WebService and socket. WebService based on http protocol is powerful and simple. Mature Web cluster can be used to improve the system load of WebService [10]; The development of socket is complex, but easy to be optimized [11].

##### **(2) Data buffer module**

The data buffer module mainly has two functions. One is to improve the throughput of cloud system. The amount of data reached per unit time is a fluctuating value, but the data processing ability of the cloud platform is relatively stable. So the cloud platform cannot process the data in time when it receives too much data and cannot make full use of its data processing ability when it receives very little data. In order to improve the throughput of the system, a data buffer module is set between data receiving module and data processing module. The other function of this module is data sharing. When other modules need the original data, they can get the data from the buffer rather than query the database, which reduces the pressure on the database.

##### **(3) Data processing module**

The data stored in the data buffer module is the original data received by the cloud platform. Because of the need for data transmission and control, these data contain both control data and encoded CNC system data. The function of the data processing module is to parse the data of the CNC system from the original data and convert it into the object data that can be used directly, then calculate and store these data according to the requirement.

##### **(4) Data storage module**

Data storage module is the end point of CNC system data. All the original data and the processed data are stored in a data storage module. These data are used to analyzing the reliability of CNC system later and provide data service to user.

#### **Data Service**

The data services provided by cloud platform include condition monitoring service and data publishing service. The data access module includes multi-dimensional data interface based on Restful software architecture [12]. And the cloud platform provides Web based, visual working condition monitoring and data access interface.

##### **(1) Condition monitoring**

The Condition monitoring service of CNC machine tools displays the multi-dimensional data related to the operation of CNC system through data visualization. The displayed data include basic information of CNC machine tools, operation data and alarm data.

##### **(2) Data publishing**

Cloud platform provides high performance, large concurrent, real-time and non-real-time data query services. Through this service, users can get all kinds of original data related to the operation of CNC machines.

### User Access Control Mechanism

Cloud platform has user access control mechanism. All terminals requesting cloud platform services need security authentication through user access control mechanism. In the data acquisition section, when the CNC system first join the cloud platform or the CNC system of machine tools is updated, the terminal system needs to send the latest version information of the CNC system to the cloud platform to complete the registration. The terminal should login the platform before sending data and logout the platform after completing its work. In the section of data services, users are only permitted to query their own data after being certified.

### System Implementation

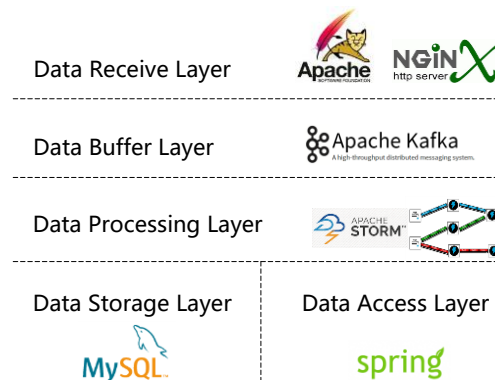


Fig. 2. System Implementation

The data cloud platform of CNC system is a big data processing framework deployed on server cluster. The frame includes data receiving module, data buffer module, data processing module, data storage module and data service module. Fig. 2 shows the open source tools used for system development.

#### (1) Data receiving module

Data receiving module accepts data in JSON format. JSON is a lightweight data interchange format. It can convert complex structured data into a string which can be transmitted between different systems. Compared to XML, JSON format is simpler and suitable for large amounts of data passing [13].

Because the terminals are cross regional. Http is a good choice for transmitting data because it can pass through the firewall between regions. Besides, the message in a single http request is a complete data. So the data receiving module is easy to be developed without the need for segmentation and reorganization of data. Therefore, the communication interface of data receiving module is implemented by WebService based http.

In order to stably receive high frequency http requests, the data receiving module uses architecture of load balancing, implemented by Nginx and Tomcat [14]. Nginx receives all requests and distributes them to Tomcats. The WebService is deployed on Tomcats.

#### (2) Data buffer module

In order to make the whole system has the ability of high throughput and disaster recovery. All received data is stored into the message queue for the first time. The message queue is used to share and protect the data. The data buffer module uses Kafka as the data buffer. Kafka is a distributed

publish subscribe system which is widely used in current big data processing system. Kafka has data partition and replication mechanism, which prevents data from being lost due to the single point of failure of servers [15].

### (3) Data processing module

The data processing module parses and analyzes the received data. It transforms the data in JSON format into object data which is convenient for processing and storage, processes data as needed, and stores data in database. The data processing module is based on Storm, which is a widely used distributed data stream processing framework. Storm divides the entire processing task into a plurality of processing units. They are uniform distribution to the servers in the cluster to perform. Storm can take full advantage of the processing capacity of the entire server cluster.

### (4) Data storage module

The data storage module stores the data received by the cloud platform into the database. These data are used to analyze the reliability of the CNC system and protect data service to user of the cloud platform. The data storage module is based on MySQL. It stores two kinds of data, one kind is historical data, which is used for analysis and mining, and the other kind is real-time data, which is used to provide data service.

### (5) Data service module

The data service module is the interface module for providing data service to users. This module is implemented on the basis of Spring. It displays the working state of CNC machine tools through the form of Web pages, and publishes data through Webservice. The Web application based on Spring is stable and efficient, which has been widely used.

## Experiment

The performance of the cloud platform determines the quality of its service. This section tests the performance of the cloud which include availability, concurrency, throughput and disaster recovery capabilities. In order to avoid the impact of the network on the experiments, all experiments are carried out in local area network. The machine configuration for the experiments is as follows in Table 1. There are two web servers, one application server and one data server. In order to simulate the server cluster, three virtual machines are set up on the application server.

Table 1. machine configuration

Server	number	software	Virtual machine	CPU	Memory
Web server	2	Nginx,Tomcat	No	4 Processors	2G
Application server	1	Kafka,Storm,Zookeeper	VM1	4 Processors	6G
			VM2	4 Processors	6G
			VM3	4 Processors	6G
Data server	1	MySQL	No	4 Processors	8G

### (1) Availability

Availability is the probability or time share expectation of a system that can run normally at an inspection time. The experiment has prepared 20 CNC machines tools, each of which sends one data per second to the cloud platform. If the data is successfully received, the system is deemed to be in a state of availability, otherwise the system is considered unavailable. The experiment lasted 24 hours, and the result was shown in Table 2. During the 24 hours of the experiment, the number of requests

gradually increased to about one million and seven hundred thousand, and all the data were successfully received by the cloud platform.

Table 2. Availability test

Duration	Request	Fail	Percentage of failure
6	432000	0	0%
12	864000	0	0%
18	1296000	0	0%
24	1728000	0	0%

## (2) Concurrency

### Concurrent Quantity Test

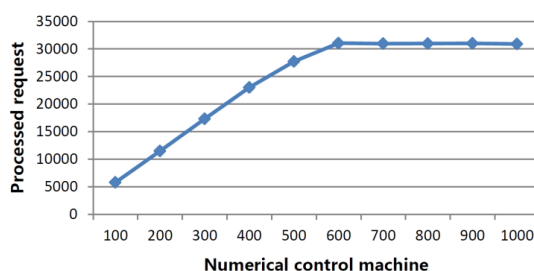


Fig.3. Concurrent Quantity Test

The experiment tests the maximum concurrent quantity of the cloud platform through stress-testing. A large number of machine tools were simulated by multithread program. Each machine tool sends one piece of data per second to the cloud platform. The experimental results are shown in Fig. 3, where the horizontal axis represents the number of machine tools, and the vertical axis indicates the number of requests that the platform handled in a minute. The experimental results show that, the number of requests processed by the cloud platform continues to increase until it reaches 600 times per second.

## (3) Throughput

### Throughput Test

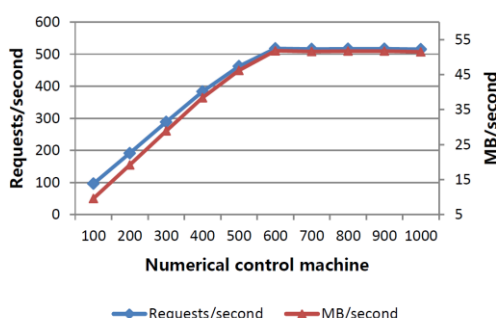


Fig.4. Throughput Test

The experiment tests the maximum throughput of the cloud platforms through stress-testing. A large number of machine tools were simulated by multithread program. Each machine tool sends one request per second to the cloud platform. The amount of data contained in each request is 100KB. The experimental results are shown in Fig. 4, where the horizontal axis represents the number of machine tools, the left vertical axis represents the number of requests processed per second, and the right vertical axis represents the throughput of the cloud platform. The throughput is approximately 500 requests per second to receive 50MB data.

## (4) Disaster recovery ability

Cloud platforms are based on distributed server cluster, and failures of some servers or software systems do not affect the normal functioning of the whole cluster. Through the experiment of man-made fault, disaster recovery ability of cloud platform is tested. Common faults can be divided into two categories: server hardware failures and software system failures.

Table 3 shows the experimental results of server hardware failures.

Table 3. Server hardware fault test

Fault	State of cloud platform	Solution
Shut down one web server	Normal	Start web server
Shut down one VM	Normal	Start VM
Shut down one data server	Normal	Start data server

- Based on the Keepalived, one web server is MASTER and the other one is BACKUP. They will continue to process the http request when one of them is shut down.
- The server cluster is formed by three VMs. Zookeeper is responsible for the management of the servers in the cluster, when a VM is shut down due to a fault, Zookeeper will distribute its tasks to other VM in the cluster.
- The received data cannot be stored in the database when the data server is shut down. Since all received data is stored in the distributed message queue of Kafka before being stored into the database, the data will continue to be stored in the database when the data server is restarted.

Table 4 shows the experimental results of software system failures.

- The impact on the cluster from operating system failure is the same with server hardware failures. So the cloud platform can run normally when the OS is shut down.
- The Nginx and Tomcat are deployed on the two web servers. When Nginx and Tomcat on one server is shut down, the http requests will be processed by the Nginx and Tomcats on the other web server.
- The Zookeeper, Kafka and Storm are widely used distributed framework. In this experiment, they are deployed on a VM cluster. These frameworks have the ability to deal with single point of failure, and problems of individual instances do not affect the normal operation of the whole cluster.
- When the MySQL is down, the data cannot be stored in the database. However, Kafka caches all received data. After MySQL is restored to normal, the data will continue to be stored in MySQL.

Table 4. Software system fault test

Fault	State of cloud platform	Solution
Shut down the OS of a random server	Normal	Start OS
Shut down Nginx	Normal	Start Nginx
Shut down Tomcat	Normal	Start Tomcat
Shut down Zookeeper	Normal	Start Zookeeper
Shut down Kafka	Normal	Start Kafka
Shut down Storm	Normal	Start Storm
Shut down MySQL	Normal	Start MySQL

In this section, the performance of the data cloud platform of CNC system is analyzed, and the experimental results are demonstrated. The analysis results show that, the cloud platform with high availability, high concurrency, high throughput and disaster recovery capability, can meet the demand of data acquisition and data services of CNC machine tools.

## Conclusion

This article designs and implements a data cloud platform system of CNC system. This system adopts the "terminal + cloud" architecture and SaaS service mode. It implements the data acquisition function for CNC system which is able to process video, picture and string; it also implements data service function including condition monitoring and data publishing. The cloud platform based on server cluster uses Kafka, Storm, and Zookeeper, to provide stable and efficient cloud services. The experiments proved the cloud platform has high availability, high concurrency and high throughput by stress-testing, and proved the cloud platform has disaster recovery capability by making artificial faults. The cloud platform can satisfy the functional and performance requirements of data acquisition and data service of CNC machine tools.

## Acknowledgement

This research was supported by the National Science and Technology Major Project (Grant No. 2016ZX04004006).

## References

- [1] C. Peng, Y.J. Meng and L.Y. Lan, Study on Software Reliability of Numerical Control System, *Applied Mechanics and Materials* (2013) 2085–2089
- [2] L. Wang, H. Huang, R.W. West and D. Wang, Intelligent manufacturing system of impeller for computer numerical control (CNC) programming based on KBE, *Journal of Central South University* 21 (2014) 4577–4584
- [3] F. Ferraz and R.T. Coelho, Data acquisition and monitoring in machine tools with CNC of open architecture using internet, *The International Journal of Advanced Manufacturing Technology* 26 (2004) 90–97
- [4] V. Kandadai, M. Sridharan, S.M. Parvathy and R. Pitchaimuthu, A comprehensive embedded solution for data acquisition and communication using FPGA, *Journal of Applied Research and Technology* 15 (2017) 45–53
- [5] X.B. Cao, X.L. Zhang and H.M. Zhang, A data acquisition system of CNC based on Zigbee and GPRS technology, *Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016 IEEE* (2016) 1188–1191
- [6] Y. Zhang and X. Chu, Design and Research of wireless sensor data acquisition system based on dynamic clamping force of hydraulic power chuck, (2016)
- [7] M.-T. Ke, C.-H. Yeh and C.-J. Su, Cloud computing platform for real-time measurement and verification of energy performance, *Applied Energy* 188 (2017) 497–507
- [8] M. Almorsy, J. Grundy and I. Müller, An analysis of the cloud computing security problem, *arXiv preprint arXiv:1609.01107* (2016)
- [9] V.C. Aiyappa, R.P. Kumar, G. Grandhi and S.K. Shapur, High Availability Cloud Maps for Cloud System using HPE Serviceguard, *International Journal of Engineering Research* 5 (2016) 439–443
- [10] H.D. Purnomo, D.A. Saputro, R. Somya and C. Fibriani, The application of restful web service and Json for poultry farm monitoring system, *journal of electrical engineering and computer sciences*, vol 1 number 1 1 (2017)
- [11] Z. Wang, Application of high performance asynchronous socket communication in power distribution automation, *AIP Conference Proceedings* (2017) 20039

- [12] E.E. Thu and T.N. Aung, Developing Mobile Application Framework by Using RESTFul Web Service with JSON Parser, Genetic and Evolutionary Computing (2016) 177
- [13] F. Pezoa, J.L. Reutter, F. Suarez, M. Ugarte and D. Vrgoč, Foundations of JSON schema, Proceedings of the 25th International Conference on World Wide Web (2016) 263–273
- [14] A. Goel, N. Bansal and S. Gupta, Comparison of Different Web Servers, Imperial Journal of Interdisciplinary Research 2 (2016)
- [15] J.G. Son, J.-W. Kang, J.-H. An, H.-J. Ahn, H.-J. Chun and J.-G. Kim, Parallel Job Processing Technique for Real-time Big-Data Processing Framework, Proceedings of the International Conference on Research in Adaptive and Convergent Systems (2016) 226–229