Japanese Virtual Observatory (JVO): implementation of VO standard protocols

Yuji Shirasaki, Masahiro Tanaka, Satoshi Honda, Satoshi Kawanomoto, Yoshihiko Mizumoto, Masatoshi Ohishi

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

Naoki Yasuda

University of Tokyo, 5-1-5 Kashiwa-no-Ha, Kashiwa Chiba, 277-8582 Japan

Yoshifumi Masunaga

Ochanomizu Univerisity, 2-1-1 Otsuka Bunkyo-ku, Tokyo, 112-8610 Japan

Yasuhide Ishihara,Jumpei Tsutsumi

Fujitsu Ltd., 4-1-1 Kamikodanaka Nakahara-ku, Kawasaki, 211-8588

Hiroyuki Nakamoto, Yuusuke Kobayashi

Systems Engineering Consultants Co. Ltd., 22-4 Sakuraoka-cho Shibuya-ku, Tokyo, 150-0031

Abstract. We developed the third prototype towards a Japanese Virtual Observatory (JVO). IVOA standards, such as Simple Image Access and ADQL, were adapted to the system for the first time. We also constructed an OAI-PMH publishing registry, a web service based searchable registry, and VO data services based on SIA and SkyNode protocols. Most of the components were built on open software, except for an XML database used for searchable registry. This paper describes the JVO proto 3 system and results of a performance measurement.

1.   Data Service

We are developing a SkyNode toolkit for building a VO compliant data service. Skynode is a VO standard data access interface. A query is described in SQL-like syntax with astronomical extensions, and is transfered as a SOAP message over http protocol. First primitive version was released in May 2005 at http://jvo.nao.ac.jp/download/ and update version will be available at the same URL. Fig.1 shows the architecture of the JVO Skynode. The toolkit mainly provides a part of "SkyNode controller", where ADQL is converted to the DB native SQL, and the result is returned in the VOTable format. The interface
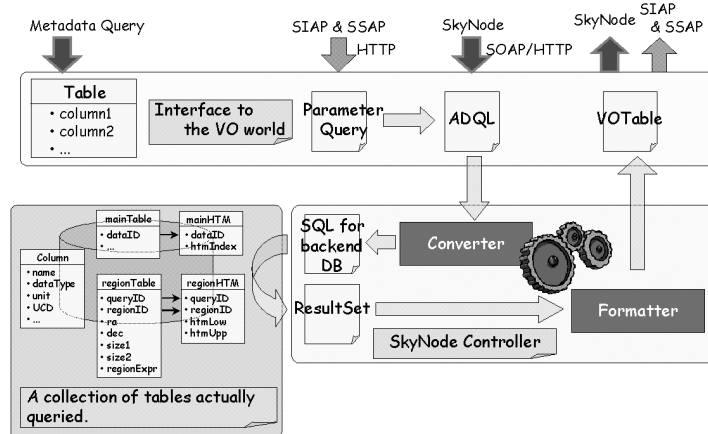
Figure 1.   Architecture of the JVO skynode

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. Space region search is performed by using HTM index, 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 are defined. Below is shown an example:

```
http://jvo.nao.ac.jp/image?POS=210,30&SIZE=1.0
```

SQL expression corresponding to the above URL query is:

```
select *
from   image img
where  region overlaps Region('Box J2000 210 +30 1.0')
```

The toolkit also provides cross-match functionality, that is, selecting data for multiple regions specified in a VOTable. Cutout images for multiple objects can be requested by the following SQL.

```
select *
from   image img, #upload vot
where  img.region overlaps BOX((vot.ra, vot.dec), 1.0)
```

Data resources currently available from JVO are: (1) Subaru Deep field survey catalogs and images, (2) Subaru Suprime-Cam Open Data Archive, (3) QSO catalog compiled by Veron et al. and copied from VizieR, (4) SDSS DR2. (5) 2Mass Catalog.

2.   Registry Service

We have constructed a publishing registry and a searchable registry. The publishing registry was made by using a NVO software, and which was slightly

modified to adapt the recent VOResource schema. Metadata of data resources are registered through the web form interface, and these data are distributed by the OAI-PMH protocol to the VO. Searchable registry collects metadata from the publishing registry not only of JVO but also of the other VO project. A native XML DB, Karearea©(SEC), is used for managing the collected metadata. Our searchable registry also collects table metadata from skynodes. These data are used for resolving the service endpoint URL from the SQL table name, and for assisting a user to make a query. The implemented WS interfaces enables a keyword-based search and an identifier-based search. ADQL search interface is under preparation.

3. VO Portal Service

A VO portal service provides a seamless access to the distributed astronomical services. JVO portal (Fig.2a) accepts JVO query language (JVOQL) (Shirasaki et al) which has similar syntax with the VO standard. JVOQL can describe a cross match query not only for the catalog services but also for image and spectrum services. As an example, JVOQL shown below describes a query to X-ray, IR, and optical catalogs and an optical image service at the same time.

```
SELECT  ir.object, ir.flux_iso_45, opt.number, opt.mag_auto_B,
        opt.mag_auto_V,  opt.mag_auto_R,  opt.mag_auto_I,
        opt.mag_auto_Z, x.number, x.rate2300,
        img.object, img.filter_id, img.access_ref, img.format
FROM    jvo:sxds.sxdsr1 AS opt,
        jvo:sxds.swire_xmm AS ir,
        jvo:sxds.xmm_epic_sxds AS x,
        jvo:subaru.spcam_mos_view AS img
WHERE   Region('Circle 34.2 -5.0 0.05') AND opt.mag_auto_R < 24
        AND distance((ir.ra, ir.dec), (opt.ra, opt.dec)) < 2 [arcsec]
        AND distance((ir.ra, ir.dec), (x.ra, x.dec)) < 2 [arcsec]
        AND distance((x.ra, x.dec), (opt.ra, opt.dec)) < 2 [arcsec]
        AND img.region = BOX((opt.ra, opt.dec), 0.02, 0.02)
        AND img.filter_id = 'W-C-RC'
```

"distance" function is used to specify the cross match condition among the three catalogs. A service identifier is used for specifying a table uniquely. For an examples, a table of "sxdsr1" provided by the Subaru deep survey catalog service whose identifier is "ivo://jvo/sxds" is identified by "jvo:sxds.sxdsr1". JVOQL is divided into queries for each data service, and each query is translated to VO standard query formats, ADQL, SIAP and SSAP, according to the service type.

4. Performance Test

Performance are measured to construct robust and reliable data services. Especially a memory usage is the most critical one to deal with a huge number of accesses, and it was expected that, in our AXIS implementation, a lot of memory are used for XML to Java object deserialization. We noticed that proto 3 is consuming a lot of memory and they are not freed. This is now fixed by
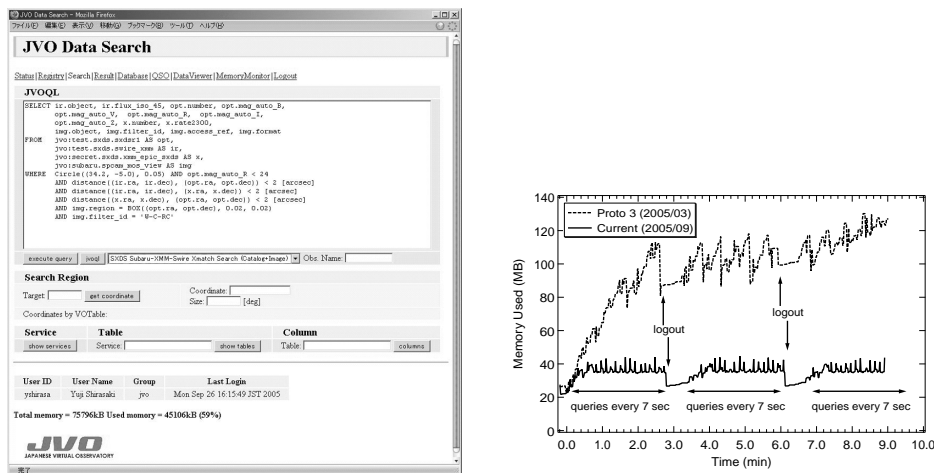
Figure 2.   (a) left: JVO portal query page. (b) right: Result of the performance test (memory usage history).

making a constraint on the memory usage of each user and enforcing an active GC (Fig.2b).

5.   Science Use Case

We applied this VO system to the study of environment of QSOs by combing the QSO catalog service and the Subaru image data service. A workflow for this study is as follows: 1. Select QSO coordinates from the QSO database. 2. Search multi-bands imaging data which covers the QSO regions. 3. Create a catalog from the imaging data by invoking a SExtractor Web service. 4. Estimate the objects' photoZ around the QSO. 5. Clustering Analysis. We succeeded to federate the catalog and image data service. By incorporating the workflow system (Tanaka et al.), which is under development, all the procedure will be done in an automatic way.

6.   Future Plan

We are now developing an operational system aiming at an alpha version release in the end of this physical year ($\sim$2006/3). More user friendly interface for query and data visualization will be provided in this version. We will start a data service of Subaru Suprime-Cam reduced image from 2006/4.

References

Shirasaki, Y. et al. 2005, will be given at a later time, [P1-1-23]
Tanaka, M. et al. this volume, [P107]