Theoretical line loss calculation based on the Spark of memory cluster technology

Wang Dewen[a], Li Jiangman[b]

[1]School ofControl and Computer Engineering, North China Elcctric Power University, BaoDing 071003, China;

[2]School of Control and Computer Engineering, North China Elcctric Power University, BaoDing 071003, China.

[a]wdewen@gmail.com, [b]lijiangman318@163.com

Keywords: Theoretical line loss calculation; distributed memory clusters; line losses; Spark.

Abstract. Theoretical line loss calculation is the basis for line loss analysis, and reduce the time of c alculating the theoretical line loss can improve the efficiency of the line loss analysis.This paper,aim ing at the wasted issue in power transmission,presents a theoretical line loss calculation framework t hat is based on distributed memory cluster, which combines the distributed technology called Hado op and memory cluster technology called Spark.It can reduce the time of theoretical line loss calcul ation and improve the ability of storing and analyzing in electric power big data. A line loss analysis of P-Q method experiment is conducted in this framework, and the comparison with the stand-alon e mode is discussed, as a consequence the feasibility and efficiency of the framework are verified.

Introduction

The application of Hadoop and other cloud technology in the power system is the current hot topic, it can improve the ability of storing and analyzing power big data [1-5], but there are deficiencies on the demand side of real-time calculation, for example theoretical line loss computation [6-8] . Currently, the real-time performance of theoretical line loss calculation is low, and it can't meet the demand for green smart grid development. Distributed memory cluster technology Spark can reduce the data read / write and program running time, transmission time by storing the computing data in resilient distributed datasets, at the same time, using directed acyclic graph to optimize program. This article combined distributed memory clusters and theoretical line loss calculation, and proposed theoretical line loss calculation framework based on the Spark. This framework make most use of the cluster disk and memory resources, can improve the ability of the power system to store and analyze power big data.

Distributed memory cluster technology Spark

Spark is based on cloud computing framework of memory cluster, the design goal is high fault tolerance, easy to expand, for rapid processing of large data, etc. Spark is suitable for analyzing big data. The principle of SparkonYarn is shown in following Figure 1. Yarn as a unified resource management system and resource scheduling mechanism is responsible for allocating HDFS and cluster resources, which include memory, CPU, disk I / O, and network I / O, etc. HDFS is responsible for big data storage and resources save on the client load. Spark is mainly responsible for completing the job submitted by the client.
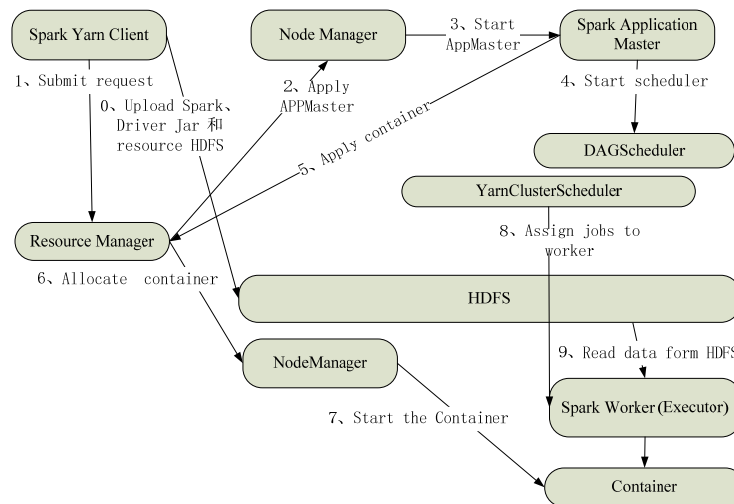
Fig 1 The schematic of Spark On Yarn

Analysis framework of theoretical line loss based on Spark

With the improvement of the economic development and people's living standards, resource consumption increases every year, building a resource-saving society, reducing energy waste, improving resource utilization relates to the sustainable development and growth of China's economy. Transformers, transmission / distribution lines and other equipment in transporting and distributing electrical energy will produce more power loss, resulting in a huge waste, which is contrary to energy-saving, environmental protection and other theories.
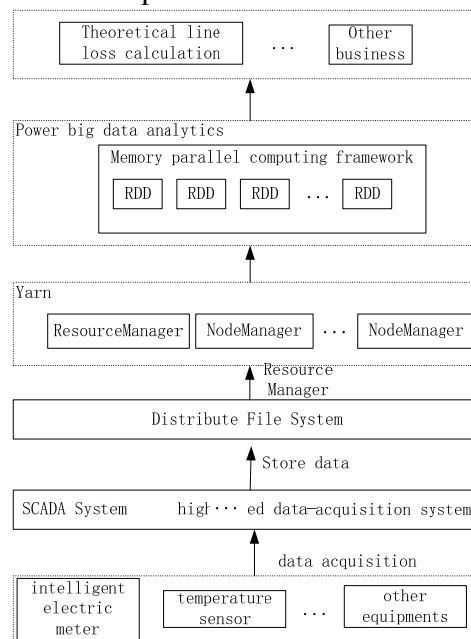

Fig 2 The theoretical line loss analysis framework based on Spark

It allow to take timely measures to achieve energy saving purposes, contribute to the construction and development of green smart grid, that quickly and accurately analyzing the distribution of the line loss. Theoretical line loss analysis is the premise and basis of line losses analysis. Improving time efficiency of theoretical line loss analysis is important. Figure2 shows theoretical line loss calculation framework based on distributed memory cluster, which can reduce the time and improve the efficiency of theoretical line loss calculation. It is seen from Figure 2, the data collected by SCADA systems stored in HDFS which is distributed file system, Spark is the analysis tools and Yarn is resource scheduling manager. Then the framework provide services for theoretical line loss computing and other business.

Experimental Results and Analysis

Theoretical line loss calculation includes a line conductor loss, component losses, transformer losses and capacitor losses, among them, There are several method to calculate line loss , for example power flow method, approximate calculation method etc. This paper analyzes the line conductor loss by P-Q method from memory footprint, computational speed and other considerations, to verify the feasibility and efficiency of this program.

The test data

The SCADA system collects the data of electric power system simulation software as experimental data, which includes two aspects: branches' resistance and nodes' voltage, power data, including branches of the resistance and reactance, node types, voltages, active / reactive power and so on. When reading the information, read the data according to the type of the node, useless data or data that does not exist is set to 0.

The experimental environment

Experimental environment is a three-node cluster, including two slave nodes and a master node, the line loss analysis experiments based on P-Q method are conducted in this environment. Each node in the cluster is the same configuration of hardware: 1G memory, 40G hard disk, 64 bits Centos operating system. At the same time, experiments are conducted in stand-alone mode which has the configuration of 1G memory, 40G hard disk, 64 bits windows7 operating system.

The experimental result

This paper conducts the experiments of line loss analysis running efficiency when test data acquisition time were 0s,13s,25s,57s,72s,86s,95s. The initial value is set closer to the correct result, and the number of iterations is no more than four times, and the amount of calculation is small. The following figure 3 ,which use the test data acquisition time as the horizontal and the line loss analysis time as vertical axis ,shows the variety of the line loss analysis running time with the data at different time points in stand-alone mode and cluster mode.
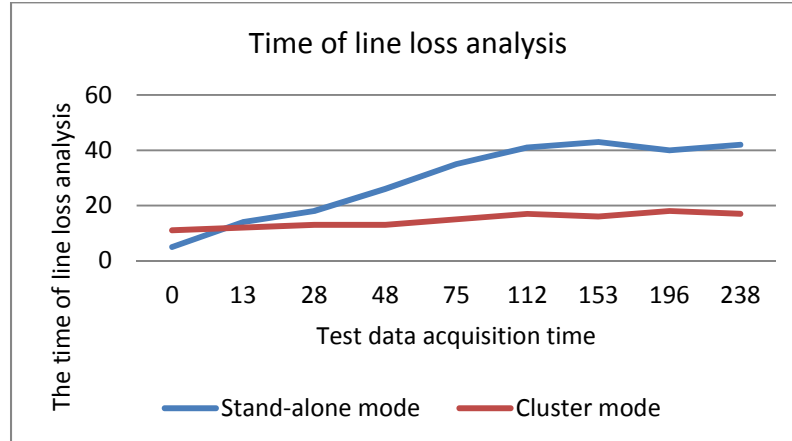

Fig3 The figure of line loss analysis time

The figure 3 show that the line loss analysis time of the cluster mode is at least 11 seconds, while the minimum time in stand-alone mode is 5 seconds, when the test data is small, the advantages of the cluster is not obvious. Because the cluster will consume some time and resources when the cluster starts and communications. But with the increasing of line loss computation complexity, the running time of Cluster mode is less than stand-alone mode, and the steady of the Cluster mode is prior.

Summary

This paper analyzes the important significance of the improved efficiency of theoretical line loss calculation to promote the development of green smart grid, and design theoretical line loss analysis framework based on distributed memory cluster, the framework combine Hadoop and spark combination, and improve the capabilities of storing and analyzing electric power big data. This paper conducts line loss analysis experiment by P-Q method, the results show that the feasibility

and efficiency of this scheme, which have a certain help and reference value in analysis of theoretical line loss online.

References

[1] Wang Dewen. Basic framework and key technology for a new generation of data center in electric power corporation based on cloud computation [J]. Automation of electric power systems, 2012, 36(11): 67-71. (in Chinese).

[2] Kenneth P Birman，Lakshmi Ganesh，Robbert　van Renesse．Running smart grid control software on cloud computing architectures[C]// Workshop on Computational Needs for the Next Generation Electric Grid, Cornell University．Ithaca，NY：DOE，2011：1-28．

[3] Wang Dewen, Song Yaqi, Zhu Yongli. The smart grid information platform based on the cloud computing[J]. Automation of electric power systems, 2010,34(22):7-12.(in Chinese)

[4] B.Genge, A.Beres, P.Haller.　A survey on cloud-based software platforms to implement secure smart grids[C]. Power Engineering Conference, 2014 49th International Universities, Cluj-Napoca:IEEE , 2014:1-6．

[5] M.Mayilvaganan,　M.Sabitha, A cloud-based architecture for Big-Data analytics in smart grid: A proposal[C]. 2013 IEEE International Conference on Computational Intelligence and Computing Research, Enathi:IEEE. 2013:1-4.

[6] Wang Yonghua, Wang Zhengfeng. Research and Application of measured load and line loss calculation on-line based on SCADA/EMS [J]. Power System Protection and Control, 2012, 40 (4): 96-100

[7] Wang Xiaowei, Hu Wenping. Provinces integrated computing platform for online line loss analysis [J] Power construction, 2014,35 (1): 62-67

[8] Zhu Ke. Design and Implementation of on-line line losses Statistical Analysis System based on visualization [D] Xi'an: Electronic science and technology university, 2012: 1-21