Application of Information Technologies on Astronomy: Japanese Virtual Observatory (JVO) Portal

Y Shirasaki[1*], M Tanaka[1], M Ohishi[1], Y Mizumoto[1], S Kawanomoto[1], S Honda[2], N Yasuda[3], and Y Masunaga[4]

*[*1]National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo, Japan*
Email: yuji.shirasaki@nao.ac.jp
[2]Gunma Astronomical Observatory: Takayama Village, Gunma, Japan
[3]University of Tokyo, ICRR, 5-1-5 Kashiwa-no-Ha, Kashiwa, Chiba, Japan
[4]Aoyama Gakuin University, 5-10-1 Fuchinobe, Sagamihara, Kanagawa, Japan

ABSTRACT

The Japanese Virtual Observatory (JVO) is a web portal to various kinds of astronomical resources distributed all over the world. We have started its official operation of the JVO portal since March 2008. The JVO provides seamless access to the Virtual Observatory (VO) compliant data services, and also access to the reduced data observed with Subaru telescope and on-line data reduction system for Suprime-Cam instrument of the Subaru telescope. The system implements standards of the International Virtual Observatory Alliance (IVOA) to communicate with the VO components in the world.

Keywords: Virtual Observatory, Database, GRID, Astronomy

1 INTRODUCTION

In the field of astronomy, a large amount of high quality digital data has been produced and will be produced at various ground- and space-based observatories. Those data are managed by data centers in each country or community individually, and they are provided to astronomical researcher through inhomogeneous interface designed by each data center independently.

Recently, astronomical study is conducted based on the multi-wavelength data. As a result, the researchers tend to use various kinds of data produced from different telescopes. In order to obtain the data, the researcher needs to know where the data that they want to use are available. For that purpose, a kind of repository that stores the information about the location and contents of each data resource is required. It is useful if all the data service implement the same access interface, as it enables to develop an automated data query system for the distributed databases.

If the data is provided without calibration, a researcher must calibrate the raw data by themselves. To calibrate the raw data, however, the researcher needs to have deep knowledge about the data and also the instrument that produced the data, but it is inefficient and even impossible for every researcher to understand them and carry out the calibration correctly. To do an astronomical research efficiently, a research environment should be provided that enables researchers to carry out the calibration in a manner that procedures specialized to each instrument are encapsulated and automatically executed in a correct way without intervention of a researcher.

To realize such an astronomical research environment, International Virtual Observatory Alliance (IVOA) has been developing standard interfaces for discovering and providing the astronomical data through the Internet, execution of data analysis on a remote server, data sharing among distributed storage, and so on. National Astronomical Observatory of Japan (NAOJ) is developing the Japanese Virtual Observatory (JVO), which is a web portal for astronomical data and analysis system, and utilizes the standards developed by the IVOA to communicate with the Virtual Observatory (VO) components in the world. We have started its official operation of the JVO portal since March 2008.

The JVO provides the following web-based services: (1) search for VO compliant data services in the world, (2) search and access to the astronomical data through the VO standard interfaces, (3) use of on-line astronomical analysis tools, (4) on-line data reduction for the Subaru telescope, and (5) use of data storage of the JVO system. As of January 23, 2009, over 1,900 resources are registered to the JVO portal. A GRID computing system is introduced to provide massive computing resource for reducing the data obtained by the Subaru telescope. In the initial operation, we offer 48 CPU cores, 4 GB memory per core, 1 TB of locally attached hard disk and 100 TB of NFS storage areas in the GRID computing system. The number of CPU incorporated to the system can be easily extended.

2 OVERVIEW OF THE JVO PORTAL

Figure 1 shows a diagram of the current JVO system. The portal is a core component of the system, which authenticates a user (LDAP), accepts a query and job submission request from a user, creates and executes a workflow (Tanaka et al. 2006) to achieve the request, and manages the user storage where search results and uploaded files are saved. SkyNode, SIAP and SSAP are VO standard data services distributed in the world. The access URL to each data service is resolved with the searchable registry (S.Reg.), which is a metadata database of the VO services.

We provide three easy-to-use interfaces for novice users, and two interfaces for advanced users to directly specify an SQL. Figure 2 shows two kinds of the graphical query interfaces. The left figure shows a quick search interface and the right one shows a SQL-based query interface. At the quick search page, one can access to the ``Digital Universe" (Tanaka et al., 2008), which is a database containing only coordinates and brightness of objects from various published catalogs, such as Subaru deep survey, TWOMASS, SDSS DR6, UKIDSS DR2, ROSAT bright catalog, and so on. At the SQL-based query page it is possible to specify more complex query conditions. More details about the JVO system can be found in Shirasaki et al. (2006, 2007, 2008).

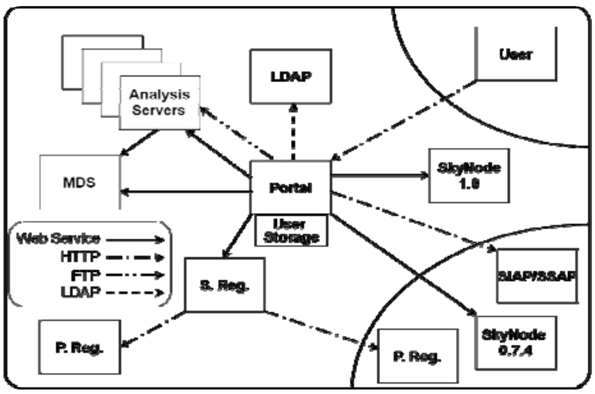

Figure 1. Architecture of the JVO system.

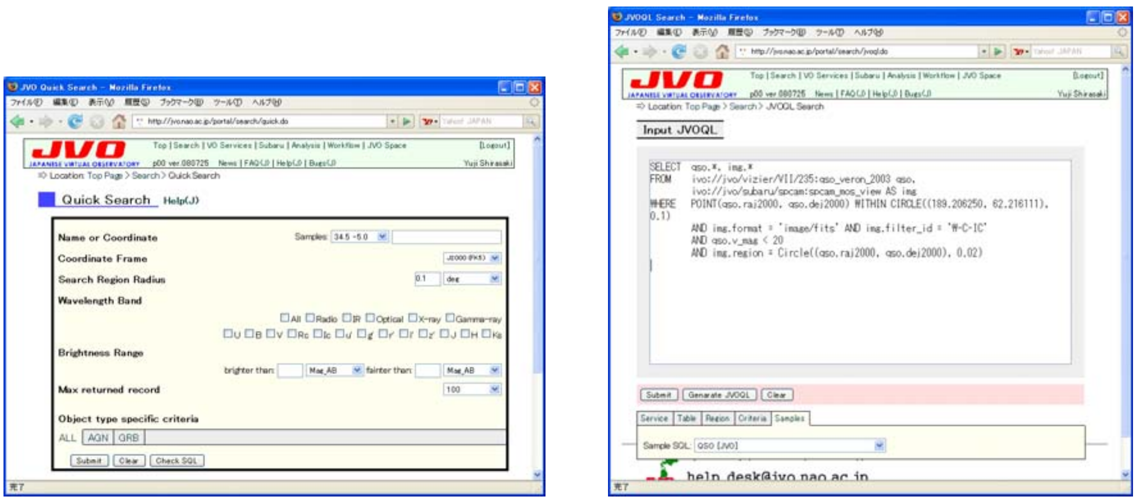

Figure 2. JVO data search GUI

3 SKYNODE

We are developing a SkyNode toolkit for building a VO compliant data service easily. The aim of this development is to reduce the time and effort for the data providers to spend for implementing a VO interface on their database. Reducing the cost for introducing the VO standard is the most important factor for widely deploying the VO services. The toolkit is available at:

http://jvo.nao.ac.jp/download/skynode-toolkit

The toolkit mainly provides a part of ``SkyNode controller'' (Figure 3), where Astronomical Data Query Language (ADQL) is converted to the DB native SQL, and the result is returned in the VOTable format that is a VO standard format to exchange the tabular data. The interface to the back-end database is provided by Java DataBase Connectivity (JDBC). A table structure that is exposed to the public is configurable, and does not necessarily need to be the same as the internal structure. Spatial region search is performed by using an Hierarchical Triangular Mesh (HTM) index (Kunszt et al., 2000), which is stored on the database in advance. It is also possible to construct an image or spectrum data service on two access protocols, SkyNode and Simple Image/Spectrum Access Protocol (SIAP/SSAP). The simple access protocol is based on an http-get parameter query, where common query parameters such as POS and SIZE, which describe center coordinate and radius of the search region respectively, are defined.

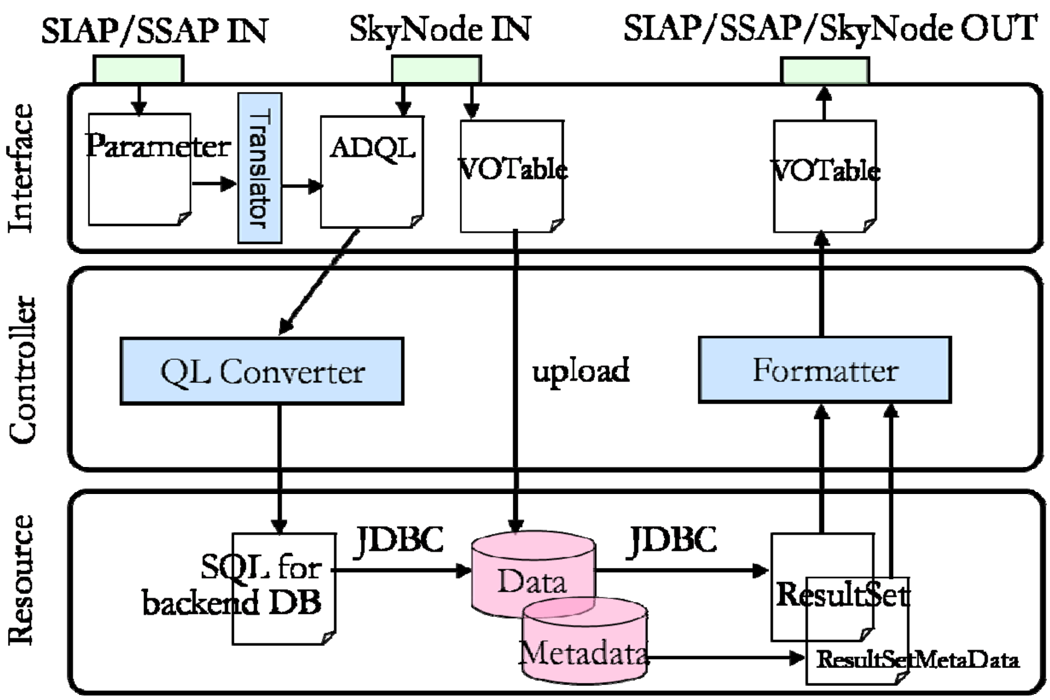

Figure 3. Architecture of the JVO SkyNode toolkit

4 JVO QUERY LANGUAGE

We defined the JVO Query Language (JVOQL) to describe a query to the distributed multiple databases in a single statement. An example of the JVOQL is shown in Listing 1, which describes a query to get Subaru Suprime-Cam images around Quasi-Stellar Objects (QSOs). The data of QSO and the Suprime-Cam image are stored at different database sites. More details can be found in Shirasaki et al. (2005).

Listing 1. An example of the JVOQL

```
SELECT qso.*, img.*
FROM   ivo://jvo/vizier/VII/235:qso_veron_2003 qso,
       ivo://jvo/subaru/spcam:spcam_mos_view img
WHERE  Point(qso.raj2000, qso.dej2000) WITHIN
       Circle((189.206250, 62.216111), 0.1)
       AND img.format = 'image/fits'
       AND img.filter_id = 'W-C-IC'
       AND qso.v_mag < 20
       AND img.region =
           Box((qso.raj2000, qso.dej2000), 0.02, 0.02)
```

5 METADATA EXCHANGE

We have constructed a publishing registry and a searchable registry. The role of the publishing registry is to expose resource metadata to the VO world (Figure 4). The publishing registry implements the OAI-PMH protocol. The searchable registry collects metadata from the publishing registries not only of the JVO but also of the other VO projects. Currently we are collecting metadata from STScI, NCSA, ESAC, CDS, HEASARC and JVO. User can access to the searchable registry on their web browser as shown in Figure 5.

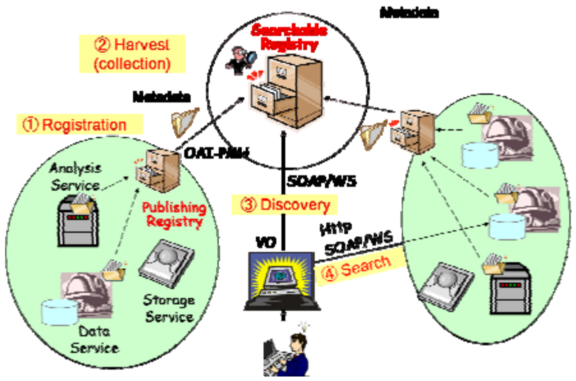

Figure 4. Metadata exchange mechanism in the Virtual Observatory

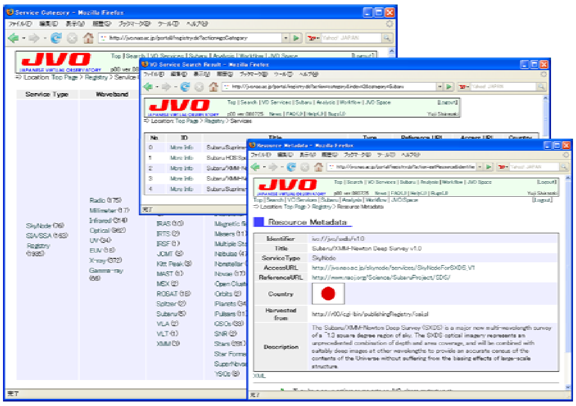

Figure 5. JVO Registry GUI.

6 GRID COMPUTING

For providing a computing service to a varying number of users, a scalable computing system is required. The Monitoring and Discovery Service (MDS) is constructed to build a grid computing system (Figure 6). Static properties such as a CPU type, number of CPU core, memory size, IP address, and so on, of each computing server are registered on the MDS. Dynamical properties such as a number of submitted jobs, load average, and so on are notified periodically by the computing server to the MDS server. Considering the status of each server, the MDS provides the best host to achieve the most efficient job balancing. This system is currently used to reduce the raw data taken by the Suprime-Cam instrument of the Subaru telescope. The details of the system are described in Shirasaki et al. (2008).

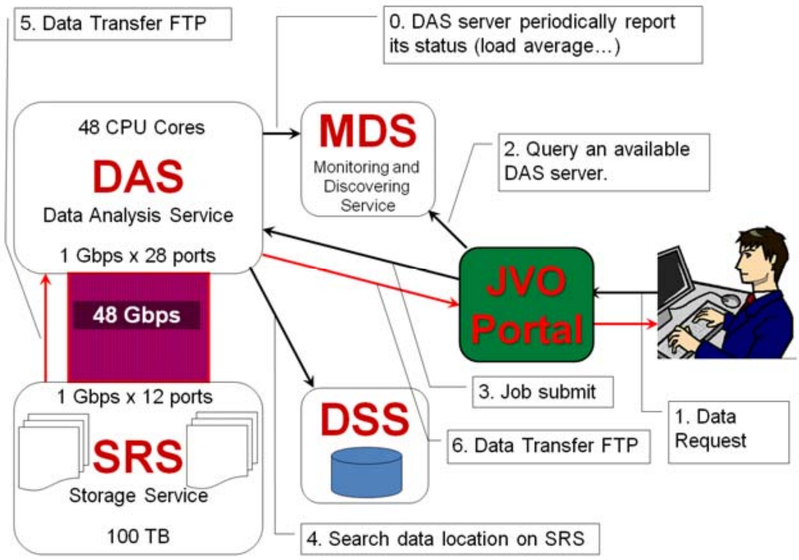

Figure 6. GRID computing system in the JVO.

7 SUMMARY

We have started its official operation of the JVO system since March 1st of 2008. Improvement of the capability and usability should be continued by taking account of use cases which many of the users want to do. Currently we are making improvements on enabling the users to do a multiple region search, to make their own multi-band catalog from the different catalogs very easily. We are also improving the content of the 'Digital Universe' by collecting the data from the VO services in an automated way. Another issue we should tackle is to increase the fraction of VO compliant data services. There are still many data resources which are not accessible through the VO interface. The success of the VO primarily depends on how quickly we can have a large fraction of the data resource accessible through the VO application.

8 ACKNOWLEDGEMENTS

This work was supported by the JSPS Core-to-Core Program and Grant-in-aid for Information Science (15017289, 16016292, 18049074 and 19024070) carried out by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. YS is grateful for support under Grant-in-aid for Young Scientists (B) (17700085) carried out by the MEXT of Japan.

9 REFERENCES

Kunszt, P. Z., Szalay, A. S., Csabai, I., & Thakar, A. R. (2000) In Manset, N., Veillet, C., & Crabtree, D. (Eds.) in *ASP Conf. Ser., Vol. 216, ADASS IX* (San Francisco: ASP), 141.

Shirasaki, Y. et al. (2008) *Fusion Engineering and Design 83*, 438.

Shirasaki, Y. et al. (2007) In Shaw, R., Hill, F., & Bell, D. (Eds.) *ASP Conf. Ser. 376, ADASS XVI* (San Francisco: ASP), 16.

Shirasaki, Y. et al. (2006) *Proc. SPIE, 6274*, 62741D.

Shirasaki, Y. et al. (2005) In Shopbell, P., Britton, M., & Ebert, R. (Eds.) *ASP Conf. Ser. 347, ADASS XIV* (San Francisco: ASP), 262.

Tanaka, M. et al. (2008) In Lewis, J., Argyle, R., Bunclark, P., Evans, D., & Gonzalez-Solares, E. (Eds.) *ASP Conf. Ser. 394, ADASS XVII* (San Francisco: ASP), 261.

Tanaka, M. et al. (2006) In Gabriel, C., Arviset, C., Ponz, D., & Solano, E. (Eds.) *ASP Conf. Ser. 351, ADASS XV* (San Francisco: ASP), 460.