Overview and Current State of the “GEO (Global Earth Observation) Grid”


Grid Technology Research Center (GTRC)
National Institute of Advanced Industrial Science and Technology (AIST)
Umezono 1-1-1, Tsukuba 305-8568, Japan.
Email: m.matsuoka@aist.go.jp

ABSTRACT

The “GEO (Global Earth Observation) Grid” is aiming at providing an E-Infrastructure to understand our Earth more insightful and more precisely, but faster and easier to worldwide Earth Sciences community. On the E-Infrastructure we are able to share data, application programs, and scientific workflows without knowing deep knowledge in IT where the grid technology is provisioned. GEO Grid provides large archives of earth observation satellite data securely and rapidly, integrated service with wide variety of geo-scientific information and GIS data, and assembles them easy-to-use formats for potential stakeholders in not only disaster management but also other several Societal Benefit Areas.

GEO Grid potentially contributes to provide IT as software package and large-scale of archived data related to the earth observation. The former enables to support securely for publishing data with incoherent owners’ policy, easy access, and distributed data provider. The latter contains the all ASTER archives and geological maps. In other words, GEO Grid has the capability to generate high-level data products (accurate geometric-, radiometric- and atmospheric-corrected, and 30-meter and/or higher resolution DEM globally) by the Computing Grid.

INTRODUCTION

Earth observations are indispensable to all of our activities, especially disaster mitigation, weather prediction, natural resource exploration, environment monitoring, and so on. In particular, satellite imagery provides a wealth of information to examine the surface of the planet Earth, however, the data rates are currently growing dramatically. For example, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) [1] mounted on the Terra earth observation mission satellite has produced more than one hundred terabytes of raw data since its launch in 1999. PALSAR (Phased Array L-band Synthetic Aperture Radar) onboard that ALOS (Japanese nicknamed “Daichi”) satellite, which started operation in 2006, will generate more than a petabyte of raw data in its five-year nominal mission period. Most of the data has been maintained in an archive system on a traditional tape library system.

In addition to such huge amounts of satellite data, researchers may be interested in accessing other databases, such as the data from land-use map, climate data, and field sensors, and then analyzing and simulating the data to know the real and future world. However, it is a difficult task for an individual user to retrieve the desired data from distributed systems and process that
data on remote computing resources. Based on the background, in this paper, in order to overcome the difficulties and to help motivated scientists advance research on earth observations, we propose a new platform, GEO (Global Earth Observation) Grid, for accessing the earth observation databases and providing computing resources easily.

**OVERVIEW OF THE GEO GRID**

The GEO Grid is an E-infrastructure for archiving and processing huge satellite observation data, such as ASTER and MODIS, and other geo-scientific dataset based on GIS by using the grid computing technologies [2]. Comprised of observation and information systems for earth observations, the GEO Grid system implements an infrastructure for flexible, secure, and coordinated sharing of resources such as satellite data, field sensor data, and computing for simulations and data analysis. In the design policy of the GEO Grid system, in order to enable security framework, the GEO Grid system introduces the concept of a virtual organization (VO) [3],[4], for its design, in which various data and computing resources are provided as services represented by standard protocols. Figure 1 shows the schematic architecture of the GEO Grid. The latter chapter describes its detail configuration by taking the ASTER data application as an example.

![Figure 1. Simple architecture of the GEO Grid](image-url)

**ASTER GRID SYSTEM**

As a sample application of the GEO Grid system, we have developed and deployed a system for accessing ASTER data using the components of the GEO Grid system. ASTER, which is a high spatial resolution multi-spectral imaging radiometer, was developed by the Ministry of International Trade and Industry of Japan (METI). The Earth Remote Sensing Data Analysis
Center (ERSDAC) of Japan has been archiving ASTER data in a traditional tape library system at the rate of approximately 70-100 gigabytes per day, then providing users the ASTER Ground Data System (GDS) [5] to search and obtain ASTER data products through a Web portal. One inconvenience issue is that it usually takes for the whole daylong when the ASTER GDS generates products in bulk requested by a user. It is desirable for users to enjoy on-demand and near-real-time delivery services of the product.

Secondary, on ASTER GDS there is no previewing of scenes in Open Geospatial Consortium (OGC) [6] compliant format before purchase and considerable latencies to get the scenes after the order. Although in a modern GIS system it is expected that any map will be searchable and accessible through a single GUI through OGC standard services, such as WMS (Web-Mapping Service), WFS (Web-Feature Service), and WCS (Web-Coverage Service). Users may want to overlay a thumbnail as previewing image of the ASTER scenes and other useful maps like geology and/or land-use map. Unfortunately as of today, users must query the ASTER data and the geology map independently through different GUIs with the ASTER GDS.

In the ASTER Grid system, ASTER data is stored not in a tape library system but in a cluster system called the GEO Grid cluster. It consists of 28 dual Intel Xeon nodes connected by Giga-bit Ethernet, each of which is equipped with a 6.3TB RAID6 hard disk drive. The total storage volume is 176.4TB as a single storage volume by Gfarm [7]. Currently, approximately 125TB of the ASTER data stored in the Gfarm file system is physically dispersed across the cluster nodes, however it can be accessed transparently by mounting the Gfarm file system in a user space via the FUSE [8] mechanism.

**ASTER DATA APPLICATION**

**DEM and orthorectified image generation**

One of the most significant characteristics of ASTER is the along-track stereoscopic data acquisition capability and accurate geolocation. The ASTER Grid system utilizes another cluster system called the F32 cluster, which has 256 dual Intel Xeon nodes connected by Giga-bit Ethernet. The F32 cluster is used for executing the image processing application which generates Digital Elevation Model (DEM), actually Digital Surface Model (DSM), and orthorectified images using Level 1A derived from Level 0 ASTER data stored in a Gfarm file system in the GEO Grid cluster storage area (Fig. 2). The spatial resolution of the ASTER DEM and orthorectified images is 15 meters in the highest level for visible and near-infrared wavelength bands. All data sets can be downloaded by on-demand and user’s request basis on the ASTER Grid portal as shown in Fig. 3.

The one of the most advanced features of ASTER DEM is to capture the detailed 3D surface in the area above 60 degrees of altitude because Shuttle Radar Topography Mission (SRTM) [9] could not cover the high-latitude regions. Surface topography should be rather stable in the absence of significant ground movements, such as earthquakes or volcanic eruptions. In other words, by using the ASTER Grid, temporal topography and coastline changing caused by geo-hazards can be examined immediately and easily.

**Integration and hazard simulation**

The GEO Grid system also supports OGC standard services like the WMS to provide the global ASTER imagery and ASTER-generated maps, which allows users integrating the ASTER data
with the data from other services (or their own data). Figure 4 shows an example of image overlaying results added by geological map through the GEON site [10]. In the future we also plan to support the WCS for more advanced scientific analysis under VO environment.

As an example of geo-hazards application, the GEO Grid contains the pyroclastic flow simulation program which provides a possible coverage map of pyroclastic flow deposits caused by a volcanic dome collapse, and can contribute to make an emergency volcanic hazard map. An energy line model is used for calculation of the maximum possible flow distance [11] at the GEO Grid (Fig. 5). Other algorithms can be applied when the programs (services) are registered. The result maps are provided by using the WMS, making it easier to overlay them on other GIS data and maps.

Figure 2. On-demand DEM and orthorectified images generated on the ASTER Grid system

Figure 3. Snapshot of the ASTER Grid portal
CONCLUSION

In this paper, we presented the architecture of the archiving and processing satellite imagery and GIS data sets, called GEO (Global Earth Observation) Grid, which is aiming at providing an E-Infrastructure to understand our Earth more insightful and more precisely, but faster and easier to worldwide Earth Sciences community. The GEO Grid project is a joint project of the Geological Survey of Japan (GSJ) and the Grid Technology Research Center (GTRC) of the National
Institute of Advanced Industrial Science and Technology (AIST), Japan. GSJ has large amounts of geospatial data and applications, and AIST stores ASTER data. GTRC provides the Grid-based e-Science infrastructure for them. Our project has only just started and the current prototype implementation is equipped only with limited functionality from the entire architecture of the GEO Grid, however, we are already offering the full set of ASTER images with our first prototype.

REFERENCES