



Contents List available at VOLKSON PRESS  
**World Symposium on Mechanical and Control  
 Engineering (WSMCE)**



## RESEARCH ON CLOUD PLATFORM OF CASTINGS QUALITY MANAGEMENT BASED ON SPARK

Xiaolan Xie<sup>1,2</sup>, Xiao Zhou<sup>3\*</sup>, Tianwei Yuan<sup>1</sup>

<sup>1</sup>College of Information Science and Engineering, Guilin University of Technology Guilin, Guangxi Zhuang Autonomous Region, China

<sup>2</sup>Guangxi Universities Key Laboratory of Embedded Technology and Intelligent Information Processing Guilin University of Technology, China

<sup>3</sup>College of Mechanical and Control Engineering, Guilin University of Technology Guilin, Guangxi Zhuang Autonomous Region, China

\*Corresponding Author Email: [2485782688@qq.com](mailto:2485782688@qq.com)

*This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited*

### ARTICLE DETAILS

### ABSTRACT

#### Article History:

Received 02 october 2017

Accepted 06 october 2017

Available online 11 november 2017

#### Keywords

Castings quality management,  
 Spark, Cloud platform, Big data

The casting quality management costs too much manpower, material and financial resources and is still lack of scientific status. In view of these conditions, this paper builds the casting quality management cloud platform based on Spark. Build and deploy Spark Standalone on the local computer. Parallel computing and cloud storage mechanism are implemented by Spark. Task management scheduling is implemented by Standalone. The foundry enterprise uses this platform to carry on the big data analysis to the influence factors of casting quality in the casting production. According to the results of the analysis, the scientific guidance and efficient scheduling of the manufacturing process are carried out. The quality management of castings is standardized and scientific. The scrap rate of castings is reduced. Through the large data analysis of various factors affecting the quality of casting, a set of scientific and efficient management scheme is provided.

## 1. Introduction

Casting quality management is the core content of foundry production. It runs through the whole process of production. The quality of casting is not easy to control because of many processes and many factors [1-4]. At present, the management level of castings is not high, which cannot meet the requirements of modern enterprises for the quality of castings. Therefore, it is especially important to standardize and scientifically manage the quality of castings by using modern technology.

In the application of modern technology for casting quality management, it also faces a difficult problem. It is mainly manifested in the operation of the casting quality management system, a large amount of calculation, large amount of data, the traditional high-performance computing cluster management in the presence of limited resources, complex parallel operations, batch jobs waiting time-consuming and other issues [5,6]. Especially for some enterprises which are not large enough, there are still some problems such as expensive infrastructure. Therefore, how to use limited resources to provide a foundry enterprise with a high-performance computing environment based on the cloud computing model and provide scalable, fast, inexpensive and dynamically locatable computing and storage resources is a key issue to be solved urgently. In this paper, a cloud platform of castings quality management based on Spark is studied [7].

## 2. RELATED INTRODUCTION

### 2.1 Castings quality management

There are many factors affecting the quality of castings. At the same time, there are many factors that interact with each other. There are five main elements, referred to as "4M1E". They are man, machine, material, method and environment. Man mainly refers to the quality of personnel and management methods. The staff can be trained to enable the employees to understand the requirements of the job guidance. The management method mainly refers to the quality management idea and the post responsibility system.

Machine mainly refers to the advanced equipment, equipment utilization, the degree of information and the use of quality information. The merits of equipment directly affect the accuracy of castings and the efficiency of production. It is very important to ensure product quality. The original and

auxiliary materials used in molding and making core cannot meet the demand of foundry production in the aspects of species, quality and performance. It greatly affects the stability and promotion of casting quality. Controlling the source and stability of raw materials is the basis of fine casting production.

The production of castings begins with the design process. The design and development process of casting is the most basic link in quality activity. More than 80% of the casting quality responsibility comes from the design and development of castings. The impact of the environment on the quality of castings can be divided into direct and indirect effects. Direct effect: If the damp environment will make the preserved sand core damp, increasing the possibility of stomatal defects. Indirect effects: such as dust, high temperature, poor ventilation, poor lighting environment.

### 2.2 Spark + Standalone mode

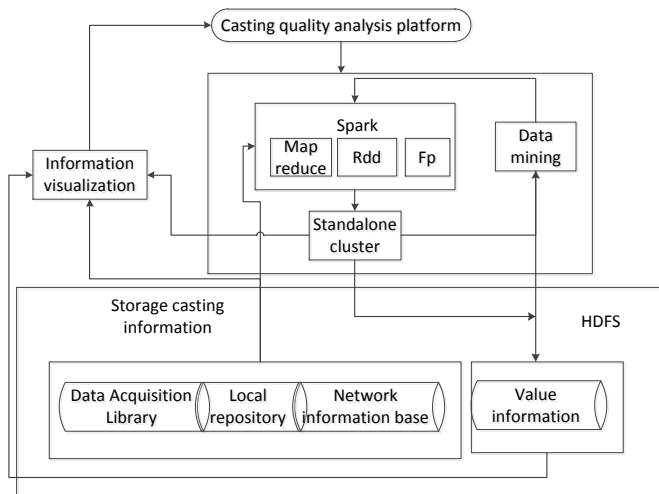
The advent of big data era, also led the big data processing platform ecosystem constantly updated. With the gradual birth of Hadoop HDFS, Hadoop MapReduce, HBase and Hive, the early Hadoop ecosystem gradually formed. As an open source project of Apache, Hadoop aims to use cheap hardware devices to build a stable and scalable distributed large data processing platform.

However, with the increase of business volume and the improvement of users' service quality, the traditional MapReduce framework needs to fall down because of processing intermediate results, so that the processing time of the job grows. As a result, a new class computing framework, Spark, was born. Spark is a memory based computing, and its processing speed is much better than that of MapReduce in the process of requiring multiple iterations. This article is related to research big data processing platform based on Spark. Spark is an iterative computing framework. It does not have the management capacity of computing tasks, the need for third-party resource management platform scheduling assignment tasks. With the release of hadoop2.0, Spark can be directly run on the second-generation Hadoop Map Reduce framework resource unified management platform Standalone. The implementation of local resource parallelization based on Spark + Standalone platform is feasible, efficient and practical.

## 3. SYSTEM DESIGN

### 3.1 System architecture

Build and deploy Spark + Standalone on your local machine to parallelize local resources and build a high-performance computing cloud platform. Parallel computing and cloud storage mechanism are implemented by Spark, and task management scheduling is implemented by Standalone. The system architecture diagram is shown in Figure 1.



**Figure 1:** Architecture of cloud platform system for casting quality management based on Spark

### 3.2 System Architecture Description

In the data mining module, the information related to the manufacturing management of the foundry obtained from the original database is analyzed through the relevant algorithms of data mining. The original database is made up of local resource database, network information database and data acquisition library. The tasks generated in the analysis process are assigned to the Spark cluster. The calculated result is processed again by the data mining module or the result is passed to the Value information. Database module includes the original database and Value information library. The database is stored in HDFS mode as a whole. The visualization module presents information related to the quality management of castings in tabular, tabular and other forms.

## 4. EXPERIMENTAL DESIGN

### 4.1 Experimental environment

The attribute information of the node in the experimental environment is shown in Table 1:

**Table 1:** Node information table

Project	Value
Number of nodes	4
Operating system	CentOS7
Spark version	2.6.1
Spark mode	Standalone mode
JDK version	1.8.0
Node CPU	Intel Xeon CPU 2. 80 GHz* 4
Node memory	Master node 2G worker node 1G

The host name of the three Worker nodes are set to Node1, Node2 and Node3, respectively.

### 4.2 Spark cluster building

- Building Spark is based on Hadoop. First, build Hadoop cluster on four machines.
- Deploy the basic software needed for Spark.
- Constructing distributed Spark cluster.

-Installing Scala:Spark framework is written by Scala language based on JVM. Spark cluster to be set up to install Scala first.

-Spark fully distributed configuration Spark cluster configuration is relatively simple, most of the settings using the system default configuration. It just needs to configure the slaves file in the conf directory.

Add worker nodes Node1, Node2 and Node3 into it.

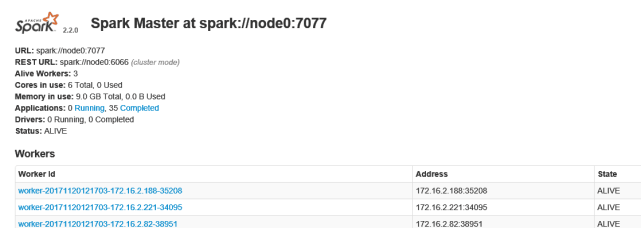
d) Start the Spark cluster.

- Start the Hadoop cluster.
- Start the Spark cluster.

- Into the spark startup file directory:/opt/spark-1.6.1-bin-hadoop2.4/sbin/.
- Start the Spark cluster: ./start-all.sh.
- Use the jps command to see the processes running in the Spark cluster.A new process "Master" has appeared in Node0.A new process "worker" appears in the Node1, Node2 and Node3.

(4)Visit "http://Node0:8080/". Enter the web user interface of the Spark cluster to see the status and usage information of the cluster

There are three worker nodes in the cluster and the information of the three nodes is shown in figure 2:



**Figure 2:** Web UI of Spark

-Close Spark cluster.

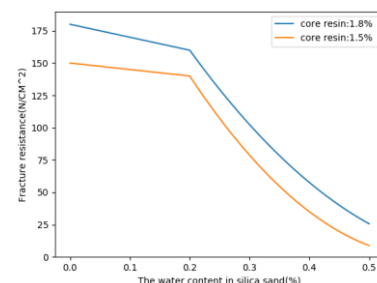
(1) Enter the directory where the closed file of Spark is entered: /opt/spark-1.6.1-bin-hadoop2.4/sbin/.

(2) Close Spark cluster: ./stop-all.sh.

-Close Hadoop cluster: stop-all.sh.

## 5. PERFORMANCE TEST OF SPARK CLOUD PLATFORM

The protection of casting material greatly affects the quality of casting. Silica sand is mainly used in casting moulding. Easy to damp the improper protection of silica sand. If the sand contains a certain amount of water, just like dust, it will reduce the degree of bonding between sand and resin, thus affecting the strength of the sand core. When the moisture in the sand is large, the sand core will generate a large amount of gas under the high temperature baking of the lead liquid during casting of the casting, resulting in the occurrence of the pores in the casting and being scrapped. Therefore, the protection of silica sand is crucial. We use a lot of relevant data in the Spark-based cloud platform to analyze the impact of water on the strength of the core as shown in Figure 3.



**Figure 3:** Analysis of the results

According to the results of big data analysis, the cloud platform is used to carry out the theoretical guidance of the late protection and the scheduling of relevant protective measures. For example, the cloud platform based on resin content and water content of sand sand core guide supplier in adding a layer of moisture-proof bag inside lining bags, and in the process of transportation with rainproof shed cloth etc.

## 6. CONCLUSION

This paper builds a cloud platform for castings quality management based on Spark. It provides an efficient solution of cloud computing for casting quality manager. Compared with the traditional methods of quality management, not only science but also save a lot of manpower, material and financial resources. Casting companies can use common PC and idle servers to build cloud computing environments. It makes effective use of local resources. In addition, the platform provided by this paper uses Spark to implement the framework of parallel computing and the big data processing mechanism to meet the requirements of timeliness in the quality management of castings. In the experimental design, the influence of moisture on the core strength is analyzed, and the schedule is given in time. It shows that the cloud platform based on Spark can deal with the quality management task of casting well. The research results of casting quality management data solution can be extended to vehicle management, biomedical engineering and other fields.

#### ACKNOWLEDGMENTS

This research work was supported by the National Natural Science Foundation of China (Grant No.61762031), Guang Xi Key Research and Development Plan, Guang Xi key Laboratory of Embedded Technology and Intelligent Information Processing.

#### REFERENCE

- [1] Trumbulovic, L., Acimovic, Z., Gulisija, Z., Andric, L. 2004. Correlation of technological parameters and quality of castings obtained by the epc method. *Materials Letters*, 58 (11), 1726-1731.
- [2] Wankhede, D. M., Narkhede, B. E., Mahajan, S. K. 2017. Identification of prominent casting defects and its impact on quality of castings. *Identification of prominent casting defects and its impact on quality of castings*, 13 (2), 1-15.
- [3] Huang, Y., Liu, J. S., Peng, Q. F., Chen, J. 2009. Casting quality control of cylinder block of wuling brand minibus. *Modern Cast Iron*.
- [4] Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., Stoica, I. 2010. Spark: cluster computing with working sets. *Usenix Conference on Hot Topics in Cloud Computing*, 15, 10-10. USENIX Association.
- [5] Barnes, J. 2003. High Integrity Software: The SPARK Approach to Safety and Security. DBLP.
- [6] Zhao, L., Liu, A., Keung, J. 2011. Evaluating Cloud Platform Architecture with the CARE Framework. *Software Engineering Conference*, pp.60-69. IEEE.
- [7] Lucovsky, M., Collison, D., Spivak, V., Chen, G. C., Laddad, R. 2016. Cloud platform architecture

