Development of Information Technologies for Storage of Data of Instrumental Observation Networks of the Far Eastern Branch of the Russian Academy of Sciences

Aleksei A. Sorokin\textsuperscript{1,*}, Sergey P. Korolev\textsuperscript{1}, and Andrey N. Polyakov\textsuperscript{2}

\textsuperscript{1} Computing Center of the Far Eastern Branch, Russian Academy of Sciences, Khabarovsk, Russia
alsor@febras.net

\textsuperscript{2} National Research Centre "Kurchatov institute", Moscow, Russia
andrew@kiae.ru

Abstract
The paper describes the development of information technologies for distributed storage of data of instrumental observation networks created by the Far Eastern Branch of the Russian Academy of Sciences (FEB RAS). The prime goal of the distributed storage design is to process a large amount of real-time seismological data and provide fast and secure access using a cloud storage platform based on OpenStack Swift. On its basis, the allocation of the corresponding data storage has been done. The interaction with storage using existing application software as well as an assessment of the effectiveness of the solution is described.

Keywords: data files, databases, cloud computing, network of instrumental observations, structured scientific datasets, archive, automated information system, application programming interface.

1 Introduction

During the past 5 years several types of instrumental observation networks have been created and developed by the Far Eastern Branch of the Russian Academy of Sciences (FEB RAS) for research and monitoring of natural hazards [4, 12, 9]. The network of seismological observations (hereinafter referred to as ”Network”) [4] is one of the most important. The hardware part of its measuring equipment is currently represented by five contemporary land-based stations equipped with 3-component broadband seismometers RefTek 151-120, digital recorders RefTek 130-01 and GPS-receivers. During observations, baseline measurements are recorded in files of 60 seconds in length in the PASSCAL binary format [14], which are then transferred...
to the Computing Center, FEB RAS (CC FEB RAS) in deferred mode or in real time via communication using the network layer protocol RTP [15].

As per day result, the average 1440 files with the total size of 50 MB are transmitted from each observation site to CC FEB RAS. The size of raw data archive accumulated over the entire period of network operation is about 200 GB, with the total number of files more than 5 million units.

With the development of the Network, its gradual transition to the operational mode and, accordingly, with the growth of transmitted and processed data, the problem arises related to developing a fault-tolerant and scalable system for instrumental seismic data acquisition and storage as well as ensuring access to the accumulated observation data archives.

2 Technologies

Today, cloud storage systems are mainly represented by proprietary commercial systems eg., Amazon S3 [10], Microsoft Azure [7], etc. [6]. To access them, a special software client or application programming interface (API) are required. Based on these systems, different public and hybrid cloud services have been created for storage of information (eg., Dropbox, OneDrive, etc.) [8, 13].

However, using systems of a like nature, you cannot build private cloud storage and provide cloud storage refinement to the needs of currently existing application information system. Besides, in the case of seismological observation network of FEB RAS, it is required to provide a guaranteed limited access to the scientific information when working with the data. Therefore, using such solutions to resolve the problems considered is impossible.

Presently, cloud storage OpenStack Swift [11] is the only software solution devoid of the deficiencies listed above. Swift allows organizing storage of information on commodity hardware, including geographically distributed storage, in which data integrity is ensured by regular replication of information. In Swift, metadata are stored together with the objects as in GlusterFS [1], and distributed storage technique uses RESTful API, which allows cloud storage integration with information systems.

In this section, we present cloud-based architecture which allows us to provide virtual computing resources with the required characteristics and an exible mechanism for their integration and interoperability with existing applications of information systems. This work involves using scalable hardware and software systems with high availability, user authorization system, and providing reliable access to data to perform low-level operations. OpenStack Swift was selected as such a tool and free software.

Based on the data center infrastructures of CC FEB RAS (Khabarovsk) and the NRC "Kurchatov Institute" (Moscow), two Swift zones were deployed with four compute nodes that had different hardware configurations and network bandwidths. The general requirements included the use of Intel processors with clock frequency of no less than 2.9 GHz, 4 GB of RAM and two hard drives to operate the system software and direct storage. Communication between the zones was implemented through existing channels to access the Internet (Figure 1). As a result, the cloud storage was organized, providing Infrastructure-as-a-Service (IaaS).

Another challenge was the integration of the cloud storage infrastructure with the information system network management observations. Currently, the network management and work with the instrumental data archive are conducted using an automated information system "Signal-S" (hereinafter referred to as "AIS") [16, 3]. It is based on the "client-server" architecture, where clients are digital recorders installed at observation sites, and the server - the relevant hardware and software components of the AIS "Signal-S".

585
3 Distributed storage of instrumental data

The basic unit of storage in the cloud is an object that represents a file with information. To systematize the notion of a group of objects, a container concept is used that is similar to a directory within the file system. For each uploaded file, Swift creates a link between the file location and its name in the SQLite database. Depending on the system configuration, at least three copies of an object are stored at different compute nodes. This ensures their safety in the case of hardware failure. Any object can be equipped with meta-information stored in the file’s extended attributes (xattr).

Using the PHP programming language and the Yii framework [2], we can create applications based on a template Model-View-Controller. A modification of the AIS has been developed and named "Signal-Cloud". It consists of five interrelated software modules designed to work with the organized cloud storage through existing AIS user interfaces.

Modules using the API HTTP RESTful for OpenStack Swift are implemented as a library for many programming languages (PHP, Java, Python, etc.). They perform the whole set of operations on creating and removing containers, loading and unloading objects and forming in a cloud identical copies of instrumental seismological data files (Figure 2).

The instrumental data obtained from the observation network are placed in the cloud as a pseudohierarchical file system. This is because the Swift container cannot implement a default directory structure of regular file system, and all operations are carried out strictly with binary objects. To move the necessary storage structure of the files and directories to the cloud, they are implemented as zero size and data type "application/directory" objects called directory markers. These markers are created at the stage of uploading the instrumental data to the cloud. They allow further making queries for the required data by AIS modules according to the specified parameters (e.g., daily observation site, etc.) with subsequent data processing and the access through user interfaces.

The described structure is required by design in AIS for the following reasons:

1. Due to default RefTek data archive, file naming conventions filenames for different stations could be the same, causing conflict and ambiguity. Prefixing by directory names helps to
2. Having filenames in the form of `<subdirectory>/<filename>` gives flexibility in metadata usage both for regular file system (where files are actually located in subdirectories) and cloud archive (where files are located as binary objects with directory markers). The relative file path will be the same in both cases.

4 Information System Testing

To test the operation of the developed data processing system, a benchmark dataset was formed. It consisted of a four-day archive of instrumental seismological data recorded at 3 stations of...
FEB RAS observation network (Chegdomyn, Vanino and Uglegorsk) from 4 to 7 June 2012 with the total number of files (objects) 17,260 and a volume of data 629399552 bytes.

In order to assess the reliability, speed of access and software system scalability, some testing methods have been developed and a study of system performance has been carried out, including operations time measurement for loading and unloading from the cloud reference dataset.

Testing was carried out by the formation of data requests through appropriate forms of AIS user interface (Figure 3), the components of which are located in CC FEB RAS.

The obtained test results are shown in Table 1:

Based on these data, the following main conclusions have been made:
<table>
<thead>
<tr>
<th>Data source</th>
<th>Data upload, seconds</th>
<th>Data download, seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRC &quot;Kurchatov Institute&quot;</td>
<td>8670</td>
<td>6527</td>
</tr>
<tr>
<td>CC FEB RAS</td>
<td>7594</td>
<td>2613</td>
</tr>
</tbody>
</table>

Table 1: Test results.

![Graph showing linear dependency between processing time and data processed.](image)

Figure 4: Test results for data fetch of CC FEB RAS cloud zone.

1. The full time of downloading the entire dataset includes the time spent on compiling a list of data files for the requested period and the overhead time associated with processing and transferring of the data within the application system and in the cloud. It can be reduced by using more efficient interaction technologies and algorithms for data processing within integrated systems. To study the performance of data in the cloud zone in CC FEB RAS, a bunch of tests were executed. Various amounts of data were fetched from the cloud using the AIS interface wizard (Figure 4). The results obtained are shown in Figure 4. The download time has almost linear dependency on the amount of data fetched from the cloud nodes.

2. The big difference (7,594 sec. vs 2,613 sec. = 2.9 times) between the data download to cloud time and upload from cloud time exists because when the first operation is executed, a large number of objects (17260 units) are created and SQLite-base of their corresponding meta-information is populated.

3. Using the remote computing resources of nodes increases the time to perform operations.
Largely, this is due to the natural time-consuming data transfer between the cities of Khabarovsk and Moscow (the observed time of data transfer attained to 143 ms, on average), where the storage nodes are located. This parameter is greatly influenced by utilization of common Internet channels. Different causes, that are not dependent on designers of distributed storage system, may exert influence on the quality of channels work (non-optimal IP-routing scheme, different bandwidth of aggregate links, etc.) As for time-critical applications, it is recommended to set the highest priority to use the nodes that are closest to the user and provide higher speed data zones when configuring repository zones. Proxy server of OpenStack Swift in interaction with the storage nodes will use these rules. In addition, one can make use of telecommunication channels layer 2 that provide a guaranteed bandwidth and for which it is possible to install QoS to perform data transfer between the given nodes.

4. To determine the resiliency of the complex during the data download from the cloud, two of three computing nodes were disabled at the same time, but the operation was not interrupted and the data download was successfully completed. This characteristic is important and necessary when working with data coming into the information system in real time.

5 Conclusion

The cloud system has been developed based on a scalable software platform OpenStack Swift, which provides robust cloud storage services and ensures access to the instrumental seismological data. Cloud system integration with the AIS is capable of effectively solving the problem of acquisition and storage of the instrumental seismological data and providing secure access to the user archives of accumulated observations.

Testing of the resulting information system has shown a high level of reliability and efficiency of the implemented cloud storage and software. This gives authors the reason to discuss the possibility of applying the established experimental model and the research results for storage and working with data from other instrumental observation networks of FEB RAS (geodynamic, video, etc.) designed for monitoring hazardous natural objects of the Russian Far East [12, 5].

In addition, the cloud server has been created containing various archives of scientific data that can be used by the applications in various research projects. Particularly, it is worth mentioning here the international project “Study of modern geodynamic processes in the northwestern Pacific and Northeast Asia and their response to the lithosphere and atmosphere based on GPS / GLONASS observations”, which is implemented under collaboration of Far Eastern Federal University and Institute of Seismology and Volcanology, Hokkaido University, Sapporo, Japan, with the participation of specialists of the Far Eastern and Siberian Branches of the Russian Academy of Sciences.

6 Acknowledgments

The study was supported by the Complex Fundamental Research Program of the Far Eastern Branch of the Russian Academy of Sciences ”Far East” (project 15-I-4-072), RFBR (projects 15-37-20269 mol a ved, 15-29-07953 ofi m), Far Eastern Federal University (project 14-08-01-05 m).
References