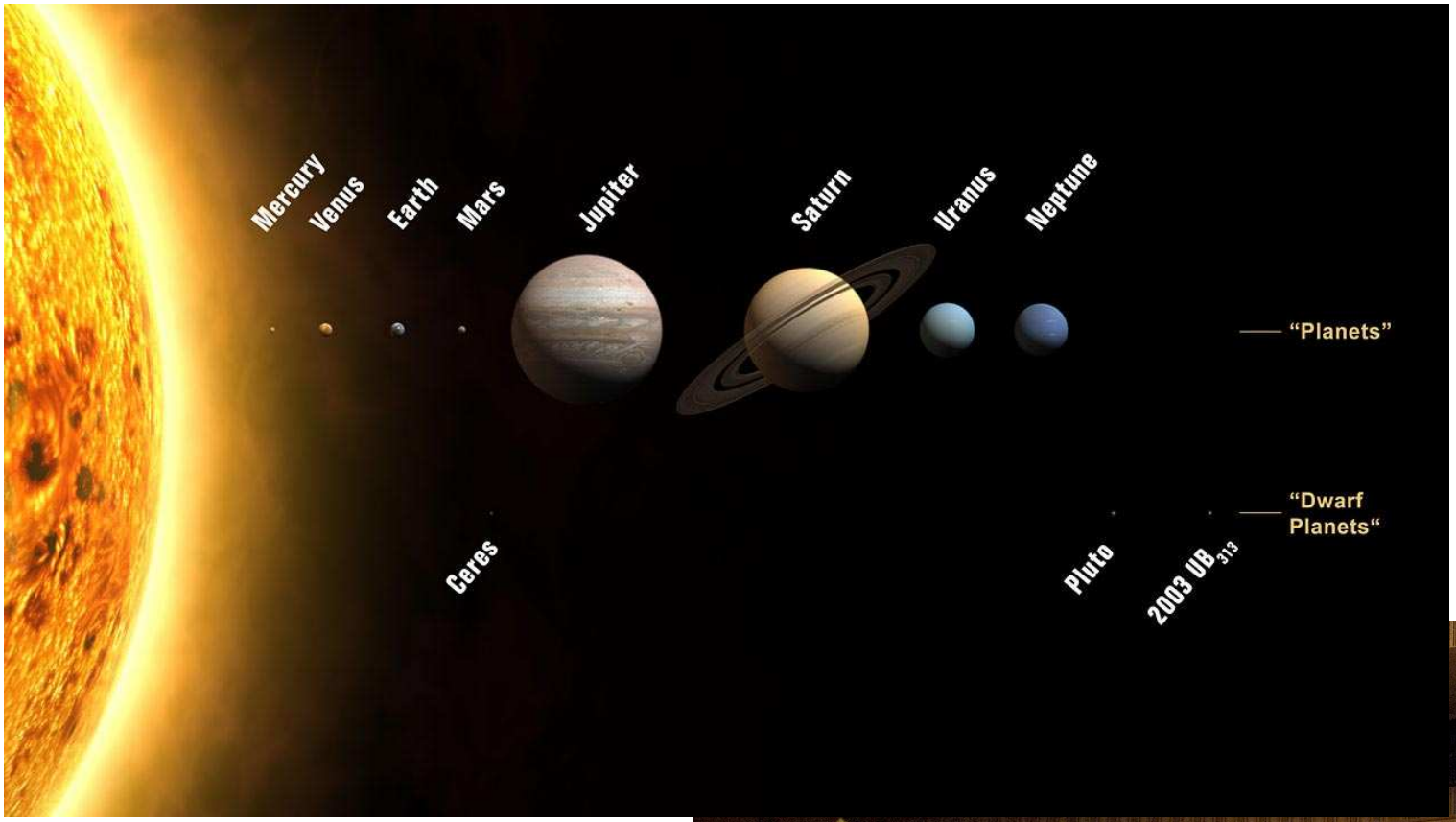
i-explosion

平成17年発足特定領域研究公募研究

最新情報技術を活用した超大規模天文データ解析機構の研究開発

Masatoshi Ohishi / NAOJ & Sokendai
大石雅寿 / 国立天文台 & 総合研究大学院大学

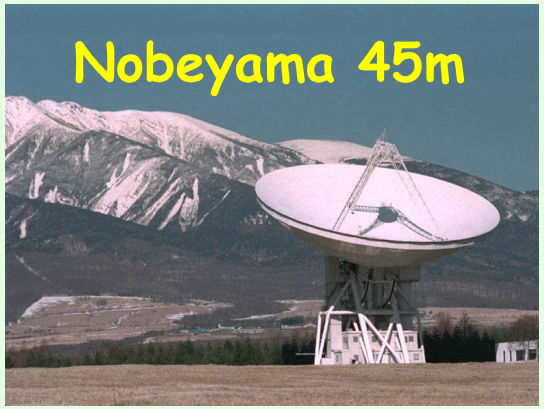
masatoshi.ohishi@nao.ac.jp

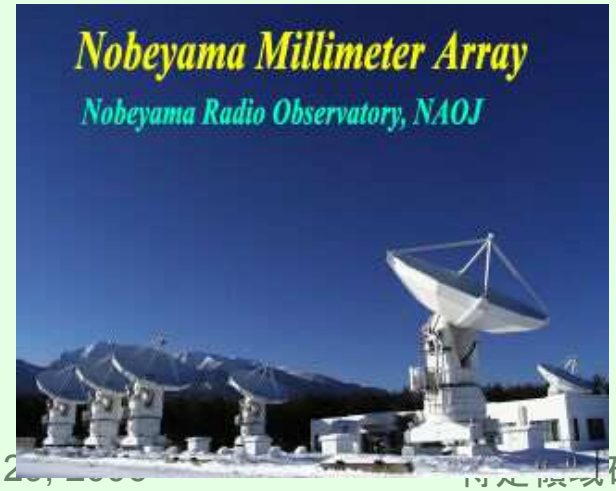


September 29, 2006

Data Resources in NAOJ

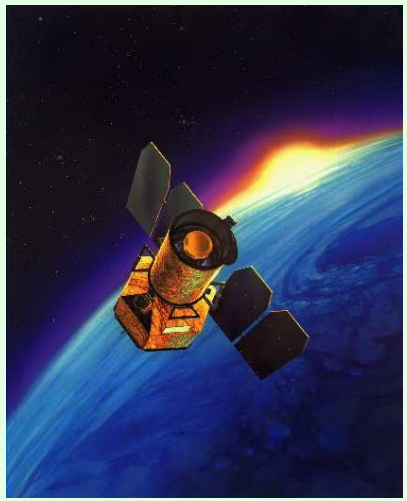
- Subaru 8.2m Optical-Infrared Telescope
- Kiso 105cm Schmidt Camera
- Okayama 188cm Optical Telescope
- Nobeyama 45m Radio Telescope
- Nobeyama Millimeter Array
- Nobeyama Radioheliograph
- VSOP
- VERA
- *ALMA*





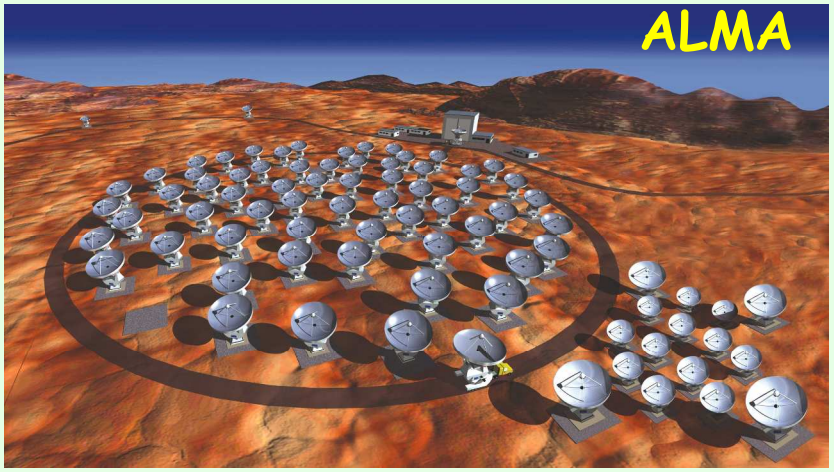
Data Resources in the World

- ESO VLT @ Paranal
- ESO telescopes @ La Silla
- Telescopes in Mauna Kea
  Subaru, Keck, JCMT, CSO, SMA,,,
- Telescopes on board satellites
  HST, GALEX,,,,

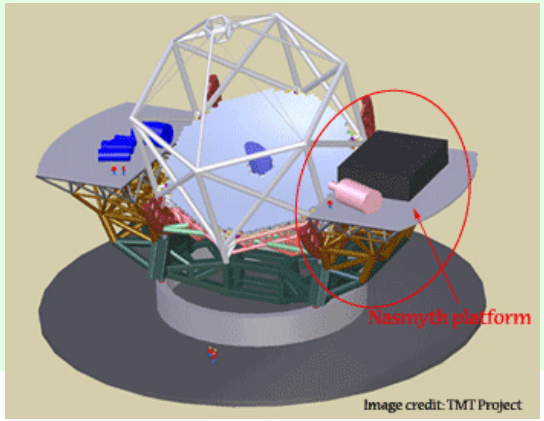


Planned Data Resources

- ALMA
- JWST
- LSST
- LOFAR
- SKA
- Thirty MetreTelescope
- Giant Magellan Telescope
- European Extremely Large Telescope



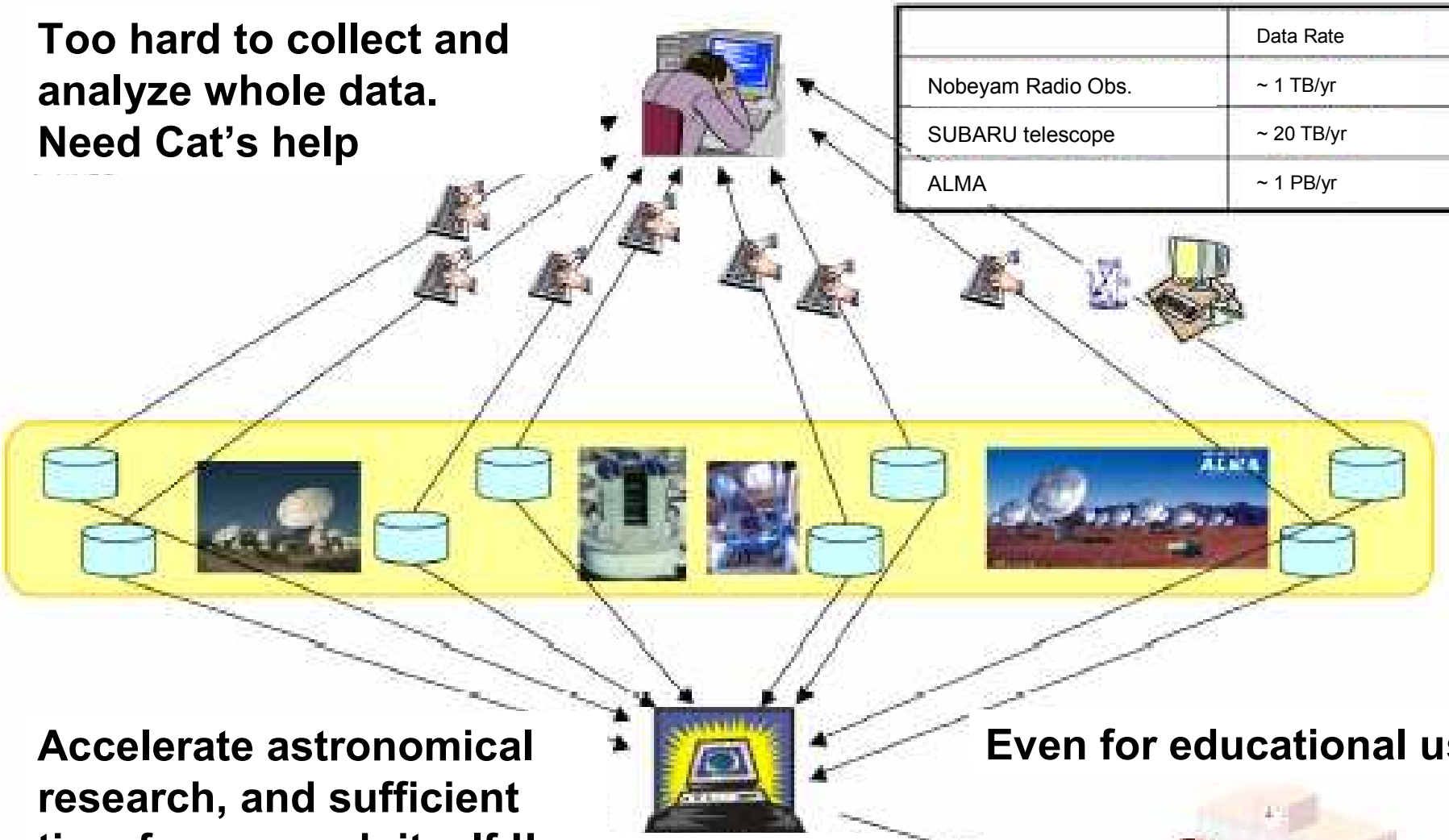
JWST



TMT

Too hard to collect and analyze whole data. Need Cat's help

	Data Rate
Nobeyam Radio Obs.	~ 1 TB/yr
SUBARU telescope	~ 20 TB/yr
ALMA	~ 1 PB/yr



Accelerate astronomical research, and sufficient time for research itself !!

Virtual Observatory

Even for educational use

Accessible from anywhere at any time

VO– New Research Infrastructure in the 21st Century

A collection of integrated astronomical data archives and software tools that utilize computer networks to create an environment in which research can be conducted.

http://www.encyclopedia.com/html/v1/virtobserv.asp

International Endorsements

- IAU XXVth GA Res. (2003 Jul.)
- OECD Rec. (‘04 Aug)
  - place archives that may be accessible via internet
  - provide adequate funding as long-term issues

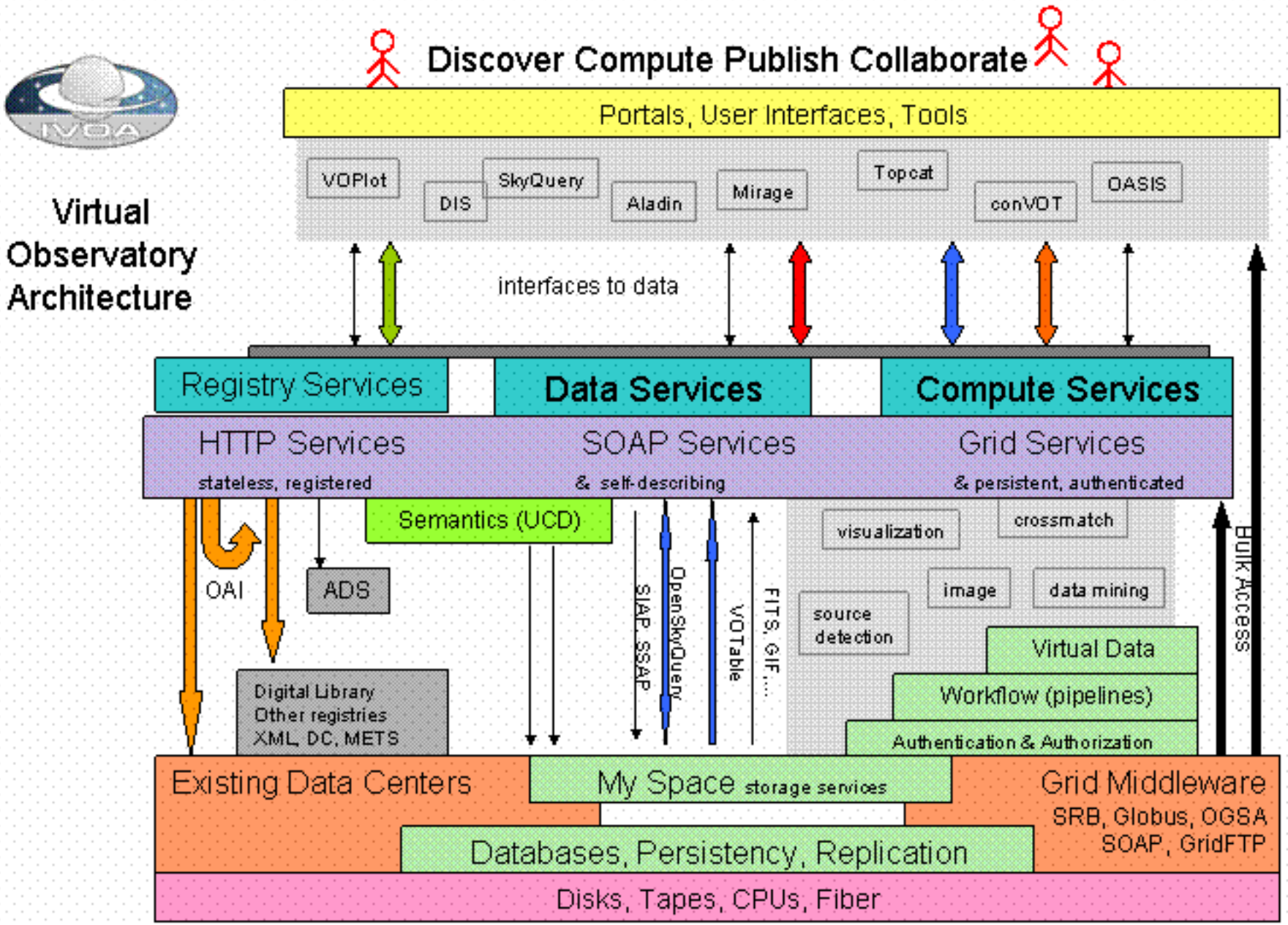
VO Projects in the world

- 16 countries and a region
- International Virtual Observatory Alliance (IVOA)
  Standards to interoperate VOs

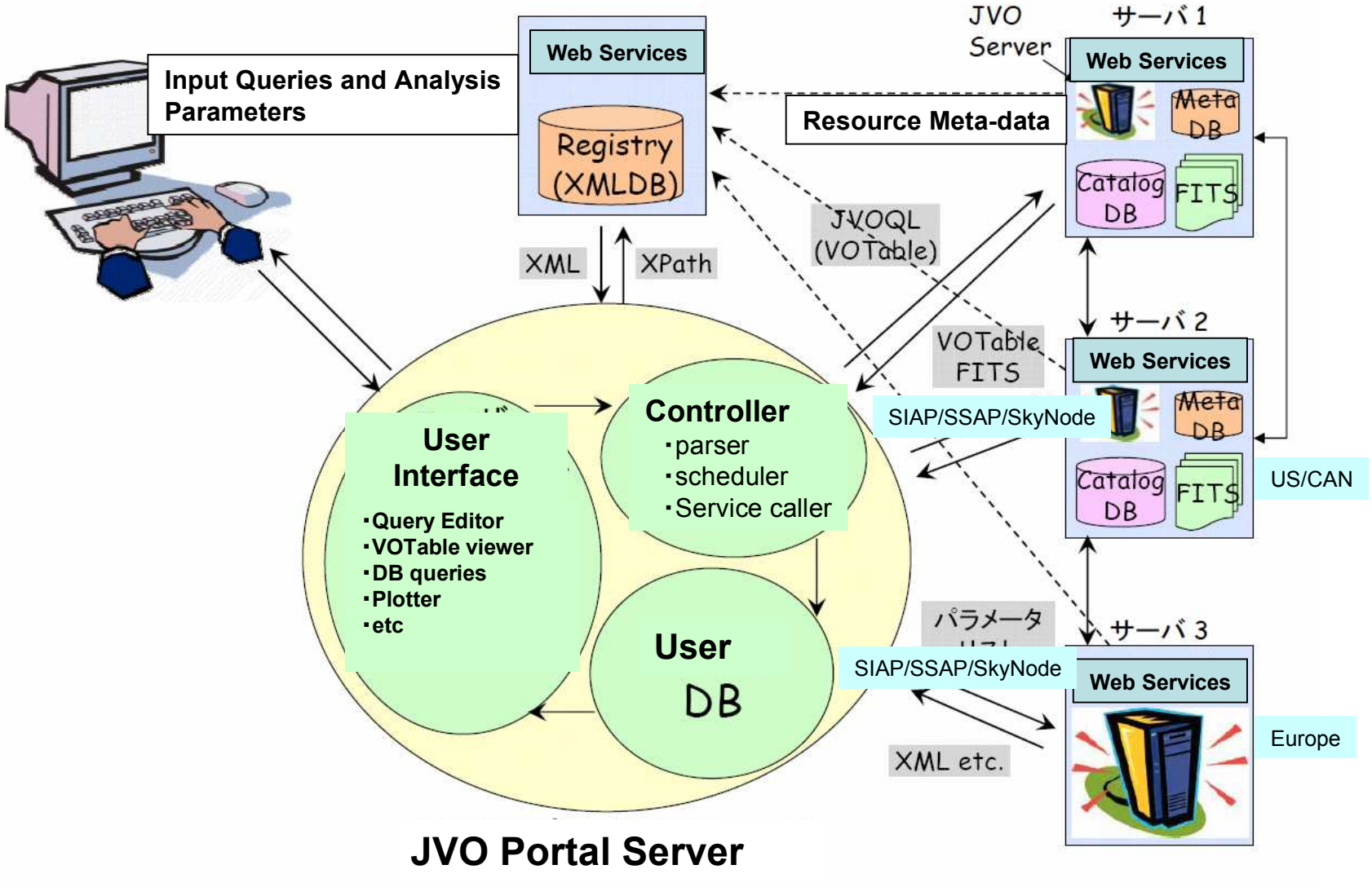
http://www.ivoa.net/

Standardization in IVOA

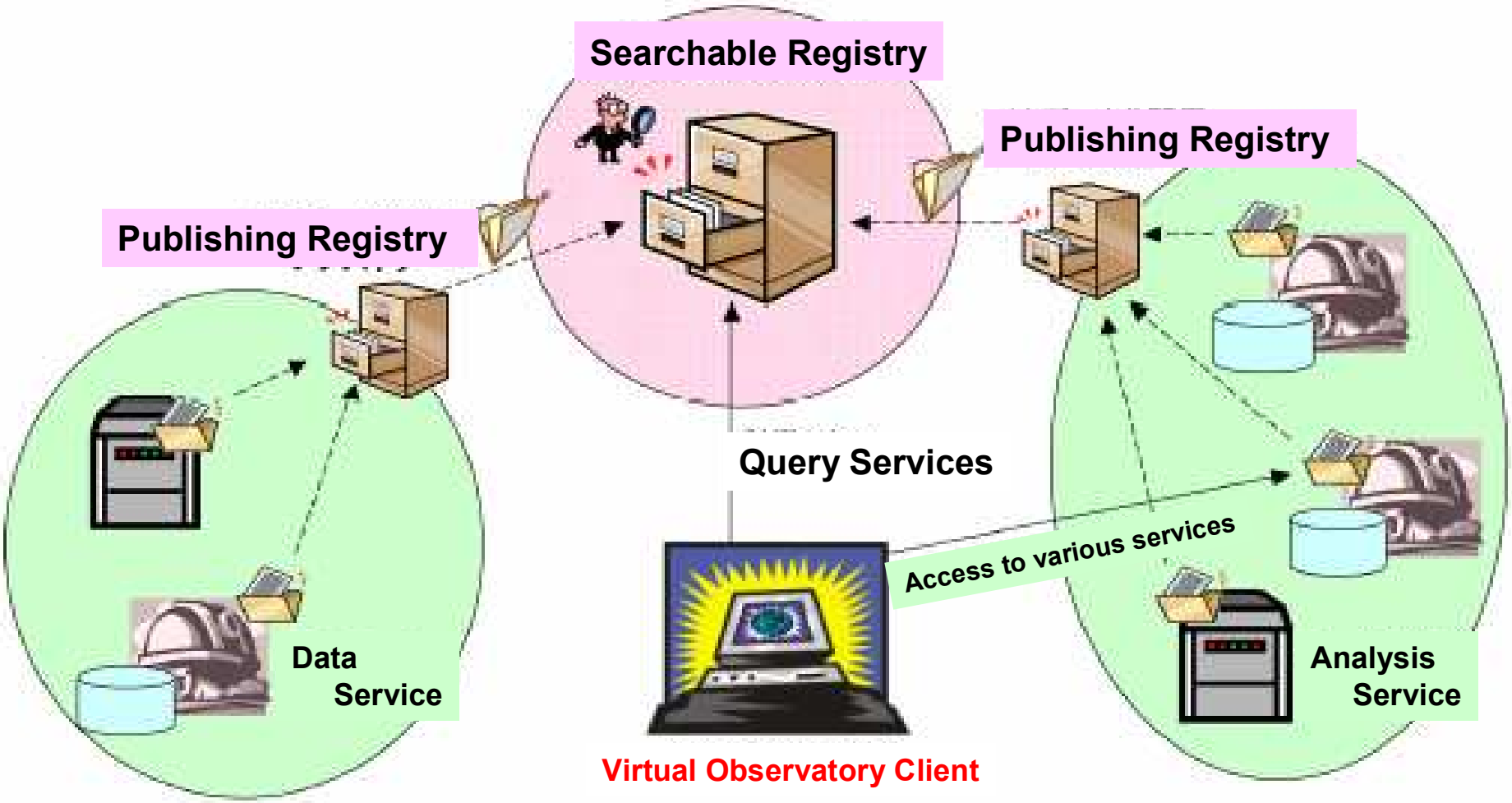
- Query Language to Federated DBs (VOQL)
- Access resource meta-data based on the OAI-PMH
- Access Images, Spectra, Catalogues, etc:
  SkyNode, SIAP, SSAP, STC, etc.
- Unified Attribute Names:
  UCD (Unified Contents Descriptions)
- Output format: VOTable (XML)
- and so on


Virtual Observatory Architecture
Discover Compute Publish Collaborate
Portals, User Interfaces, Tools
VOPlot
DIS
SkyQuery
Aladin
Mirage
Topcat
conVOT
OASIS
interfaces to data
Registry Services
Data Services
Compute Services
HTTP Services
stateless, registered
SOAP Services
& self-describing
Grid Services
& persistent, authenticated
Semantics (UCD)
OAI
ADS
Digital Library
Other registries
XML, DC, METS
SIAP, SSAP
OpenSkyQuery
VOTable
FITS, GIF, ...
visualization
crossmatch
image
data mining
source detection
Virtual Data
Workflow (pipelines)
Authentication & Authorization
BULK Access
Existing Data Centers
My Space storage services
Grid Middleware
SRB, Globus, OGSA
SOAP, GridFTP
Databases, Persistency, Replication
Disks, Tapes, CPUs, Fiber

Schematic diagram of VOs



Exchange of Meta Data:OAI-PMH



JVO Simple Data Search

Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

⇒ Simple | JVOQL

Find Data Service

X-ray AND Search Service

Region Selection Criteria

Object Name				
Crab Name Resolver				
Coordinate	**Frame**	**Size**	**unit**	**Shape**
83.633212 22.014460	J2000	1	deg	box size

Selected Services

Chandra X-Ray Observatory Data Archive

http://cxc.harvard.edu/cda

table name	Data Selection	Query Condition	Submit Query	description
N/A	Data	Condition	Search	

Workflow Status

Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

⇒ All | Detail

Workflow Name : work_20060123185359073

Activity Name	Host	Elapsed Time (s)	Flag	Status
1_executeQuery		0.0		waiting
		0.0		waiting
2_storeVOTable		0.0		waiting

Cancel Results

User ID	User Name	Group	Last Login
ohishi	Masatoshi Ohishi	jvo	Mon Jan 23 18:28:14 JST 2006

Total memory = 91992kB Used momory = 37743kB (41%)

JVO
JAPANESE VIRTUAL OBSERVATORY

Workflow Status

Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

⇒ All | Detail

Workflow Name : work_20060123185359073

Activity Name	Host	Elapsed Time (s)	Flag	Status
1_1_executeQuery	cda.harvard.edu	2.265		success
		0.0		success
2_2_storeVOTable		0.041		success

Results

User ID	User Name	Group	Last Login
ohishi	Masatoshi Ohishi	jvo	Mon Jan 23 18:28:14 JST 2006

Total memory = 91992kB Used memory = 35776kB (38%)

JVO JAPANESE VIRTUAL OBSERVATORY

Workflow Results

Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

⇒ Workflows | Results

Workflow Name : work_20060123185359073

File Name	File Type	Action
result_votable0	VOTable	Viewer Download

User ID	User Name	Group	Last Login
ohishi	Masatoshi Ohishi	jvo	Mon Jan 23 18:28:14 JST 2006

JAPANESE VIRTUAL OBSERVATORY

Total memory = 91992kB Used memory = 38219kB (41%)

VOTable Viewer

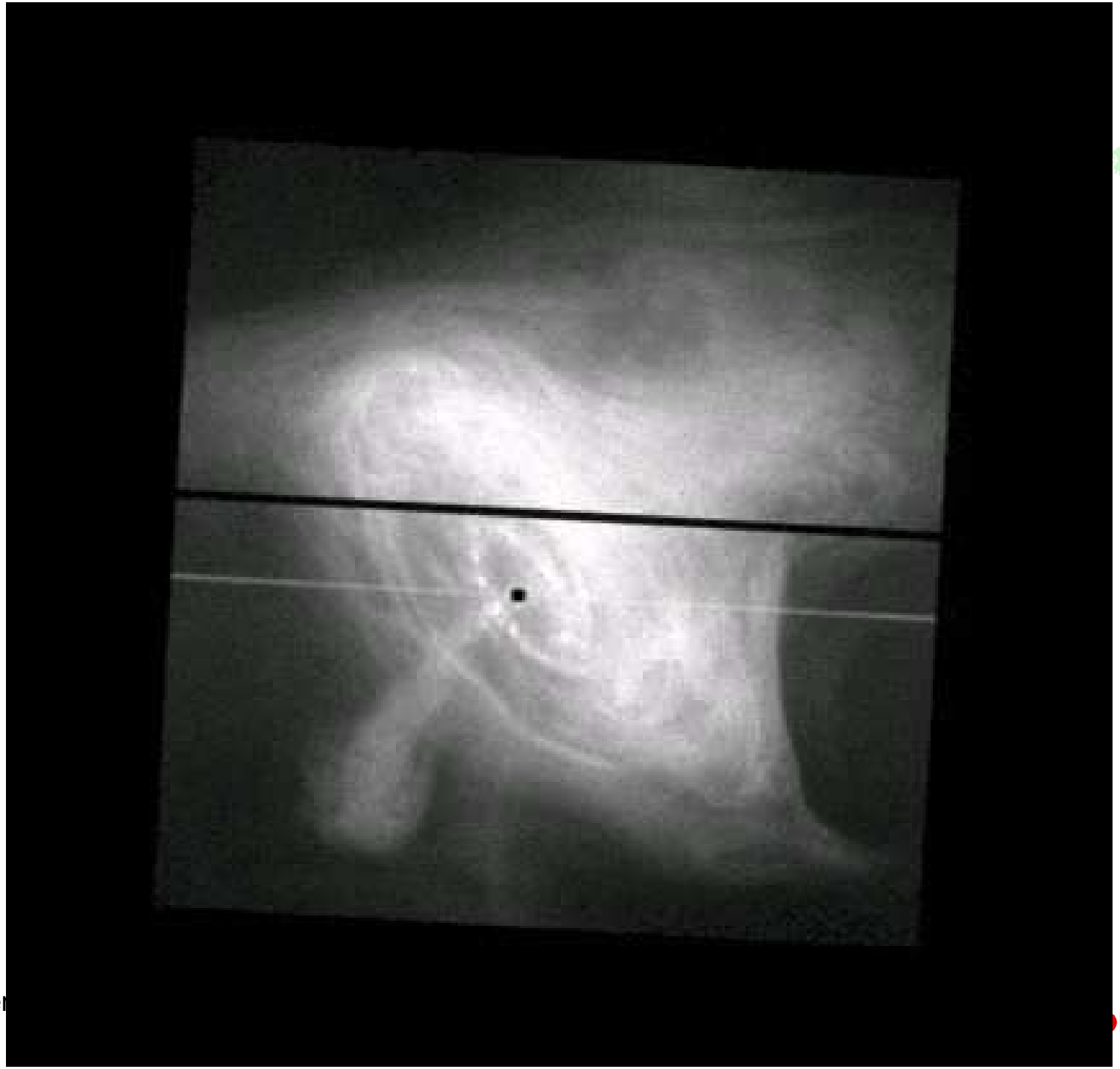
Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

Workflow Name : work_20060123185359073

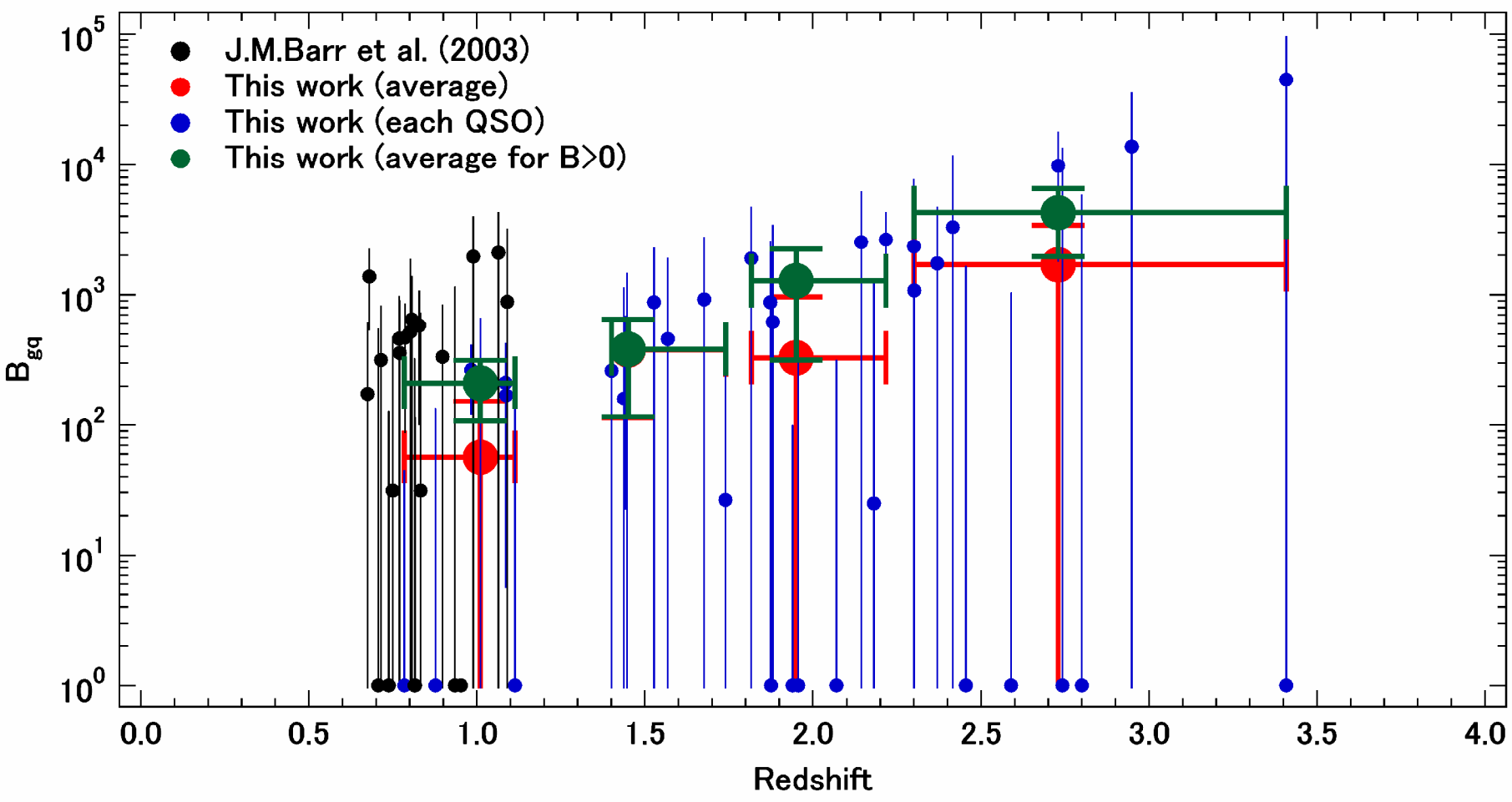
File Name : result_votable0

data viewer | Reset | VOTable download

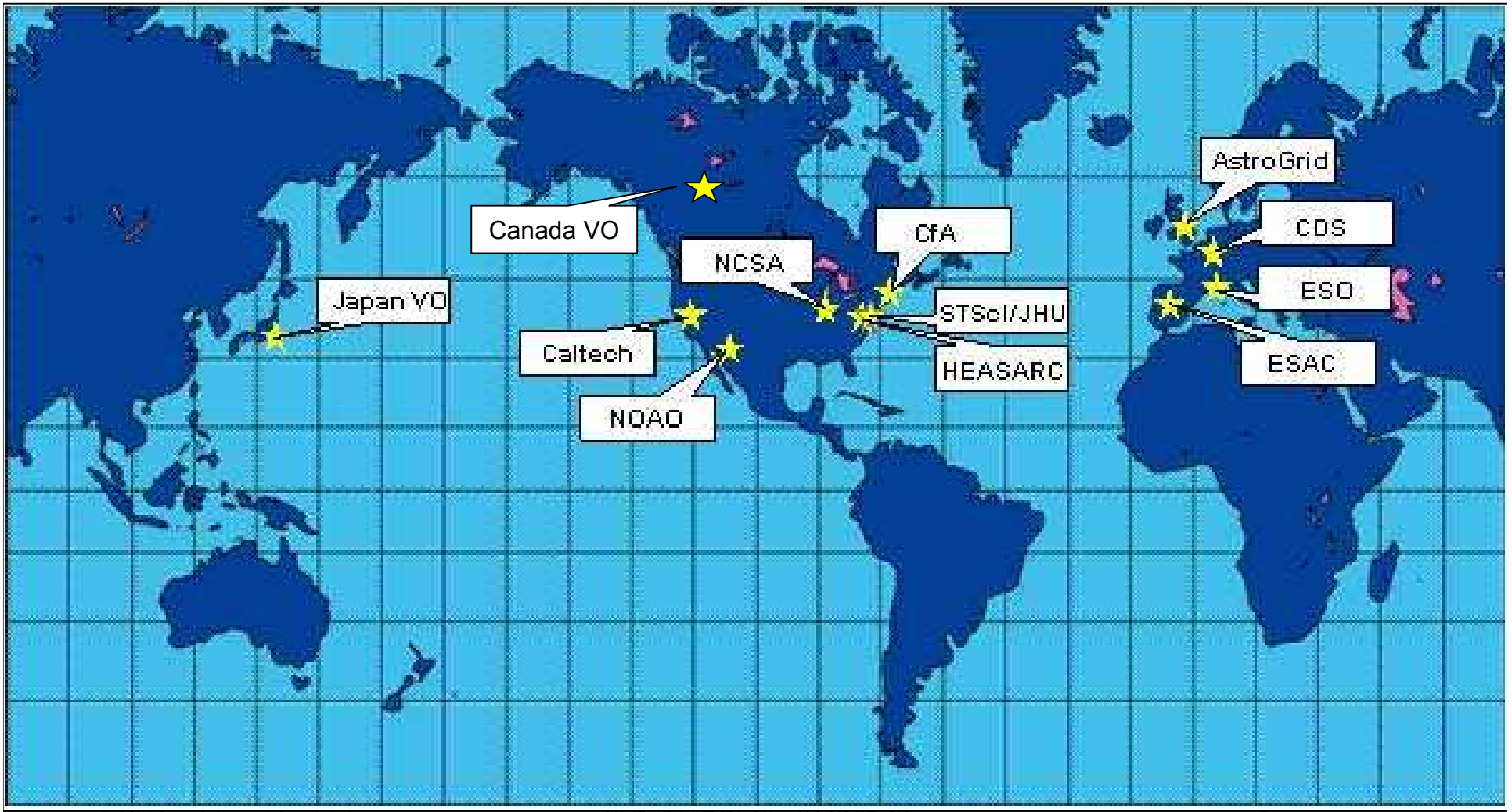
#	check	download	VOX:Image_Title	VOX:DateObs	POS_EQ_RA_MAIN	POS_EQ_DEC_
			Object	DATE-OBS(TT)	RA	Dec
0		Download	Crab Nebula	Tue Jan 27 01:11:43 EST 2004	83.63333333333333	22.01444444444
1		Download	Crab Nebula	Tue Jan 27 01:11:43 EST 2004	83.63333333333333	22.01444444444
2	☐	Download	Crab Nebula	Tue Jan 27 01:11:43 EST 2004	83.63333333333333	22.01444444444
3	☐	Download	Crab Nebula	Tue Jan 27 01:11:43 EST 2004	83.63333333333333	22.01444444444
4		Download	Crab Nebula	Tue Jan 27 01:11:43 EST 2004	83.63333333333333	22.01444444444



Study on Environment around QSOs



Interconnected VO Data services in the World



Latest Science outputs

- IAU GA – VO Special Session (SPS3)
  - August 17, 18, 21, 22 & 23
- 500+ participants registered
  - ~30 science talks out of ~50
  - ~100 posters


INTERNATIONAL VIRTUAL OBSERVATORY ALLIANCE
IVOA
EURO VO
AstroGrid
EURO VO

VO Summer School (August)

Work Flow Description Language (XML tags)

- 変数定義
  - 〈variables〉 ～ 〈/variables〉
    - 〈variable〉を要素にとる
- 外部サービス呼出
  - 〈invoke〉 ～ 〈/invoke〉
    - 〈input〉・〈output〉を要素にとる
    - 属性値でサービスを指定
- 内部コマンド実行
  - 〈command〉 ～ 〈/command〉
    - 〈input〉・〈output〉を要素にとる
- 逐次実行
  - 〈sequence〉 ～ 〈/sequence〉
    - 〈command〉などを要素にとる
- ループ処理
  - 〈for〉 ～ 〈/for〉
    - 要素を順次実行する
  - 〈parfor〉 ～ 〈/parfor〉
    - 要素を並列実行する
- 条件判定
  - 〈if〉 ～ 〈/if〉

BPEL4WS※をベースにして天文学解析フローを記述するためのタグを実装

※BPEL4WS = Business Process Execution Language for Web Services
複数のウェブサービスを連係させ、複雑なフローを記述できる。

Workflow

Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

Upload your workflow Upload

参照...

Edit workflow Execute | clear | Sequential Query | Brown Dwarf Search #1 | Brown Dwarf Search #2 | QSO Study

```
<?xml version="1.0" encoding="UTF-8"?>
<Workflow xmlns="http://jvo.nao.ac.jp/workflow/v0.4" xmlns:xsi="http://www.w3.org/2001/XMLSchema-ins
  <!-- ************************************************************ -->
  <!-- Search for Brown Dwarf candidates -->
  <!-- 01. Simplest Case -->
  <!-- ************************************************************ -->
  <name>SequentialQuery</name>
  <author>M.Tanaka</author>
  <identifier>ivo://jvo.nao.ac.jp/workflow/Y.Shirasaki/SequentialQuery</identifier>
  <create_date>2006/01/22 00:20:00</create_date>
  <update_date>2006/01/22 00:20:00</update_date>
  <description>Query to Mutple Data Services</description>
  <status>wating</status>

  <variables>

    <variable name="query1" type="String">
       <value>SELECT * FROM ivo://cxc.harvard.edu:cda WHERE region = BOX( (83.633212,22.014460), 0.2
    </variable>

    <variable name="query2" type="String">
       <value>SELECT ra, dec, dered_z  FROM   ivo://jvo/sdss:photoobjall WHERE  Region('Circle 20. 0
    </variable>
```

Edit workflow Execute | clear | Sequential Query | Brown Dwarf Search #1 | Brown Dwarf Search #2 | QSO Study

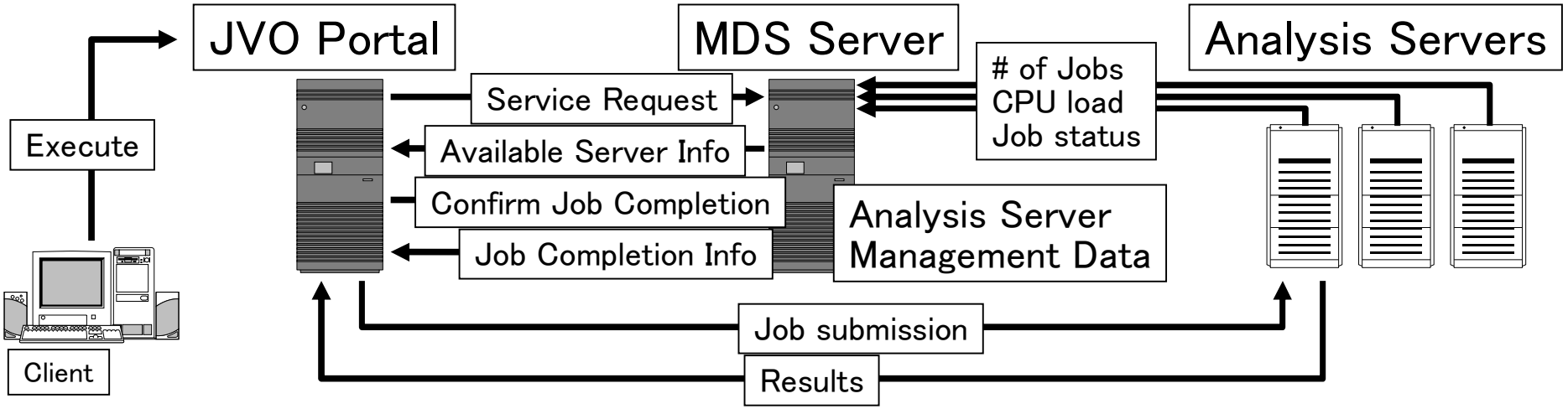
```
<command xsi:type="builtin" proc_name="noname" name="executeQuery">
  <activity_status>
    <status>wating</status>
    <elapsed_time>0.0</elapsed_time>
    <log>none</log>
    <error_info>none</error_info>
  </activity_status>
  <input>
    <varRef>query1</varRef>
  </input>
  <output>
    <varRef>result1</varRef>
  </output>
</command>

<command xsi:type="builtin" proc_name="noname" name="executeQuery">
  <activity_status>
    <status>wating</status>
    <elapsed_time>0.0</elapsed_time>
    <log>none</log>
    <error_info>none</error_info>
  </activity_status>
  <input>
    <varRef>query2</varRef>
  </input>
  <output>
    <varRef>result2</varRef>
  </output>
</command>
```

Overview of workflow Update

Management of Multiple Servers

- Data Analysis services under JVO
  - SExtractor (Image → Catalog)
  - HyperZ (Catalog → photometric redshift)
    - In operation on multiple servers

→ Monitor and Discovery Service (MDS) server



Parallel Executions on Multiple Servers are available

並列実行の実装

- 各解析サーバは、実行中のジョブ数やロード情報を定期的にMDSサーバに報告する。
- クライアントは実行したサービスの実行場所 (URL) を解決するためにMDS サーバに問い合わせる。その結果にもとづきジョブを実行（Web サービス）。
- MDS サーバは各解析サーバ上で走っているジョブの数等を管理しており、それに基づきジョブの投入が可能であるかを判断する。ジョブの投入が可能であるサーバの中から、ラウンドロビンロジックによって投入先を 決定し、クライアントに通知する。
- 実行可能なサーバが見付からない場合、クライアントは見付かるまで定期的にMDS サーバに問い合わせる。
- ジョブは長い場合があるので、非同期実行になっている。ジョブの終了時に解析サーバー側から MDS サーバにジョブの終了を報告する。
- クライアントは定期的に MDS に自分が実行したジョブが終了したかを問い合わせる。終了している場合は、解析サーバーから結果の取得を行う。データの転送は ftp をつかっている。

Workflow Status

Status | Registry | Search | Workflow | Result | Database | QSO | DataViewer | Link | MemoryMonitor | Logout

⇒ All | Detail

Workflow Name : work_20060123200255167

Activity Name	Host	Elapsed Time (s)	Flag	Status
1_executeQuery	cda.harvard.edu	1.93		success
		0.0		success
2_executeQuery	jvo.nao.ac.jp	2.143		success
		0.0		success
3_executeQuery	pma.iso.vilspa.esa.es	4.632		success
		0.0		success
4_executeQuery	www.cadc-ccda.hia-iha.nrc-cnrc.gc.ca	0.0		executing
		0.0		waiting
5_executeQuery		0.0		waiting

Cancel Results

User ID	User Name	Group	Last Login
ohishi	Masatoshi Ohishi	jvo	Mon Jan 23 19:56:18 JST 2006

Presentations

- Invited talks
  - COSPAR GA in Beijing (Ohishi)
  - IAU GA in Prague (Yasuda, Ohishi)
  - ADASS in Tucson (Shirasaki)
  - RIKEN (Ohishi)
- Contributed papers
  - SPIE in Orland (Shirasaki)
  - IVOA InterOp meetings (Shirasaki, Tanaka)
  - IAU GA in Prague (Shirasaki)

New Moves

- Some VO projects move from R&D phase to operations phase
  - UK, Japan, US, EU,,,
- New VO projects
  - Brazil, Chinese Taipei, Blugaria, Tagikistan, Ireland, Czech,,,

Items to be done

- Complete WFDL / WF execution mechanism
  → IAU/IVOA standard
- Distributed Storage to store query/analysis results
- Secure access to VOs : single-sign-on
- Other Standardizations
  - Standard application interface
- Advertisement to Data Centers / Astronomers
  - data centers need implement VO interfaces