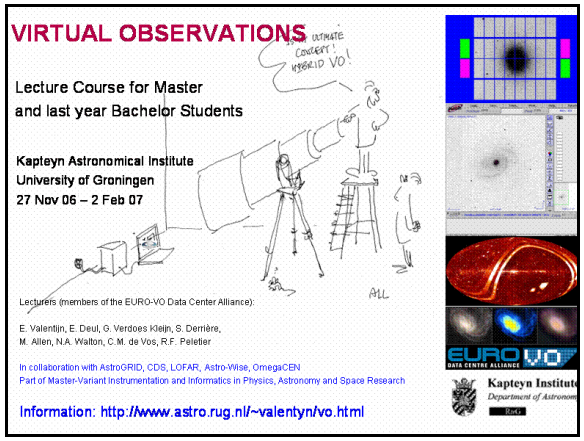
VIRTUAL OBSERVATIONS



Lecture Course for Master
and last year Bachelor Students

Kapteyn Astronomical Institute
University of Groningen
27 Nov 06 – 2 Feb 07

Lecturers (members of the EURO-VO Data Center Alliance):

E. Valentijn, E. Deul, G. Verdoes Kleijn, S. Derrière,
M. Allen, N.A. Walton, C.M. de Vos, R.F. Peletier

In collaboration with AstroGRID, CDS, LOFAR, Astro-Wise, OmegaCEN
Part of Master-Variant Instrumentation and Informatics in Physics, Astronomy and Space Research

Information: http://www.astro.rug.nl/~valentyn/vo.html

This course

- A novum
- Programme
  http://www.astro.rug.nl/~valentyn/VO.html
  - Basis
    - intro
    - concepts – systems
    - Astrometry – photometry
  - CDS – VO
    - EURO-VO tools
  - E – Lofar
  - AstroWise
  - AstroGRID
  - Students VO

E-science

- Beyond “workstation science” of the 80-90’s
- Distributed services
- Distributed communities
- Distributed archives
- p2p networks – KAZAA- NAPSTAR
  - Share cpu
  - Share storage
  - Share info / meta data /knowledge

different views on

- Surveys
- Templates
- Pipelines
- Virtual Observatories

Surveys

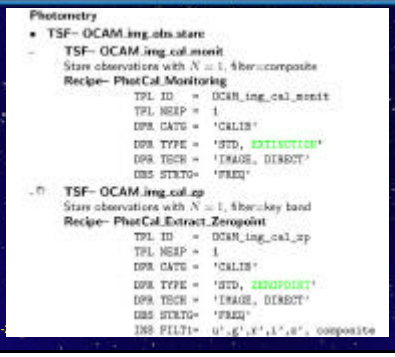
- Defined area on sky
- Homogeneous
  - Survey limit
    - Flux (magnitude)
    - Size
    - Surface brightness
    - distance
- Quality control

Templates

Standards very important for VO

- Observing templates
  - Astronomical Observing Templates at ESA
  - Templates / Template signature files at ESO

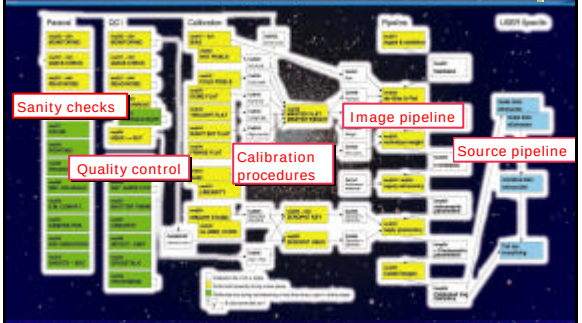
ESO parse info via headers



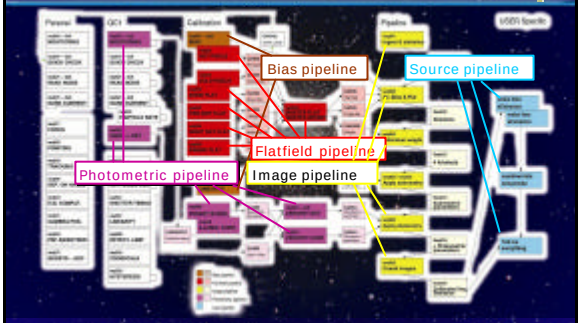
pipelines

- Workflow
- What triggers a pipeline?
  - Data items
  - Operators
  - users

Data Model / flow



Astro-Wise Pipelines



Virtual observatories

- ????????????

GigaPort seminar for astronomers

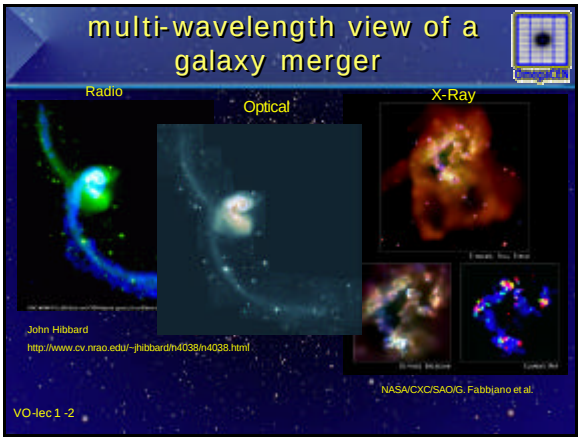
link to VO

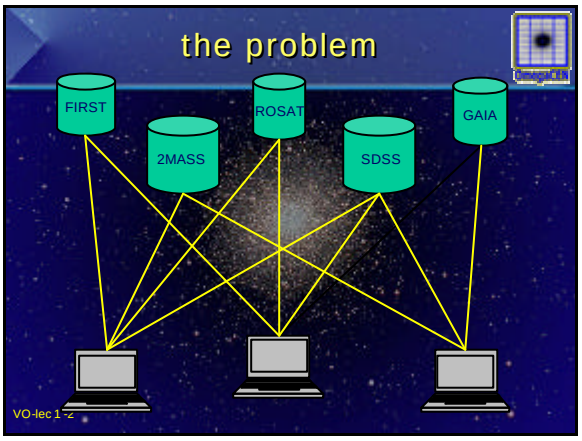
lecture by EV

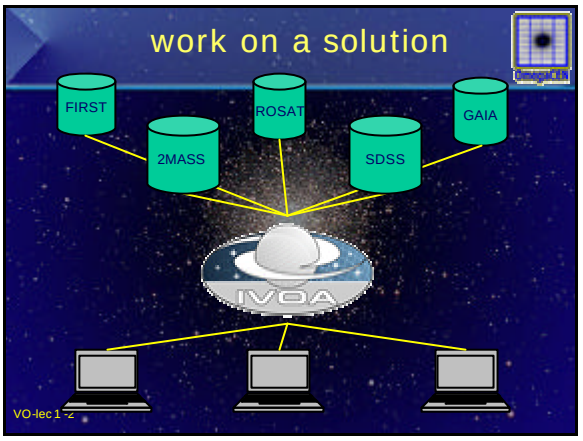
Virtual observatories

- Broad VOs
  - IVOA
  - Euro VO
- EuroVO DataCenter Alliance
- Focussed VOs
  - AstroWise


Astro Grid
Aus-VO
EURO VO
CVO
IVOA
ArVO
JVO
HVO
GAVO
france


multi-wavelength view of a galaxy merger
Radio
Optical
X-Ray
John Hibbard
http://www.cv.nrao.edu/~jhibbard/n4038/n4038.html
NASA/CXC/SAO/G. Fabbjano et al.
VO-lec 1 -2


the problem
FIRST
ROSAT
GAIA
2MASS
SDSS
VO-lec 1 -2


work on a solution
FIRST
ROSAT
GAIA
2MASS
SDSS
IVOA
VO-lec 1 -2


AVO-> Euro-VO
• Standards
– FITS
– Universal Content descriptor UCD
– VO table – VOT
• Communities – workshops/training
• Registry
• Connecting archives
– Cone search
– X-match
– All kinds of tools/ web services
• Relatively static
VO-lec 1 -2


standards: Universal content Descriptor - UCD
• http://www.ivoa.net/Documents/REC/UCD/UCD-20050812.html#ToC8
Examples of UCD1+ and how they are built:
http://www.ivoa.net/Documents/PR/UCD/UCDlist-20061006.html
VO-lec 1 -2

Standards:VO table VOT

- VOtable_SourceList.xml
- http://www.ivoa.net/twiki/bin/view/IVOA/IvoaVOTable

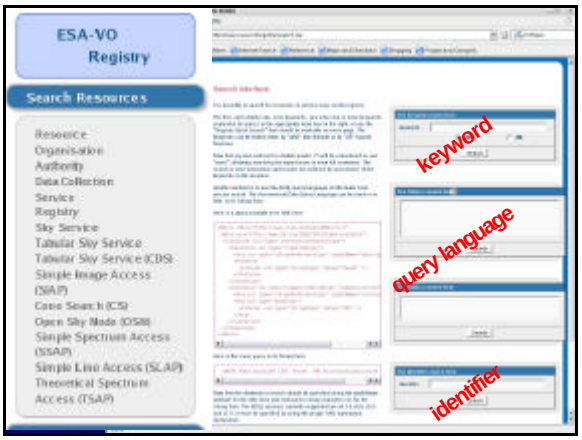
VO-lec 1 -2

Finding information in the VO

Registries are here

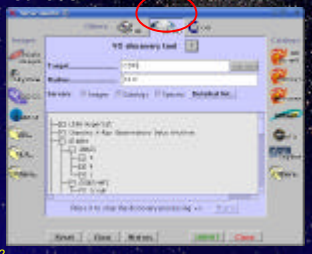
- multiple interfaces
  - human readable
  - machine readable
- simple/advanced search

VO-lec 1 -2

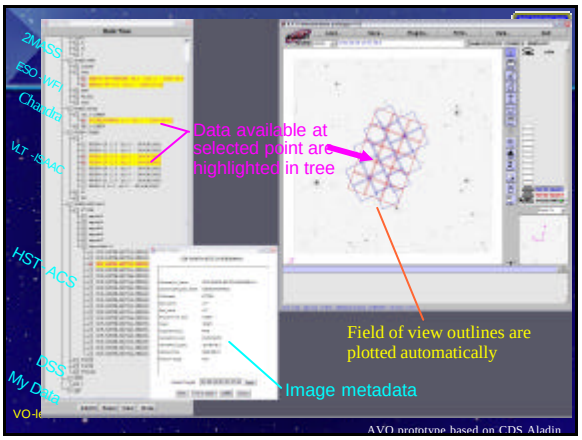


Examples

CDS Aladin: combines registry discovery with other access methods

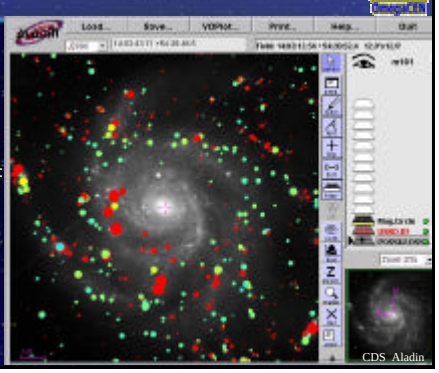


VO-lec 1 -2

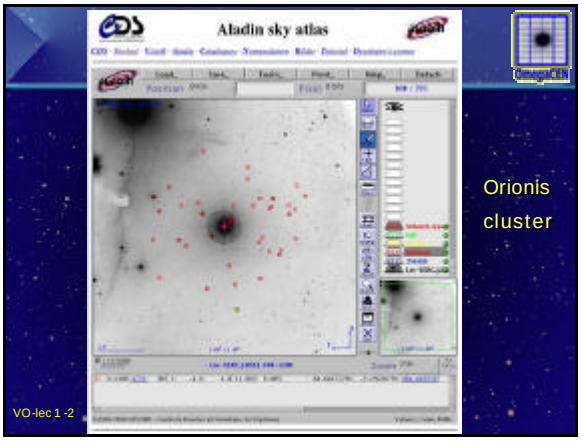


ALADIN

- Manipulation
- X-match
- Visualization
- Direct links to:

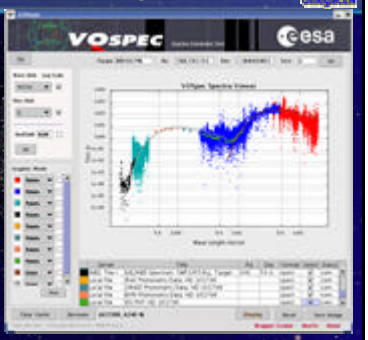


VO-lec 1 -2



Multi archive spectra

emerging conventions
and standards

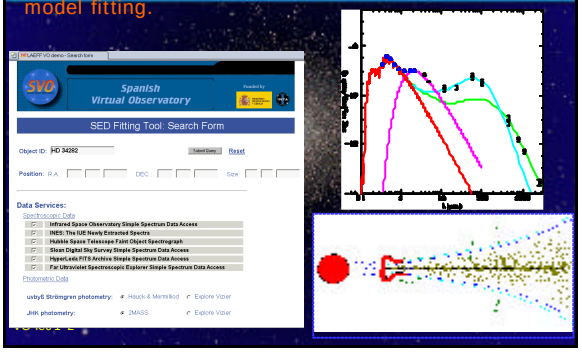


VO-lec 1 -2

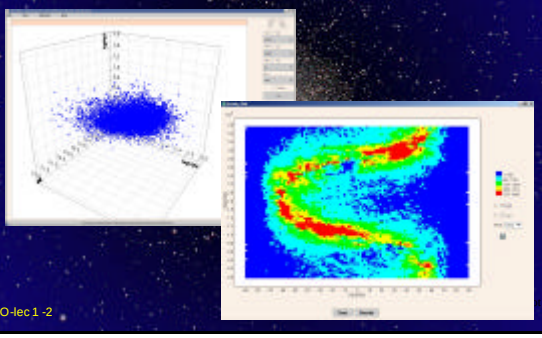
ESA VOSpec

VO applications and analysis tools

• VOSED: a Spectral Energy distribution builder and model fitting.

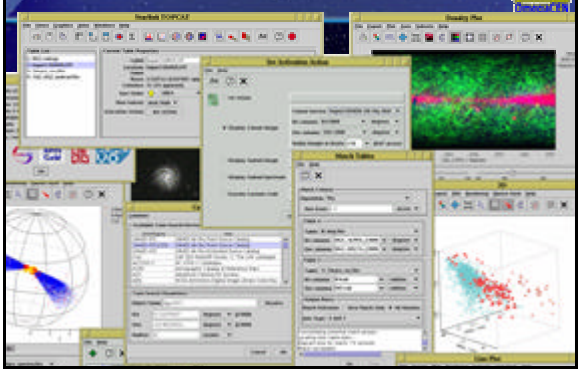


VOPlot



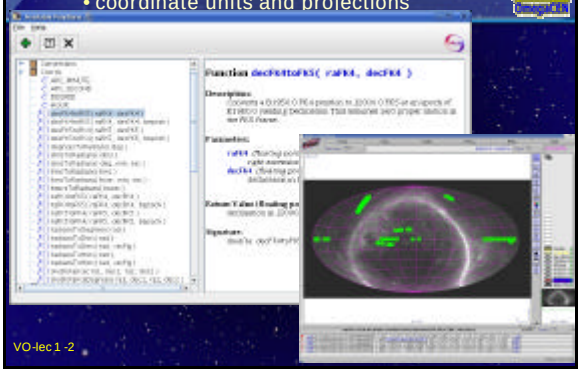
VO-lec 1 -2

TOPCAT VOTable view



• Astronomy functionality

• coordinate units and projections

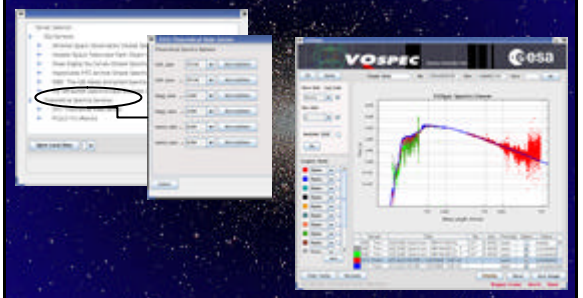


VO-lec 1 -2

Astronomical archives: take-up of VO standards.

Theory in the VO

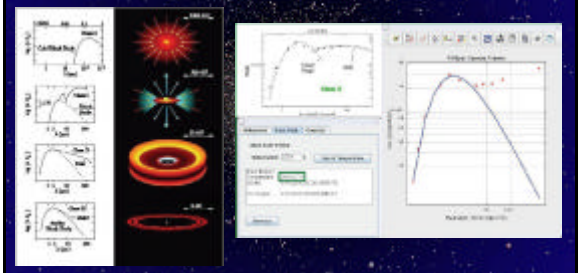
- Theoretical model web server + TSAP



VO-lec 1 -2

ASTRID-CM: VO Science

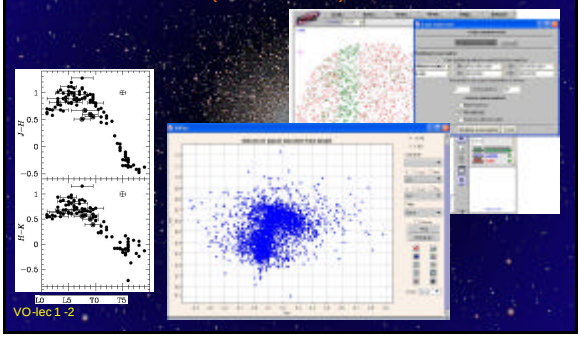
- From Class 0 to Class III: Search and classification



VO-lec 1 -2

SVO Thematic Network: VO Science

- Discovery of ultracool (T-type) brown dwarfs (IAC+LAEFF)



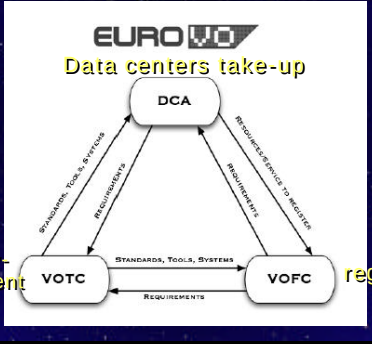
VO-lec 1 -2

applications

- Astro-GRID → Euro-VO
  - Workflow
    Taverna 2 (Bioinformatics-life sciences)
    client side workflow editor and engine
  - Connecting webservices – grid
  - my space

VO-lec 1 -2

Euro –VO connect and mine archives



Technology-development

registry

VO-lec 1 -2

Data Center Alliance

- Consortium Agreement
  F, I, G, UK, Sp, NL, ESO, ESA
- Now EU funded: FP6 1.5 M Euro
  COMMUNICATION NETWORK DEVELOPMENT
  COORDINATION ACTION
  - start 1/9/2006 2.5 year (startup project)
  - ~ 1 fte NL, + travel
- NOVA -NL

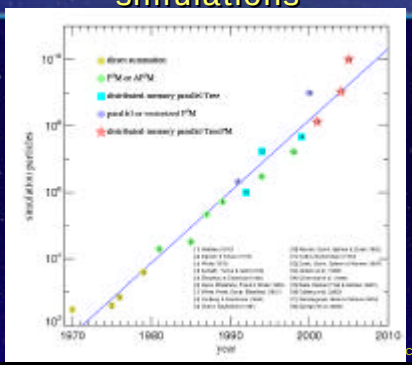
DCA tasks – WPs take-up VO technology

- WP3: support of take-up and implementation of VO technology
  - NL: AstroWise/OmegaCAM, Lofar
  - Census
- WP 4: inclusion theoretical data VObs
- WP 5: coord with computational GRIDS
  - NL: I-GrID- F.Pasian
- WP 6:support other EU countries

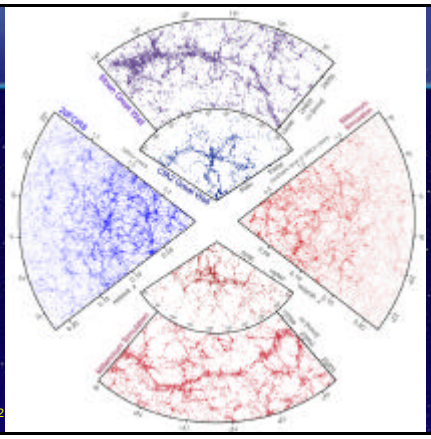
Theory in the VO: issues Gerard Lemson

- Simulations not so simple
  - complex observables
  - no standardisation (not even HDF5)
  - archiving ad hoc, for local use
- Moore's law makes useful lifetime relatively short: few years later can do better
- Current IVOA standards somewhat irrelevant
  - no common sky
  - no common objects
  - requires data models for content, physics, code

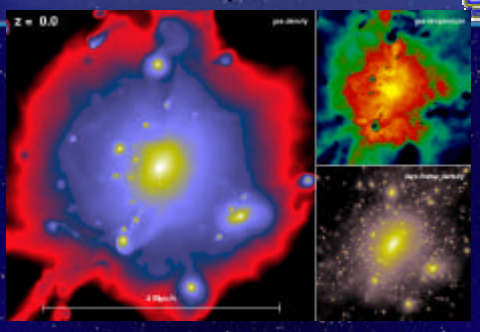
"Moore's law" for N-body simulations



Courtesy Simon White



Detailed predictions

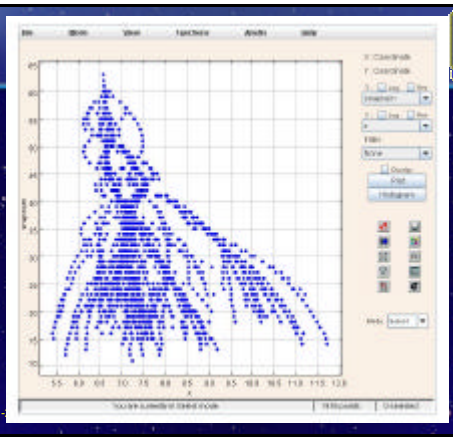


Courtesy Volker Springel

The Virgo consortium's *Millennium* simulation

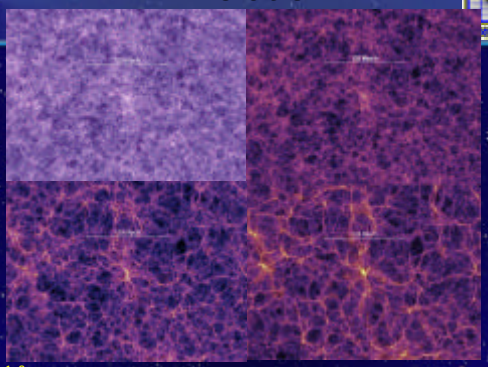
- Millennium simulation
  - 10 billion particles, dark matter only
  - 500 Mpc (~2Gly) periodic box
  - "concordance model" (as of 2004) initial conditions
  - 64 snapshots
  - 350000 CPU hours
  - O(30Tb) raw + post-processed data
- Postprocessing:
  - dark matter density fields smoothed at various scales (45 * 256^3 grid cells)
  - dark matter cluster merger trees (~750 million)
  - galaxy merger trees (~1 billion/catalogue)
    - DeLucia & Balizot, 2006
    - Bower *et al*, 2006

VO-lec 1 -2



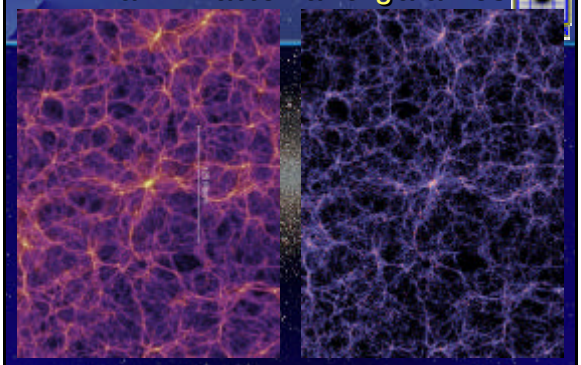
VO-lec 1 -

Evolution



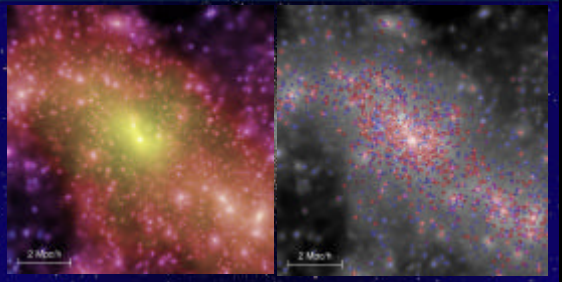
VO-lec 1 -2

Dark matter and galaxies



VO-lec 1 -2

Halos and galaxies

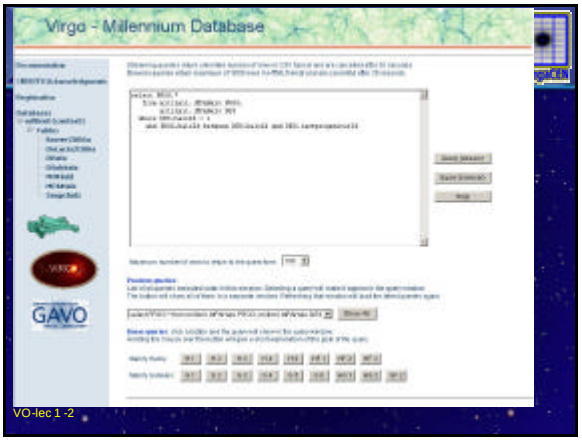


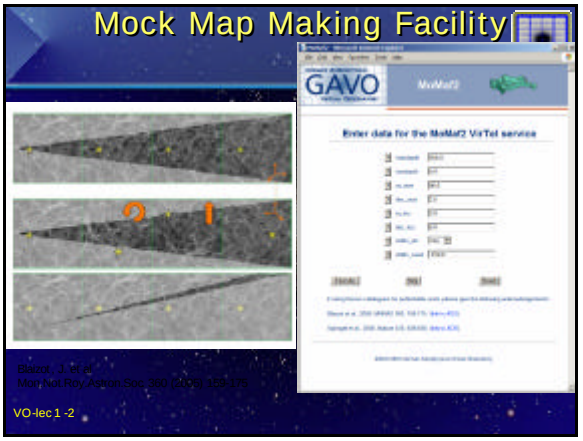
VO-lec 1 -2

the *Millennium* database + web server

- Post-processing results only
- SQLServer database
- Web application (Java in Apache tomcat web server)
  - portal: http://www.mpa-garching.mpg.de/millennium/
  - public DB access: http://www.g-vo.org/Millennium
  - private access: http://www.g-vo.org/MyMillennium
  - MyDB
- Access methods
  - browser with plotting capabilities through VOPlot applet
  - *wget* + IDL, R
  - TOPCAT plugin

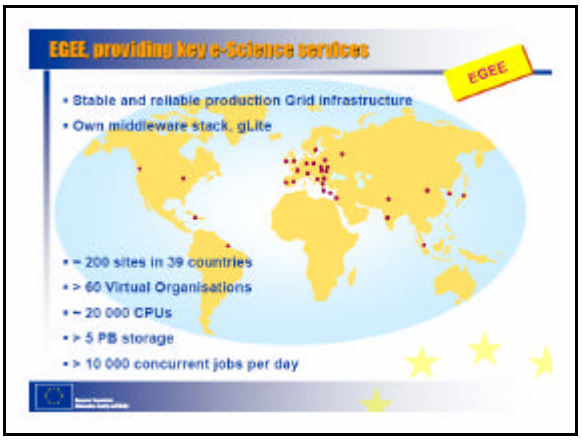
VO-lec 1 -2

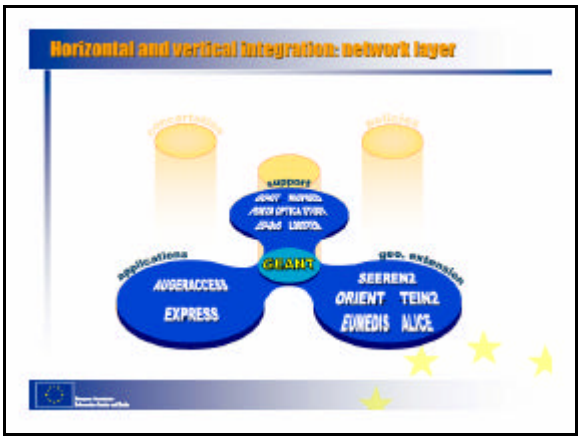

Virgo - Millennium Database
GAVO
VO-lec 1 -2

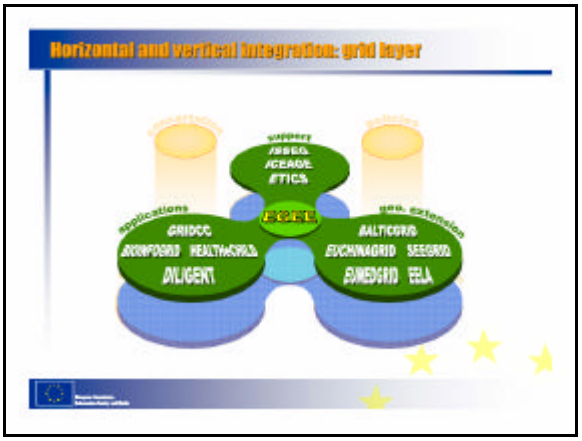

Mock Map Making Facility
GAVO
Enter data for the MoMaf2 VirTel service
Balázs, J. et al.
Mon.Not.Roy.Astron.Soc.380 (2007) 159-175
VO-lec 1 -2


EAGE Conference
in May 2006
Moving e-Infrastructure into a new era
the FP7 challenge
Mário
Campolargo
European Commission - DG INFSO
Head of Unit Research Infrastructures


EGEE, world leading grid infrastructure
EGEE
• Stable and reliable production Grid infrastructure for ERA
• Own middleware stack, gLite
• Scientific communities
High Energy Physics
Astrophysics
Computational Chemistry
Fusion
Life Sciences
Biomedics
Earth Sciences
Finance
Geophysics
Multimedia
• Industrial policy
• Standardisation, training and education

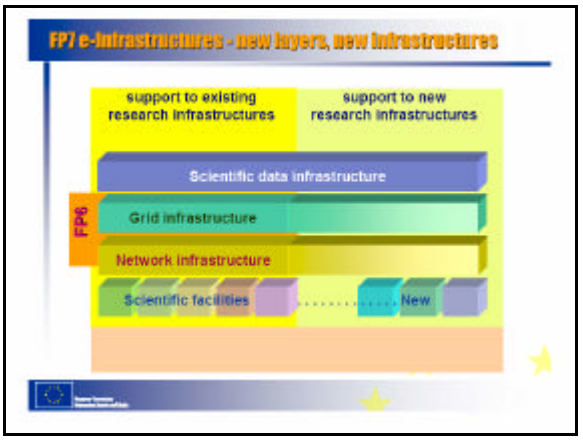

EGEE, providing key e-Science services
EGEE
• Stable and reliable production Grid infrastructure
• Own middleware stack, gLite
• ~ 200 sites in 39 countries
• > 60 Virtual Organisations
• ~ 20 000 CPUs
• > 5 PB storage
• > 10 000 concurrent jobs per day

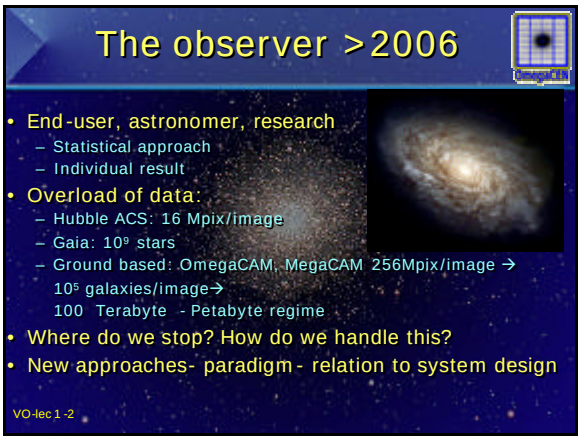

Horizontal and vertical integration: network layer
GÉANT
AUSTRALIA
EXPRESS
SEERENE
ORIENT
EUMEDIS
ALICE


Horizontal and vertical integration: grid layer
support
applications
geo. extension
EGEE


FP7 – addressing the new scientific data layer
DATA


FP7 – addressing the new scientific data layer
DATA
• Foster the emergence of Europe as a world scientific data hub.
• Deploy data repositories for the scientific community and future generations of scientists supporting, in a coordinated way, digital libraries, archives, data storage, data curation, access to information and the necessary pooling of resources.

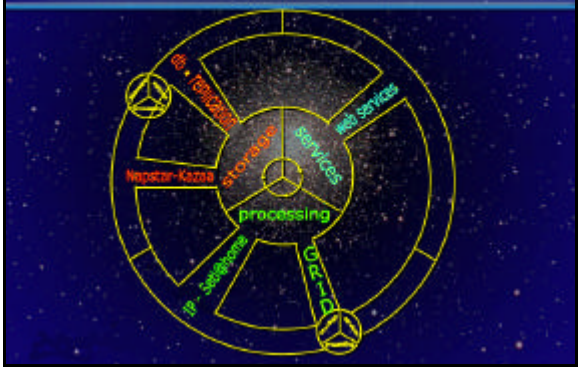

FP7 e-Infrastructures - new layers, new infrastructures
support to existing research infrastructures
support to new research infrastructures
Scientific data infrastructure
FP6
Grid infrastructure
Network infrastructure
Scientific facilities
New


The observer >2006
• End-user, astronomer, research
– Statistical approach
– Individual result
• Overload of data:
– Hubble ACS: 16 Mpix/image
– Gaia: 10⁹ stars
– Ground based: OmegaCAM, MegaCAM 256Mpix/image →
10⁵ galaxies/image→
100 Terabyte - Petabyte regime
• Where do we stop? How do we handle this?
• New approaches- paradigm - relation to system design
VO-lec 1 -2


Courtesy Canon

Lofar

VO and the world - communities



from Design-> deliver

- Scientific requirements - SRD
- User requirements - URD
- Architectural design - ADD
- Detailed design - DDD
  Implementation
- Quantify
- Build
- Qualify – unit tests

New approaches new balances

Anarchy ← →coordinated
Freedom ← → fixed system

Standard data products ← →user tuned products
Data releases ← → user defined hunting

DESIGN

5 Essential STEPS:

1- calibration plan integrated up-link /down link

NOVA - Kapteyn Institute	OmegaCAM	ID	internal
		Issue	Version 2.11
USM – OaPd	DFS	Date	7 Oct 2004
		Page	42

Selected items from Odoco file system	
Summary sections	•
Items:	
Objective	•
Fulfilling or fulfilled by	•
When performed/frequency	•
Sources, obs.,…	•
Inputs	•
Outputs	•
Required accuracy, constraints	•
Estimated time needed	•
Priority	•
Template Signature File.	•
Recipe	•
Calibration Analysis spec's	•
Needed Functionalities	•
CA implementation (pseudo code)	•
Status of Req.	
FLAG	

www.astro.rug.nl/~omegacam

This document is prepared by the Odoco Document Control System.

NOVA - Kapteyn Institute	OmegaCAM	ID	internal
		Issue	Version 2.11
USM – OaPd	DFS	Date	7 Oct 2004
		Page	49

```
bias1, bias2        : the two raw bias images
MAXIMUM_ITERATIONS  : maximum number of iterations for statistics
                      measurement (integer).
                      Range of allowed values: 2 - 10. Default: 5
REJECTION_THRESHOLD : rejection threshold for bad pixels in sigma (float).
                      Range of allowed values: 1.0 - 10.0. Default: 5.0
```

Before applying this recipe, use **Recipe– Split**—which is documented in **seq.– 631**—with the **-t bias** option to split the raw multi-extension FITS input files.

CA:

Process (make):

1. Subtract the two bias frames to produce a difference image.
2. Iteratively reject outliers in the difference image.
3. Compute the mean, median and rms of the remaining pixels of the difference image.
4. The readout noise is the rms divided by $\sqrt{2}$.

CAP:

```
diff_image = eclipse.image_sub(bias1, bias2)
stats = eclipse.iter_stat(diff_image,
                          MAXIMUM_ITERATIONS,
                          REJECTION_THRESHOLD)
read_noise  = stats.stdev / sqrt(2)
mean_diff   = stats.avg_pix
median_diff = stats.median
```

QC Parameters:

```
read_noise  : the readout noise in ADU
mean_diff   : mean difference between the two input raw bias frames
median_diff : median difference between the two input raw bias frames
```

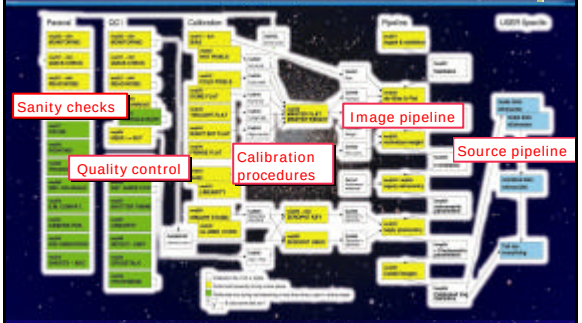
VO-lec 1 -2

2 -Procedurizing

- Procedurizing
  - Data taking at telescope for both science and calibration data - Templates
    - Observing Modes: —Stare —Jitter —Dither —SSO
    - Observing Strategies: —Stan —Deep —Freq —Mosaic
  - Full integration with data reduction
  - Design- ADD
  - Data model (classes) defined for data reduction and calibration
  - View pipeline as an administrative problem

VO-lec 1 -2

3 Data Model



VO-lec 1 -2

4 Integrated archive and Large Data Volume

- Handling of the data is non-trivial
  - Pipeline data reduction
  - Calibration with very limited resources
  - Things change in time:
    - –Physical changes (atmosphere, various gains)
    - –Code (new methods, bugs)
    - –Human insight in changes
  - Working with source lists

Science can only be archive based

VO-lec 1 -2

4 archive and Virtual Survey System

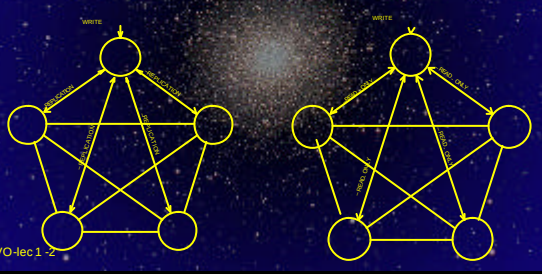
- Environment that provides systematic and controlled
  - Access to all raw and calibration data
  - Execution and modification image/calibration pipelines
  - Execution of source extraction algorithms- catalogues
  - Archiving or regenerate on the fly dynamically
  - Paradigm: no static data releases but dynamic on request data
  - federated to link different data centers
- Dynamical archive continuously grows, can be used for
  - small or large science projects
  - generating and checking calibration data
  - exchanging methods, scripts and configuration

raw pixel data ⇔ pipelines/cal files ⇔ catalogues

VO-lec 1 -2

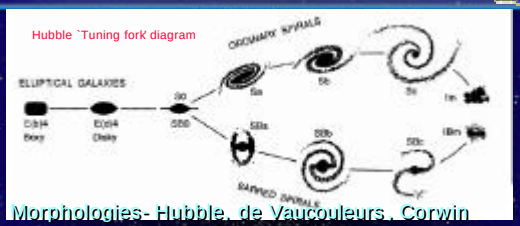
5 Intra-operability peer to peer

- CVS
- "Advanced Replication" evolving to pointers



VO-lec 1 -2

Case: morphologies
parameters vs theory - degeneracy



Morphologies- Hubble, de Vaucouleurs, Corwin

- U, B, V, R , I, Z, K
  - m_tot, m_25, m_26, D_26, D_25, D_90%, r_eff, sb_0, sb_eff
- Structure/radial
  - a/b, pa, B/D, N, exp scale length , delta (B-R)
- colours
  - Central (B-R), (B_R)_tot,,