Construction of the Japanese Virtual Observatory (JVO)

Yoshihiko MIZUMOTO, Masatoshi OHISHI, Naoki YASUDA, Yuji SHIRASAKI and Masahiro TANAKA

National Astronomical Observatory of Japan, 2-21-1, Osawa, Mitaka, Tokyo 181-8588, JAPAN

Yoshifumi MASUNAGA

Ochanomizu University, 2-1-1, Otsuka, Bunkyo, Tokyo 112-8610, Japan

Ken MIURA, Hirokuni MONZEN, Kenji KAWARAI, Yasuhide ISHIHARA, Yasushi YAMAGUCHI and Hiroshi YANAKA

Fujitsu Ltd., 1-9-3, Nakase, Mihama, Chiba 261-8588, Japan

Abstract.
The National Astronomical Observatory of Japan (NAOJ) has been operating several large astronomical facilities, such as the SUBARU telescop, the 45 m radio telescope and the Nobeyama Millimeter Array, and plans to construct the ALMA under close collaborations with the US and the EU. Since January 2002, the NAOJ has been connected to the Super-SINET with 10 Gbps, and it has become possible to provide huge amount of observed multi-color data and analyses facilities to other astronomical institutions. Thus we have started the Japanese Virtual Observatory (JVO) project since April 2002. JVO utilizes the Grid technology to combine several remote computational facilities. We have completed to define the query language for the JVO (JVOQL), and have been designing the JVO components. We plan to construct a JVO-prototype by the end of 2002. Further details are available at `http://jvo.nao.ac.jp/english/`.

1. Introduction

The National Astronomical Observatory of Japan operates the Subaru telescope (optical and infrared) in Hawaii and large radio telescopes in Nobeyama. All the observed data are digitally archived and are accessible via internet, and the data archives have strongly supported many researchers on astronomy and astrophysics. The radio telescopes of Nobeyama produce $\sim$ 1 TBytes per year, and the Subaru telescope outputs $\sim$ 20 TBytes per year. However, until recently, there was a severe restriction to access such amount of data – the network bandwidth betwen the NAOJ and the Internet was only 10 Mbps.

The National Institute of Informatics has started to operate a new network, the Super SINET (`http://www.sinet.ad.jp/english/`) since January 4, 2002, and the NAOJ has become an important node for the Super SINET.

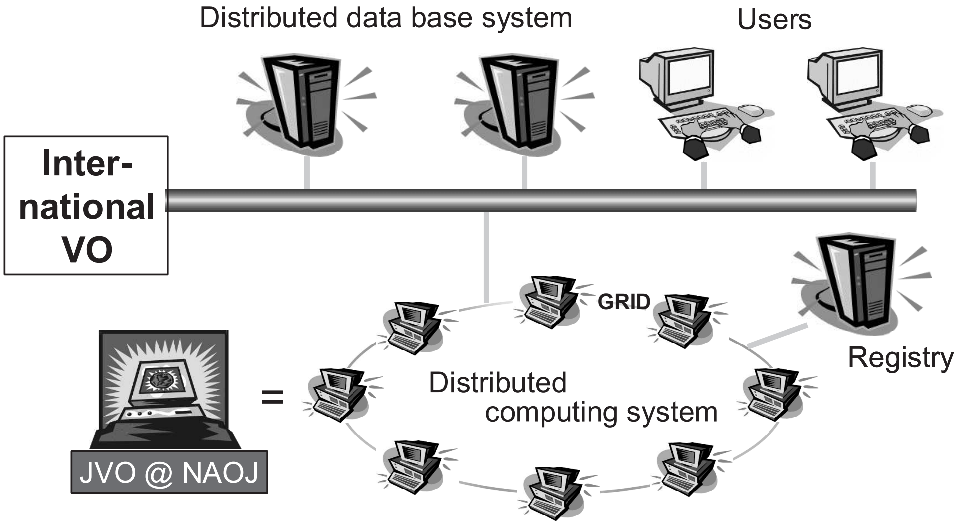

Figure 1. The Schematic Diagram of the JVO System

The Super SINET is an ultrahigh-speed network intended to develop and to promote Japanese academic researches by strengthening collaborations among leading academic research institutes. The backbone network connects research institutes with the bandwidth of 10 Gbps, and the leading research facilities in Japan are directly connected with 1 Gbps network.

The JVO is designed to seamlessly link the distributed database (DB) and data analyses system for Subaru, Nobeyama observatories, and other observational data for astronomers in research institutes by utilizing the state-of-the-art GRID technology through the 10Gbps Super SINET.

This paper briefly describes our overall concept and future plans of the Japanese Virtual Observatory.

2. Structure of the JVO system

Figure 1 shows the schematic diagram of the JVO system. JVO consists of a distributed computing system (DCS) which is deployed over the high-speed network such as the Super SINET by utilizing the GRID technology. The registry plays quite an important role in the JVO system, which provides information required for DCS to resolve the URLs of distributed DB systems, data analysis servers, and so on. All the computers of the DCS may have independent functions. However many of them need to have redundant functions with others to guarantee robustness of the JVO system. The Resource Manager automatically selects the most appropriate machine for a given task requested by the JVO users through GRID Resource Information Service. It is inevitable that the JVO has an interoperability with the other VOs, such as the NVO, the AVO, the AstroGrid, and so on, to enable researchers to access other databases in the world.

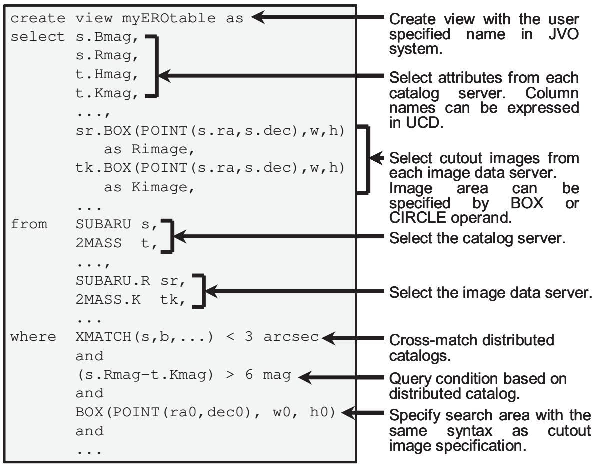

Figure 2. Samples of the JVO Query Language showing queries for catalog searches.

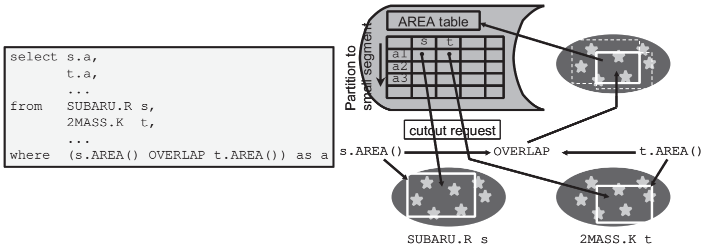

Figure 3. Samples of the JVO Query Language showing queries to search for images.

3. The JVO Query Language

The JVO Query Language (JVOQL) is used in JVO as a language to specify a variety of user queries. The samples are illustrated in Figures 2 and 3.

The JVOQL is designed to keep upward-compatibility with the standard relational database language, Structured Query Language (SQL), to enable in handling image data and cross-matching among distributed databases. ! ! The interpreter of JVOQL communicates with the registry of available databases and issues query sequences to distributed databases.

JVOQL has an ability to query image data without referring to catalogs. This function is useful for multi-color or multi-epoch analyses. The JVOQL example (Figure 2) shows how to obtain R-band images taken by SUBARU and K-band images by 2MASS in an area where both SUBARU and 2MASS observed. The operand "OVERLAP" returns overlapped area of the two data. Figure 3 shows an example to search for required images. Similar to Figure 2, the operand "X.AREA()" returns the observed area of server X.

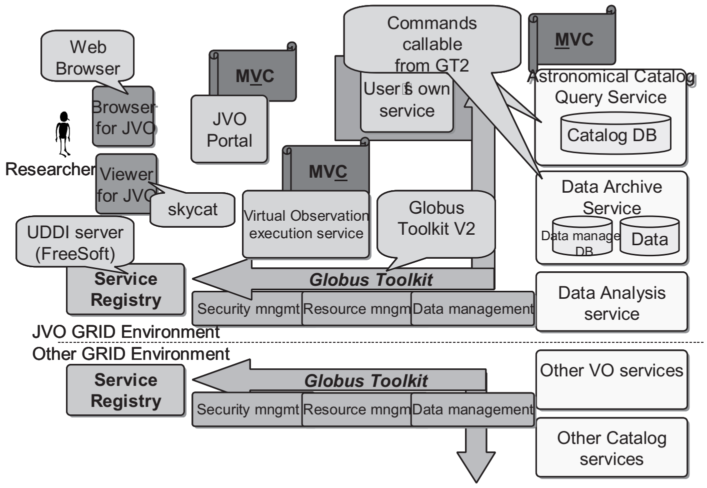

Figure 4. The Schematic diagram showing interrelations among components of the JVO system

4. Three-Tiered Design of the JVO Prototype

The JVO prototype is now under development. The first prototype is scheduled to be completed by the end of 2002. The design of the JVO prototype is shown as a schematic diagram (Figure 4). We adopted to use the Globus Toolkit V2 for our prototype. However we also take into account the Web service concept which will be included in the Globus Toolkit V3.

First of all, researchers provide the JVO with simple instructions how they plan to use their own "Virtual Observation" through the JVO portal. The JVO portal interprets them and generates a work-flow through consulting the UDDI servers, where available JVO services are registered. Based on the work-flow, built-in or user-defined services are called. The GRID framework is used for dynamical assignment of distributed resources according to their availabilities. Execution results of the work-flow are transferred through GridFTP, and are presented to the researchers with the skycat, etc.

5. Future Plans

When our first prototype have been completed, we will assess it and then modify, reimplement the second prototype. Toward the second prototype we plan to federate with more data of the SUBARU open use / Nobeyama Radio Observatory, to implement the interoperability with other VOs toward International VO, and CPU intensive image analysis tools such as deconvolution, image subtraction, and so on.

We also plan to implement the data mining / visualization tools to manage huge amount of data for new discoveries through statistical data analyses on the operational version of the JVO system. The JVO is intended to be a core system of the Regional Support Center in Japan for the ALMA in the near future.