Cloud computing based mountain flood disaster monitoring and forewarning platform

Yunjia Niu, Hua Yan*

College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China

Received 1 October 2014, www.cmnt.lv

Abstract

Aiming at the disadvantages of undeveloped data disaster recovery and inefficient data sharing measures of existing mountain flood disaster monitoring and forewarning platforms, this paper analyzes the defects and designs a new platform architecture based on cloud computing. From the perspective of practical application, a customized data disaster recovery plan is provided. To integrate platform resources, a design of SaaS architecture and applications is presented afterwards. The introduction of cloud computing provides a new way for the construction of mountain flood monitoring and forewarning platform.

Keywords: cloud computing, mountain flood disaster, data disaster recovery, data sharing

1 Introduction

Mountain flood disaster is an abrupt geological hazard with strong destructive power, which is difficult to predict and forewarn. Triggered by rainfall, mountain flood and debris flow directly threat people's lives and property every year in China. Enormous damage to public facilities causes serious losses to local economy. However, existing mountain flood disaster monitoring and forewarning platforms are becoming inadequate for the flood control departments as the amount of information grows hugely with the development of information technology.

With the advent of the big data era, cloud computing is becoming a popular choice for massive data storage, processing and sharing. Cloud computing can easily handle TB-level data with its reliable, scalable and easily maintainable services. The purpose of this study is to provide a design that applies cloud computing to existing mountain flood disaster monitoring and forewarning platforms, in order to enable the flood control departments to make decisions and deployments more timely and accurately.

2 Analysis of existing mountain flood disaster monitoring and forewarning platforms

Hydrological and meteorological services involve a huge amount of data and various applications and it requires service diversification with high reliable data storage. Presently, there are mountain flood disaster prevention projects under construction in 175 cities and counties in Sichuan Province, China, but resources such as databases, servers, monitoring stations, etc. are run separately by each county or city, which exposes obvious deficiencies as follows.

1) Undeveloped data disaster recovery measures.

As the reliance of business requirements on information technology grows, it becomes even harder for traditional data storage schema in separated databases to deal with a blooming amount of mountain flood information. Sudden occurrence of serious natural disasters, equipment failure or incorrect operation may all result in loss of data or even worse. Simple data backup or local protection can hardly meet the needs for data recovery or the continuity of platform services.

2) Inefficient data sharing mechanism

The adequateness of data sharing has decisive influences on platform performance. However, the scattered platform distribution creates data gaps and information asymmetry between flood control departments. Only with certain standards for database format and message transfer mode, the technology implementations from different developing contractors are not unified. The communication quality between departments is reduced. The inappropriate mechanism of data sharing and configuration of software and hardware facilities lead to difficulties in platform performance, maintenance and management, which results in a huge waste of human and mater resources.

3 Mountain flood disaster monitoring and forewarning platform based on cloud computing

3.1 PLATFORM ARCHITECTURE

Cloud computing is the delivery of computing as a service rather than a product, whereby shared resources, software, and information are provided to computers and other devices as a utility (like the electricity grid) over a network (typically the Internet) [1]. Resources in the cloud can be invoked and released on demand, which complies with the seasonal periodicity and exigency of scaling the size of

* ***Corresponding author's*** e-mail: yanhua@scu.edu.cn

data and services in monitoring and forewarning business.

Based on cloud computing, the conceptual architecture design of mountain flood disaster monitoring and forewarning platform (Figure 1) includes:

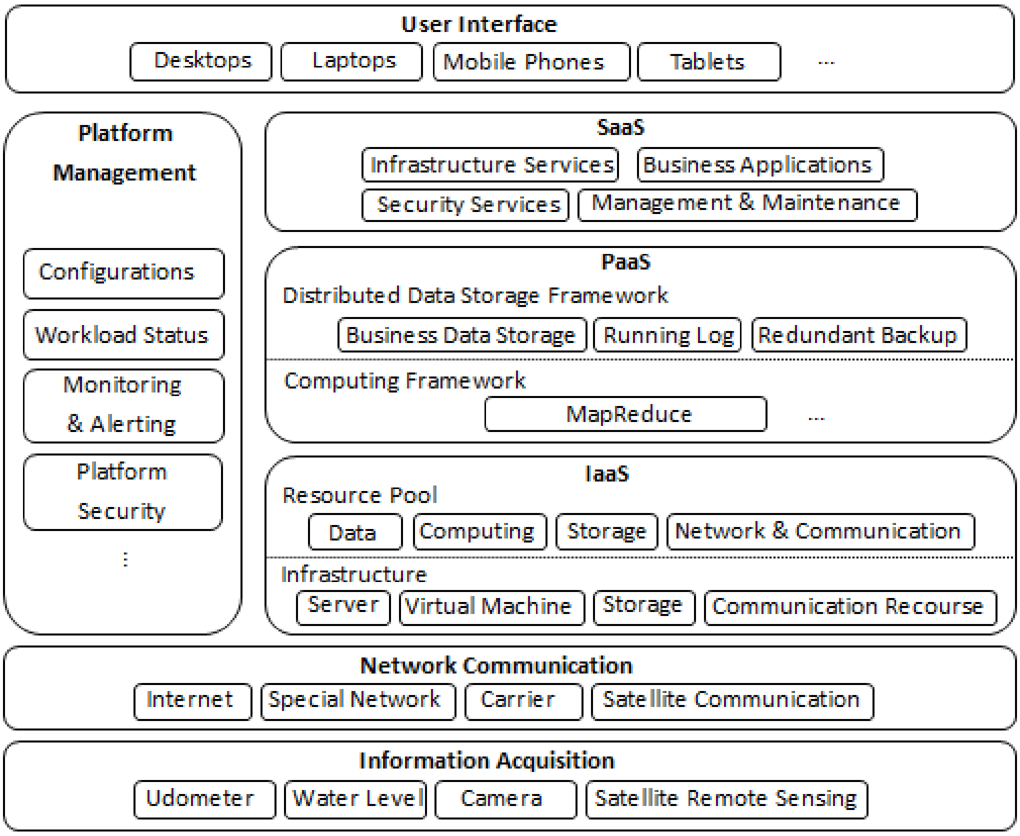

FIGURE 1 Conceptual architecture design of mountain flood disaster monitoring and forewarning platform based on cloud computing

1) Information Acquisition. Multiple data sources of mountain flood.

2) Network Communication. The connection between the platform and required information through networks from multiple providers.

3) Cloud Computing Service. Cloud computing presents the hardware devices, software environment, development tools, applications, etc. as services, by realizing an abstract encapsulation of all the mentioned elements. The taxonomy of cloud computing services typically includes:

a) Infrastructure as a service (IaaS). On-demand virtualized resource pools of infrastructure.

b) Platform as a service (PaaS). On-demand virtual environment for software running and platform development, along with reliable storage and computing framework for further development.

c) Service as a service (SaaS). On-demand customized software applications.

4) Platform Management. Operations and capabilities to keep the platform running in order and functioning properly.

5) User Interface. Ubiquitous interface for the platform users over networks with run-time software services for cloud execution.

3.2 RELEVANT KEY TECHNOLOGIES OF CLOUD COMPUTING

3.2.1 Virtualization and Elastic Computing Service

Virtualization builds a virtual environment for operating system, application programing interface, storage device and network resources, etc. to run like in a real computer [2]. Upper features will no longer rely on underlying hardware implementation. Only a small amount of physical machines are able to support large numbers of virtual machines. Virtualization can provide the departments integrated massive heterogeneous platform resources on-demand in one “cloud”, which greatly improves platform resource utilization.

Elastic computing service refers to the dynamic variation in the use of computer resources to meet a varying workload [3, 4], which depends largely on the virtualization technology. Performing as “one” cloud, cloud computing centers can be built according to local business size and geographical condition. All the computing resources are transferable among these centers through communication networks. It is much more flexible in responding to the sharp increase of workload during flood season, and reduces unnecessary costs of hardware facilities.

3.2.2 Apache Hadoop

Apache Hadoop is an open-source software framework for storage and large-scale processing of data-sets on clusters of commodity hardware [5], which fits the requirement of data scalability from the monitoring and forewarning services and applications. The Apache Hadoop framework composing modules includes Hadoop Distributed File System (HDFS) and Hadoop MapReduce, etc.

A direct solution for enhancing the capacity and improving the service quality of a platform with a great amount of heterogeneous mountain flood data is to adopt a highly efficient and reliable schema of data storage. HDFS provides distributed file system for massive data, whereby big data blocks are divided into small ones and multiple copies of each block can be created for redundant backup to ensure data availability [6]. Therefore, monitoring and forewarning systems are able to run upon low-cost, consumable hardware facilities. At the same time, the fast error detecting and resuming mechanism of HDFS gives the platform data disaster tolerant ability.

MapReduce is a programming model with a parallel, distributed algorithm on a cluster. Briefly, it breaks large-scale data process into two steps - Map and Reduce. Map task divides the input into smaller parts and processes them in parallel. Reduce task collects the results and combine them properly to form the output. Running status of MapReduce is monitored by its own program, and failed tasks can be resumed [7].

4 Applications of key technologies

4.1 DATA BACKUP AND DISASTER RECOVERY

Considering the budget limits and particularity of mountain flood disaster monitoring and forewarning business, the local backup and remote disaster recovery of data is essential. By adopting the strength of cloud computing, as of virtualization, certain data back and disaster recovery plans are purposed.

1) Storage-Virtualization based Local Data Backup. A lower-cost data backup solution on local virtual machines, especially for data with lower priority and longer expected recovering time.

2) Remote Data Disaster Recovery in Different Areas. A solution for datacenters with unstable geographical conditions. A data disaster recovery center will be set at a safer place. Over communication networks, data replicates asynchronously. Platform services will be switched to its disaster recovery center, if its original datacenter is damaged. This plan also suits data with shorter expected recovering time and higher security requirements.

3) Mirror Data Disaster Recovery in the Same Area. A solution for core platform services to run continuously where data is highly synchronized. The original and disaster recovery datacenter are built within certain distance and deployed under one cloud for fast switching.

Especially for departments with higher strategic needs, an integrate plan of three-center-two-area is recommended. Remote and mirror data disaster recovery center can both be constructed, with the support of local storage-virtualization backup. Specific data synchronization strategy accords to data priority.

4.2 PLATFORM RESOURCES INTEGRATION AND SHARING

As the user interface of the cloud computing platform, SaaS not only relies on, but also provides access to the underlying platform resources. Data is not local for the users, which expands the workplace, simplifies the office procedure, and to some extent, strengthens data security. For users, SaaS integrates the whole platform with fully shared resources and presenting them as services.

1) Infrastructure Services. The underlying computing and network capabilities provided as services.

2) Monitoring and Forewarning Applications. Software applications integrated or developed as services upon the business request of mountain flood disaster monitoring and forewarning. Overlapping application development is avoided.

3) Operation Management and Maintenance. The capabilities for SaaS applications running efficiently and stably on track.

4) Security Services. The capabilities needed to keep SaaS applications running safely and used within permission.

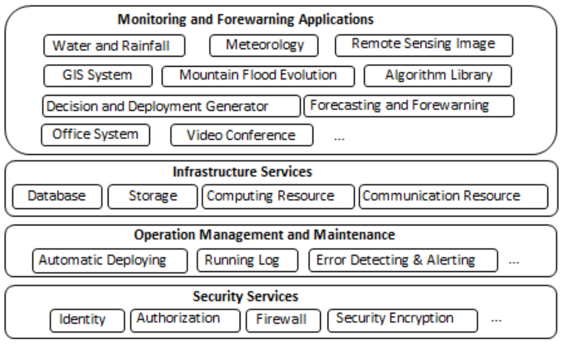

FIGURE 2 Conceptual SaaS architecture and applications

5 Conclusion and prospect

From angles of data disaster recovery and sharing, cloud computing is able to enhancing capabilities and stability of existing mountain flood monitoring and forewarning platforms. Scattered heterogeneous platform resources can be integrated, and provided by SaaS as serial services. Together with customized data disaster recovery plans, the platform will surely be more efficiently and powerful. As users of the cloud computing services, flood control departments can be more focused on dispatches and commands of flood prevention and evacuation, rather than being distracted by irrelevant burdens such as upfront infrastructure costs, management and maintenance.

Considering the geological condition of the facilities of mountain flood preventing, the stability of communication link demands prompt solutions. And to improve the platform capability, the optimization of information security and variation of application integration are expected.

References

[1] Mell P, Grance T 2011 The NIST definition of cloud computing *Press: NIST Special Publication* 800 145

[2] Zhu Z, Gu Z, Wu J, Gui S 2012 Application of cloud computing in electric power system data recovery *Power System Technology* **36**(9) 43-50

[3] Buyya R, Broberg J, Goscinski A 2011 CLOUD COMPUTING Principles and Paradigms *Press: John Wiley and Sons*

[4] Perez et al 2009 Responsive Elastic Computing *ACM/IEEE Conference on International Conference on Autonomic Computing (2009)* 55-64

[5] Apache Hadoop, http://zh.wikipedia.org/wiki/Apache_Hadoop

[6] Jin S, Yang S, Fan H, Liu F 2012 Research on data security policy of hadoop cloud computing platform for large-scale critical business *Netinfo Security* 90-3

[7] Luo J, Jin J, Song A, Dong F 2011 Cloud computing: architecture and key technologies *Journal on Communications* **32**(7) 3-21

Authors

Yunjia Niu, September 22.1989, Sichuan, China.

Current position, grades: graduate student of Pattern Recognition and Intelligent System (Sichuan University).
University studies: Bachelor of Electronic and Information Engineering (Sichuan University), 2008.
Scientific interest: cloud computing; pattern recognition.

Hua Yan, March 13.1971, Sichuan, China.

Current position, grades: associate professor of pattern recognition and intelligent system at Sichuan University.
University studies: BSc of Electronic engineering (Sichuan University) in 1993, MSc of Radio electronics (Sichuan University) in 1996, PhD of Machinery Manufacturer (Sichuan University) in 2008.
Scientific interest: water conservancy informatization.