i-explosion

平成17年発足特定領域研究公募研究　A01-23

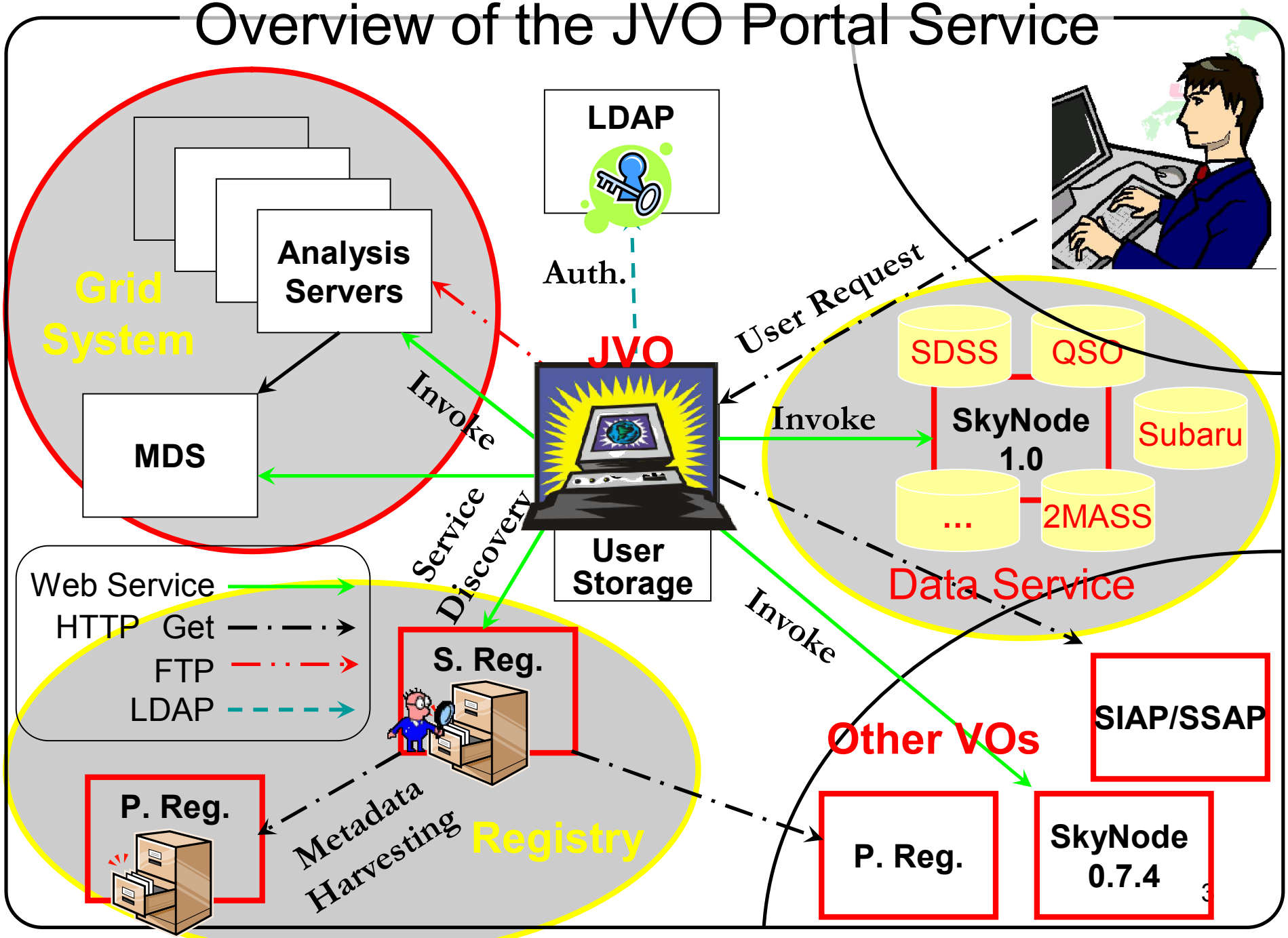
最新情報技術を活用した超大規模天文データ解析機構の研究開発

研究代表者:
Masatoshi Ohishi / NAOJ, Sokendai & NII
大石雅寿 / 国立天文台，総合研究大学院大学，
国立情報学研究所

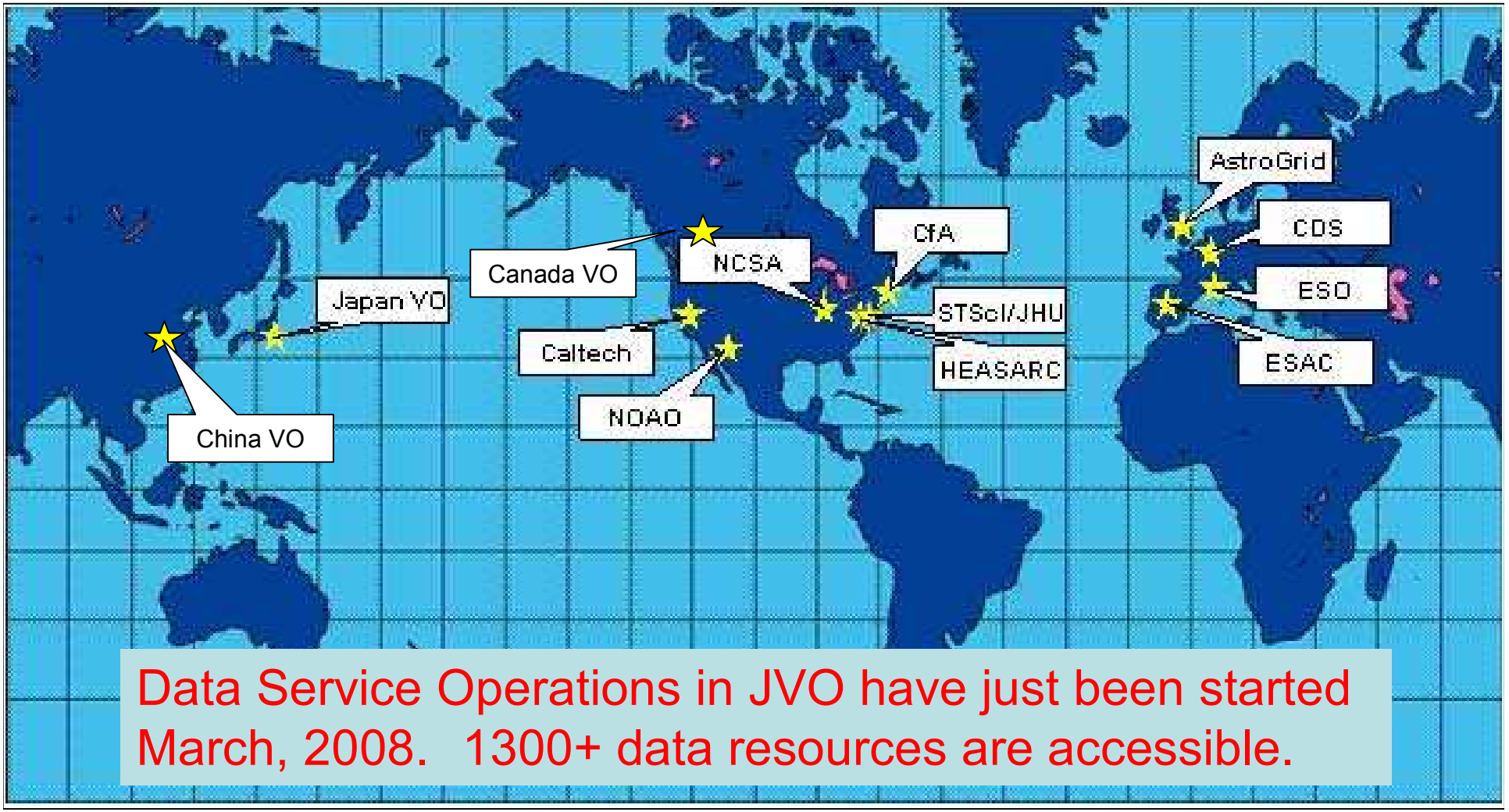
masatoshi.ohishi@nao.ac.jp


大量のデータをどう処理したらよいか悩む天文学者。猫の手も借りたい状況。
データ生成率
野辺山宇宙電波望遠鏡
~ 1TB/年
すばる望遠鏡
ALMA 電波望遠鏡
Universe on Your Desktop
VO の
効率的に研究を進める天
者、研究のアイデアも豊富に浮かぶ。
バーチャル天文台
教育の教材としても利用できる。
いつでもどこでも天文データにアクセスできる。

Overview of the JVO Portal Service



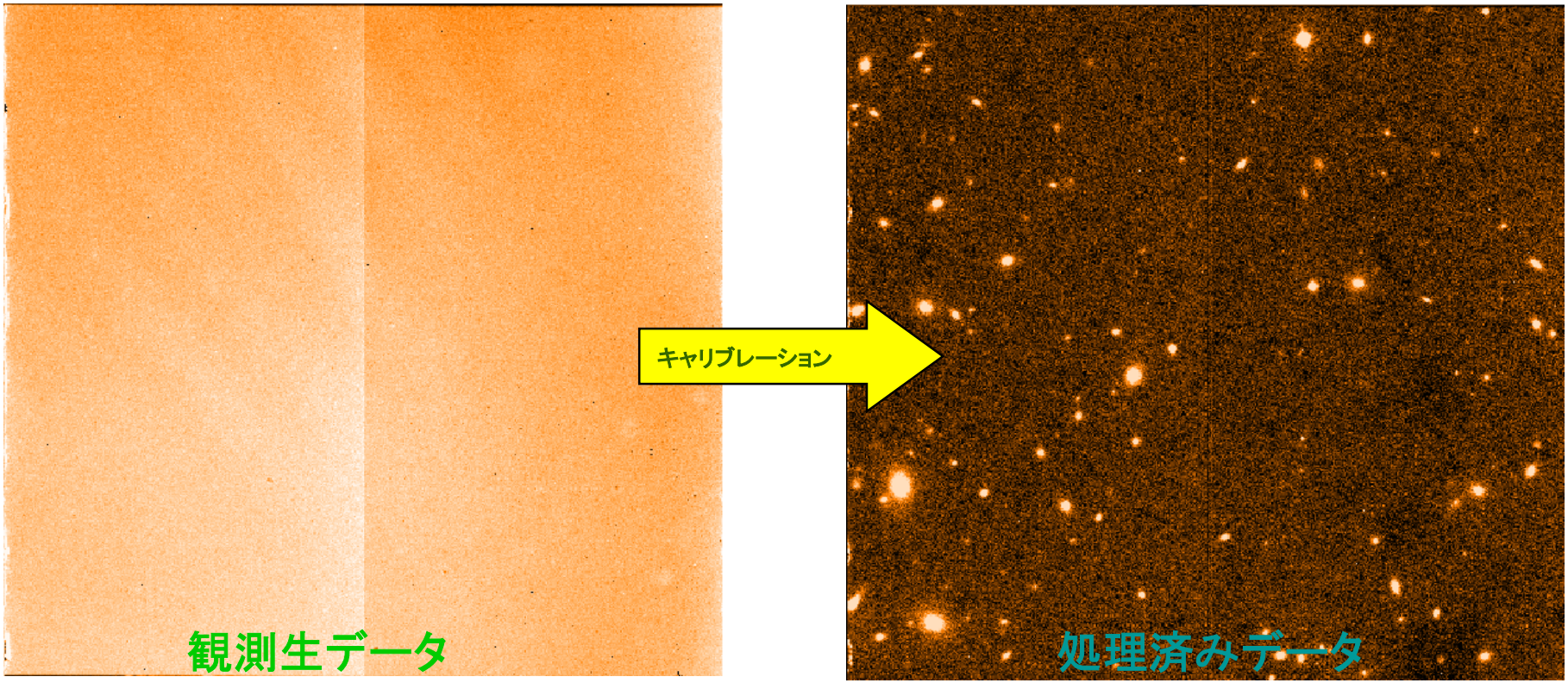
Astronomical Virtual Observatories
~ A Data Grid Environment ~



Data Service Operations in JVO have just been started March, 2008. 1300+ data resources are accessible.

観測データの解析処理

- 銀河団Cl0939+47の近赤外線画像
  - 生データでは何が撮像できたか分からない



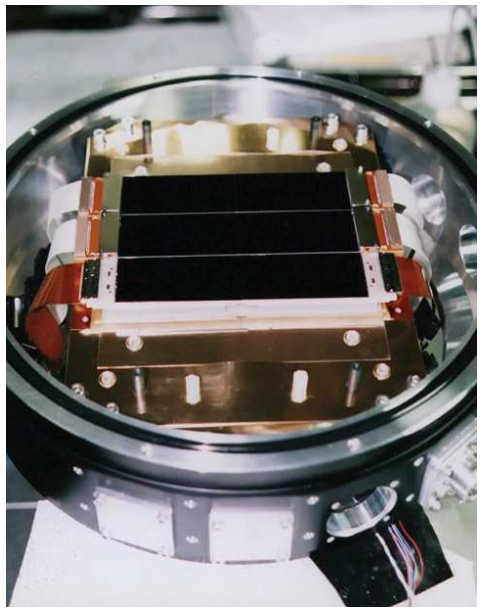
電気ノイズ、熱ノイズを補正し、相関をとりながら画像40枚を積分処理

SuprimeCam Data from Subaru

•SuprimeCam Data ~ 10TB. It takes 154 days to analyse all the data. → parallel processing is crucial

• Detailed analysis of all the SuprimeCam data could lead to:

- ✓ detection of the Cosmic Strings by means of the gravitational lens effect
- ✓detection of outliers
- ✓etc

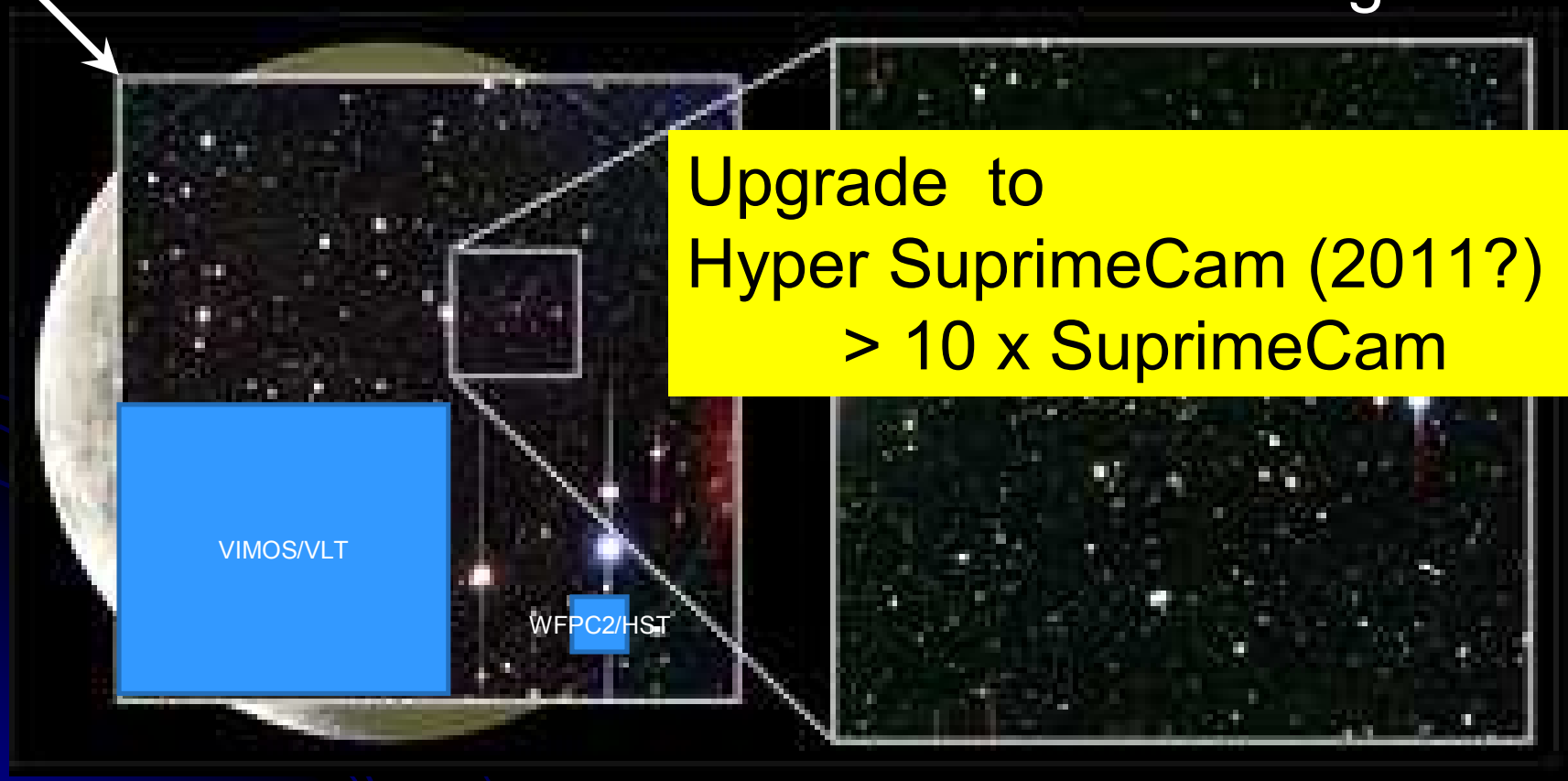


SuprimeCam Field of View

SuprimeCam FOV

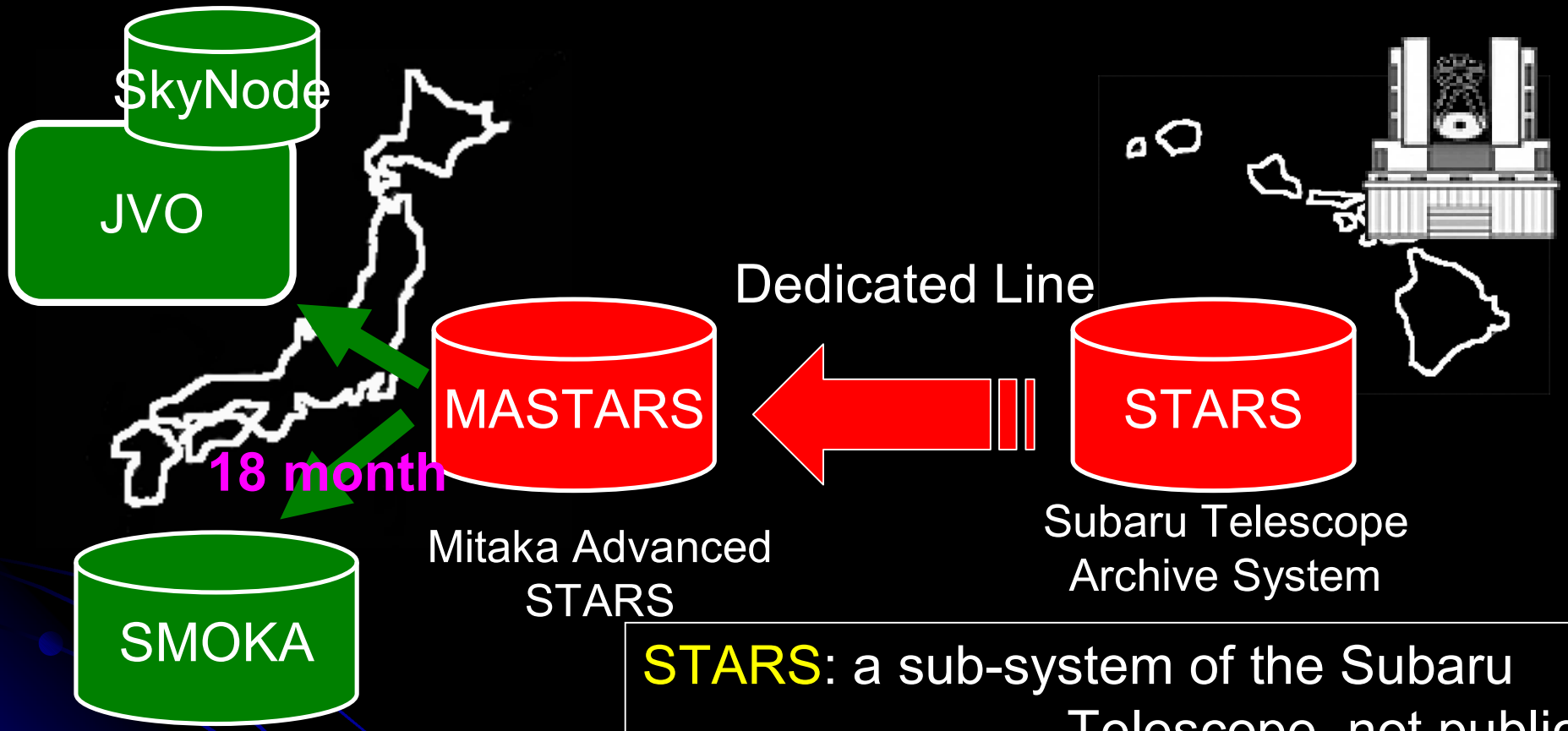
0.2”/pix

160MB/shot
~ tens GB/night



Upgrade to
Hyper SuprimeCam (2011?)
> 10 x SuprimeCam

Subaru Data Archive



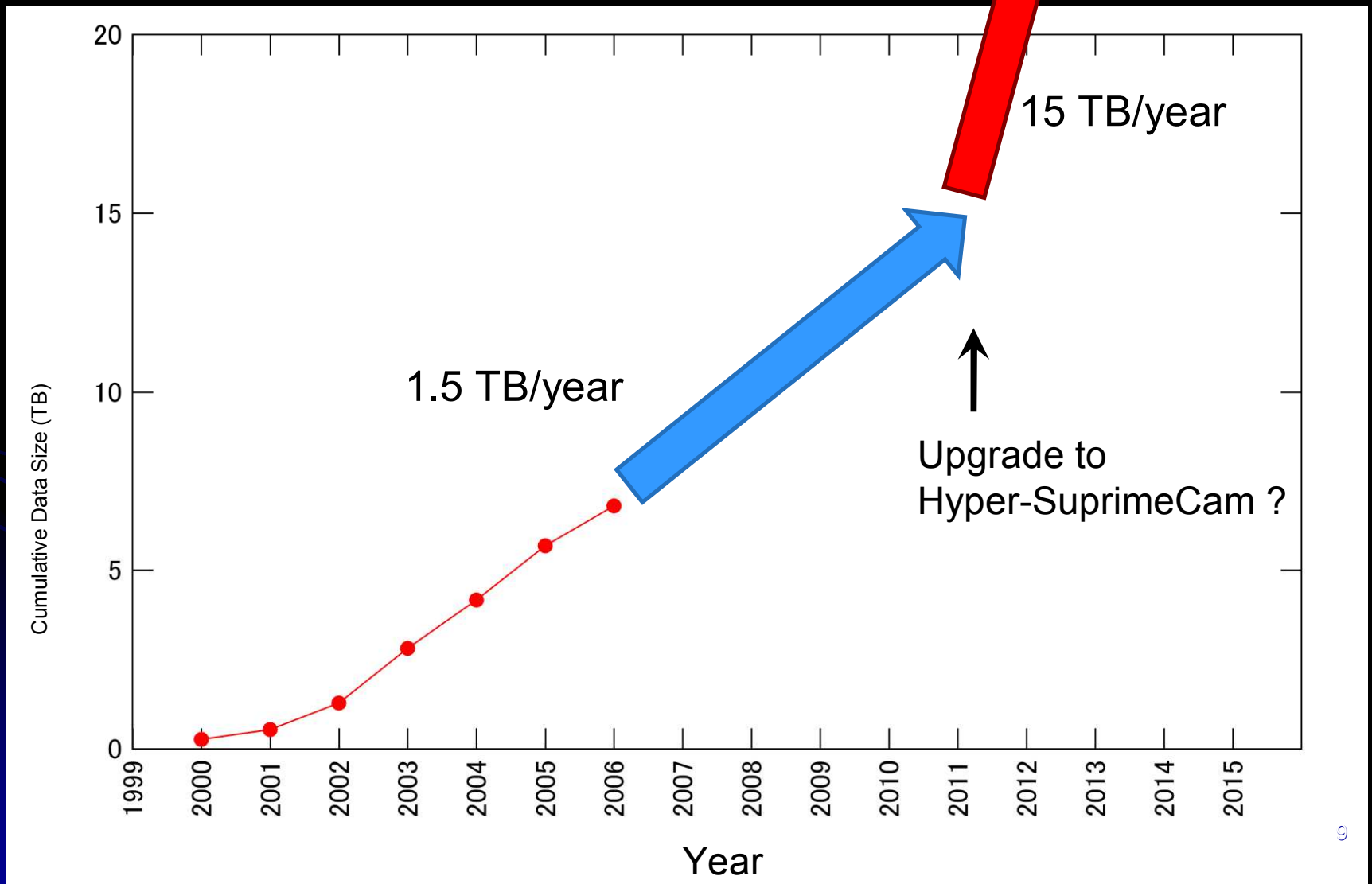
STARS: a sub-system of the Subaru Telescope, not public

MASTARS: a mirror of STARS

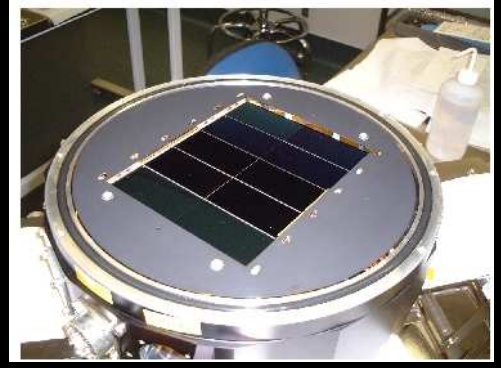
SMOKA: public archive system

JVO: Virtual Observatory Web Portal

Cumulative Data Volume of SuprimeCam



SuprimeCam CCD



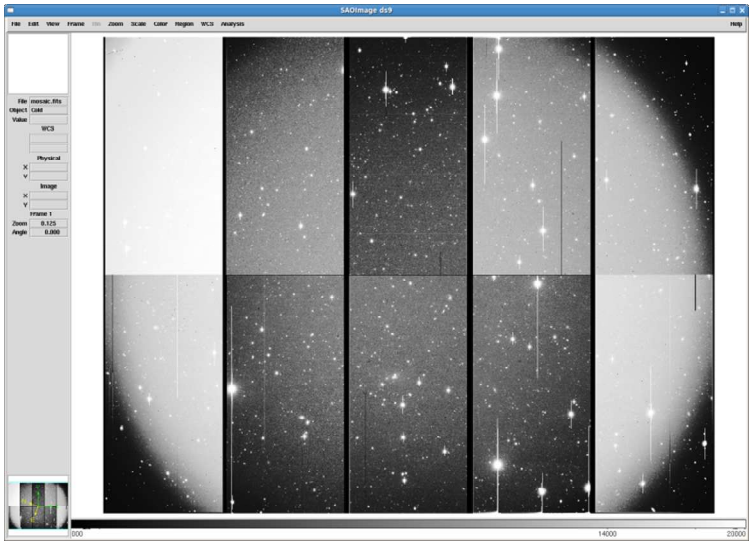
CCD gap

Bad pixel /
Cosmic Ray

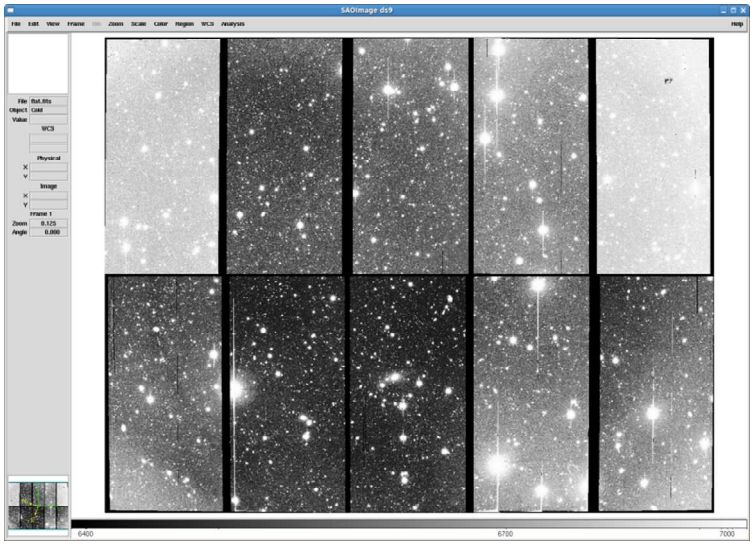
Non-uniformity

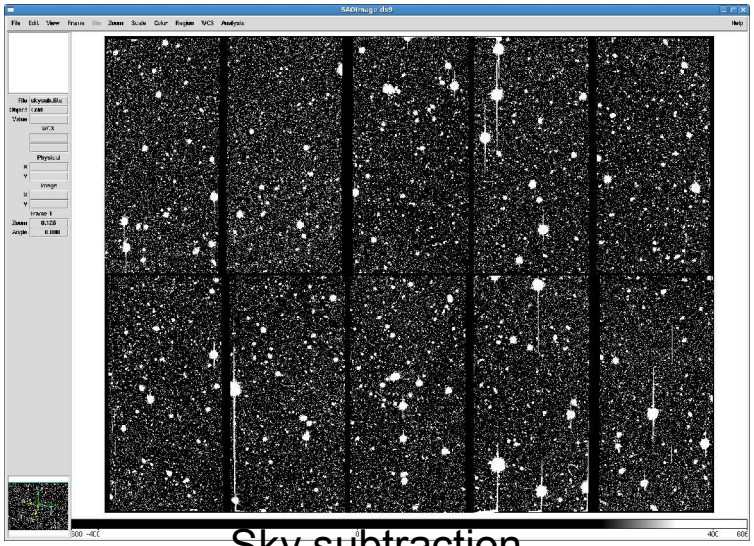
An Example of Image Data Processing

Raw Data



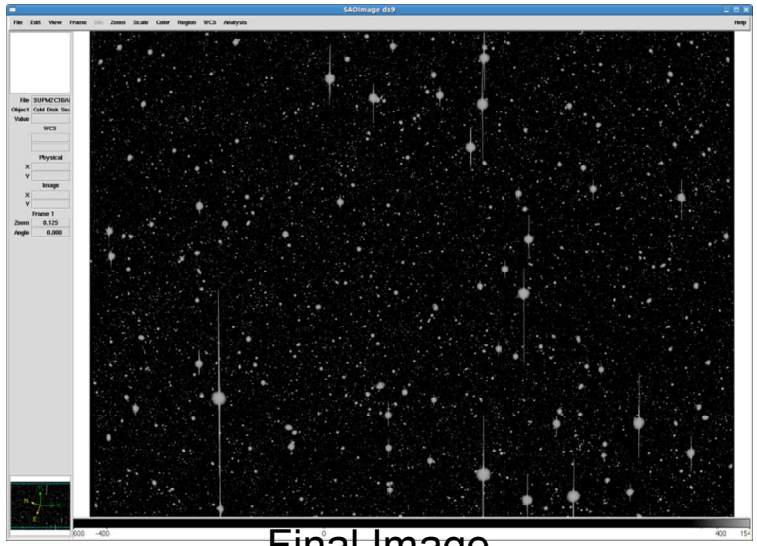
Flat correction





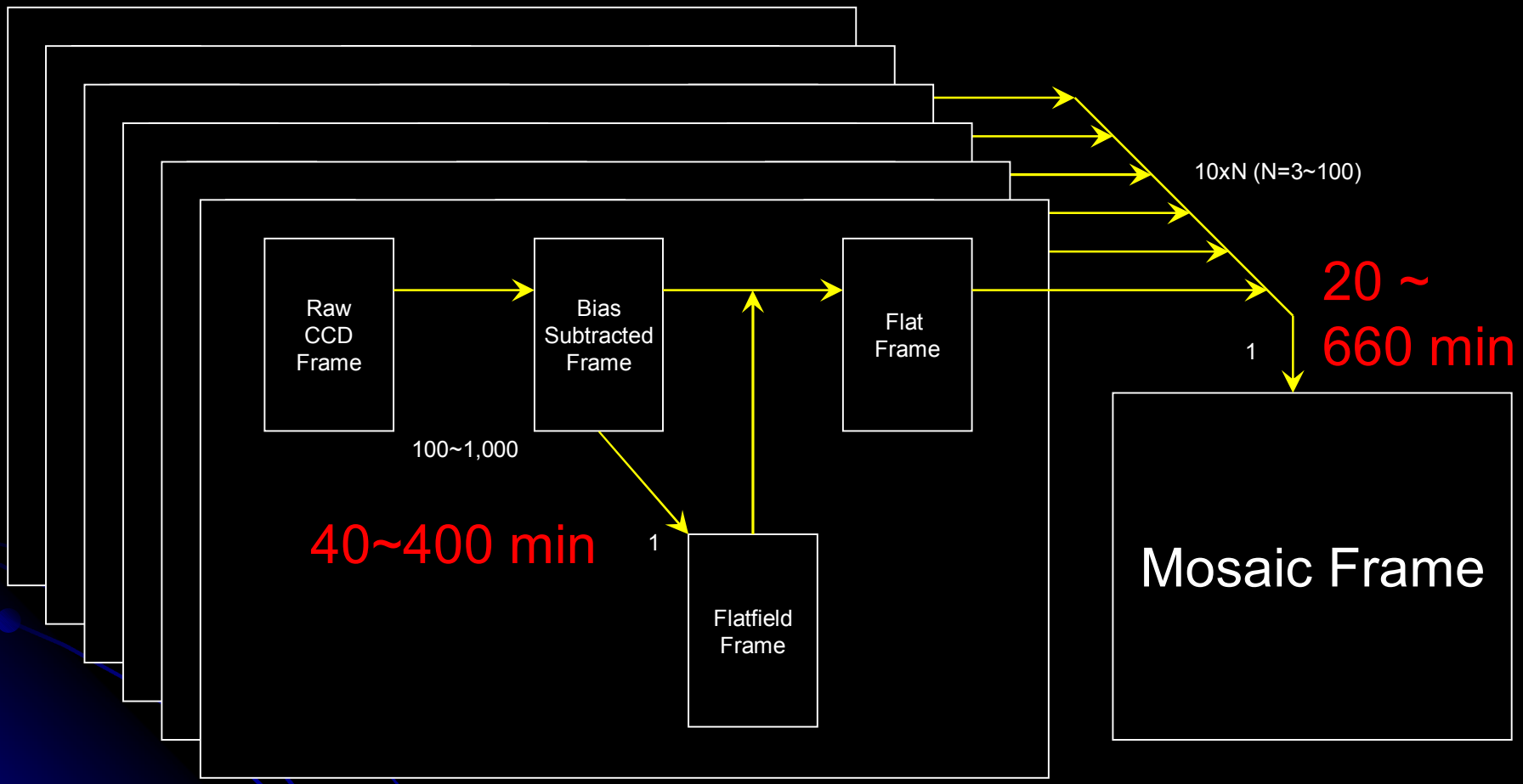
Sky subtraction

Coadd images



Final Image

CCD Image Calibration / Reduction



One day for each observation night

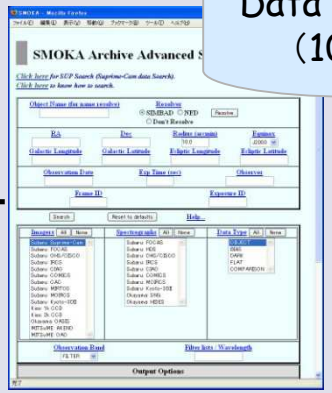
Current situation in astronomy

Manual (Instr.A)

Manual (Instr.B)

Manual (Instr.C)

Subaru Telescope Data Archive (10TB)

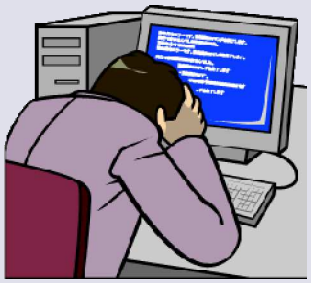


Download by hand

Software for instr. A

Software for instr. B

Software for instr. C



Subaru Data Analysis and Virtual Observatory

User disk space

Subaru Telescope Data Archive (10TB)

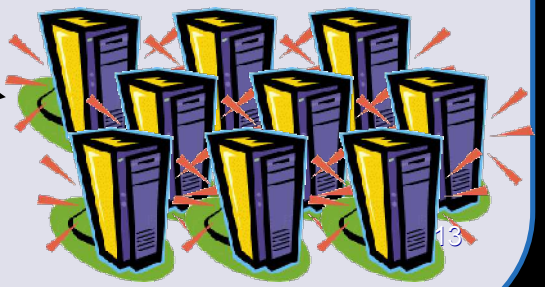
4. Retrieve reduced data

3. Save the reduced data

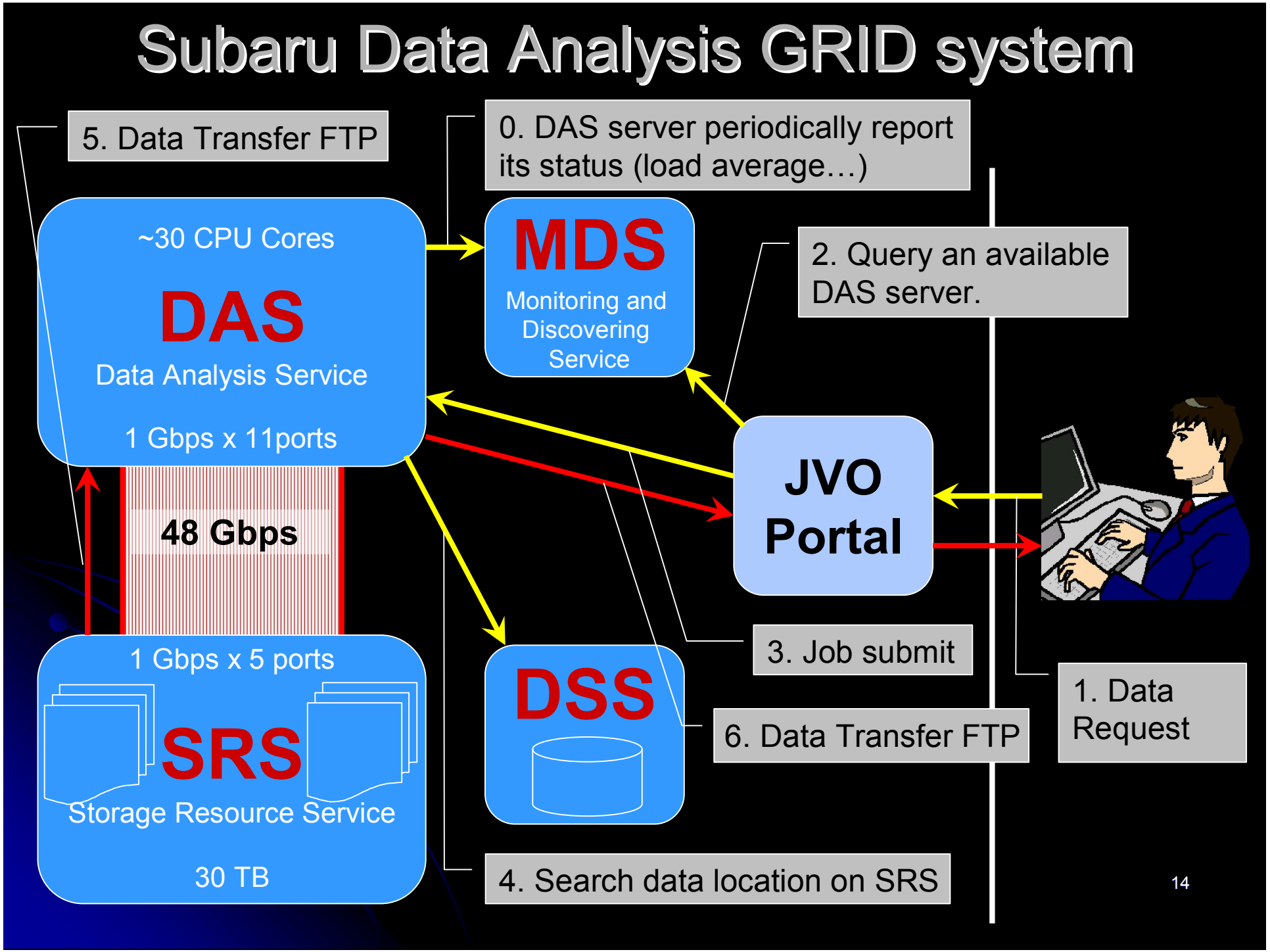
1. Request Data

JVO Portal Service

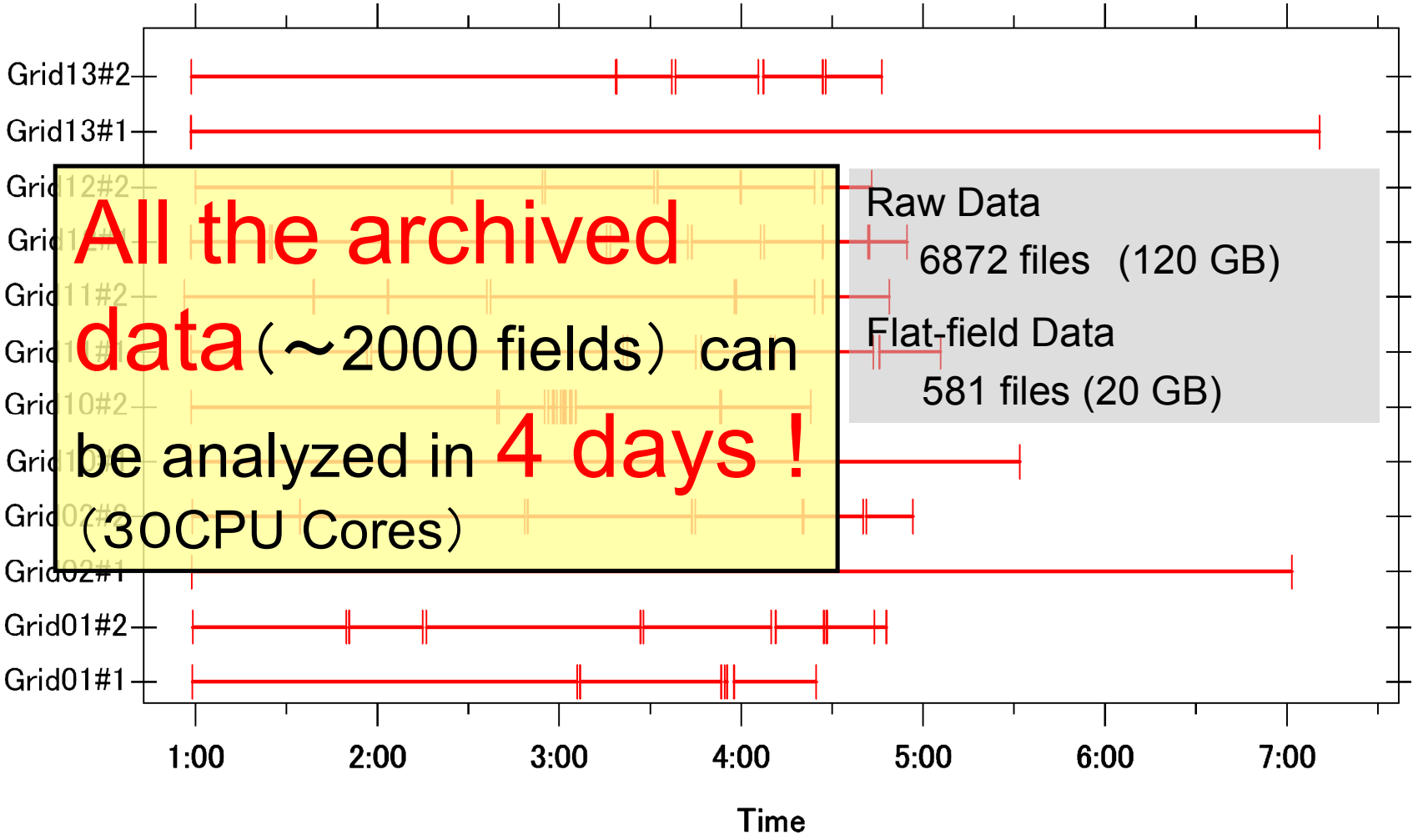
2. Reduce raw data



Subaru Data Analysis GRID system



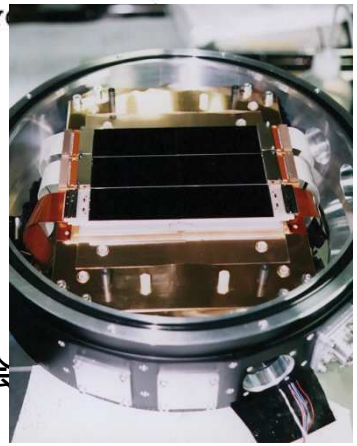
Experiment (58mosaic/12CPU Core)

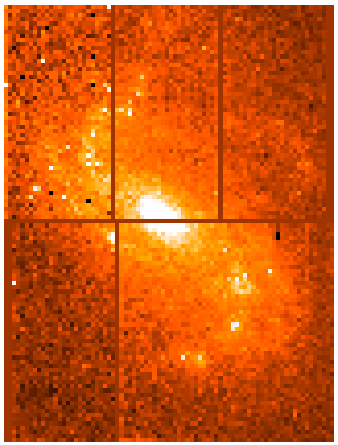


Access to the SuprimeCAM Data

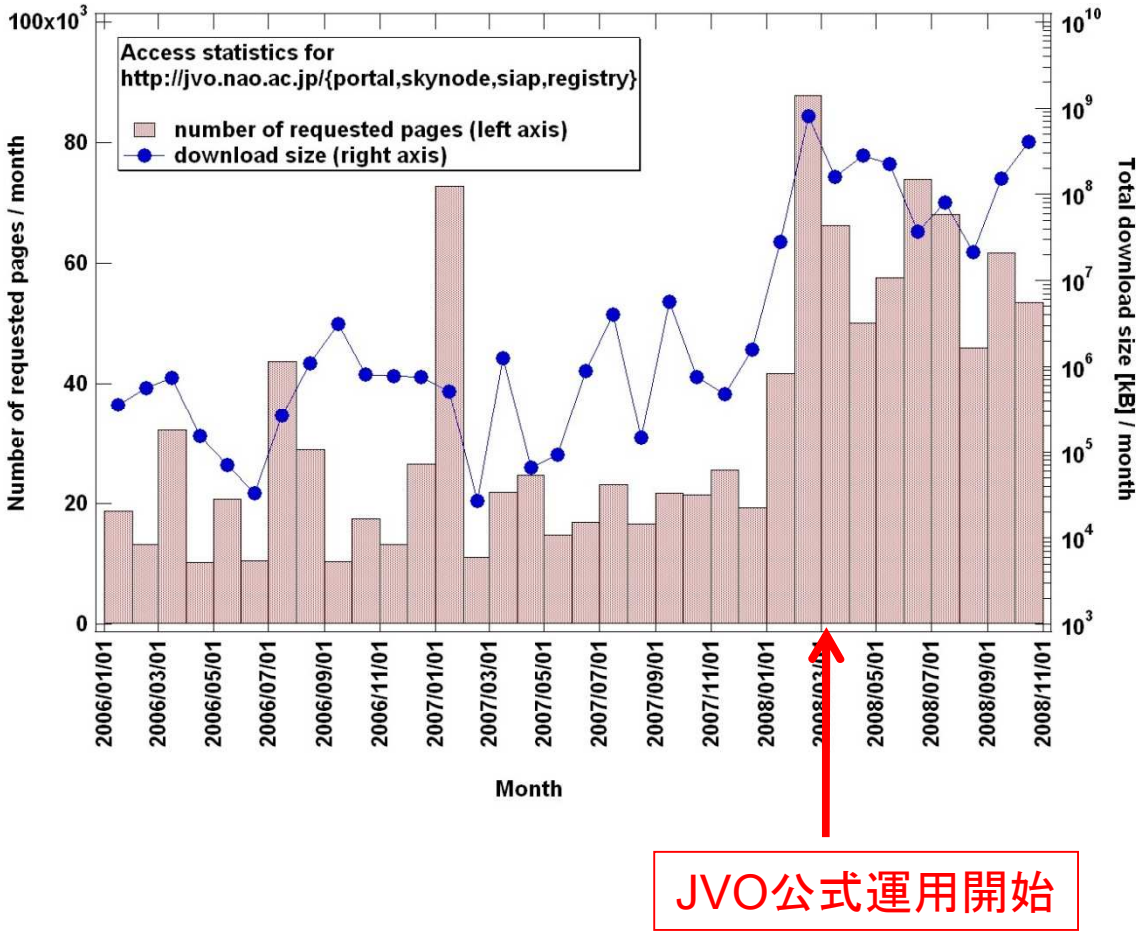
- On-Demand Mosaicing of images and calibrations
- Pre-processed mosaic images are also accessible
- ~ 10 Tera bytes of data

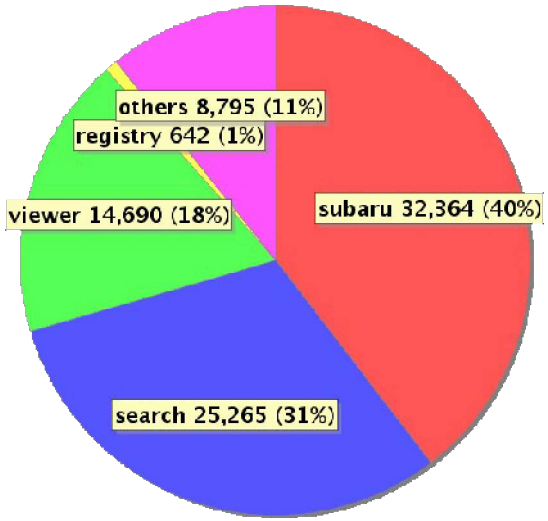
Mostly used/downloaded from the JVO portal





JVOアクセス統計 (2008 Nov)





Computing Grid Services in VO

- Subaru data reduction pipeline system
  - SuprimeCamデータを高速処理
  - On-Demand データ処理機構
  - …
- Backend computing system of the JVO
  - 画像やスペクトルからのカタログ生成
  - photo-z 計算 → 遠方銀河の距離を推定
  - 各種画像演算
  - …

各地の天文台に解析プログラムが存在

Testbed for NAOJ-KEK federation

User Browsers

User Browsers

NAOJ Site

KEK Site

Portal

SS

IS-NAS

IS-CDAS

GridVM Ser

MDS

AMS

Gfarm

NAT/DNS

UMS/VOMS

GridVM Eng

Compute nodes

SINET3

FireWall

VOMS/KEK

Portal

SS

IS-NAS

IS-CDAS

GVM Ser

MDS

AMS

Gfarm

NAT/DNS

UMS/VOMS

Portal

SuperSched

Infosys-NAS

Infosys-CDAS

GridVM Serv

MetaData Serv

AccessManage

Gfarm Storage

DataGrid part

NAT/DNS

UMS/VOMS

Private network

192.168.2.101~

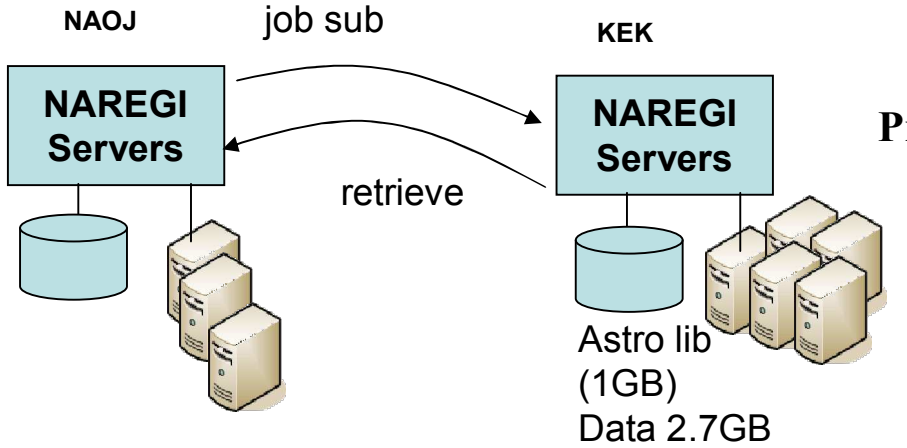
GridVM Eng

Compute nodes

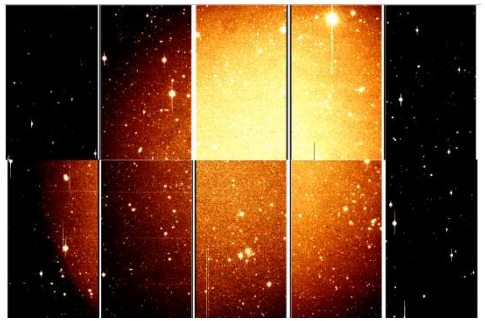
Test VO: naokek

Federation Experiment between NAOJ-KEK

- Setup astro libraries at KEK site
- Job submission to KEK with Work Flow Tool(WFT) at the NAOJ Portal
- Input data are transferred from NAOJ and Output data are staged-out to NAOJ portal



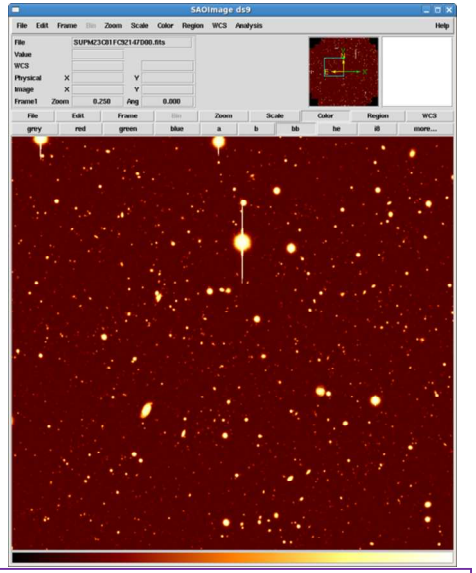
SUBARU Telescope in Hawaii



Input Data:(2.7 GB)
10 CCD mosaic images
160MB x 17

Process: 10 Hours
flat, bias
correct distortion
astrometry
mosaicing
coadd 17 frames

Visualization



- Output data was processed with vis. software as shown in the right picture.

50k+ objects identified in this frame.

Summary

- A GRID computing system for Subaru data analysis
  - All the archived data (8 years)

    More than one year (1CPU) → One week
  - Web-based operational system has been completed → most requested
- Successful experimental federation between NAOJ and KEK through the NAREGI Grid middleware
- More Data Contents added