EAI Endorsed Transactions
on Scalable Information Systems
Research Article ICST.ORG

Design Private Cloud of Oil and Gas SCADA System

Liu Miao[1,*], Mancang Yuan[2] and Guodong Li[3]

[1] China Petroleum Longhui Automation Engineering Co., Ltd., China Petroleum Pipeline Bureau, Langfang, China, The 10 layer, No.1 building, Guanganmennei Street 311, Beijing, China
[2] China Petroleum Longhui Automation Engineering Co., Ltd., China Petroleum Pipeline Bureau, Langfang, China, The 10 layer, No.1 building, Guanganmennei Street 311, Beijing, China
[3] China Petroleum Longhui Automation Engineering Co., Ltd., China Petroleum Pipeline Bureau, Langfang, China, The 10 layer, No.1 building, Guanganmennei Street 311, Beijing, China

Abstract

SCADA (Supervisory Control and Data Acquisition) system is computer control system based on supervisory. SCADA system is very important to oil and gas pipeline engineering. Cloud computing is fundamentally altering the expectations for how and when computing, storage and networking resources should be allocated, managed and consumed. In order to increase resource utilization, reliability and availability of oil and gas pipeline SCADA system, the SCADA system based on cloud computing is proposed in the paper. This paper introduces the system framework of SCADA system based on cloud computing and the realization details about the private cloud platform of SCADA system.

Keywords: SCADA system; cloud computing; private cloud platform.

Received on 09 January 2014, accepted on 02 April 2014, published on 27 May 2014

doi: 10.4108/ sis.1.3.e5

1. Introduction

SCADA (Supervisory Control and Data Acquisition) system is computer control system based on supervisory. The SCADA system of oil and gas pipeline can manage sequential control transmission of petroleum pipeline, equipment monitoring, data synchronization transmission record and monitoring operation conditions of every station control system. In addition, the SCADA system has more features, such as leaking detection, system simulation, water hammer protection in advance, and so on. SCADA system can continuously monitor equipments which are scattered over the wide regions and it can operate remote devices from the control center so that the operating efficiency is improved, the energy is saved and the cost is reduced. And, the SCADA system can guarantee the integrity of pipeline by continuous monitoring the key parameters of system, such as pressure, flow, oil tank liquid level and so on [1]. So, SCADA system is very important to oil and gas pipeline engineering.

Cloud computing, as a current commercial offering, started to become apparent in late 2007 [2]. It was intended to enable computing across widespread and diverse resources, rather than on local machines or at remote server farms. Although there is no standard definition of Cloud Computing, most authors seem to agree that it consists of clusters of distributed computers (Clouds) providing on-demand resources or services over a network with the scale and reliability of a data centre [3]; notions familiar from resource virtualization and Grid computing. Where these clusters supply instances of on-demand Cloud computing; provision may be comprised of software (e.g. Software as a Service, SaaS) or of the physical resources (e.g. Platform as a Service, PaaS). The Amazon Elastic Compute Cloud (Amazon EC2) [4] is an example of such an approach, where a computing platform is provided. In common with many commercial approaches *provision* is the primary objective; management and governance handled via redundancy or replication, scaling capacity up or down as required. In contrast the authors proposed a Cloud Coordination framework in 2005 with the notion of a Cloud being a system of loose boundaries, which interacts and merges with other systems [5]. This definition of a Cloud is refined to a federation of interacting services and resources, which share and pool resources for greater efficiency. Thus governance, in general, and scalability are handled as part of the separated coordination framework.

*Corresponding author. Email: lhbj_liumiao1@cnpc.com.cn

In order to increase resource utilization, reliability and availability of oil and gas SCADA system, the SCADA system based on cloud computing is proposed in the paper. The remainder of this paper is organized as follows. In Sect. 2, traditional structure of SCADA system is described. The framework of cloud computing and the key features and characteristics of cloud computing are detailed in Sect. 3. Section 4 the system architecture of SCADA system based on cloud computing is proposed. In Sect. 5, the realization details about the private cloud platform of SCADA system is described. Finally, the paper concludes in Sect. 6.

2. Traditional structure of SCADA system

The SCADA system adopts distributed supervisory control and centralized management. The traditional structure of SCADA system is described in figure1.The system is organized by the supervision center, a lot of site control systems and communication medium. The supervision center (master station) is the core of system and in charge of controlling and managing system running. It is organized by a lot of data processing servers and database servers. The outside site is the intelligent measure and control module by microprocessor or DSP. It can collect and process the data of remote sites, control local sites and communicate with remote supervision center.

The SCADA system includes: information acquisition subsystem, information transmission subsystem and information processing subsystem. The information acquisition subsystem includes all kinds of sensors and controllers. It takes charge of data collection and devices controlling of local site. The information transmission subsystem can transport collected data and system controlling signal of local site by efficient communication technologies. It is the bridge of information acquisition subsystem and information processing subsystem. It includes RTU (Remote Terminal Unit), modem and central communication controller. The information processing subsystem is located at supervision center and it can gather, process and analysis the data of remote sites. It includes master computer and all kinds of expert software.

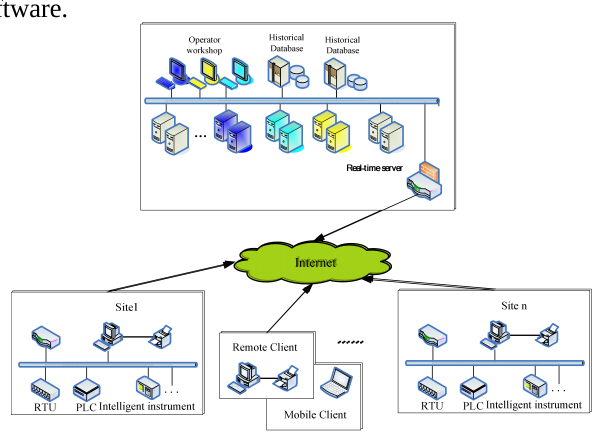

Figure1. The structure of SCADA system

However, the current SCADA system has problems as follow:

(1) Reliability problem

Servers of supervision center are the command center of system. They can gather the pipeline running data, state and warning information and provide the real-time database for all operating sites. Now, the SCADA system of oil and gas pipeline adopts Hot Standby as the way of redundant configuration to ensure the reliability of system. When master server is working online, the slave server monitors work state of master server and get the data from master server for keeping the data consistency between the master server and slave server. Once the master server breaks down, the slave server takes over the work from the master server immediately, and the slave server becomes the master server, the repaired master server becomes the slave server. As the controlling core of SCADA system, the central servers are responsible for data collecting and controlling of all line. It is important to ensure the data transmission between the servers and site control system for all SCADA system network normal operation. The single backup server cannot ensure the reliability of system completely.

(2) Source wasting problem

Hot Standby is used for ensuring the reliability of system. With system scaling out constantly, the number of server increases. Redundant configuration can cause wasting server source and the workload of operation administration and maintenance increasing. So, it is necessary to find more advanced resource configuration mode to increase resource utilization rate.

(3) Load balancing problem

With SCADA system of oil and gas pipeline scaling out constantly, the number of server increases. There is lots of operation and data are processed on different servers. However, CPU load is different with different site data. So, different servers have different load, some servers overload but some servers just are idle. It is necessary to study new load balancing strategy for improving the performance of system.

3. The framework and the key features of cloud computing

Generally, the architecture of a cloud computing environment can be divided into 4 layers, as shown in Fig. 2. [6].

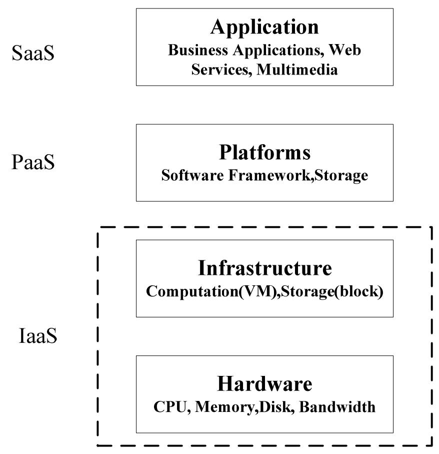

Figure2. The architechture of cloud computing

The hardware layer: This layer is responsible for managing the physical resources of the cloud, including physical servers, routers, switches, power and cooling systems. In practice, the hardware layer is typically implemented in data centers. A data center usually contains thousands of servers that are organized in racks and interconnected through switches, routers or other fabrics. Typical issues at hardware layer include hardware configuration, fault tolerance, traffic management, power and cooling resource management.

The infrastructure layer: Also known as the virtualization layer, the infrastructure layer creates a pool of storage and computing resources by partitioning the physical resources using virtualization technologies such as Xen [7], KVM [8] and VMware [9]. The infrastructure layer is an essential component of cloud computing, since many key features, such as dynamic resource assignment, are only made available through virtualization technologies.

The platform layer: Built on top of the infrastructure layer, the platform layer consists of operating systems and application frameworks. The purpose of the platform layer is to minimize the burden of deploying applications directly into VM containers. For example, Google App Engine operates at the platform layer to provide API support for implementing storage, database and business logic of typical web applications.

The application layer: At the highest level of the hierarchy, the application layer consists of the actual cloud applications. Different from traditional applications, cloud applications can leverage the automatic-scaling feature to achieve better performance, availability and lower operating cost. Compared to traditional service hosting environments such as dedicated server farms, the architecture of cloud computing is more modular. Each layer is loosely coupled with the layers above and below, allowing each layer to evolve separately. This is similar to the design of the OSI model for network protocols. The architectural modularity allows cloud computing to support a wide range of application requirements while reducing management and maintenance overhead.

4. The system framework of SCADA system based on cloud computing

Compare to traditional structure of SCADA system, the SCADA system based on cloud computing used cloud computing technologies to integrate server resource of supervision center, manage and assign this resource uniform, but not allocate a certain server to be in charge of a certain pipe line. In addition, the local site control systems are redeployed at the cloud server of remote private cloud platform. The structure of SCADA system based on cloud computing is shown in Fig.3.

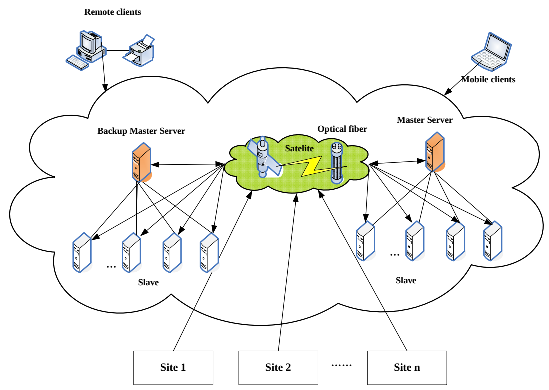

Figure3. The structure of SCADA system based on cloud computing

The virtualization technology enables the abstraction or decoupling of the application payload from the underlying physical resource. The physical resource can then be carved up into logical or virtual resources as needed. This is known as provisioning. With server virtualization, we can create complete logical (virtual) servers that are independent of the underlying physical infrastructure or their physical location. We can specify the computing, network and storage resources for each logical server (virtual machine) and even move workloads from one virtual machine to another in real-time (live migration). Every practitioner of server virtualization is aware of how virtualization can result in “Virtual Machine (VM) Sprawl” and the associated management burden it creates. VM Sprawl is a result of the ease with which new VMs can be created and proliferated on virtualized servers. This is however not the only factor affecting management complexity. Since VMs are cheaper to setup and run than a physical server, load balancers, routers and other applications that required physical servers are all now being run as in VMs within a physical server. Consequently, we now have to manage and route network traffic resulting from all these VMs within a server as well as the network traffic being routed across server. Adding to this confusion are the various OS vendor, each offering the other vendor's OS as a "guest" without providing the same level integration services. This makes the real life implementations, very management intensive, cumber some and error-prone to really operate in a heterogeneous environment. It is our contention that the cost of this additional management may yet offset any cost benefits that virtualization has enabled through consolidation. A traditional management system with human-dependency is just an untenable solution for cloud computing.

The SCADA system model of oil and gas pipeline is created as follow:

The real-time data processing tasks of SCADA system are expressed by $T=\{T_1, T_2, ..., T_n\}$. Each task of SCADA system is independent and the task is non-preemptive.

Each real-time task can be defined as

$$T_i=(d_i, T_i^1, T_i^2, T_i^3) \quad \textbf{(1)}$$

d_i is the deadline of real-time task; T_i^1, T_i^2, T_i^3 is the one primary copy and two backup copies, respectively; the code of three tasks is same completely.

$$T_i^1=(C_i, s_i^1, \rho_i^1) \quad \textbf{(2)}$$

$$T_i^2=(C_i, s_i^2, \rho_i^2) \quad \textbf{(3)}$$

$$T_i^3=(C_i, s_i^3, \rho_i^3) \quad \textbf{(4)}$$

There, s_i^1 , s_i^2 and s_i^3 is the time for tasks to be began running; ρ_i^1 , ρ_i^2 and ρ_i^3 are processors assigned to three task copies.

There are different execution time for different processors so defining a computing time vector for each task T_i as follow,

$$C_i=[c(i,1), ..., c(i,m)] \quad \textbf{(5)}$$

There c_{ij} represents the execution time for T_i^1, T_i^2, T_i^3 on processor P_j.

If SCADA system can tolerate 2 processors disable, the sum of primary copy execution time and backup copy execution time should be less than or equal to deadline.

$$\forall i \in [1,n], (c(i,\rho_i^1)+c(i,\rho_i^2)) \le d_i \wedge (c(i,\rho_i^1)+c(i,\rho_i^3)) \le d_i \quad \textbf{(6)}$$

For each processor P_i of SCADA system

$$\forall i \in [1,m], \sum_{T_i^1 \in \Delta_j} c(i,j)+\sum_{T_i^2 \in \Delta_j} c(i,j) \le d_i, \sum_{T_i^1 \in \Delta_j} c(i,j)+\sum_{T_i^3 \in \Delta_j} c(i,j) \le d_i \quad \textbf{(7)}$$

5. Realization details about the private cloud platform of SCADA system

The physical architecture of Private cloud includes: controller node, compute nodes, NFS server and network devices. It is shown in Fig.4.

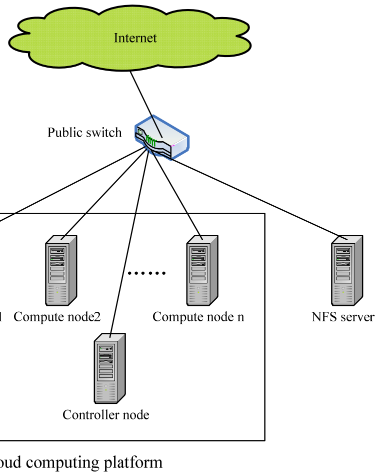

Figure4. The physical architecture of private cloud

The function of Controller node includes showing, monitoring, management and scheduling. The controller node provide a database and web server to support the management interface of system. The users issue control commands and operate the system by the management interface of the controller node. Information of other nodes must register to the controller node and the controller node schedule the resource constantly.

Compute nodes is the operating foundation of virtual machine. They are in charge of the specific operating work of private cloud system. Compute nodes compose the resource pool together and provide the resource with form of virtulization. The virtual machines which privat cloud system provide for outside users to use run in compute nodes. Compute nodes are dominated and managed by the controller node.

NFS server provides NFS service which supports dynamic extension function of private cloud system.

Network devices include routers and switches. They connect every devices of private cloud system and access to network. In addtion, they take charge in network address assignment.

The logical architecture of Private cloud is shown in Fig.5.

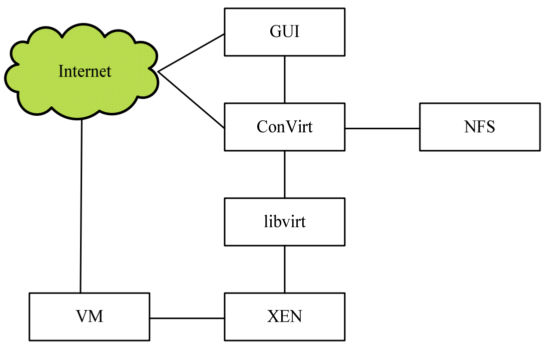

Figure5. The logical architecture of private cloud

Graphic interface management is the external interface of system. The users manage and operate the system by GUI. It is the web service based on web server and database. Users can enter the management interface by accessing to the port of related server via internet without setting up client. In addition, users can access and manage all system, anytime and anywhere, by all kinds of network terminal to master running state of system and find and solve the problem in time.

ConVirt is the core of system. It provides all kinds of data to management interface and monitor all system. Also, it is the real manager of system. It converts all operations on management interface by users to computer instruction and sends the instruction to every execution parts of system. ConVirt is programmed by python and it is fit for all kinds of computer system. NFS is network file system. The private cloud can copy and migrate virtual machine conveniently without redeploying by the function of shared catalog and files provided by NFS.

Libvirt provides a set of reliable and convenient interface for all kinds of virtualization tools to shield the difference of instruction of virtualization tools.

Xen is a virtualization solution. It can realize virtualization function of x86 framework. The virtual machine of private cloud system is made by Xen directly. Users operate the virtual machine on graphic management interface is passed on layers and layers by ConVirt converting to computer instructions. The running condition of virtual machine is collected and fed back to users by ConVirt.

6. Conclusion

Oil and gas pipeline project is an industrial system with highly systemic. The running data of pipelines can be obtained real-time, processed scientifically and made decision by SCADA system. For improving the reliability and resource utilization of traditional oil and gas pipeline SCADA system, the application framework of oil and gas pipeline SCADA system based cloud computing is proposed and the system model is designed in the paper. This application framework can realize real-time data multi-copy execution and history data multi-restore. So, the new application framework is with more reliability and higher resource utilization. Next, the real-time scheduling algorithm for the SCADA system based on cloud computing can be researched.

References

[1] Li Yang, Xiedong Cao, Research on fnn-based security defense architecture model of SCADA network, Proceedings of IEEE CCIS.2012.pp.1829-1833.
[2] Armbrust M, Fox A, Griffith R, et al. A view of cloud computing. Communications of the ACM, 2010, 53(4): 50-58.
[3] Ekanayake J, Fox G. High performance parallel computing with clouds and cloud technologies. Cloud Computing. Springer Berlin Heidelberg, 2010: 20-38.
[4] Ostermann S, Iosup A, Yigitbasi N, et al. A performance analysis of EC2 cloud computing services for scientific computing. Cloud Computing. Springer Berlin Heidelberg, 2010: 115-131.
[5] Zhang Q, Cheng L, Boutaba R. Cloud computing: state-of-the-art and research challenges. Journal of internet services and applications, 2010, 1(1): 7-18.
[6] Qi Zhang, Lu Cheng and Raouf Boutaba, Cloud computing: state-of-the-art and research challenges, J Internet Serv Appl, 2010, pp. 1:7-18
[7] XenSource Inc, Xen, www.xensource.com.
[8] Kernal Based Virtual Machine, www.linux-kvm.org/page/MainPage.
[9] VMWare ESX Server, www.vmware.com/products/esx.