SEISMIC VISUALIZATION AND DEVELOPMENT PLATFORM ON THE ARMENIAN GRID INFRASTRUCTURE

Vladimir, Sahakyan
Institute for Informatics and Automation Problems of NAS RA
Yerevan, Armenia
e-mail: svald@sci.am

Lilit, Sargsyan
Western Survey for Seismic Protection
Yerevan, Armenia
e-mail: lilit_geo@yahoo.com

Mikayel, Gyurjyan
Institute for Informatics and Automation Problems of NAS RA
Yerevan, Armenia
e-mail: Mikayel_Gyurjyan@ipia.sci.am

Hayk, Poghosyan
Western Survey for Seismic Protection
Yerevan, Armenia
e-mail: nssp_haykp@yahoo.com

Raffie, Durgaryan
Institute of Geological Sciences of NAS RA
Yerevan, Armenia
e-mail: raffie_d@yahoo.com

Jon, Karapetyan
Institute of Geophysics and Engeneering Seismology after A. Nazarov of NAS RA
Gyumri, Armenia
e-mail: jon_iges@mail.ru

ABSTRACT

In this paper the main components of a seismology visualization and development platform are introduced based on the Armenian National Grid Infrastructure (ARMNGI). The main core of the platform is the seismic data server that organizes and gives performance access to distributed seismic data.

The Seismic Data Server is the rendering layer of an integrated GRID Infrastructure that enables the research community to have access to a broad range of earthquake data from Armenia and its surroundings. This brings together distributed Seismic Stations to provide a single access point from which researchers can search for and download selected data and data products.

Keywords

Armenian grid infrastructure, seismic visualization, metadata, seismic portal.

1. INTRODUCTION

Located in one of the world's most active seismic zones, Armenia frequently experiences earthquakes. Seismic observation and seismic monitoring in Armenia are concentrated in Survey for Seismic Protection of the Republic of Armenia (SSP). It is a Governmental organization established in 1991, after the catastrophic Spitak earthquake (December 7, 1988). The seismicity of Armenian Upland is related to the Arabian-Eurasian plate's collision, which is characterized by diffusive distribution of shallow small and moderate earthquakes.

There are 4 organizations collaborated in ARMNGI project: Institute for Informatics and Automation Problems NAS RA, Survey for Seismic Protection of MES, Institute of Geological Sciences NAS RA, Institute of Geophysics and Engineering Seismology after A.Nazarov NAS RA. The seismic data are received in the Data Center on a near real-time basis from whole stations homogeneously distributed in Armenia. When an earthquake occurs, SSP immediately (for local earthquakes during 20 minutes) issues information on its hypocenter, magnitude and observed seismic intensity. The information is provided to disaster prevention authorities and reaches the public (if it is necessary) through local governments and the media. The data are archived at the Seismology Department and stored in database system: as waveforms, seismological bulletins and catalogues for research work and other activities. SSP also provides data to international organizations such as ESMC, ISC est. as near real time information about earthquake and seismological bulletins.

To study and analyze very large scale and complex seismic data seismologists require to have seismological data sets from different regional areas, experimental results from different earthquake research laboratories, some good data processing, analyzing software and above all high performance computational resources. Nowadays seismologists need both computational resources to solve equations that arise in the mathematical modeling of seismic phenomena as well as storage resources in order to store and access historical earthquake information and massive seismic data that are collected either in continuous or discrete data from several geographically distributed sensors. Grid infrastructures [1] solve these problems by offering a platform where computational storage resources and other miscellaneous resources are available all connected by high speed networks. Grid computing [2] is an advanced technology of the distributed parallel calculations recently intensively developed in Europe, America, and Asia and in other regions of the world. Grid computing technology assumes a collective shared mode of access to network

resources and to the services, connected to them using frameworks of globally distributed virtual organizations consisting of the enterprises and the separate experts. Actual Grid-networks consist of large-scale systems of calculations, monitoring, management, complex analysis services and globally distributed sources of the data capable to support structures of scientific, education, government organizations and industrial corporations and forming powerful e-Infrastructure for e-science.
Taking into account the importance of this direction, the interested parties established an Armenian National Grid Initiative that represents an effort to establish a sustainable grid infrastructure in Armenia [3] to expand the high performance computing resources with collaboration of academic and commercial participants and to improve national applications.

2. SEISMOLOGY PLATFORM

The suggested seismology platform consists of the following levels:

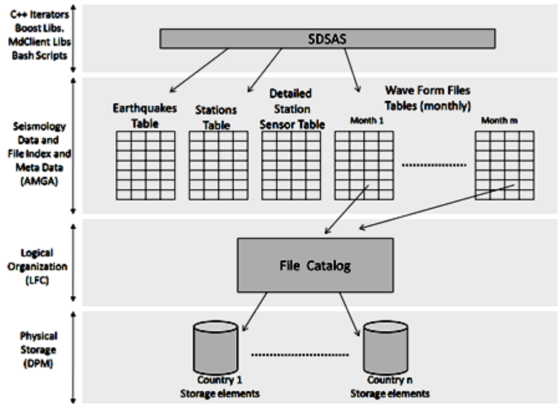
- The AMGA Metadata Catalog
- Seismic Data Server Application Services
- Seismic Data Portal
- Earthquake Location Finding (ELF) Application

2.1 THE AMGA METADATA CATALOG

File and Metadata Catalogs are essential services of a Data Grid, allowing users and applications to discover and locate data among the numerous sites of the Grid. File Catalogs map logical filenames to the physical location of one or more replicas of a file, while Metadata Catalogs store metadata describing the contents of the files, allowing users to search for files based on their description. A Metadata Catalog stores entries corresponding to the entity being described, typically files. These entries are described by user-definable attributes, which are key/value pairs with type information. Entries are not associated directly with attributes. Instead they are grouped into schemas, with the schemas holding the list of attributes that are shared by all their entries. AMGA [5, 6] structures metadata as a hierarchy, similar to a file-system. Directories play the role of schemas; they may contain both entries and other schemas. This hierarchical model has the advantages of being natural to users as it resembles a file-system, and of providing good scalability as metadata can be organized in sub-trees that can be queried independently.
In our model earthquake and station data are kept in AMGA tables whereas waveform data are stored as files. There can be millions of waveform data files. Therefore, an index of these waveform data files is also kept in AMGA tables. The AMGA tables can be queried directly by using various AMGA interfaces (shell, Perl, java, C/C++).

2.2 SEISMIC DATA SERVER APPLICATION SERVICES

Seismic Data Server Application Services (SDSAS) is a set of tools for storage, indexing and providing high level access to massive seismic data. The objective of the tool is to hide the details of where the data files reside and map high level user specification (dates, hours, location etc.) to appropriate pathnames automatically. Available seismic data is massive residing in hundreds of thousands of files.
SDSAS tool consists of two main components. The first component is the data collection component for uploading data using various scripts. The second one is a programming tool in the form of C++ classes and iterators that can be used by programmer for accessing seismic and earthquake data. The aim of the SDS C++ iterators is to provide a higher level interface to seismic data that does not necessitate seismologist users to learn AMGA and which provides additional functionalities. It allows querying AMGA catalogue automatically using high level specifications such date ranges and locations for accessing station data, waveform data files and earthquakes.



Architecture of the SDSAS

2.3 SEISMIC DATA PORTAL

The Seismic Data Portal provides a single point of access to diverse distributed Armenian seismological data. Based on internet-standard portlet and web services technologies, it enables the scientists/users to integrate and combine different data services.
Seismic data is stored on storage element of the ARMNGI and made available through the File Catalogue Service. The SAC (Seismic Analysis Code) [7] file format is used for seismic data. The data consists of static and dynamic parts. The static data includes information about stations (general information about stations, detailed information about sensors in the stations, response file of stations) and earthquake catalogues including main parameters of earthquakes.
The SDSAS is including services providing event and waveform information. These data services are available through internet web standards. The services currently include:

1. **Event Information** (event search, latest event, event detail)
2. **Waveform Services** (provides information about networks, stations, channels and more in general a complete inventory of the available resources.)
3. **User Tools** based on the SDS:
    - *Waveform Data Explorer* (creation of a dedicated waveform metadata, definition and implementation of parametric data exchange procedures, specification of protocols for easy and unified access to waveforms databases in Armenia).
    - *Event Selection* (allows selecting multiple events from Event Data Catalogue, which has to be populated with the Events Explorer, visualizes selected events and main information about the event on the map).
    - *Time Selection* (it is possible to specify the time window you`d like to request data for).

Web data services provide a simple and straight-forward mechanism to expose access to data center holdings to the broader community. Using simple internet technologies, one is able to build applications that directly query data center holdings.

2.4 EARTHQUAKE LOCATION FINDING (ELF) APPLICATION

An earthquake location finding application is ported on seismology platform. A web interface was also developed that displayed the locations of the earthquakes and the stations using the Google Maps interface.

Earthquake Location Finding (ELF) application finds the hypocenter of an earthquake by using the seismic waveform data generated by seismic stations. ELF is based on the widely known HYPO71 application in the seismology area. Applications that find earthquake locations by processing waveform data collected from seismology stations are quite important since they are used daily by seismologists. Most of the other applications in seismology are also based on location finding. Earthquake location finding application can also be used on archived data in order to find historical earthquakes.

We used parallelized version of ELF application. The main aim of parallelization was to speed-up the application. In order to achieve this, different parallelization strategies were investigated. These are: (i) the use of MPI message passing libraries, (ii) the use of OpenMP multithreading directives and (iii) to express the parallel application as a workflow using Job Description Language (JDL). To implement spatial decompositions, stations are divided among the worker nodes. A node is assigned a task which is responsible for accessing a specific station's waveform file and computing picks. The results from all worker nodes are then sent to a collector node which runs the HYPO71 program to locate hypocenter of the earthquake. We present timing results from several tests. The results indeed show that parallelization of file accesses improves performance.

Output of the ELF application can be visualized by using SDSAS application service kml generation routines. These routines will directly read results stored on the AMGA table and generate the kml file.

An example kml file on Google Map

2.5. CONCLUSION

Armenian National Grid Infrastructure will give the following benefits:

- Serve seismic data that is mirrored from national seismology centers using a high level interface that is easy to use/adapt.
- Use the grid platform for its seismology data and applications, which include models for detecting earthquake locations, fault-planes solution and seismic hazard assessment.
- Logical organization, indexing and update of distributed seismic data, large volume data analysis
- Performance aspects - (high performance computing, high performance access to massive data).
- High level Web based interface that will provide easy access to seismic data
- Orchestrated workflows across service components.

Running seismic wave propagation simulations requires huge amount of parallel computation, heavy runtime input/output loading as well as massive data analysis and mining tasks. Grid could be an essential research-supporting infrastructure for earthquake simulation to meet the emerging needs of disaster mitigation.

3. DETAILS

3.1. Contact address

Institute for Informatics and Automation Problems,
1, P. Sevak str., Yerevan, 0014, Armenia
E-mail: csit@sci.am
Phone: +(374 10) 527090, 282030, 282040
Fax: +(374 10) 569281, 285812

4. ACKNOWLEDGEMENT

The authors would like to thank various sponsors for supporting their research.

REFERENCES

[1] Grid Infrastructures: http://www.gridipedia.eu/grid-infrastructures.html

[2] Grid Computing Info Centre: http://www.gridcomputing.com/

[3] Armenian National Grid Initiative: www.grid.am

[4] B. Koblitz, N. Santos, V. Pose, "The AMGA Metadata Service", Journal of Grid Computing, Springer Netherlands, 2008, 61-76.

[5] N. Santos and B. Koblitz, "Distributed Metadata with the AMGA Metadata Catalog", in Workshop on Next-Generation Distributed Data Management, HPDC-15, Paris, France, June 2006

[6] N. Santos and B. Koblitz, "Metadata services on the grid", in Proc. of Advanced Computing and Analysis Techniques (ACAT'05), Zeuthen, Berlin, May 2005.

[7]SAC - Seismic Analysis Code: http://www.iris.edu//manuals/sac/SAC_Home_Main.html