

Open SkyQuery

VO Compliant Dynamic Federation of Astronomy Archives

Tamás Budavári¹, Alex Szalay¹, Tanu Malik¹, Ani Thakar¹, William O'Mullane¹, Roy Williams², Jim Gray³, Bob Mann⁴, Naoki Yasuda⁵

¹The Johns Hopkins University, ²California Institute of Technology, ³Microsoft Research, ⁴Royal Observatory Edinburgh, ⁵Nat'l Astronomical Observatory of Japan

ABSTRACT

We discuss the redesign of the SkyQuery architecture, originally built as a simple proof of concept for dynamic federation of astronomical archives. In keeping with the Virtual Observatory philosophy of hierarchical services, the design of Open SkyQuery is based upon higher level services extending the basic functionality of the current VO standard, the ConeSearch. Open SkyQuery implements the VO specifications for data access, retrieval and spatial join. Data are published via Web Services called SkyNodes providing a rich functionality including footprint coverage. SkyNodes are discovered through the VO registry. We propose to have at least two levels of SkyNode compliance (Core and Advanced). We will also provide templates for publishing data into a SkyNode.

Keywords: SkyQuery, VOTable, ConeSearch, VOQL, ADQL, VORegistry

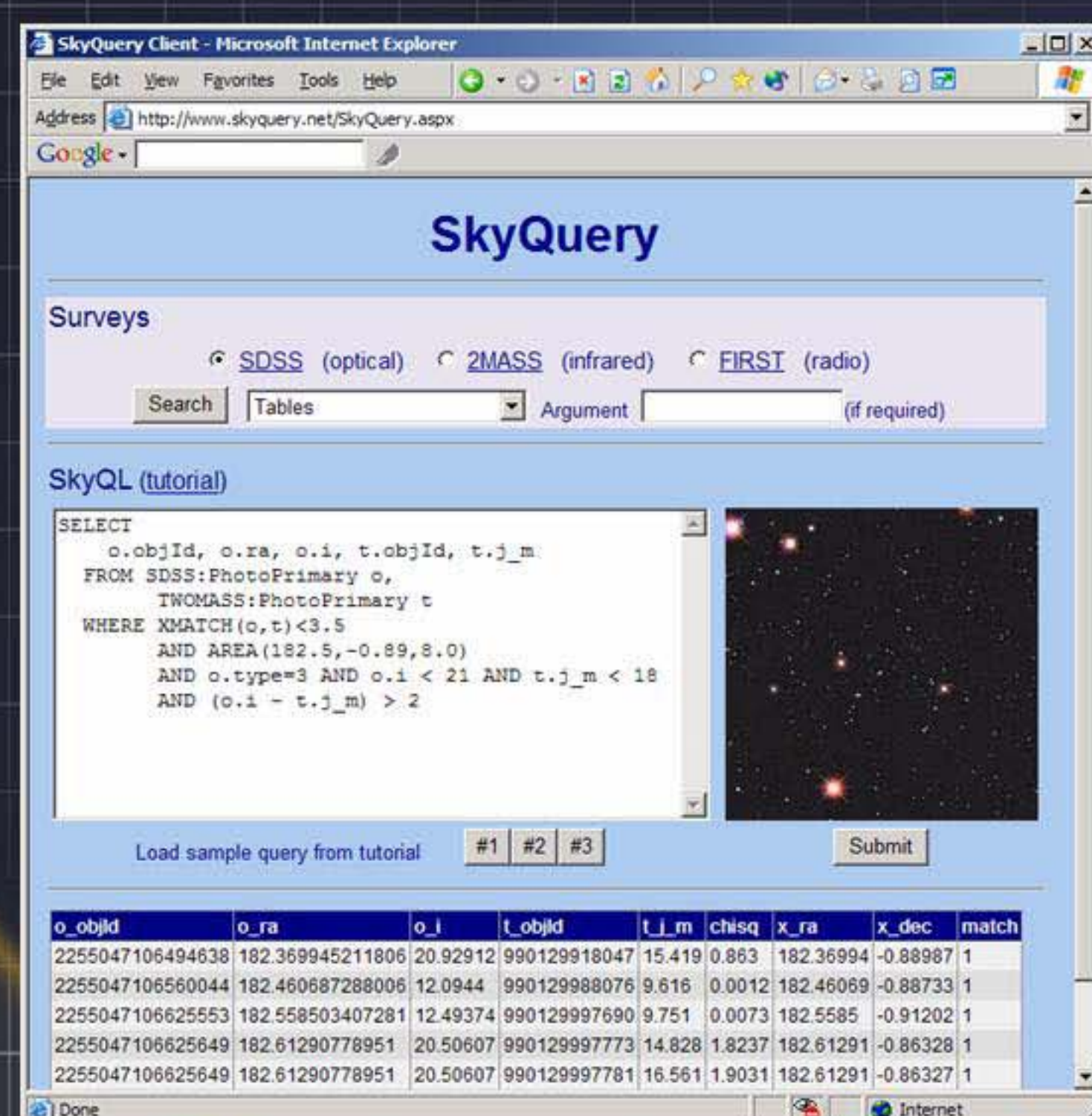
MOTIVATION

With the advent of large CCD detectors, the way astronomy is done changes rapidly. Because of the exponential growth in the size and speed of the silicon chips, new surveys are expected to have significantly higher data rates. These survey projects become both the authors and publishers of their data [Szalay et al. 2002]. In this exponential world at every given time only 10% of all astronomical information is available in central archives. In order to have access to all up-to-date observations, we need to find a way to federate geographically separated astronomical archives.

Current sky surveys such as SDSS, 2MASS, DPOSS have proven that discoveries are always made at the boundaries when going deeper or using more colors. By covering different wavelength ranges, surveys can very well complement one another if one finds a way to combine them. The Virtual Observatory initiative was launched to help this process.

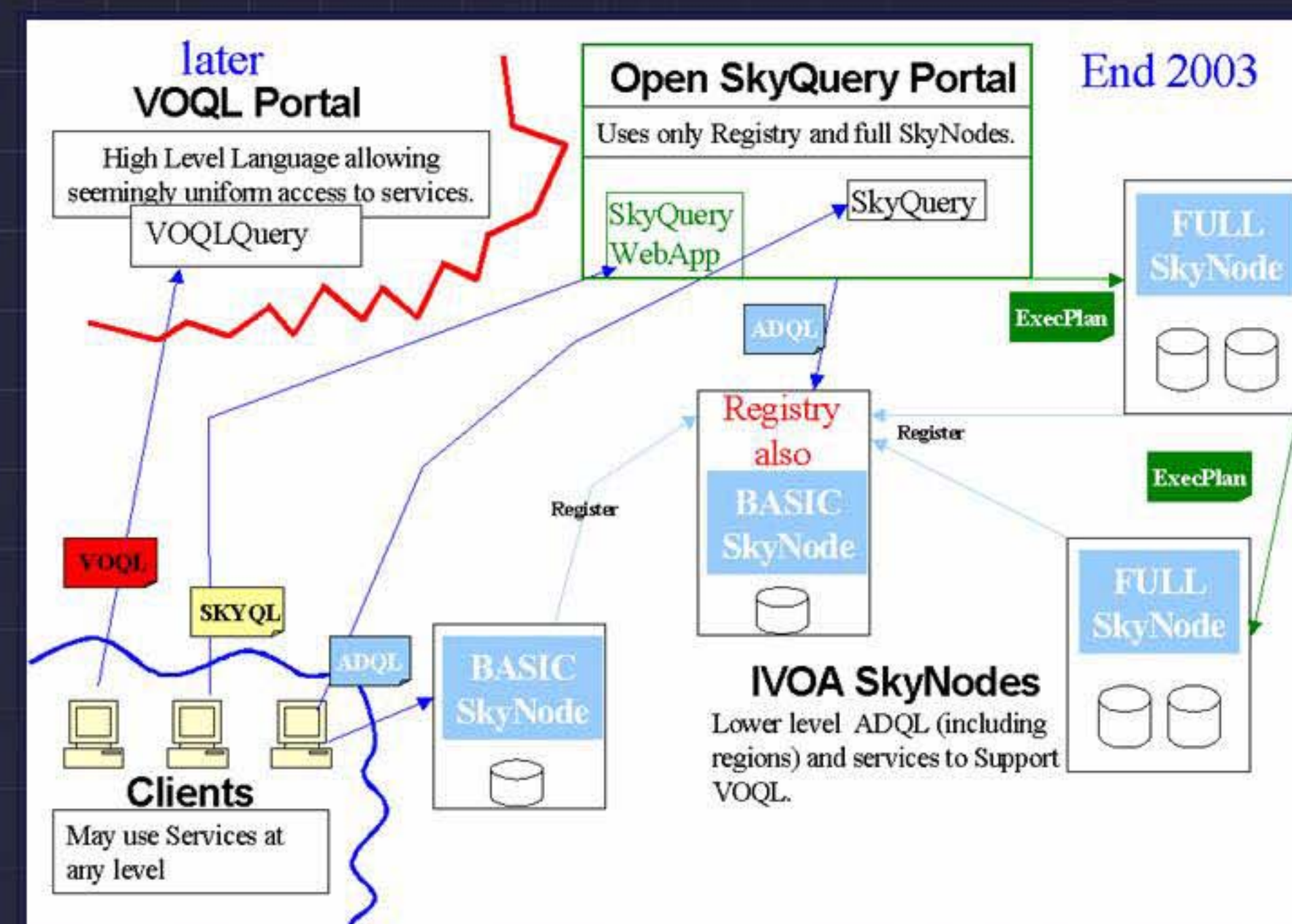
THE PROTOTYPE SKYQUERY

was built last year in six weeks as a feasibility study [Budavári et al. 2002; Malik et al. 2002]. It used a hierarchy of XML Web Services to implement a distributed query system that provided seamless access to SDSS, 2MASS and FIRST data. Since the launch of the SkyQuery web site, many other catalogs have become available including the Isaac Newton Telescope's Wide Field Survey (INTWFS), IRAS, NVSS, 2dF, PSCz, 2QZ and Rosat.



VIRTUAL OBSERVATORY STANDARDS

The SkyQuery architecture is being redesigned to utilize the recently emerging VO standards such as the VOTable, the Astronomical Data Query Language (ADQL), the VO Query Language (VOQL) and the VO Registry services. The data are going to be published by the SkyNodes that implement XML Web Services to extend the basic functionality of the ConeSearch.



THREE LAYERS OF VOQL

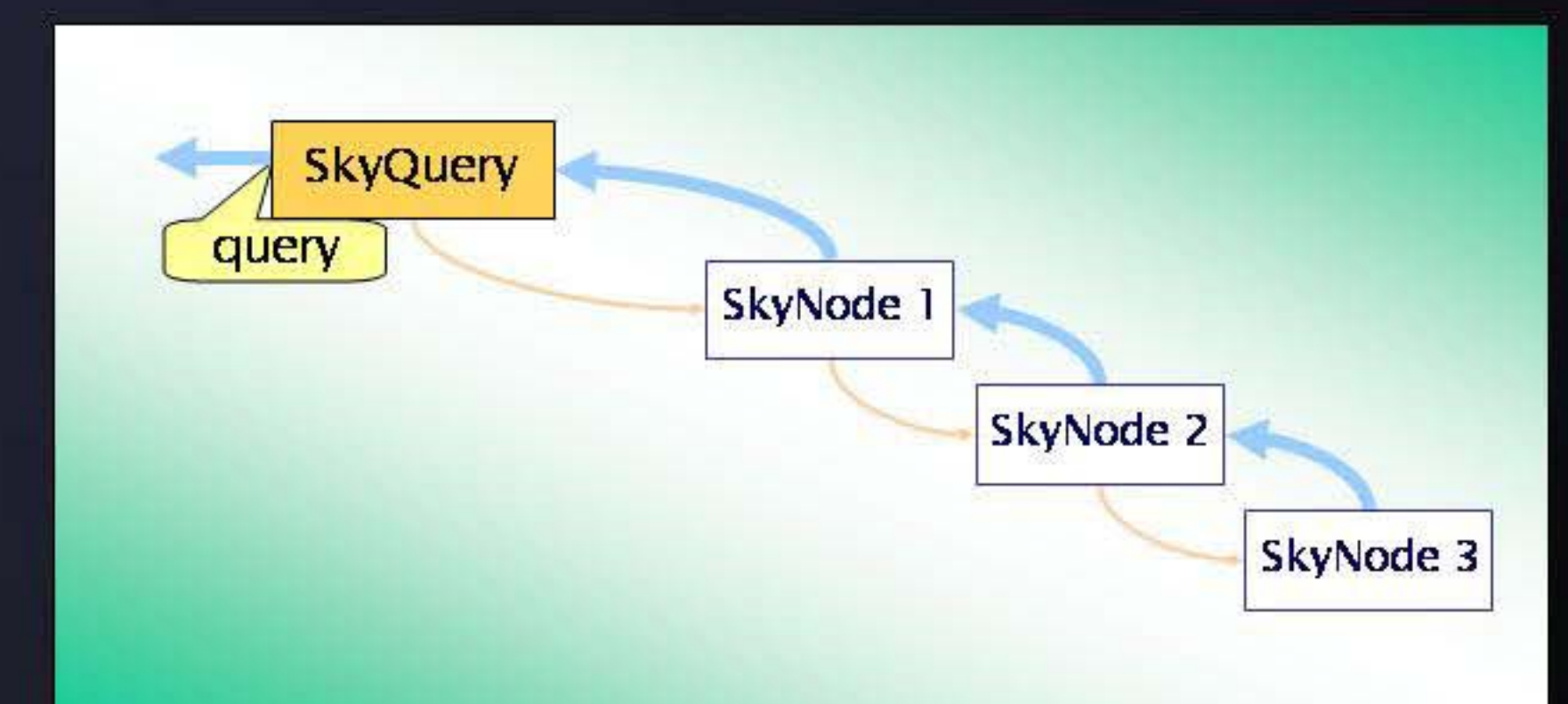
We propose to have three layers of VOQL building on top of one another:

- **VOQL1 – Web Services:** ADQL and VOTABLE to exchange information between machines
- **VOQL2 – SkyQL:** SQL-like query language and federation system, i.e. combination of SkyQuery, JVOQL and VO standards
- **VOQL3 – SkyXQuery:** future XML-based query language

Figure on the left illustrates the Open SkyQuery architecture and shows the relations between the components.

SKYQUERY STRATEGY

In order to ensure fast response, one needs to optimize the query plan. Our simulations show that the simple sequential execution proves to be optimal because today the wire speed is the limiting factor. One needs to arrange to SkyNodes in the reverse order by number of returned records so that the least amount of data is transferred. This simplifies the logic of the portal significantly. However, the SkyNodes are designed to deal with more complicated query plans, so that the system may be enhanced easily later on.



REFERENCES AND LINKS

1. Budavári, T. et al. 2002 ADASS XII. 12, 31, 'SkyQuery – A Prototype Distributed Query Web Service for the VO'
 2. Malik, T., et al. 2002 CIDR '03, 17, 'SkyQuery: A WebService Approach to Federate Databases'
 3. Szalay, A. S., et al. 2002 SPIE, 'Web Services for the Virtual Observatory'
- <http://www.SkyQuery.net>
 ➤ <http://skyservice.pha.jhu.edu/develop/vo/adql/>
 ➤ <http://www.ivoa.net/twiki/bin/view/IVOA/IvoaVOQL>

THE US NATIONAL VIRTUAL OBSERVATORY

Supported by the National Science Foundation's Information Technology Research Program under Cooperative Agreement AST0122449 with The Johns Hopkins University

