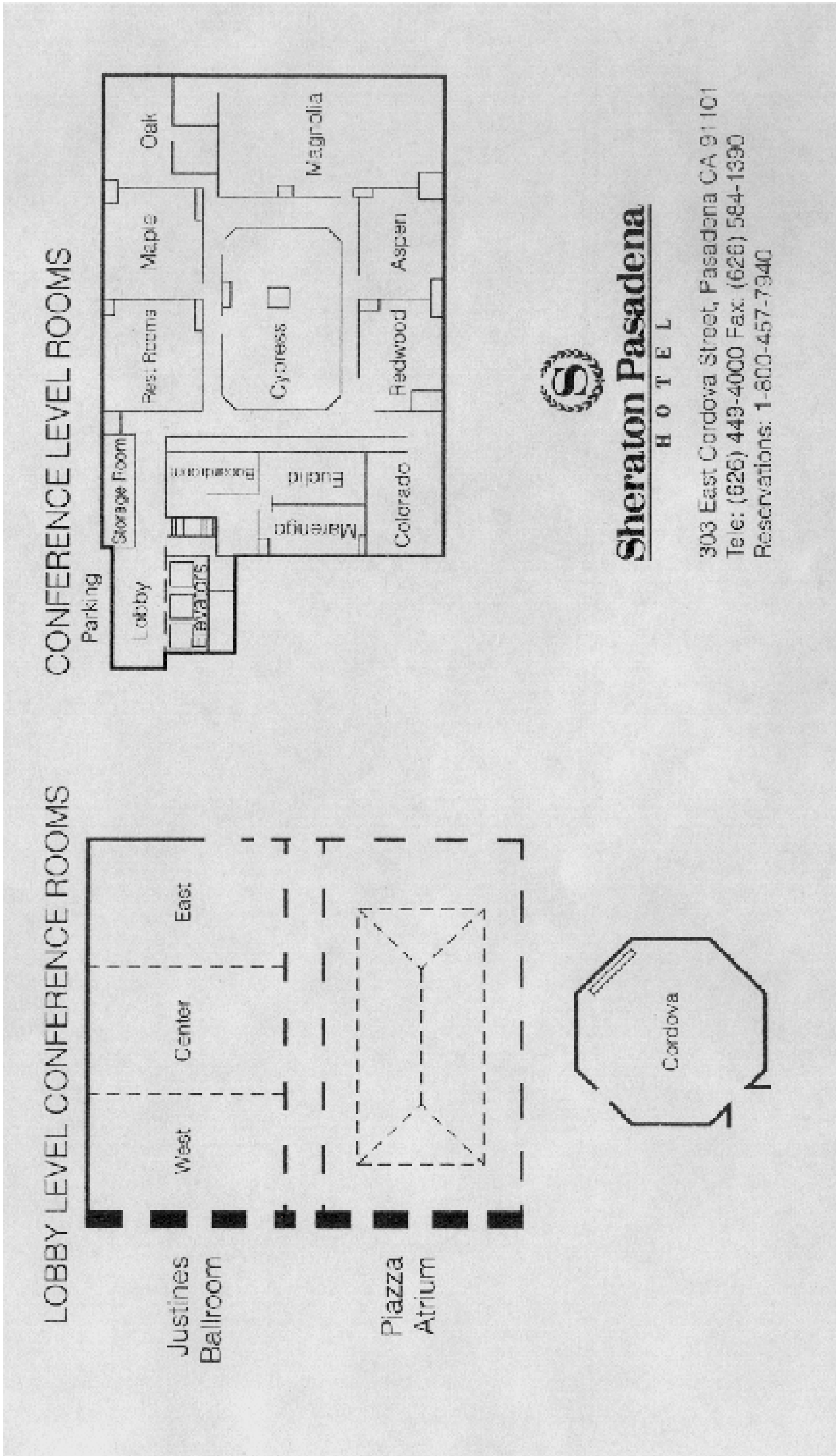
ADASS XIV

Astronomical Data Analysis Software & Systems XIV

October 24 - 27, 2004

Pasadena California

Sheraton


LOBBY LEVEL CONFERENCE ROOMS
West
Center
East
Justines Ballroom
Piazza Atrium
Cordova
CONFERENCE LEVEL ROOMS
Parking
Lobby
Elevators
Storage Room
Res. Rooms
Maple
Oak
Magnolia
Cypress
Aspen
Redwood
Euclid
Marengo
Colorado
Sheraton Pasadena
HOTEL
303 East Cordova Street, Pasadena CA 91101
Tele: (626) 449-4000 Fax: (626) 584-1390
Reservations: 1-800-457-7940

Sponsors

ADASS XIV is pleased to express our appreciation for the support of its sponsoring organizations.

CALIFORNIA INSTITUTE OF TECHNOLOGY

CANADA-FRANCE-HAWAII TELESCOPE

CENTRE DE DONNEES ASTRONOMIQUES DE STRASBOURG

EUROPEAN SOUTHERN OBSERVATORY

INFRARED PROCESSING AND ANALYSIS CENTER

JET PROPULSION LABORATORY

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

NATIONAL OPTICAL ASTRONOMY OBSERVATORIES

NATIONAL RADIO ASTRONOMY OBSERVATORY

NATIONAL SCIENCE FOUNDATION

Smithsonian Astrophysical Observatory

SPACE TELESCOPE SCIENCE INSTITUTE

Committees

Program Organizing Committee

Tim Cornwell (NRAO)
Dick Crutcher (UIUC/NCSA)
Daniel Durand (CADC)
Daniel Egret (CDS)
Carlos Gabriel (ESA-VILSPA)
Tom Handley (IPAC/Caltech/JPL)
Richard Hook (ST-ECF)
Jim Lewis (IoA)
Glenn Miller (ST ScI)
Koh-Ichiro Morita (NAOJ)
François Ochsenbein (CDS)
Michele Peron (ESO)
Arnold Rots, POC Chair (SAO)
Dick Shaw (NOAO)
Betty Stobie (Univ. of Arizona)
Christian Veillet (CFHT)

Local Organizing Committee (Caltech/IPAC)

Mary Ellen Barba
Rick Ebert
Bill Green
Tom Handley
Eloise Kennedy
Mike Khuong
Jake Llamas
Steven Lo
Rosanne Scholey
Patrick Shopbell
Carl Smay

Table of Contents

Confernce at a Glance

	SUNDAY OCTOBER 24, 2004	
10:00 – 13:00 And 13:00 – 16:00	Huntington Gardens and Museum Tour (1 hour guided, 1½ hour self-guided)	Huntington Gardens and Museum
16:00 – 21:00	Registration	Cordova
16:30 – 18:00	FITS (BoF.1)	Maple
19:00 – 21:00	Reception	Piazza
	MONDAY OCTOBER 25, 2004	
8:30 – 9:00	*Plenary Session* Dr. George Helou, IPAC Executive Director	Justine
9:00 – 10:15	*Imaging Algorithms (O1)*	Justine
10:30 – 11:00	*VOSpec: A tool for handling Virtual Observatory compliant Spectra (FM.1)*	Aspen
11:15 – 12:15	*Ground-based Observatories I (O2)*	Justine
12:15 – 13:45	Lunch	
12:30 - 13:45	*Focus Demonstration*	Aspen
13:45 – 15:15	*Great Space Observatories (O3)*	Justine
16:00 – 17:30	*Future Astronomical Data Analysis Environments* (BoF.2) *Sky Indexation, Pixelization, and the VO* (BoF.3)	Maple Cypress
18:00 – 21:15	JPL Tour, Museum, Multimission Image Processing Laboratory	JPL
	TUESDAY OCTOBER 26, 2004	
8:30 – 9:00	*Plenary Session* Dr. David Baltimore, President Caltech	Justine
9:00 – 10:30	*Great Space Observatories I (O4)*	Justine
10:30 – 11:00	*Focus Demonstration*	Aspen
11:15 – 12:15	*The Virtual Observatory: Handling the Data (O5)*	Justine
12:15 – 13:45	Lunch	
12:30 – 13:00	*Focus Demonstration)*	Aspen
13:45 – 15:15	*The Virtual Observatory: Grid (O6)*	Justine
15:30 – 16:00	*Spitzer Pride – Science User Tools (FT.4)*	Aspen
16:00 – 17:15	*The Virtual Observatory: Mining and Analysis (O7)*	Justine
18:30 – 21:30	Banquet	FrontRunner
	WEDNESDAY OCTOBER 227, 2004	
8:30 – 9:00	*Plenary Session* Dr. Charles Elachi, Director JPL	Justine
9:00 – 10:30	*Ground-based Observatories II (O8)*	Justine
10:30 – 11:00	*Focus Demonstrations*	Aspen
11:15 – 12:15	*Detection Algorithms (O9)*	Justine
12:15 – 13:45	Lunch	
12:30 – 13:00	*Focus Demonstration)*	Aspen
14:15 – 16:00	*More Algorithms (O10)*	Justine
1545 – 1600	Closing Remarks	Justine
16:00	Conference Closed	

Program

Sunday October 24, 2004

TIME	ACTIVITY	LOCATION
10:00 – 13:00	Huntington Gardens and Museum Tour (1 hour guided, 1½ hour self-guided)	Huntington Gardens and Museum
13:00 – 16:00	Huntington Gardens and Museum Tour (1 hour guided, 1½ hour self-guided)	Huntington Gardens and Museum
16:00 – 21:00	Registration	Cordova
16:30 – 18:00	FITS (BoF.1)	Maple
19:00 – 21:00	Reception	Piazza

Monday October 25, 2004

TIME	ACTIVITY	LOCATION
8:00 – 17:00	Registration	Cordova
8:00 – 17:00	Poster Session (P1)	Piazza
10:00 – 17:00	Exhibit Demonstrations	Magnolia
8:30 – 9:00	*Plenary Session* Dr. George Helou, IPAC Executive Director	Justine
9:00 – 10:15	*Imaging Algorithms (O1)*	Justine
	9:00 Blind Deconvolution and Super Resolution Imaging Algorithm (O1.1) - Tony Chan	
	9:30 Visual Data Mining of Astronomic Data With Virtual Reality Spaces: Understanding the Underlying Structure of Large Data Sets (O1.2)- Julio Valdes	
	10:00 Hires: Super-resolution for the Spitzer Infrared Telescope (O1.3)- Charles Backus	
10:15 – 11:15	Break	Piazza
10:30 – 11:00	*VOSpec: A tool for handling Virtual Observatory compliant Spectra (FM.1)*	Aspen
11:15 – 12:15	*Ground-based Observatories I (O2)*	Justine
	11:15 Pan-STARRS (O2.1)- Nick Kaiser	
	11:45 The 2MASS Extended Mission and Ancillary Data Products (O2.2)- Roc Cutri	
12:15 – 13:45	Lunch	
12:30- 13:00	*XSA : XMM-Newton Science Archive (FM.2)*	Aspen
13:15 – 13:45	*TOPCAT & STIL: Starlink Table/VOTable Processing Software (FM.3)*	Aspen

13:45 – 15:15	*Great Space Observatories (O3)*	Justine
	13:45 The Great Observatories Origins Deep Survey (GOODS)(O3.1) - Anton Koekemoer	
	14:15 INTEGRAL and its Science Data Centre (O3.2)- Roland Walter	
	14:45 XMM-Newton: Approaching 5 years of successful science operations (O3.3) - Carlos Gabriel	
	15:00 GaiaGrid: Its Implication and Implementation (O3.4)- S. G. Ansari	
16:00 – 17:30	*Future Astronomical Data Analysis Environments* (BoF.2)	Maple
	Sky Indexation, Pixelization, and the VO (BoF.3)	Cypress
17:45 – 19:15	*Python* (BoF.4)	Maple
	Customizing and Extending DS9: Community discussion (BoF.5**)**	Cypress
18:00 – 21:15	JPL Tour, Museum, Multimission Image Processing Laboratory	JPL

Tuesday October 26, 2004

TIME	ACTIVITY	LOCATION
8:00 – 17:00	Registration	Cordova
8:00 – 17:00	Poster Session (P2)	Piazza
8:00 – 17:00	Exhibit Demonstrations	Magnolia
8:30 – 9:00	*Plenary Session* Dr. David Baltimore, President Caltech	Justine
9:00 – 10:30	*Great Space Observatories I (O4)* 9:00 The Spitzer Space Telescope – Science Results and Data Analysis Challenges (O4.1)– Michael Werner 9:30 Spitzer Space Telescope Data Processing and Algorithmic Complexity (O4.2)– Mehrdad Moshir 10:00 The Mars Exploration Rover Imaging System (O4.3)– Justin Maki	Justine
10:15 – 11:15	Break	Piazza
10:30 – 11:00	*Taming the Measurement Equation with MeqTrees (FT.1)*	Aspen
11:15 – 12:15	*The Virtual Observatory: Handling the Data (O5)* 11:15 Science with Virtual Observatory Tools (O5.1) – Paolo Padovani 11:45 Interoperability in action: the Aladin Experience (O5.2) – F. Ochsenbein 12:00 VOSpec: A Tool for handling Virtual Observatory complaint Spectra (O5.3) – Pedro Osuna	Justine
12:15 – 13:45	Lunch	
12:30 – 13:00	*Montage: An Astronomical Image Mosaic Service for the NVO (FT.2)*	Aspen
13:15 – 13:45	*A Demo of Python Tools for Analysis and Visualization of Astronomical Data (FT.3)*	Aspen

13:45 – 15:15	*The Virtual Observatory: Grid (O6)*	Justine
	13:45 Using Grid Technologies to Support Large-Scale Astronomy Applications (O6.1) – Eva Dellman 14:15 Features of the AstroGrid approach to Virtual Observatory Architecture (O6.2) – Tony Linde 14:30 Deploying the AstroGrid: Science Use Ready (O6.3) – Nicholas A. Walton 14:45 Reliable, Automatic Transfer and Processing of Large Scale Astronomy Datasets (O6.4) – Tevfik Kosar 15:00 A Framework for Parallel Data Analysis on a Distributed Grid (O6.5) – Jeffrey P. Gardner	
15:15 – 16:00	*Break*	Piazza
15:30 – 16:00	*Spitzer Pride – Science User Tools (FT.4)*	Aspen
16:00 – 17:15	*The Virtual Observatory: Mining and Analysis (O7)*	Justine
	16:00 Registries and Publishing in the Virtual Observatory (O7.1) – Ray Plante 16:30 OpenSkyQuery & OpenSkyNode – the VO Framework to Federated Astronomy Archives (O7.2) – William O'Mullane 16:45 VO-Enabling a major astronomical analysis and reduction software system (O7.3) – David Giaretta 17:00 An O(NlogM) Algorithm for Catalogue Matching (O7.4) – Drew Devereux	
18:30 – 21:30	Dr. Raymond Arvidson, Deputy Principal Investigator, Mars Exploration Rover Mission Spirit and Opportunity: The Mars Exploration Rover Mission	FrontRunner – Santa Anita Racetrack

Wednesday October 27, 2004

TIME	ACTIVITY	LOCATION
8:00 – 12:00	Registration	Cordova
8:00 – 16:00	Poster Session (PM)	Piazza
8:00 – 16:00	Exhibit Demonstrations	Magnolia
8:30 – 9:00	*Plenary Session*	Justine
8:30	Dr. Charles Elachi, Director JPL	
9:00 – 10:30	*Ground-based Observatories II (O8)*	Justine
9:00	Simulating the Performance of the Square Kilometer Array (O8.1) – Colin Lonsdale	
9:30	The Gemini Science Archive: Current Status and Future Projects (O8.2) – Colin Aspin	
9:45	Simulation of the Future LST Data Pipeline (O8.3) – Ghaleb Abdulla	
10:00	Linking and tagging initiative at the Astrophysical Journal (O8.4) – Greg Schwarz	
10:15	Development of SAOImage DS9: Lessons learned from a small but successful software project (O8.5) – William Joye	
10:15 – 11:15	Break	Piazza
10:30 – 11:00	*New Software for Ensemble Creation in the Spitzer-Space-Telescope Operations Database (FW1.)*	Aspen

11:15 – 12:15	*Detection Algorithms (O9)*	Justine
	11:15 SEXtractor, ten years after (O9.1) – Emmanuel Bertin	
	11:45 Fast Algorithms for Massive-Scale Classification Problems: Toward 1 Million Quasars (O9.2) – Alexander Gray	
	12:00 Detection of Rare Objects in Massive Astrophysical Datasets Using Innovative Knowledge Discovery Technology (O9.3) – Alvaro Soto	
12:15 – 13:45	Lunch	
12:30 – 13:00	*RedShift 5- Virtual Planetarium for Amateurs and Professionals (FW.2)*	Aspen
14:15 – 16:00	*More Algorithms (O10)*	Justine
	14:15 Model-Based Count-Limited Image Restoration (O10.1) – David Van Dyk	
	14:45 SAADA: Astronomical Database Made Easy (O10.2) – L. Michel	
	15:00 Mosaicking with MOPEX (O10.3) – David Makovoz	
	15:15 W projection: a new algorithm for wide field imaging with radio synthesis arrays (O10.4) – Tim Cornwell	
	15:30 Hardware Acceleration for Astronomical Data Analysis (O10.5) – Eric Sessoms	
1545 – 1600	Closing Remarks	Justine
16:00	Conference Closed	

Birds of a Feather (BoFs)

BoF.1 FITS

William Pence and the IAU FITS Working Group

This Birds-of-a-Feather session will present a summary of current activities related to the FITS data format and will provide a forum for the discussion of current issues. The following topics will be covered:

- Recent changes to the FITS committees and their membership
- New streamlined procedures for approving FITS proposals
- Status of current FITS proposals
  - FITS MIME types
  - World Coordinate Systems Papers III and IV
  - TDIMn and Variable Length array conventions
  - Registry of FITS conventions and keyword data dictionaries
- The role of FITS in the VO: Are any changes needed?
- Open forum for short presentations on FITS topics by attendees.

All ADASS attendees are encouraged to subscribe to the FITS newsgroup at

http://listmgr.cv.nrao.edu/mailman/listinfo/fitsbits

to keep abreast of the latest FITS issue

BoF.2 Future Astronomical Data Analysis Environments

P. Grosbol, D. Tody

Most of the systems currently used to analyze astronomical data were designed and implemented a decade or more ago. Although they are still very useful for analysis, one would like better support for newer concepts like archives, Virtual Observatories and GRID. Furthermore, incompatibilities between most of the current systems with respect to control language and semantics make it cumbersome to intermix applications from different origins.

An OPTICON Network, funded by EU FP6, started this year to discuss high-level needs for an astronomical data analysis environment which could provide flexible access to both legacy applications and new astronomical resources. At the same time the VO community is developing the infrastructure to enable distributed multi-wavelength data analysis, and any future data analysis environment will need to be fully integrated with VO. In this BOF we will review current activities in both the OPTICON network and in the VO, and discuss requirements and approaches for such a future environment.

BoF.3 Sky Indexation, Pixelization, Cypress and the VO

Gorski, Szalay, William O’Mullane, Banday

No abstract available.

BoF.4 Python: A Better Development Language for Astronomical Data Analysis

Lee Rottler, Paul Barrett, Perry Greenfield

Python continues to gain interest as a basis for scripting and writing data analysis applications for astronomy. It is already being used for operational software by a number of astronomical institutions and for significant data analysis applications such as MultiDrizzle, JConsole, and JIDE. The last 2 are applications written in Jython, a version of Python running in a Java environment, which are used as the interactive data reduction and analysis environment for the Herschel project. What this Bird of a Feather session will focus on is what remains to be done to make Python a suitable replacement for IDL(TM) as the interactive data analysis environment for everyday astronomers.

To that end we will begin by summarizing the current state of tools available as well as outline what we view is needed to be done to make Python competitive with IDL(TM) and Matlab(TM). We will also present a short summary of why Python is the language of choice for scientific data analysis and reduction in general and astronomy in particular. The remainder of the session will be devoted to discussion and feedback, particularly with regard to the requirements and tools astronomers feel are most important towards achieving this objective. Given the current state of Python and its many add-on packages, it is hoped that this objective can be met in the next year.

We particularly seek the opinions of those with an interest in Python but for one reason or another have yet to start using it. For instance, what are the current barriers, perceived or otherwise, preventing one from jumping in and using it now? What features are missing that if added would give one the incentive to drop IDL and start using Python for doing development and interactive astronomical data reduction?

The feedback received at the BoF will help prioritize the features needed to be worked on to finally make Python the interactive data analysis and reduction language for astronomers. We also hope that it will help spur efforts to coordinate such work between different institutions and research groups, and different areas of astronomy.

BoF.5 Customising and Extending DS9: community discussion

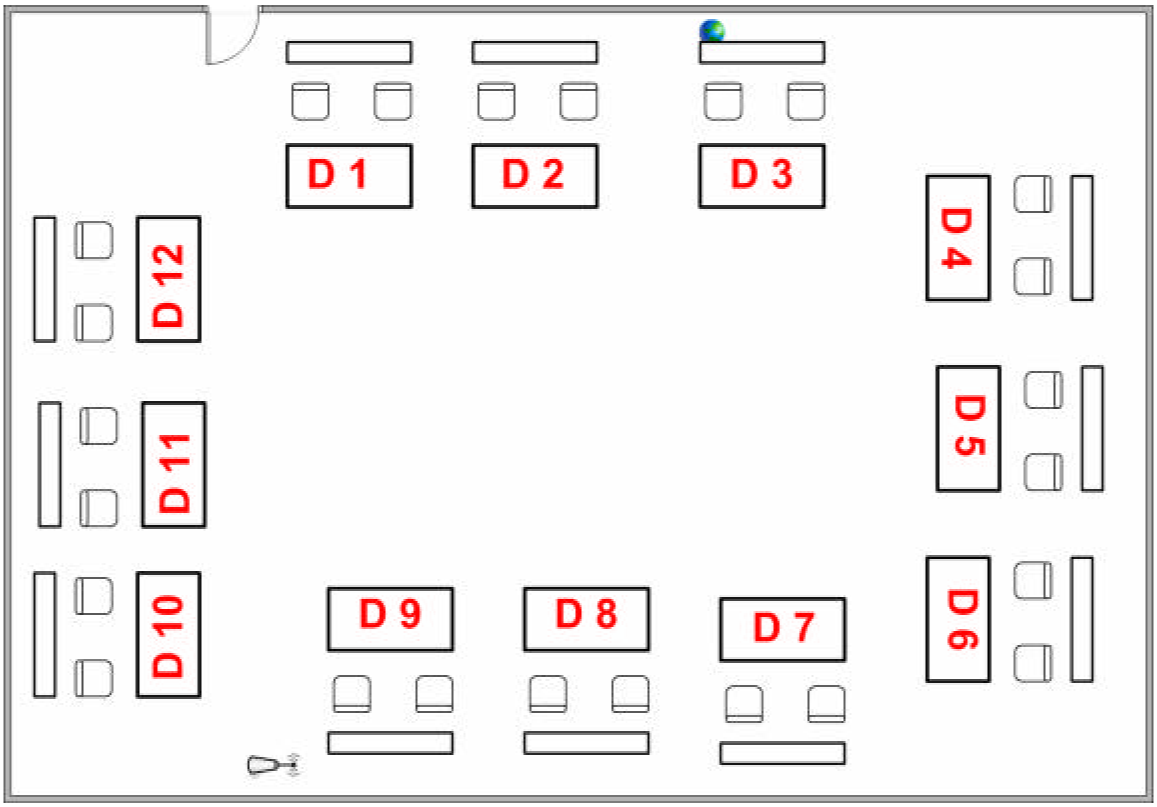
De Clarke, William Joye, Steven Allen

CFA's display tool DS9 is becoming a de facto standard at many astronomical institutions. The tool is highly flexible and adaptable. DS9 hackers Clarke and Allen, and author Bill Joye, would like to invite the user community (beginners, experts, and prospective users) to a discussion about the nuts and bolts of DS9 enhancement and extension, DS9 as a display tool for IRAF, DS9 for live readout, and similar topics.

Our objective is to discover how extensive the DS9 user community is, to share useful information (possibly tutorials and demos), and if there is sufficient interest, to start an ongoing community effort with a view to code sharing, API specifications, and distributed development. We'd also like to find out what others are doing with DS9.

We feel that DS9 -- with its open source, unusual flexibility, hooks for VO capability, and so on -- has (potentially) a bright future as the quick-look, live-capture, and image- reduction display of choice for astronomers. Now seems like a good time to lay the foundation of a strong ongoing user community to support and enhance the toolkit. ADASS seems like a good place to start. (UCO/Lick also plans to set up a mailing list and wiki, to be announced at ADASS.)

Demonstrations (Magnolia Room)



D.1 *New Content and Functionality of the NASA/IPAC Extragalactic Database (NED)*

Marion Schmitz, IPAC, Caltech/JPL, Olga Pevunova, IPAC, Caltech/JPL, Marianne Brouty, IPAC, Caltech/JPL, Kay Baker, IPAC, Caltech/JPL, Harold Corwin, IPAC, Caltech/JPL, Cren Frayer, IPAC, Caltech/JPL, Cheryl LaGue, IPAC, Caltech/JPL, Barry Madore, IPAC, Caltech/JPL, Joseph Mazzarella, IPAC, Caltech/JPL, George Helou, IPAC, Caltech/JPL

New content and functionality of NED will be demonstrated.

Highlights include diameter data from major survey catalogs and a new mode for output in VOtable XML.

D.4 *NOAO and IRAF Project Demonstrations*

NOAO Data Products Program

Members of the NOAO Data Products Program will be on hand to discuss recent and near-term developments in:

- IRAF/X11IRAF Development -- CL error handling and new user features, new tasks, VO integration and major XImtool enhancements;
- The Mosaic Pipeline -- An IRAF-based, science driven, modular, highly configurable, and distributed reduction pipeline system for mosaic detectors;
- Gemini/IRAF Software -- Cooperative development of reduction software and core system functionality;
- Survey Program data access and next-generation archive;
- An LSST Observation simulator -- A general-purpose survey observation planning tool.

Demonstrations and further information on these and other projects will be available at the demo table.

D.5 *ACS GTO Science Data Archive Visualization Tools*

Terence Allen, Kenneth Anderson, W. Jon McCann

The ACS GTO Science Data Archive (SDA) offers access to all products generated by the ACS GTO data processing pipeline (APSIS) through a database-driven web interface. It enables science team members to construct ad hoc and deeply structured queries through visually intuitive navigation and search tools. The current capabilities include: color and magnitude constrained object searches, cross-correlated searches with other online catalogs, direct download access to all FITS images, download of complete object catalogs and search results in a variety of formats (including VOT, HTML, XML, CSV).

The web interface comprises a suite of tools that have been developed for the SDA to supplement the core functionality of the archive. We will demonstrate tools that provide such functionality as on-the-fly color image generation, clickable object maps, image cutouts of cataloged objects, object aperture and wcs data overplotting, arbitrary image zooming, scaling and color bias, user-selectable image intensity scaling/stretching, data product pedigree charting and dynamically generated data product inventories. The recently added functionality of allowing a user to fine tune the full color maps which are generated on the fly has proven its scientific worth: a selection of scales and scale stretches has aided in the visual identification of AGNs and other.

D.6 *Development of VO-enabled Applications in Italy*

Claudio Gheller, CINECA, Ugo Becciani, INAF Catania, Giuliano Castelli, INAF Trieste, Marco Comparato, INAF Catania, Ciro Donalek, Naples University, Giuseppe Longo, Naples University, Fabio Pasian, INAF Trieste, Alberto Pepe, CINECA, Riccardo Smareglia, INAF Trieste, Giuliano Taffoni, INAF Trieste, Roberto Tagliaferri, Salerno University, Claudio Vuerli, INAF Trieste

We will present the results obtained by several Italian research institutions in order to provide new Virtual Observatory enabled applications and tools. The Italian Institute for Astronomy and Astrophysics (INAF) coordinates the effort. The academic supercomputing center CINECA and the Universities of Naples and Salerno have largely contributed to the results.

-AstroMAF

AstroMAF is software of visualization and analysis of astronomical multi-dimensional data (observational and numerical).AstroMAF can connect to a VO-enabled web service and retrieve data in the VOTable, FITS, binary or ASCII formats. Data can be visualized using various advanced graphical techniques and can be analysed in different ways. Multidimensional visualization is fully supported.

-AstroMining

The AstroMining software is a package written using MatLab to perform a large number of data mining and knowledge discovery tasks in multiparametric astronomical datasets.

AstroMining handles VOTables or ASCII tables. Via interactive interfaces, it is possible to perform a number of operations, from data manipulation to parameters selection etc. The package accounts for large set of visualization and statistical tools.

-Grid-Enabled Databases Access

The Italian research Grid infrastructure (Grid.it) is built on top of GT-2 which does not provide any mechanism to access databases. To overcome this limitation a database client has been built and installed on the WorkerNode of Grid.it site and connects directly with the Oracle server. Data are extracted from the Oracle database and made available in VOTables stored on the WorkerNode.
A user-friendly interface allows to submit queries, submit grid jobs to process data and display results.

D.7 *Pipeline support tools, ESO Vimos IFU pipeline, Midas*

Carlo Izzo, ESO, Ralf Palsa, ESO, Nick Kornweibel, ESO, Derek McKay, ESO and RAL, Klaus Banse, ESO

We will demonstrate Gasgano and EsoRex, two tools for executing recipes within the currently installed VLT pipelines as well as for the new CPL-based pipelines.
Gasgano is a GUI-based Data File Organiser to help ESO's user community to manage and organise in a systematic way the astronomical data observed and produced by all VLT compliant telescopes. EsoRex is a command line tool to list, configure and execute CPL-based recipes. Recipes for the CPL pipelines run as plugins (dynamic libraries). Therefore, it is not possible to run them directly from the command line; one has to use a wrapping application. EsoRex is such a wrapper, saving recipe developers the need to write such an application themselves. Furthermore, the ESO-IFU pipeline for VIMOS will be demonstrated.

Finally, the 04SEP release of ESO-Midas will be shown.

D.8 *The NASA/IPAC Infrared Science Archive (IRSA): The Demo*

Anastasia Alexov, G. Bruce Berriman, Nian-Ming Chiu, John C. Good, Thomas H. Jarrett, Mih-seh Kong, Anastasia Clower Laity, Serge Michel Monkewitz, Naveed D. Tahir-Kheli, Saille Warner-Norton, Angela Zhang

This demonstration shows the user services available at the NASA/IPAC Infrared Science Archive (IRSA). Companion presentations describe the architecture that supports this archive.

Currently there are nearly 250,000 data requests a month, taking advantage IRSA's data repository which includes 660 million sources (60 catalogs), 10 million images (22 image sets; 10.4 TB) and over 70,000 spectra (7 spectroscopic data sets).These data are the science products of The Two Micron All Sky Survey (2MASS), The Infrared Astronomical Satellite (IRAS),The Midcourse Space Experiment (MSX),The Submillimeter Wave Astronomy Satellite (SWAS), The Infrared Space Observatory (ISO), The Infrared Telescope in Space (IRTS), The Spitzer First Look Survey (FLS), Spitzer Legacy & Ancillary data, Spitzer Reserved Observations (ROC) and the Spitzer Space Telescope data.

IRSA is also seamlessly interoperable with 10 remote archives and services: GOODS, ISO, MAST, VizieR, DSS, NVSS, FIRST, HEASARC, NED and JPL, which help expand the available data set wavelength range from X-ray to radio. The majority of IRSA's image collections are Simple Image Access (SIA) compliant and are available through the Virtual Observatory (VO) data mining tools.

The IRSA demo will include: IRSA's inventory service RADAR, IRSA's data fusion service OASIS and IRSA's general search service for complex data collections Atlas.

D.9 *ESA VO Services*

Iñaki ORTIZ, ESA / ESAC, Christophe ARVISET, ESA / ESAC, Isa BARBARISI, ESA / ESAC, John DOWSON, ESA / ESAC, Jose HERNANDEZ, ESA / ESAC, Pedro OSUNA, ESA / ESAC, Jesús SALGADO, ESA / ESAC, Guillermo SAN MIGUEL, ESA / ESAC, Aurèle VENET, ESA / ESAC

Following ESAs decision to increase its participation in the VO, more emphasis and dedicated manpower has been setup to develop ESA VO services at the European Space Astronomy Centre (previously VILSPA) located near Madrid, Spain.

All ESA VO services available from the ESAC Archive Group will be demonstrated, in particular:

- ESA-VO Portal
- SIAP access to the ISO Data Archive
- SIAP access to the XMM-Newton Science Archive
- SSA access to the ISO Data Archive
- ESA-VO Registry
- ESA VOSpec tool

Additionally, existing scientific archives located at ESAC will be
accessible for demonstration:
- The ISO Data Archive (IDA),
- The XMM-Newton Science Archive (XSA),
- The Integral SOC Science Archive (ISDA),
- The Planetary Science Archive (PSA),

D.10 *Extending the Capabilities of CIAO with S-Lang-based Tools*

F. Primini, Smithsonian Astrophysical Observatory, M. Noble, Massachusetts Institute of Technology CXC Science Data Systems Group

We demonstrate a number of tools that extend the capabilities of the CIAO x-ray astronomy analysis software package. These tools were developed using S-Lang (http://www.s-lang.org), an interpreted language and multi-platform programmer's library that may be easily embedded into other applications. The S-Lang interpreter's C-like syntax and sophisticated vector and array manipulation capabilities allow rapid development by scientists of programs that meet their individual analysis needs. The tools demonstrated here include simple point source aperture photometry tools launched from DS9, and interactive data visualization and filtering tools. All are available for downloading via the CIAO web pages (http://cxc.harvard.edu/ciao/slang/).

D.11 *The National Virtual Observatory*

Robert Hanisch, Space Telescope Science Institute
on behalf of the NVO Project Development Team

We will demonstrate the most recently developed capabilities of the US National Virtual Observatory project.

D.12 *Starlink Software Developments*

David Giaretta, Starlink/RAL, Malcolm Currie, Starlink/RAL, Stephen Rankin, Starlink/RAL, Mark Taylor, Starlink/U of Bristol, Alasdair Allan, Starlink/U of Exeter, Norman Gray, Starlink/U of Glasgow, Peter Draper, Starlink/U of Durham, David Berry, Starlink/U of Central Lancashire (normally abbreviated to U C Lancs)

We shall demonstrate the latest Starlink software and its applicability to the Virtual Observatory. This includes Java analysis tools such as FROG for time-series, ORAC-DR pipelines for more ESO instruments, automated astrometric calibrations in GAIA, and distributed pipeline processing. For the first time we shall give some presentations under Mac OS X.

Focus Demonstrations (Aspen Room)

FM.1 *VOSpec: A tool for handling Virtual Observatory compliant Spectra*

Pedro OSUNA (ESAC/ESA), Christophe ARVISET (ESAC/ESA), Isa BARBARISI (ESAC/ESA), Jesús SALGADO (ESAC/ESA)

The VOSpec tool can superimpose spectra coming from VO-compatible resources that conform to the VO Simple Spectral Access Protocol (SSAP) with three added extra fields in the VOTable through the use of an algorithm based on dimensional analysis.

The Focus Group Demo will consist of a general presentation of the tool, how it interacts with the VO through the usage of VO protocols and already existing tools (like VOTable, VOTable parsers, Registries, etc.) and how it solves the problem of superimposition.

VU-Graph presentation will be accompanied by live demonstration of the capabilities of the tool. A rough timeline for the presentation/demo follows:

- General Layout of the tool
- Target/Coordinate Search and Connection to SSAP registries
- Superimposition of Spectra: Dimensional Equation solution
- Display of superimposed spectra for different instruments with different units: the case of the Short Wavelength Spectrometer (SWS) and the Long Wavelength Spectrometer (LWS) of the Infrared Space Observatory (ISO).
- Display of spectra from different wavelengths: the Infrared Space Observatory (ISO) and International Ultraviolet Explorer (IUE) cases
- Display of local VOTable compatible resources to allow users to compare their data with VO-obtained spectra: the XMM-Newton and Chandra cases.
- Display of local VOTable data: HST, SPITZER and VLT cases.

- A spectrum display of the whole wavelength range using different VO-compatible spectra.

FM.2 *XSA : XMM-Newton Science Archive*

Christophe ARVISET, ESA / ESAC , John DOWSON, ESA / ESAC, Jose HERNANDEZ, ESA / ESAC, Iñaki ORTIZ, ESA / ESAC, Pedro OSUNA, ESA / ESAC, Jesús SALGADO, ESA / ESAC, Guillermo SAN MIGUEL, ESA / ESAC, Aurèle VENET, ESA / ESAC

The XMM-Newton Science Archive (XSA) is located at ESAC (European Space Astronomy Centre, previously called VILSPA), near Madrid, Spain. It has been available to scientific community since mid April 2002 at http://xmm.vilspa.esa.es/xsa/.

In this demo, we will show all the facilities which makes the XSA one of the most user friendly and powerful scientific archives, in particular:

- Powerful, complex and easy to build queries against the observations and exposures catalogues
- Unique search capabilities against the XMM-Newton source catalogue
- User configurable query panels and results display
- Product visualization tools
- Product direct retrieval with a single click
- Customizable product retrieval via a shopping basket
- Link to the articles related to an observation
- On-the-Fly Reprocessing capabilities

FM.3 *TOPCAT & STIL: Starlink Table/VOTable Processing Software*

Mark Taylor (Starlink, Bristol University)

The Starlink Tables Infrastructure Library (STIL) is a pure-Java, open source library for I/O and manipulation of tabular data such as astronomical catalogues. It is written to be high-performance, and to cope with large tables (not limited by available memory).The core library is format-neutral, with the work of serialization and deserialization performed by pluggable format-specific I/O handlers. This means that the programmer sees a high-level abstraction of a table which is easy to work with, and also that support for new data formats can be added easily. A corollary is that conversion between any of the supported formats is trivial. Supplied handlers provide support for VOTables, FITS table extensions, relational databases via SQL and plain text tables, amongst others. The VOTable handler, which has VOTable-specific as well as generic table features, is believed to be the only existing library capable of reading or writing all the defined VOTable serialization formats (TABLEDATA, FITS, BINARY).

TOPCAT, based on STIL, is a user-friendly graphical program for viewing, analysis and modification of tables. It has facilities for plotting, cross matching, row selection, sorting and metadata manipulation. Synthetic columns can be created and row selections made using a powerful and extensible algebraic expression language.

We will demonstrate TOPCAT's analysis, matching and editing facilities, and describe how STIL can be used to read or write VOTables or other tables in your own programs. We will also demonstrate command-line utilities which provide scriptable access to some of this functionality.

FM.4 *The National Virtual Observatory*

Robert Hanisch, Space Telescope Science Institute on behalf of the National Virtual Observatory Project Team

We will show highlights of the most recently developed US NVO science applications and tools.

FT.1 *Taming the Measurement Equation with MeqTrees.*

J.E. Noordam, ASTRON , O.M. Smirnov, ASTRON

A Measurement Equation is used to predict values of the data measured with a particular instrument, e.g. a radio telescope. It is a combined model of the instrument and the observed object(s). One way of implementing an arbitrary M.E. is by means of 'MeqTrees', which can also be used to solve for arbitrary subsets of its parameters. We will demonstrate how the use of a more complex M.E. gives better results than using the ones that are implicit in existing radio astronomy reduction packages.

FT.2 *Montage: An Astronomical Image Mosaic Service for the NVO*

Anastasia Clower Laity, Infrared Processing and Analysis Center, Caltech, Nate Anagnostou, Infrared Processing and Analysis Center, Caltech, Bruce Berriman, Infrared Processing and Analysis Center, Caltech, John Good, Infrared Processing and Analysis Center, Caltech, Joseph C. Jacob, Jet Propulsion Laboratory, Caltech, Daniel S. Katz, Jet Propulsion Laboratory, Caltech

Montage is a software system for generating astronomical image mosaics according to user-specified size, rotation, WCS-compliant projection and coordinate system, with background modeling and rectification capabilities. Its architecture has been described in the proceedings of ADASS XII and XIII. It has been designed as a toolkit, with independent modules for image reprojection, background rectification and co-addition, and will run on workstations, clusters and grids. The primary limitation of Montage thus far has been in the projection algorithm. It uses a spherical trigonometry approach that is general at the expense of speed. The reprojection algorithm has now been made 30 times faster for commonly used tangent plane

to tangent plane reprojections that cover up to several square degrees, through modification of a custom algorithm first derived by the Spitzer Space Telescope. This focus session will describe this algorithm, demonstrate the generation of mosaics in real time, and describe applications of the software. In particular, we will highlight one case study which shows how Montage is supporting the generation of science-grade mosaics of images measured with the Infrared Array Camera aboard the Spitzer Space Telescope.

FT.3 *A Demo of Python Tools for Analysis and Visualization of Astronomical Data*

Perry Greenfield, STScI, Jay Todd Miller, STScI, Jin-Chung Hsu, STScI, Paul Barrett, STScI, Warren Hack, STScI

We will demonstrate a number of the Python modules and packages that we have developed (or helped to develop) as part of the Space Telescope Science Institute's efforts to make Python a productive data analysis environment for both interactive use and application development. We will demonstrate the use of the multi-dimensional array package numarray, PyFITS, numdisplay (a DS9 display package for arrays), and matplotlib, with particular emphasis on matplotlib, a new 2D plotting package.

FT.4 *Spitzer Pride - Science User Tools*

` Trey Roby, SSC Caltech, Xiuqin Wu, SSC Caltech

Spitzer(formerly SIRTF) Science Center has supplied astronomy community with a set of science user tools called Spitzer Pride. They are Spot(the planning tool), Leopard(the archive client), and Subscriber(the file downloading application). We will discuss the whole system design of client-middleware-server briefly. The session will focus on how the set of tools working together, how they were designed to use the common set of Java classes to give the familiar look and feel, how they communicate to keep each other auto-updated at the same time. The tools can be run on Solaris, Windows, Linux, and Mac OSX platforms.

FW.1 *New Software for Ensemble Creation in the Spitzer-Space-Telescope Operations Database*

Russ Laher, John Rector

Some of the computer pipelines used to process digital astronomical images from NASA's Spitzer Space Telescope require multiple input images, in order to generate high-level science and calibration products. The images are grouped into ensembles according to well-documented ensemble-creation rules by making explicit associations in the operations Informix database at the Spitzer Science Center(SSC). The advantage of this approach is that a simple database query can retrieve the required ensemble of pipeline input images. New and improved software for ensemble creation has been developed. The new software is much faster than the existing software because it uses pre-compiled database stored-procedures written in Informix SPL (SQL programming language). The new software is also more flexible because the ensemble-creation rules are now stored in and read from newly defined database tables. This table-driven approach was implemented so that ensemble rules can be inserted, updated, or deleted without modifying software.

FW.2 *RedShift 5 - Virtual Planetarium for Amateurs and Professionals*

Sergey Kudryavtsev, Maris Technologies Ltd., Nick Maris, Maris Technologies Ltd., Oleg Kalitsev, Maris Technologies Ltd., Michael Fleissner, United Soft Media Verlag GmbH, Kurt Tomaszewski, United Soft Media Verlag GmbH, Claus Vester, cc-live

RedShift is a desktop planetarium developed over more than 10 years by Maris Technologies, Ltd. The latest, 5th version of the program includes:

- the complete Hipparcos and Tycho-2 stellar catalogues (2,5 mln stars) as well as some 15 mln stars of up to the 14th-15th mag from the Hubble Guide Star Catalog, ver.2.2;
- the newest version of the General Catalogue of Variable Stars and New Catalogue of Suspected Variable Stars;
- Catalogue of Principal Galaxies (70,000+ objects) plus latest catalogues of optically visible open clusters and nebulae, radio and X-ray sources, quasi-stellar objects and BL Lacertae;
- updated catalogues of more than 50,000 asteroids and 1,500 comets;
- expanded collection of images of galactic and extragalactic objects from Messier and Caldwell catalogues;
- revised high resolution maps of planets obtained from the latest space missions;
- all natural moons of planets including recently discovered moons of Jupiter, Saturn, Uranus, Neptune;
- night vision option (screen dimming and reddening) useful for observers;
- macro-player to view the collection of guided tours of space;
- solar eclipses maps;
- quick set-up of the sky view with newly integrated Open GL technologies.

The database of asteroids, comets and spacecraft can be easily updated through importing the latest orbital elements from NORAD, MPC and other Internet resources.

The family of RedShift 5 will be expanded by release of a special version for Beginners (Autumn 2004) and the College Edition version (Spring 2005).

Banquet

Speaker: Dr. Raymond Arvidson

Deputy Principal Investigator, Mars Exploration Rover Mission

James S. McDonnell Distinguished University Professor Department of Earth and Planetary Sciences Washington University in St. Louis

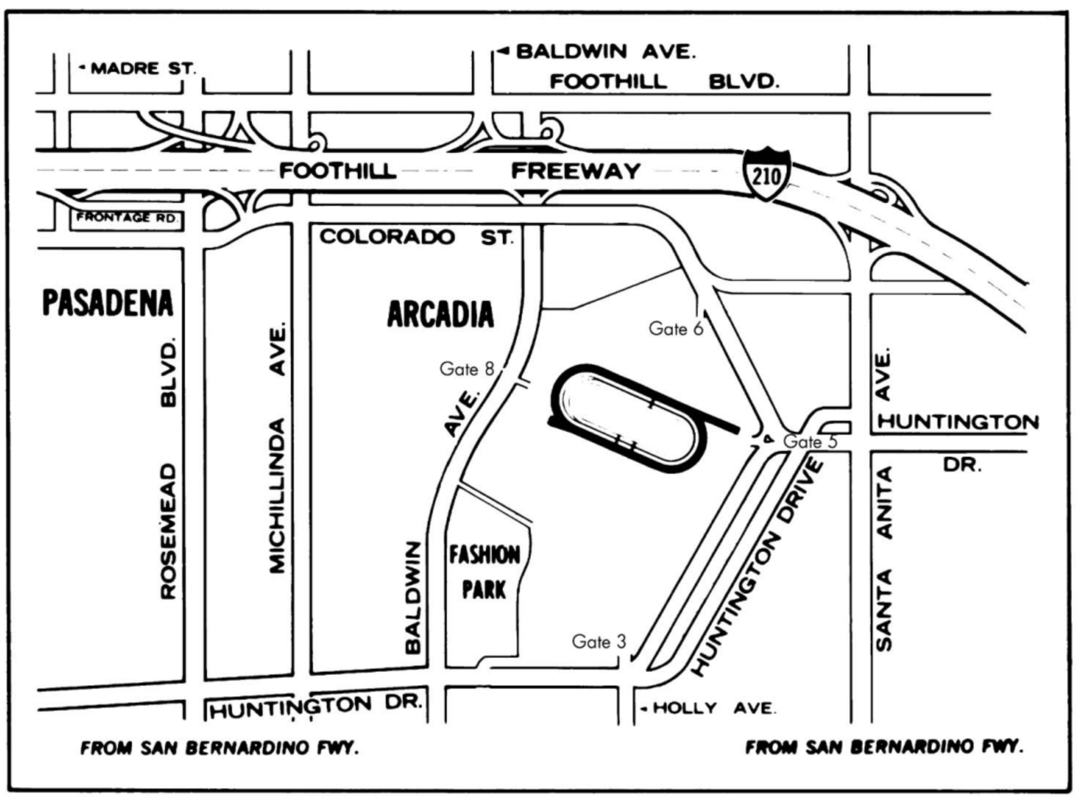
Title: Spirit and Opportunity: The Mars Exploration Rover Mission

Abstract: The two rovers touched down on Mars in January 2004 and have been exploring the surface and making measurements designed to understand the geologic histories and the role of water for the plains within Gusev Crater (Spirit) and the hematite-bearing deposits in Meridiani Planum (Opportunity). The basaltic lava flows that Spirit landed on have been altered by impact, wind, and perhaps deposition of salts during moderately wet period, e.g., by vapor deposition. The Columbia Hills, which Spirit drove onto on Sol 157, contain older rocks that have been highly modified by corrosive ground waters. Opportunity found evidence for layered, cross bedded rocks formed in shallow water and extensive chemical precipitates (sulfates) that formed by evaporation of a water body and/or alteration by later corrosive ground waters. The Mars Exploration Rover Mission has revolutionized our understanding of the red planet and demonstrated that water was present on and beneath the surface, a necessary ingredient for the origin and development of life.

Directions to FrontRunner at Santa Anita

Driving Directions from the Pasadena Sheraton to Santa Anita…

From the hotel,
    turn right (east) on Cordova
    turn right (south) on South Los Robles
    turn left (east) on East California Blvd
      at Michillinda, the street's name changes to Sunset
      and curves gently south to Huntington
    turn left (east) on Hungtington Drive
      at Baldwin Avenue
      go into Gate # 5. From Gate #5, there will be signs
      directing you to the valet lot
      where you will find self-parking.



Oral Presentations

Session 1: Imaging Algorithms

O1.1 ***Blind Deconvolution and Super-Resolution Imaging Algorithm***

Tony Chan, UCLA , Fred Park, UCLA , Andy Yip, UCLA

I'll review some work that I have been involved in on blind deconvolution and super-resolution imaging algorithms. The techniques include (1) a total-variation based blind deconvolution method capable of recovering discontinuities in both the image and the blurring function, (2) two recent extensions of (1) which include auto-focusing and the blind recovery of blurred images further degraded by occlusion, and (3) a super-resolution method for obtaining higher-resolution images from a set of sub-pixel shifted lower resolution images. It is hoped that the modeling and computational techniques may also be useful for applications in astronomical imaging.

O1.2 ***Visual Data Mining of Astronomic Data with Virtual Reality Spaces: Understanding the Underlying Structure of Large Data Sets*.**

Julio Valdes

The information explosion in astronomy, partially related to the developments in data gathering, sensors, and other technologies, requires the development of data mining procedures that enable the faster discovery and more in-depth understanding of the internal structure of the data. This is crucial for the identification of valid, novel, potentially useful, and understandable patterns and regularities.

There are different kinds of data (relational, graphical, symbolic, etc.), and patterns (geometrical, logical, behavioral, etc.). The studied objects are usually described in terms of large collections of heterogeneous properties. That is, variables in the nominal, ordinal, interval, and ratio scales, and variables of a more complex nature, such as images, graphs, time-series, videos, documents, and others. In addition to the complexity of the information, different degrees of imprecision, uncertainty, and incompleteness (missing data) may be present.

A Virtual Reality (VR) approach for large heterogeneous, incomplete and imprecise information is introduced for the problem of visualizing and analyzing astronomic data. This VR-based visual data mining technique speeds up the process of discovery and understanding the hidden structure of data, and allows the incorporation of the unmatched geometric capabilities of the human brain into the knowledge discovery process.

The method is based on mappings between one heterogeneous space representing the data, and a homogeneous virtual reality space.

This approach has been applied successfully to a wide variety of real-world domains, and it has a large potential in astronomy. Examples are presented from the domain of galaxy research.

O1.3 ***Hires: Super-resolution for the Spitzer Infrared Telescope***

Charles Backus, JPL , Thangasamy Velusamy, JPL , Tim Thompson, JPL , John Arballo, JPL

We present a fast deconvolution program based upon the R-L algorithm and its generalization to the case of redundant sky coverage addressed by Aumann, Fowler, and Melnyk. The program makes extensive use of FFT's, resamples using a drizzle method, and accounts for projection and distortion effects that are important in Spitzer data. We address some of the tactical issues involved, such as spatial frequency response, background removal and ringing near sharp transitions, and the importance of accurate psf

images. We conclude with a few examples of processed Spitzer observations.

Session 2: Ground-based Observatories I

O2.1 *Pan-STARRs*

Nick Kaiser

No abstract provided.

O2.2 *The 2MASS Extended Mission and Ancillary Data Products*

Roc M. Cutri, IPAC/Caltech

The objective of the primary 2MASS mission was to carry out a highly uniform digital imaging survey of the entire sky in three near infrared bands. Three fundamental data products were released to the astronomical community in March of 2003 as the realization of this goal: a digital Image Atlas containing 4.1 million calibrated FITS images; 2) a Point Source Catalog (PSC) containing accurate astrometry and photometry for ~471 million sources; and 3) an Extended Source Catalog (XSC) that contains accurate positions, photometry and basic shape information for ~1.6 million resolved sources, most of which are galaxies.

Considerably more data were acquired during survey operations than were ultimately used in the generation of the All-Sky Release products. These included repeated scans of selected Tiles to raise the quality of the Survey data, or as part of targeted validation observations, thousands of scans of 8.5' x 1° calibration fields used to derive nightly photometric transformations, and special six times longer exposure scans (6x) of ~500 sq.deg. in selected regions of the sky to address specific scientific objectives proposed by the 2MASS Science Team.

I will discuss the Ancillary Data Products being prepared for release as part of the 2MASS Extended Mission. The goal of the Extended Mission is to exploit the full potential of all data acquired during the survey's operational period. The 2MASS Ancillary Products

complement the highly reliable and uniform all-sky Catalogs by pushing to the faintest possible flux levels from the Survey, calibration and

Session 3: Great Space Observatories I

O3.1 *The Great Observatories Origins Deep Survey (GOODS)*

Anton M. Koekemoer, STScI

We are currently in a unique era with the availability of several great observatories in space covering the widest range yet of observable wavelengths, including the Hubble Space Telescope in the optical, ultraviolet and near-infrared, Chandra and XMM in the X-rays, and most recently Spitzer in the mid- to far-infrared, as well as several other space missions either planned or currently underway. The Great Observatories Origins Deep Survey (GOODS) has been specifically designed to take advantage of the deepest pointings ever of the Universe by a number of these telescopes, in particular by intensively studying two regions of the sky that have each been imaged for 1 - 2 million seconds by the Chandra X-ray telescope. The GOODS project aims to address several fundamental questions in astrophysics, including how galaxies form and evolve, and the nature of their relationship with black hole formation and growth over cosmic time. This is achieved by combining detailed imaging and spectroscopic observations over as many wavelength regimes as possible from X-rays through to radio wavelengths, and also by using a wide array of space-based and ground-based telescopes, in order to obtain the most complete information possible on several tens of thousands of normal galaxies and several hundred quasars and other active galaxies. I will present an overview of the GOODS project, with particular emphasis on the multi-wavelength software, and how this has been used by the GOODS team to carry out a variety of investigations on this unique and extensive set of observational data.

O3.2 *INTEGRAL and its Science Data Centre*

Roland Walter

Data from the International Gamma-Ray Astrophysics Laboratory are processed at the INTEGRAL Science Data Centre and made available worldwide. I will present challenges and results of the mission as well as a summary of the ISDC activities, performances and technical developments.

O3.3 *XMM-Newton: Approaching 5 years of successful science operations*

Carlos GABRIEL , Matteo GUAINAZZI , Leo METCALFE

The major ESA X-ray observatory, XMM-Newton, is approaching 5 years of operations in flight. The design concepts driving the activities of the Science Operations Centre at the European Space Astronomy Centre in Spain have shown their validity for this astrophysics long-term mission.

Scientific exploitation of data by the astronomical community is facilitated especially through a system resting on 4 pillars:

- scientific data dissemination: from raw telemetry up to processed and calibrated high-level science products, such as images, spectra, source lists, etc;
- dedicated science analysis software: development and distribution of mission specific software, as well as of continuously updated instrument calibration;
- scientific archive: access to data but also to high level information on data contents through state-of-the-art, in-house developed archival facilities;
- documentation: continuously updated documentation on all aspects of spacecraft and instrument operations, data reduction and analysis, with particular focus on the most important scientific results obtained by

XMM-Newton, reachable through a comprehensive set of project web pages.

We intend to review all these and related aspects forming the basis of a modern astronomical observatory. Our aim is to show how the innovative but solid elements in the approach to each the mentioned points contribute to forefront science in the light of some of the outstanding scientific results achieved by XMM-Newton.

O3.4 ***GaiaGrid : Its Implication and Implementation***

S. G. Ansari, Directorate of Science, European Space Agency , M.A.C. Perryman, Directorate of Science, European Space Agency , M. ter Linden, DutchSpace

Gaia is an ESA space mission to determine positions of 1 billion objects in the Galaxy at micro-arcsecond precision. The data analysis and processing requirements of the mission involves about 20 institutes across Europe, each providing specific algorithms for specific tasks, which range from relativistic effects on positional determination, classification, astrometric binary star detection, photometric analysis, spectroscopic analysis etc. In an initial phase, a study has been ongoing over the past 3 years to determine the complexity of Gaia's data processing. Two processing categories have materialised: core and shell. While core deals with routine data processing, shell tasks are algorithms to carry out data analysis, which involves the Gaia Community at large. For this latter category, we are currently experimenting with use of Grid paradigms to allow access to the core data and to augment processing power to simulate and analyse the data in preparation for the actual mission. We present preliminary results and discuss the sociological impact of distributing the tasks amongst the community.

Session 4: Great Space Observatories II

O4.1 *The Spitzer Space Telescope – Scientific Results and Data Analysis Challenges*

Michael W. Werner

The Spitzer Space Telescope, launched in August 2003, is returning outstanding images and spectra at infrared wavelengths. Spitzer is the infrared component of NASA's family of Great Observatories. It promises to revolutionize our understanding of the infrared Universe during its lifetime, which is now projected to be greater than five years before cryogen depletion. This talk will overview the initial scientific results from Spitzer [see also the September issue of Astrophysical Journal Supplement] and describes the structure of the Spitzer science program and upcoming opportunities [the Cycle 2 Call for Proposals will be issued in November, 2004].

The data analysis needs of specific scientific programs of particular significance for Spitzer will be summarized. The Spitzer data archive is now open [see spitzer.caltech.edu/SSC/ for information about access], and interested participants are invited to form their own assessment of the challenges and opportunities which the Spitzer data present for innovations in astronomical software and data analysis.

Based on research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

O4.2 *Spitzer Space Telescope Data Processing and Algorithmic Complexity*

Mehrdad Moshir, Spitzer Science Center/Caltech

Data reduction for the Spitzer Space Telescope requires an intricate and flexible software system as well as use of a variety of algorithms. We will discuss the software environment and examples of the algorithms that support the processing of data from three suites of instruments on Spitzer.

O4.3 *The Mars Exploration Rover Imaging System*

Justin N. Maki

In January 2004, NASA's Mars Exploration Rover (MER) mission landed a pair of rovers onto the Martian surface. The vehicles carry a total of 20 cameras - more than all previous Mars-landed missions combined. The Spirit and Opportunity rovers have since driven several kilometers across the Martian surface and have acquired over 50,000 images. The mobile nature of the mission requires an image processing system that produces mosaics and 3-D stereo products within minutes of data receipt here on Earth. This talk will provide an overview of the MER ground image processing system and describe how it is used during MER surface operations.

Session 5: The Virtual Observatory: Handling the Data

O5.1 *Science with Virtual Observatory tools*

Paolo Padovani, ST-ECF/ESO, Garching, Germany

The Virtual Observatory is now mature enough to produce cutting-edge science results. The exploitation of astronomical data beyond classical identification limits with interoperable tools for statistical identification of sources has become a reality. I will present the discovery of 68 optically faint, obscured (i.e., type 2) active galactic nuclei (AGN) candidates in the two Great Observatories Origins Deep Survey (GOODS) fields using the Astrophysical Virtual Observatory (AVO) prototype. Thirty-one of these sources have high estimated X-ray powers ($>10^{44}$ erg/s) and therefore qualify as optically obscured quasars, the so-called QSO 2. The number of these objects in the GOODS fields is now 40, an improvement of a factor > 4 when compared to the only 9 such sources previously known. By going ~ 3 magnitudes fainter than previously known type 2 AGN in the GOODS fields the AVO is sampling a region of redshift - power space previously unreachable with classical methods. I will also discuss the AVO plans to move to our next phase, the EURO-VO, and how that will make the scientific utilization of astronomical data through VO tools routine.

O5.2 *Interoperability in action: the Aladin Experience*

F. Ochsenbein, P. Fernique, F. Bonnarel, M. Allen, T. Boch, F. Genova, A. Schaaff

Aladin was born 8 years ago as a small applet named "Aladin-lite", devoted to a visual correlation of sky images with the astronomical objects described in catalogues and databases. Since then, Aladin has been continually developed to provide good visualisation of catalog data overlaid on the observed pixels, and to be

easily applicable to the holdings of existing image and data servers. Recently, Aladin was used as the main component of the Astrophysical Virtual Observatory prototype, demonstrating the use VO interoperability standards for enabling real science. But interoperability is not just a set of standards used to describe accurately the data -- it is also a means by which programs and tools can interact and cooperate. Several ways of communicating with and from Aladin have been developed, including a command language (scripting), java interfaces, and plug-ins. This presentation will focus on the lessons learned during the Aladin experience in several aspects: the strengths and weaknesses of the programming language (Java); the importance of the standards and protocols, their surprises and their challenges; the difficulties of actual interoperation; the necessity of keeping proper references to the data producers and maintainers -- and maintaining compatibility with the users' material.

O5.3 ***VOSpec: A tool for handling Virtual Observatory compliant Spectra***

Pedro OSUNA ESAC/ESA , Christophe ARVISET ESAC/ESA , Isa BARBARISI ESAC/ESA , Jesús SALGADO ESAC/ESA

The European Space Agency (ESA) has been heavily involved in the development of the Simple Spectral Access Protocol (SSAP) of the International Virtual Observatory Alliance (IVOA) by providing access to the Infrared Space Observatory (ISO) mission spectral data.

The European Space Astronomy Centre (ESAC) at Villafranca del Castillo near Madrid, Spain, will be the centre for Space Based Virtual Observatory activities within the European Space Agency. The VOSpec tool is part of the already started VO activities of ESA.

There are lots of tools for spectra analysis and display in the astronomical community. However, there is a need for a “VO-enabled” tool that can superimpose spectra coming from different projects within the VO and in VO-format.

One of the main reasons for the current absence of such a tool is the fact that astronomical spectra are not as well defined as, e.g., images. The IVOA has made a significant effort in trying to define a Simple Spectral Access Protocol by which part of this problem is solved, giving the rules to be able to handle spectral data in the VO context.

By defining an algorithm based on dimensional analysis for the superimposition of spectra coming from VO resources, ESAC has been able to create a tool that can superimpose VO spectra that declare their metadata in a specific way. To do this, VO SSAP-compatible resources need only give three new extra fields in their VOTable SSAP.

At the time of writing this abstract, ESAC is planning to propose these three extra metadata fields for acceptance at the IVOA as optional inputs in the definition of the Simple Spectral Access Protocol.

Session 6: The Virtual Observatory: Grid

O6.1 *Using Grid Technologies to Support Large-Scale Astronomy Applications*

Ewa Deelman

This talk presents work done to date on describing and running large-scale astronomy applications on distributed systems such as the grid. One of the applications targeted in this work is the Montage application that constructs mosaics of the sky on demand. Montage is used to reproject, background match, and finally mosaic many image plates into a single image. Montage has been used to mosaic image plates from synoptic sky surveys, such as 2MASS in the infrared wavelengths. Montage is being developed by a team that includes Caltech IPAC, Caltech CACR, and JPL. To generate and map applications such as Montage onto grid resources, technologies such as Chimera and Pegasus can be used. Chimera allows the user to construct an abstract workflow representing the application in a resource independent way. Pegasus is a workflow management system designed to map abstract workflows onto the Grid resources. Pegasus consults various Grid information services to determine the !
available resources and data. Given all the information, Pegasus generates an executable workflow that identifies the resources where the computation will take place, the data movement for staging data in and out of the computation, and registers the newly derived data products. To date Pegasus has enabled Montage to run on a variety of distributed resources such as individual machines, condor pools and TeraGrid sites.

O6.2 *Features of the AstroGrid approach to Virtual Observatory architecture*

Tony Linde, University of Leicester , Andrew Lawrence, University of Edinburgh , Nicholas Walton, University of Cambridge

AstroGrid is a UK e-Science project. Its aim is to develop a complete Virtual Observatory infrastructure and facilitate its deployment at key astronomical data centres throughout the UK. The project began in 2001 and is now nearing the end of its first phase, with a complete (V1.0) release of all software components planned for Dec 2004. A second phase has been approved and funded and will run to end 2007.

This presentation will describe the key conceptual features of the AstroGrid approach to VObs architecture, including:

service-oriented architecture
browser-based access to all services
compilation of astronomical tasks into complex workflows
access to distributed file storage (MySpace)
configurable approach to dataset access and application execution
full management of users and groups and their access authorities

We will show how the AstroGrid suite of services fully incorporate current and pending recommendations from the standards working groups of the International Virtual Observatory Alliance (IVOA), including:

common web service interface
both 'local' and 'full' registries for resource discovery
data exchange using the VOTable standard

We will describe the advanced features of AstroGrid which allow developers to:

make complex registry queries using XQuery
read and write directly to MySpace
deploy existing command line tools

In conclusion, we will describe developments planned for AstroGrid-2, including:

client-side "astronomer's workbench"
command-line interface and common scripting language
semantically-rich resource discovery
advanced data mining and visualisation services
full rollout across the UK

O6.3 *Deploying the AstroGrid: Science Use Ready*

Nicholas A Walton, Institute of Astronomy, University of Cambridge, UK , Paul Harrison, Jodrell Bank Observatory, Victoria University of Manchester, UK , Martin Hill, Institute for Astronomy, University of Edinburgh, UK , Anita Richards, Jodrell Bank Observatory, Victoria University of Manchester, UK

AstroGrid (see http://www.astrogrid.org), a UK eScience project with collaborating groups drawn from the major UK data archive centres, is creating the UK's first virtual observatory. We describe the current state of the AstroGrid testbed deployment system. Related presentations at this meeting cover AstroGrid and the international context and technical details of the software components of AstroGrid.

In order to ensure rapid user feedback AstroGrid is fully deploying each iteration release, connected to relevant data and application products, in such a fashion as to allow scientific use of that release. The early users are primarily the science advisory group and now the beta tester communities of the project.

The scientific functionality of the current deployment is highlighted. This includes access to a sophisticated workflow capability. Its use in allowing image extraction from multiple image datasets, input photometry file creation, redshift determination, and visualisation for the outputs to allow discovery of high redshift objects is described. Further examples show the use of the latest Astrophysical Data Query language standard and how it is being used to enable large data queries of remote databases

in searching for low mass objects in the Pleiades with results returned to a virtual 'MySpace' user storage area, where further visualisation and processing can be performed.

We note how the deployed system is being tested by the science community, and how that comment and feedback is vital in informing the project as to future releases including the December 2004 'Release 1'.

The latest AstroGrid deployment release is available from http://www.astrogrid.org/release .

O6.4 *Reliable, Automatic Transfer and Processing of Large Scale Astronomy Datasets*

Tevfik Kosar, University of Wisconsin-Madison , George Kola, University of Wisconsin-Madison , Robert Brunner, NCSA , Miron Livny, University of Wisconsin-Madison , Michael Remijan, NCSA

Astronomers are increasingly obtaining larger datasets, particularly in the optical and near-infrared. Unfortunately, the technologies to process large amounts of image data and share the data and the results with collaborators spread around the globe, have not kept pace with the data flow. In the past, this type of software has also required significant human involvement to deal with failures. We have designed and implemented a fault-tolerant system that can process large amounts of astronomy images using idle CPUs on desktops, commodity clusters and grid resources. It reliably replicates data and results to collaborating sites and performs on-the-fly optimization to improve throughput. It is highly resilient to failures and can recover automatically from network, storage server, software and hardware failures. To demonstrate the capabilities of this framework, we have successfully processed three terabytes of DPOSS images using idle grid resources spread across three organizations. This system is freely available to others groups who wish to significantly decrease the time and effort required to perform large scale data transfer and processing.

O6.5 *A Framework for Parallel Data Analysis on a Distributed Grid*

Jeffrey P. Gardner, Pittsburgh Supercomputing Center , Andrew , University of Pittsburgh

Virtual observatories will give astronomers easy access to an unprecedented amount of data. In many cases, mining these data will require the power of parallel computers. These machines may be a small Beowulf cluster, a large massively parallel platform, or a collection of parallel machines distributed across a Grid.

Harnessing the power of these machines can be difficult, since parallel programs often demand a great deal of time and expertise to develop. However, nearly all analysis of large astronomical datasets use a limited number of data structures: trees and grids. Our framework will provide a flexible, extensible, and easy-to-use way of using these data structures for data analysis on parallel and grid-distributed machines. By minimizing development time for these platforms, we will enable wide-scale knowledge discovery on massive datasets.

Session 7: The Virtual Observatory: Mining and Analysis

O7.1 *Registries and Publishing in the Virtual Observatory*

Raymond Plante, NCSA

A registry, as a repository containing descriptions of available data and services, is where data discovery and data publishing first meet in the Virtual Observatory (VO).In the last year, the IVOA framework for registries has advanced from a working prototype system to a comprehensive set of standard interfaces. I review the emerging standards for registries and the metadata they contain and discuss how registries around the world can work together to create an up-to-date picture of available VO resources. I illustrate how the standards--in particular, the metadata standards--allow the registries to be used in a variety of ways. I highlight some of the current challenges being examined. One such issue is that of registry curation: ensuring sufficient accuracy, consistency, and completeness to make the descriptions useable. Finally, I focus on the NVO approach to registries and how it fits into the larger process of publishing in the VO.

O7.2 *OpenSkyQuery & OpenSkyNode - the VO Framework to Federate Astronomy Archives*

William O'Mullane, JHU , Tamas Budavari, JHU , Vivek Haridas, JHU , Nolan Li, JHU , Tannu Malik, JHU , Masatoshi Oishi, NAOJ , Alexander Szalay, JHU , Aniruddha Thakar,JHU, Ramon Williamson,NCSA

OpenSkyNode and ADQL are the major new steps in the Data Access layer of the Virtual observatory. Presented as a poster last year OpenSkyQuery(OSQ) now has a functioning portal which allows cross matches between catalogs on thirteen nodes. The portal also allows a user to upload

their own list of source to be cross matched with the registered SkyNodes. This is one of the first systems utilizing of the IVOA's nascent standard Astronomical Data Query Language(ADQL). We shall demonstrate the functioning of the portal as it communicates with the nodes and the NVO searchable registry. We shall present the OSQ architecture and explain its use of WebServices as an open standard, Java and C# implementations of OpenSkyNode demonstrates that the open standard does indeed work. We shall also discuss the newest ADQL0.8 specification which will allow ADQL to be used to query registries in the VO indeed the NVO searchable registry is already accepting ADQL0.7.4. and implements OpenSkyNode0.7.4.

O7.3 ***VO-enabling a major astronomical analysis and reduction software system***

David Giaretta, Starlink/RAL , Malcolm Currie, Starlink/RAL , Stephen Currie, Starlink/RAL , Mark Taylor, Starlink/U of Bristol , Norman Gray, Starlink/U of Glasgow , Peter Draper, Starlink/U of Durham , David Berry, Starlink/U C lancs , Alasdair Allan, Starlink/U of Exeter

This talk will describe the Starlink experience in its process of VO-enabling its existing comprehensive collection of astronomical analysis and reduction software packages. Starlink has been involved from the early stages of the VO in the UK and internationally. It has been working actively in the IVOA process, for example the contributing to the VOTable standard development as well as creating the most complete VOTable library, as part of the STIL.
While not tracking every twist and turn, Starlink has incorporated, in pilot studies, several of the favoured Grid techniques of the day. Each time the aim has been to explore options for making minimum modifications while Grid/VO enabling the bulk of the software.

As the GRID/VO standards (appear) to be stabilising and providing greater functionality, VO-enabled Starlink applications are being released, having adequate functionality, robustness, level of documentation and quality.

This talk will describe some of the design choices we have made, the VO capabilities we are incorporating and contributing and the lessons learned.

O7.4 *An $\mathcal{O}$(N \log M) Algorithm for Catalogue Matching*

Drew Devereux , David J. Abel , Robert A. Power, Peter R. Lamb

A basic problem in data fusion in Virtual Observatories is that of how to determine which source records from different catalogues actually refer to the same source, on the basis of spatial co-location. This is the {\it catalogue matching} problem, and it is inherently costly to solve. Our algorithm applies filter-refine and plane sweep techniques. Pre-processing consists of a sort by declination, and the active list is a queue indexed by a binary tree. The algorithm is $\mbox{\cal O}(N \log M)$ in both I/O and processor costs, with only moderate memory requirements. Empirical assessment on catalogues of up to a billion records suggests that the algorithm performs at least an order of magnitude better than the techniques in current use.

Session 8: Ground-based Observatories II

O8.1 *Simulating the Performance of the Square Kilometer Array*

Colin Lonsdale, MIT Haystack , Sheperd Doeleman, MIT Haystack

The Square Kilometer Array (SKA) seeks performance levels that far exceed those of current instruments. Such performance levels require an understanding of, and new algorithms to correct for, a variety of effects that are generally ignored for current instruments. In the near term, the only way to generate data with which to systematically attack these problems in a controlled environment is through simulation. The MIT Array Performance Simulator (MAPS) has been constructed with the ability to accurately incorporate a wide variety of effects as a design requirement. Multiple such effects are already implemented, and MAPS is substantially more capable in this regard than other available simulation packages. The package is currently in routine use, and is available to the SKA community. In this talk, the architecture and capabilities of the simulation package will be described, and plans for future enhancements will be reviewed.

O8.2 *The Gemini Science Archive: Current Status and Future Prospects*

Colin Aspin, Gemini Observatory, David Bohlender, CADC, HIA, Victoria

We will present details of the design, development and operation of the Gemini Science Archive, the GSA. The full basic archive was released to the public in July 2004 and has many innovative features. We will additional discuss the Phase III development of the GSA including many advanced capabilities which will allow the GSA to publish extensive VO-enabled data products.

O8.3 *Simulation of the Future LSST Data Pipeline*

Ghaleb Abdulla, David Liu, Jim Garlick, Sergei Nikolaev, Marcus Miller, Michael Franklin, Kem Cook, Jim Brase ,

In this paper we describe our approach to build a pipeline simulator for the future Large –scale Synoptic Survey Telescope (LSST). The simulated pipeline will be used to research and evaluate software architectures that are efficient and flexible. It will also be used to define the real-time software and hardware requirements needed to support the anticipated LSST data rates. The LSST data pipeline requirements are still being defined, however, previous surveys can provide a good source for data requirements. Our approach is to use the SuperMacho data pipeline as a prototyping tool to identify a framework for building Modular Data-Centric Pipeline (MDCP) architectures. The prototyping is done in a hierarchical fashion to help capture and define the general data attributes (schema) first. We then model other necessary components based on science and performance requirements. We use identified schemas or data attributes as a way to define a data model for LSST. !

O8.4 *Linking and tagging initiatives at the Astrophysical Journal*

Greg Schwarz, AAS Journals Staff Scientist

At the start of 2005 the Astrophysical Journal will be offering three new linking and tagging options for the electronic edition. The linking projects are in conjunction with partnerships with ADS, CDS, and NED.

Authors may use new AASTeX mark up to tag data sets of participating data centers. A common name resolver hosted at ADS verifies the data set tag during copy editing and forwards the reader who follows the links to the appropriate data center. A demonstration of this

capability can be seen in the September Spitzer ApJ Supplement. In a similar vein authors we will support object linking to SIMBAD and NED. For this project we will also query authors to supply the most important objects of their paper during peer review. These object lists will be supplied to the data centers to help them integrate the information into their databases and potentially check for errors. The SIMBAD and NED object web pages that are now created weeks and months after publication at CDS will be also be dynamically added to the navigation bar once they are made available. The final endeavor is a controlled set of major astronomical facility keywords to help organizations track the effectiveness of their telescopes. The ultimate goal of each project is to allow authors make their papers more useful and to help researchers seamlessly navigate between the journal and the data archives.

O8.5 *Development of SAOImage DS9: Lessons learned from a small but successful software project*

William Joye, Smithsonian Astrophysical Observatory , Eric Mandel, Smithsonian Astrophysical Observatory

We will present a discussion of lessons learned during the development of SAOImage DS9. Starting with general observations on scientific software development, we will discuss our design and implementation cycle, allocation of effort and resources, keys to our success, overall strategies that have worked well (and those that have not), and future issues and challenges. We hope our experience will be of use to other small software development projects.

Session 9: Detection Algorithms

O9.1 *SExtractor, ten years after*

Emmanuel Bertin, IAP and Obs. Paris

SExtractor has become a popular software tool amongst astronomers for quickly extracting catalogs from deep-sky images. In the ten years since SExtractor was released both the observational context and available computer hardware have changed dramatically, but the core engine of SExtractor has remained essentially unchanged.

I will review lessons learned in the field of automated source extraction and outline the new directions taken for the development of version 3.0 as a component of the TERAPIX pipeline software suite.

O9.2 *Fast Algorithms for Massive-Scale Classification Problems: Toward 1 Million Quasars*

Alexander Gray, Carnegie Mellon University , Gordon Richards, Princeton University , Robert C. Nichol, University of Portsmouth , Robert Brunner, NCSA/University of Illinois Urbana-Champagne

Quasar detection from SDSS data is an example of a classification problem which must simultaneously achieve as high accuracy as possible while somehow retaining computational tractability in the face of datasets with up to millions of objects in three or more measurements. We present a new algorithm for the highly-accurate Nonparametric Bayes Classifier (the Bayes-optimal classifier for arbitrary distributions), which makes it computable for such massive datasets. We have used it to achieve the largest quasar catalog to date, and believe this fundamental statistical/computational technology can enable similar large-scale astronomical data mining projects.

O9.3 *Detection of Rare Objects in Massive Astrophysical Datasets Using Innovative Knowledge Discovery Technology*

Alvaro Soto, Pontificia Universidad Catolica de Chile , Antonio Cansado, Pontificia Universidad Catolica de Chile , Felipe Zavala, Pontificia Universidad Catolica de Chile

This work presents new knowledge discovery in databases (KDD) technology that aids the scientist in the detection of rare types of astronomical objects. Our main idea is that while computers have the power to search huge amounts of data, the expert has the domain knowledge to efficiently lead this search. Our system builds upon two main components: a probabilistic model able to scale to large datasets and a set of modules to interact with the scientist.

The probabilistic model represents the joint uncertainty of objects attributes in an astrophysical catalog. The model consists of a combination of Bayesian networks, and Gaussian mixture models, trained through an accelerated version of the expectation maximization algorithm (EM). The model provides the system with three main capabilities: detection of rare objects, through the identification of points with low probability, explanation of the sources of the anomalies, through the identification of the unusual attributes of an object, and scalability to large datasets, through the use of efficient data structures and algorithms.

The modules to interact with the scientist incorporate innovative active learning techniques that use the expert feedback to progressively improve the performance of the probabilistic model. Using a modified version of the Distance-Weighted K-nearest neighbor algorithm, the system performs an interactive exploration of the space of potential rare objects. Areas where the system receives a positive feedback from the scientist, meaning that a detected object is indeed a rare object, are fully explored.

Our system is currently being tested using data from the release 1 of the Sloan Digital Sky Survey.

Session 10: More Algorithms

O10.1 *Model-Based Count-Limited Image Restoration*

David A van Dyk, University of California, Irvine
The California-Harvard AstroStatistics Collaboration

In recent years, there has been a quantum leap in the quality and quantity of new data in observational high-energy astrophysics. Recently launched telescopes allow for very high resolution imaging, spectral analysis, and time series analysis. The complexity of these instruments, the complexity of the astronomical sources, and the complexity of the scientific questions leads to a subtle inference problem that requires sophisticated statistical tools. In this talk we discuss the use of highly structured statistical models that are designed to capture the complexity of both the data collection and the cosmic sources. To fit these models, we use Bayesian statistical methods that are well suited to answer relevant scientific questions. Bayesian techniques allow us to combine information in the data with scientific data outside the data such as smoothness constraints. Thus, we combine a Poisson likelihood that accounts for the low-count nature of the data with a multi-scale prior that encourages smooth reconstructions. Although such prior are generally formulated in terms of a number of user specified tuning parameters, we show that these parameters can also be fit to the data. The statistical computation that is necessary to fit such highly structured models can be formidable. We propose both EM algorithms for mode finding and MCMC methods to fully explore the posterior distribution. Because the computational efficiency of such methods can be highly sensitive to the choice of Markov chain, they must be designed and implemented with care. Thus, we discuss both our choice of implementation and convergence criterion.

O10.2 *SAADA: Astronomical Databases Made Easy*

L. Michel, H. Nguyen Ngoc, C. Motch

Many astronomers wish to share datasets with their community but have not enough manpower to develop databases having the functionalities required for high-level scientific applications. The SAADA*) project aims at automatizing the creation and deployment process of such databases. A generic but scientifically relevant data model has been designed which allows to build databases by only providing a limited number of product mapping rules. Databases created by SAADA rely on a relational database supporting JDBC and covered by a Java layer including a lot of generated code.

Such databases can simultaneously host spectra, images, source lists and plots. Data are grouped in user defined collections whose content can be seen as one unique set per data type even if their formats differ. Data sets can be correlated one with each other using qualified links. These links help, for example, to handle the nature of a cross-identification (e.g., a distance or a likelihood) or to describe their scientific content (e.g., by associating a spectrum to a catalogue entry). The SAADA query engine is based on a language well suited to the data model which can handle constraints on linked data, in addition to classical astronomical queries. These constraints can be applied on the linked objects (number, class and attributes) and/or on the link qualifier values. Databases created by SAADA are accessed through a rich WEB interface or a Java API. We are currently developing an inter-operability module implanting VO protocols.

*) French acronym for Auto-configurable Archiving System for Astronomical Databases.

O10.3 *Mosaicking with MOPEX*

David Makovoz, Caltech , Iffat Khan, Caltech

We present MOPEX - a software package for image mosaicking and point source extraction. MOPEX has been developed for the SPITZER Space Telescope. This presentation concentrates on the mosaicking aspects of the package. MOPEX features the use of several interpolation techniques, coaddition schemes, and robust and flexible outlier detection based on spatial and temporal filtering.

A number of original algorithms have been designed and implemented in MOPEX. Among them is direct plane-to-plane coordinate transformation, which allows at least an order of magnitude speed up in performing coordinate transformation by bypassing the sky coordinates. The dual outlier detection makes possible outlier detection in the areas of even minimal redundancy. Image segmentation based on adaptive thresholding is used for object detection, which is part of outlier as well as point source detection. Efficient use of computer memory allows mosaicking of data sets of very deep coverage of thousands of images per pointing, as well as areas of sky covering many square degrees. Although designed for SPITZER data, MOPEX does not require any SPITZER-specific fits header keywords to run , and can be applied to other data, that have standard header information on the image geometry and pointing.

The package is available for distribution at
http://ssc.spitzer.caltech.edu/postbcd/ .

O10.4 *W projection: a new algorithm for wide field imaging with radio synthesis arrays*

Tim Cornwell, NRAO , Kumar Golap, NRAO , Sanjay Bhatnagar, NRAO

Wide field imaging with low frequency synthesis arrays is limited by a number of troublesome effects. First amongst

these is the "non-coplanar baselines" distortion whereby the integral relationship between sky brightness and measured visibility function is not a simple Fourier transform. A piece wise approximation to the integrals can be used and forms the basis of the facet approaches used for the last 15 years. These approaches are difficult to program and perform relatively poorly. We have developed a novel, high performance algorithm based upon convolution of the visibility samples with a Fresnel kernel. We interpret the Fresnel kernel as being required to propagate the electric field to a common reference plane. The role of Fresnel diffraction in radio inteferometry seems to have been unrecognized previously.

O10.5 ***Hardware Acceleration for Astronomical Data Analysis***

Eric Sessoms, NRAO

The floating-point performance of commercial Graphics Processing Units (GPUs) currently exceeds that of more costly CPUs by a factor of 3 and the parallel architecture of GPUs is so well suited to many physical problems that researchers have reported practical performance gains of 6- to 10-times with some algorithms. As GPU performance is on an exponential growth curve doubling every 6-months, instead of doubling every 18-months as main line CPUs, GPUs are an increasingly cost-effective computational resource.

We present an overview of GPU architecture and discuss how that architecture makes these performance gains possible. We then explore the programming model for GPUs in greater detail, introduce programming frameworks that have been developed to take advantage of GPU performance, and identify a number of data processing packages already available for GPUs.

Posters

Monday Poster Session (P1.x.x)

P1.1.1 ESA-VO: European Space Based Astronomy through the VO

Christophe ARVISET, ESA / ESAC, Isa BARBARISI, ESA / ESAC, John DOWSON, ESA / ESAC, Jose HERNANDEZ, ESA / ESAC, Iñaki ORTIZ, ESA / ESAC, Pedro OSUNA, ESA / ESAC, Jesús SALGADO, ESA / ESAC, Guillermo SAN MIGUEL, ESA / ESAC, Aurèle VENET, ESA / ESAC

In June 2004, ESA decided to rename the VILSPA centre into ESAC for European Space Astronomy Centre. All astronomy missions (ISO, XMM-Newton, Integral, Herschel, Planck, Gaia, etc..) will be located at ESAC. ESAC will also host most of the ESA astronomy and planetary Scientific Archives, becoming the main space science data provider in Europe.

ESAC Archive Group had already been involved in some VO activities. The ISO Data Archive and the XMM-Newton Science Archive have already been VO-enabled via the SIAP (Simple Image Access Protocol). Furthermore, the IDA has been the first archive to be SSA (Simple Spectra Access) compliant. All these features were demonstrated at the AVO first science demonstration in January 2004.

Thanks to these successes, ESA has decided to have its own ESA-VO project to become the European Space based astronomy VO actor. Dedicated manpower and computer resources have been allocated to the ESA-VO project to participate more actively in all IVOA initiatives.

ESA-VO main axes of development are the following:

- Keep on making all ESA astronomy archives VO compliant (SIAP, SSA, VOQL. …)
- Build an ESA-VO Portal including a multi-missions search interface based on VO protocols

- Inter-operability from the existing ESAC archives to remote archives via VO protocols
- Development of specific VO applications such as VOSpec, a tool for retrieval and overlay display of Spectra from SSAP VO-enabled archives, including various unit conversions as per metadata specified in the SSA
- Building a ESA VO registry
- Building a ESA VO Grid

P1.1.2 Data Mining in Distributed Databases for Interacting Galaxies

Kirk Borne, George Mason University

We will present results from an exploratory data mining project to identify classification features of special classes of interacting galaxies (for example, infrared-luminous galaxies) within distributed astronomical databases. Using a variety of data mining techniques, interaction-specific features are learned, to distinguish this class of galaxies from a control sample of normal galaxies. Subsequently, the corresponding rule-based feature model of that class of galaxies is then applied to the large multi-wavelength astronomical databases that are now becoming available. This distributed data mining activity is a prototype science use case for the Virtual Observatory. We specifically apply multi-archive multi-wavelength data to the problem. We will present examples of both successful and unsuccessful data mining attempts.

P1.1.3 Integrating Legacy Code Into Virtual Observatories - A Test Case

Joerg M. Colberg, University of Pittsburgh, Ryan Scranton, University of Pittsburgh, Andrew J. Connolly, University of Pittsburgh

As virtual observatories are being built, the integration of existing astronomical analysis codes is becoming increasingly important. This is particularly true for standard applications that are widely and often used. In these cases, creating what amounts to on-site data-mining facilities is clearly beneficial for a wide variety of

users, including, but not limited to, users with very little current knowledge of VO technologies or users interested in using the machinery for educational purposes. Given the nature of most legacy codes, there is a number of problems that have to solved. Here, we present an application that makes a very fast and efficient n-point correlation function code interact with the Sloan Digital Sky Survey. We discuss the basic ideas and setup, technical problems encountered, and possible extensions.

P1.1.4 Data Mining in the Era of Virtual Observatories

Andrew Connolly, University of Pittsburgh, Joerg Colberg, University of Pittsburgh, Simon Krughoff, University of Pittsburgh, Ryan Scranton, University of Pittsburgh, Andrew Moore, Carnegie Mellon University

With the development of the Virtual Observatories astronomers now have access to massive, distributed data sets covering many decades of the electromagnetic spectrum. To fully realize the scientific potential of the VO we require analysis tools that can scale to the size and dimensionality of these collections of data and that are interfaced directly onto the Virtual Observatories (i.e. to make the analysis of VO data as seamless and transparent as our access to the data themselves).

In this talk we present the initial implementation of a toolbox for data mining in the era of the Virtual Observatory. We will discuss a range of tools designed to analyze the large, multidimensional data sets that are currently available through the VO. We will specifically address analyses that cover a broad range of questions that astronomers might want to ask of the VO data: how to measure the clustering of galaxies from large photometric and spectroscopic surveys using a single and parallel computing environment, how to identify anomalous or unusual sources in large multidimensional data sets, how to track moving sources in time domain data and how to integrate the multi-frequency data available through the VO with external data sets. We will discuss how the algorithms for addressing these questions are designed and implemented and how these tools are integrated into a webservice environment together with the associated challenges for

undertaking statistical astronomy on large heterogeneous data sets.

P1.1.5 GALEX Services for the Virtual Observatory

Alberto Conti, Space Telescope Science Institute,
Bernie Shiao, Space Telescope Science Institute,
Myron Smith, Space Telescope Science Institute

The Galaxy Evolution Explorer (GALEX) will observe galaxies in the ultraviolet across 10 billion years of cosmic history. The GALEX mission has the primary science goals if investigating how star formation in galaxies evolved from the early Universe up to the present. The Multi mission Archive at Space Telescope (MAST) will deliver GALEX data (~5TB) to the entire astronomical community and to the general public. GALEX data products include near-UV and far-UV medium, deep and all-sky surveys in imaging mode and partial sky surveys in spectroscopic mode.

The GALEX archive at MAST is built around a core set of web applications and xml web services that mine GALEX data stored in a Microsoft SQL Server database. Interoperability with other data centers is achieved by offering many VO compliant services. Together with other MAST surveys such as GOODS, UDF and HDFN/S, the GALEX archive is an integral part of OpenSkyQuery: the VO SkyNode portal.

Herein we provide examples of search and retrieval capabilities for the first release of observations (DR1). We also show the GALEX Multi-mission Archive Cross-Correlation tool (GMAX): a proof of concept application specifically developed to utilize the large collection of services available under VO. GMAX uses the VO Registry to locate ConeSearch and SIAP VO services from surveys that overlap the GALEX footprint. The GALEX image cutout service uses the results of the ConeSearch to display objects from these surveys onto a GALEX image. Cross-correlation with GALEX data is achieved using the Hierarchical Triangular Mesh (HTM) index or the VO SkyNode service.

P1.1.6 VOTFilter, Bridging VO Services to Industrial Desktop Application

Chenzhou CUI, National Astronomical Observatory of China, Kui WU, Beijing Institute of Astronautical System Engineering

VOTable is an XML format defined for the exchange of tabular data in the context of the Virtual Observatory. It is the first Proposed Recommendation defined by International Virtual Observatory Alliance, and has obtained widely support from both VO communities and many astrophysics projects. Taking advantage of XML file format of OpenOffice, and XML filter VOTFilter is developed by the Chinese Virtual Observatory project. Using the VOTFilter, one can read and write files in VOTable format, edit and analyze astronomical data using OpenOffice Calc. In the paper, we will describe the importance of the toolkit, introduce its usage and discuss possible extension for the current version. Some technical problems met during its development will be also discussed.

P1.1.7 MySpace: distributed data storage for the VO

A. C. Davenhall, D. Morris, K. T. Noddle, N. A. Walton

MySpace is a component of AstroGrid (see http://www.astrogrid.org and related presentations at this conference), the UK developed infrastructure for Virtual Observatories . MySpace is a co-operating network of web services providing both temporary and long-term storage for VO users. It can be used to hold various types of data items including: data files, documents, database tables, work-flow segments, etc.

Although individual data items are geographically dispersed, users access and navigate the work space as though all the items were stored in a single location. In other words, the design hides from the user the fact that the work space is not entirely local.

MySpace is a fully integrated component of the AstroGrid system, but its modular nature means that it can be installed and used in isolation or, in principle, in conjunction with components from other IVOA compliant VO projects.

We show the architecture of the MySpace component, it's interaction with other Virtual Observatory components and the benefits to users of the system. The current state of development and deployment is examined and plans for future expansion considered.

P1.1.8 The Multimission Archive at the Space Telescope Science Institute in the context of VO activities

Inga Kamp, STScI, Randall Thompson, STScI, Alberto Conti, STScI, Dorothy Fraquelli, STScI, Tim Kimball, STScI, Karen Levay, STScI, Bernie Shiao, STScI, Myron Smith, STScI, Rachel Somerville, STScI, Richard L. White, STScI

In the past year, the Multimission Archive at the Space Telescope Science Institute (MAST) has taken major steps in making MAST's holdings available using VO-defined protocols and standards, and in implementing VO-based tools. For example, MAST has implemented the Simple Cone Search protocol, and all MAST mission searches may be returned in the VOTable format, allowing other archives to use MAST data for their VO applications. We have made many of our popular High Level Science Products available through Simple Image Access Protocol (SIAP), and are implementing the VO Simple Spectral Access Protocol (SSAP). The cross correlation of VizieR catalogs with MAST missions is now possible, and illustrates the integration of VO services into MAST. The user can easily display the results from searches within MAST using the plotting tool VOPlot.

P1.1.9 ROME (Request Object Management Environment)

Mih-seh Kong, Infrared Processing and Analysis Center, Caltech, G. Bruce Berriman, Infrared Processing and Analysis Center, Caltech, John C. Good, Infrared Processing and Analysis Center, Caltech

Most current astronomical archive services are based on an HTML/CGI architecture where users submit HTML forms via a browser and CGI programs operating under a web server process the requests. Most services return an HTML result page with URL links to the result files or, for longer jobs, return a message indicating that email will be sent when the job is done. This paradigm has serious limitations when applied to time-intensive jobs, in that it provides no mechanism for managing, queuing, and controlling multiple jobs.

ROME is a collection of middleware components being developed under the National Virtual Observatory Project to provide mechanism for managing long jobs such as computationally intensive statistical analysis requests or the generation of large scale mosaic images. Written as EJB objects within the open-source JBOSS applications server, ROME receives processing requests via a servelet interface, stores them in a DBMS using JDBC, distributes the processing (via queuing mechanisms) across multiple machines and environments (including GRID resources), manages real-time messages from the processing modules, and ensures proper user notification.

The request processing modules are identical in structure to standard CGI-programs -- though they can optionally implement status messaging -- and can be written in any language. ROME will persist these jobs across failures of processing modules, network outages, and even downtime of ROME and the DBMS, restarting them as necessary.

P1.1.10 A Source Extraction Web Service with Cross Matching Capability

K. Simon Krughoff, U. of Pittsburgh, Andrew J. Connolly, U. of Pittsburgh, Joerg M. Colberg, U. of Pittsburgh

Web services provide an excellent means for users to take advantage of algorithms and computing time on machines that are physically displaced from their location. By using XML, specifically SOAP, communication between client and server it is possible to build sets of tools that are essentially platform and language independent. This allows for a large set of utilities to be constructed for use by the entire astronomical community via simple client programs.

We present a case study in which we have deployed the popular source extraction code ``SExtractor" in a web service. Our web service provides the additional option of cross matching the resultant source catalog with other catalogs (i.e. SDSS, 2MASS, FIRST, HDF, etc.) through the SkyPortal web service provided by OpenSkyQuery. This is an important step in the development of web services as we begin to use pre-existing web services to build increasing utility into the set of tools available through web services.

P1.1.11 VO Software: Overview of the past twelve months of Prototypes & Tools

Marco Leoni, ESO, Thomas McGlynn, NASA HEASARC, Roy Williams, Caltech

Thanks to the increased effort in the development and spread of the "VO novel", the past year saw a spur in the creation of software for the Virtual Observatory: including both prototypes and released tools, produced by the national VO project members of the International Virtual Observatory Alliance (IVOA).

This work describes completely new developments as well as existing software adapted to the VO. It highlights their main features and the way they interact and interoperate using common industry standards as well as astronomy-specific ones developed within the IVOA.

P1.1.12 XML data in the Virtual Observatory

Bob Mann, U. Edinburgh, Rob Baxter, U. Edinburgh, Peter Buneman, U. Edinburgh, Robert Carroll, U. Edinburgh, Byron Choi, U. Edinburgh, Rob Hutchison, U. Edinburgh, Ted Wen, U. Edinburgh

XML is the lingua franca of the Web services world and so will play a major role in the construction of the Virtual Observatory. Its great advantages are its flexibility, platform-independence, ease of transformation and the wide variety of existing software that can process it. An obvious disadvantage in its use as an astronomical data format is its verbosity; the number of bytes taken up writing the XML tags can easily outnumber those constituting the actual astronomical data. This becomes prohibitively inefficient when large amounts of data are stored in XML, and the developers of VOTable sought to circumvent this, by allowing for the use of binary data, either in the VOTable document itself or in an external file linked from it. The verbosity of XML in this regard is a problem in many other disciplines, and computer scientists are developing more generic solutions to that found in the VOTable specification. In this paper we describe several of these projects currently underway in Edinburgh, which focus on the compression and querying of XML, and a technology for representing the structure of a binary file in XML, enabling it to be read as if it were XML.

P1.1.13 Distributed Data Mining for Astrophysical Data

Sabine McConnell, David Skillicorn

We present a data mining technique in which the datasets are partitioned by attributes such that datasets contain subsets of information for all available objects . Our approach is evaluated on a variety of astrophysical datasets and is suitable for a distributed setting, where

the introduction of real time data analysis imposes an upper bound on the size of the data that can be processed. Using a decision tree technique, we show that this ensemble technique can achieve the same or better classification accuracy compared to an approach where the complete dataset is contained in a single location. At the same time, this substantially increases the amount of data that can be processed.

P1.1.14 Spectral Energy Distribution Data Model for the VO

Jonathan McDowell, SAO, Doug Tody, NRAO, Tamas Budavari, JHU

We present formalism for describing and handling spectral energy distributions in the Virtual Observatory, and discuss some of the technical problems in combining photometric and spectrophotometric information from heterogeneous archives. We present an abstract data model for SEDs compatible with the Simple Spectral Access Protocol being developed by the VirtualObservatory and present example XML, VOTABLE and FITS serializations. JCM acknowledges support from NSF grant no. AST 0121296 and Cooperative Agreement AST 0122449, as well as the Chandra X-ray Center under NASA contract NAS8-39073.

P1.1.15 Creating Data that Never Die: Building a Spectrograph Data Pipeline in the Virtual Observatory Era

D.J. Mink, SAO, W.F. Wyatt, SAO, J.B. Roll, SAO, S.P. Tokarz, SAO, M.A. Conroy, SAO, N. Caldwell, SAO, M.J. Kurtz, SAO, M.J. Geller, SAO

Data pipelines for modern complex astronomical instruments do not begin when the data is taken and end when it is delivered to the user. Information must flow between the observatory and the observer from the time a project is conceived and between the observatory and the world well past the time when the original observers have extracted all the information they want from the data. For the 300-fiber Hectospec low dispersion spectrograph on the MMT, the

SAO Telescope Data Center is constructing a data pipeline which provides assistance from preparing and submitting observing proposals through observation, reduction, and analysis to publication and an afterlife in the Virtual Observatory. We will describe our semi-automatic pipeline and how it has evolved over the first nine months of operation.

P1.1.16 MaxBCG: The Importance of Database Systems for the Virtual Observatory and the Grid

María A. Nieto-Santisteban, Johns Hopkins University, Alexander S. Szalay, Johns Hopkins University, Jim Gray, Microsoft Research, Aniruddha R. Thakar, Johns Hopkins University, William J. O'Mullane, Johns Hopkins University, James Annis, Fermi National Accelerator Laboratory

The Maximum-likelihood Brightest Cluster Galaxy (MaxBCG) application searches for galaxy clusters that span a large dynamic range of cluster masses and provides good redshift estimates using the cluster red sequence galaxies. MaxBCG was developed originally in tcl using the SDSS Astrotools package and ran on the Terabyte Analysis Machine (TAM), a 10-CPU cluster specially tuned to solve this type of problem. The same application code was grid-enabled and used to test the Chimera Virtual Data System created by the GriPhyN project. As is common in astronomical file-based applications, the TAM and Chimera implementations used hundreds of thousands of files for the computations, files served from the SDSS Data Archive Server. The data required to run the MaxBCG is also available in the SkyServer database. At Johns Hopkins University, we have implemented a version using SQL on a SQL Server 2000 cluster that processes the same sky area an order of magnitude faster than TA!M. We are also working on an efficient grid-enabled system that does not require moving large volumes of data across the network. When the user submits the application, upon authentication and authorization, the SQL code is deployed and executed on the available data-grid nodes hosting the SkyServer database system.

In this paper we describe why and how the Grid and the Virtual Observatory can take advantage of Database Management Systems.

P1.1.17 VWhere: An Extensible Filtering and Data Exploration Guilet

Michael S. Noble, Massachusetts Institute of Technology

In this paper we describe VWhere, an S-Lang/Gtk guilet which combines graphical ease of use with powerful data exploration and filtering. Analysis in VWhere amounts to manipulating sets of region filters applied to one or more plots. Input vectors originate from either files or in-memory arrays, with plots and regions created via mouse. A text window is also provided, in which additional vectors may be fabricated from arbitrary mathematical expressions. A list of points passing all filters is output, which can be applied to other datasets, create output files, etcetera.

In contrast with the "file in/file out" model of command line tools, in which an entire set of filters is conceptually applied to an entire dataset in one pass, arbitrary cuts of axes may be visualized in VWhere, then applied while exploring others. This incremental filtering helps discern non-obvious relationships, and is considerably faster and more fluid for iterative analysis than file-based alternatives (multiple file reads are avoided, no intermediate file litter is created, no syntax to learn/forget, results are immediately visualized), particularly as datasets scale up in size.

This approach also far surpasses file-oriented tools in expressive power and extensibility, since external S-Lang modules and scripts may be loaded at runtime, giving VWhere the ability to call virtually any S-Lang or C/C++/FORTRAN function (e.g. GNU Scientific Library wrappers). Moreover, since plots and regions may also be customized and printed, again with point/click ease, VWhere can also serve as a reasonably capable, general purpose plotting GUI.

P1.1.18 New NED XML/VOtable Services and Client Interface Applications

Olga Pevunova, IPAC/Caltech/JPL, John Good, IPAC/Caltech/JPL, Joseph Mazzarella, IPAC/Caltech/JPL, G. Bruce Berriman & IPAC/Caltech/JPL, Barry Madore, OCIW & IPAC/Caltech/JPL

The NASA/IPAC Extragalactic Database (NED, http://nedwww.ipac.caltech.edu) provides cross-identifications and data for over 7 million extragalactic objects fused from thousands of survey catalogs and journal articles. The data cover all frequencies from radio through gamma rays and include positions, redshifts, photometry and spectral energy distributions (SEDs), sizes, and images. NED services have traditionally supplied data in HTML format for connections from Web browsers, and a custom ASCII data structure for connections by remote computer programs written in the C programming language. We describe new services that provide responses from NED queries in XML documents compliant with the international virtual observatory VOtable protocol. The XML/VOtable services support cone searches, all-sky searches based on object attributes (survey names, cross-IDs, redshifts, flux densities), and requests for detailed object data. Initial services have been inserted into the NVO registry, and others will follow soon. The first client application is a Style Sheet specification for rendering NED VOtable query results in Web browsers that support XML. The second prototype application is a Java applet that allows users to compare multiple SEDs. The new XML/VOtable output mode will also simplify the integration of data from NED into visualization and analysis packages, software agents, and other virtual observatory applications.

This work is the result of collaboration between NED, IRSA and the NVO consortium. O.P., J.G. and G.B.B. were supported by NSF/ITR NSF Cooperative agreement AST 0122449.J.M. and B.M. were supported by the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

P1.1.19 K-12 Education with the National Virtual Observatory

Jordan Raddick, Johns Hopkins University, William O’Mullane, Johns Hopkins University

We have created the first K-12 education activity using data provided by the National Virtual Observatory (NVO). The activity uses NVO’s “Datascope,” which delivers multiwavelength data on single objects requested by the user.

The activity, “Adopt an Object,” was suggested by Heidi Kaiter, a middle school science teacher from Concord, MA. It is designed for middle school students but could be adapted for high school and Astro 101 students as well. Each group of students selects a well-known sky object (star, galaxy, or nebula) to “adopt” for detailed study. Students use the datascope to look up multiwavelength images, observations, and catalog data for their object. The activity ends with a brief oral presentation and written report summarizing what each group has learned about their object.

The activity includes a complete teacher’s guide, with a lesson plan, additional resources, and correlations to national education standards. We are currently looking for K-12 teachers to field test the activity in their classes. We are also developing several other activities for K-12 and college teachers using NVO data.

P1.1.20 Development of a Virtual Observatory Prototype in ALMA-Japan

Tsuyoshi Sawada, Nobeyama Radio Observatory, Ken'ichi Tatematsu, National Astronomical Observatory of Japan, Ryo Kandori, Graduate University for Advanced Studies, Kouichiro Nakanishi, Nobeyama Radio Observatory, Koh-ichiro Morita, Nobeyama Radio Observatory, Kazuyoshi Sunada, Nobeyama Radio Observatory

The Japanese side of the Atacama Large Millimeter/submillimeter Array(ALMA-Japan) is developing a Virtual Observatory, ALMA-JVO, which will be one of the key functionalities of the ALMA Regional Center in Japan. We have summarized our science requirements and science cases for ALMA-JVO on the basis of our expertise in radio astronomy at Nobeyama Radio Observatory, National Astronomical Observatory of Japan (NAOJ). We are developing data processing pipelines for existing radio and near-infrared telescopes to evaluate system design and interfaces. Although present astronomical databases generally store reduced and calibrated images, the data produced with ALMA will be delivered to the observer after a pipeline process by default. Thus the pipeline data processing is quite essential to ALMA-JVO. We have also included an interactive operation in our prototype, because we regard it as an essential feature for astronomical data reduction. When a user requests the data, he/she will receive results through "default" pipeline and a log, which is reusable for next requests. If he/she is not satisfied with the results, he/she edits the reusable log, and re-submit a request. When necessary, he/she can request interactive operations (data flagging, selection of data area, and so on).We are also interested in data abstraction (i.e., making catalogue) of 3 dimensional data, because it will be a key technology for efficient VO astronomy with ALMA, which produces radio data at a high rate. We report present status of our activity. The program is being carried out in collaboration with Astronomical Data Analysis Center, NAOJ.

P1.1.21 A metadata layer to enable VO-compliant access to the ESO/ST-ECF Archive

Diego Sforna, ESA/ESO Space Telescope European Coordinating Facility, Alberto Micol, ESA/ESO Space Telescope European Coordinating Facility

The metadata layer is build around a centralized repository from where metadata can be retrieved. The existence of this repository is transparent to the users, but it is a key point for offering archive services in the VO frame, allowing us to provide a better data characterization in a comprehensive and homogeneous way.

At the same time we are also building for our services a web service interface, together with the traditional HTML pages designed for human consumption. The ST-ECF Instrumental Characteristics Service (see A. Micol "Database of Instrumental Characteristics in a Real Observatory") is one example where both the metadata layer and the new VO interface are already in place.

The service provides a match between characteristics of standard, well-known filters and the ones of real filters mounted on ESO/HST instruments.

The output for matching filters is offered both via a web service and in HTML format. The web service is used to retrieve at once in VOTABLE format all the data for the matching filters together with a set of standardized metadata.

We will show usage examples for the Instrumental Characteristics and other ST-ECF services, focusing on the new access possibilities and on the advantages of a hybrid interface.

P1.1.22 Merger of VO Data Model, Ontology, and High Level Query

Ed Shaya, U. of Maryland, Brian Thomas, U. of Maryland

We propose a VOCatalog data model as a list of sourceObjects that aggregates Quantities (property objects with errors, units, values). The quantities should be allowed to be of arbitrary depth allowing for QuntitySets of QuantitySets built up into a tree-like structure. For instance, a galaxyCatalog provides, at a minimum, basic data about a set of galaxies by aggregating simple quantities: magnitudes, ra, dec, morphological class etc. In addition, one may desire certain Observations of each galaxy, such as Images. One may want to directly provide crucial metadata about each image (exposure time, ra, dec, filter) and perhaps a URL to the actual data. One may wish to group these images into various regions. The structure so far is /galaxy/region/observation/image. Region may specify, besides the location on the celestial sphere, information on the type of region, e.g. spiral arm, interarm, open cluster region, outerhalo.

The real power of this schema is derived when the schema is accepted by all data centers as a query language. It accesses the scientific objects within datacenters, but completely hides each datacenter's internal organization. A query for galaxies with supergiant stars in the interarm region is a simple XPath:

```
//galaxy//region[@type="interarm"]//star//spectralType/valu
e="supergiant"
```

Sent to all relevant datacenters, it can be decomposed into a set of internal SQLs to retrieve the appropriate data and then construct a galaxyCatalog for output. Used inside of an XQuery, the request could compose an alternate structure for the output such as a starCatalog Special Requests

---

P1.1.23 Structured Query Language for Virtual Observatory

Yuji Shirasaki, NAOJ, Masahiro Tanaka, NAOJ, Satoshi Honda, NAOJ, Yoshihiko Mizumoto, NAOJ , Masatoshi Ohishi, NAOJ, Naoki Yasuda, ICRR, Yoshifumi Masunaga, Ochanomizu Univ., Masafumi Oe, NAOJ

Currently three query languages are defined as standards for the Virtual Observatory (VO). Astronomical Data Query Language (ADQL) is a query language for catalog data, Simple Image Access Protocol (SIAP) is for image data, and

Simple Spectrum Access Protocol (SSAP) is for spectrum data.

As a result, when we query each data service, we must know in advance which language is supported on the service and then need to construct a query language accordingly. The SIAP and SSAP are simple, but they have a limited capability. For example, there is no way to specify multiple regions in one query, and it is difficult to specify complex query conditions. If both the image and spectrum queries are written in SQL syntax, multiple regions query can be written with usual manner of SQL and complex conditions can be specified in WHERE clause.

In this paper, we propose a unified query language for any kind of astronomical database on the basis of SQL99. SQL is a query language optimized for a table data, so to apply the SQL to the image and spectrum data, the data structure need to be mapped to a table like structure.
We defined a virtual table whose columns are related with the parameters of SIAP and SSAP, and the corresponding image and spectrum data. We present detailed specification of this query language and an example of the architecture for the database system.

P1.1.24 Japanese Virtual Observatory (JVO) prototype 2

Masahiro Tanaka, NAOJ, Yuji Shirasaki, NAOJ, Satoshi Honda, NAOJ, Yoshihiko Mizumoto, NAOJ, Masatoshi Ohishi, NAOJ, Naoki Yasuda, U. Tokyo, Yoshifumi Masunaga, Ochanomizu U., Yasuhide Ishihara, Fujitsu Ltd., Katsumi Abe, Fujitsu Ltd., Jumpei Tsutsumi, Fujitsu Ltd., Hiroyuki Nakamoto, SEC Ltd., Yuusuke Kobayashi, SEC Ltd., Tokuo Yoshida, , SEC Ltd., Yasuhiro Morita, SEC Ltd.

We describe the architecture of the Japanese Virtual Observatory (JVO) prototype system version 2.JVO aims at seamless access to astronomical data archives stored in distributed data servers as well as data analysis environment. For this purpose, it is important to build a framework for access to remote servers, including remote procedure calls (RPCs) and data transfer. A data request for distributed database is written in the JVO Query

Language. The JVO system parses the query language, decomposes it into individual remote procedures such as retrieval of catalog, image and spectrum and cross matching, and generate a work flow. Based on this work flow, remote procedures are called. For RPCs of JVO prototype system 1, we employed Globus toolkit 2 (GTK2).However, latency time of GTK2 RPCs was too long for successive short-time jobs. Therefore, we employ Globus toolkit 3 (GTK3) for JVO prototype system 2.As the result, we find that Grid Service in GTK3 improves performance of Racine addition to Grid Service, Reliable File Transfer (RFT) is used for efficient data transfer. Astronomical data stored in distributed servers are discovered through a registry server which provides metadata discussed in the IVOA registry working group and is built using a XML database.

P1.1.25 Theory in a Virtual Observatory: some demo

Peter Teuben

The current status of a demo of the TVO (Theory in a Virtual Observatory) presented at the AAS earlier this year will be highlighted, as well as future plans as the TVO matures within the VO.

In this demo a suite of complex theoretical simulations of globular clusters, including both stellar dynamics and stellar evolution, have been made available through a web interface.

Sample observed color-magnitude diagrams are also available through the same interfaces, allowing users to select combinations of age, metallicity, initial binary fraction, etc., and see how well the associated model compares with the data.

A second demo makes galactic hydro simulations available to the user, and allows the construction of datacubes and derived products for a comparison with observational data.

P1.1.26 A quantity-based object-relational hybrid system for the VO: the NOAO NX system implementing a catalog model

Brian Thomas, University of Maryland, Edward Shaya, University of Maryland, Zenping Huang, University of Maryland, Rafael Hirart, NOAO, Irene Barg, NOAO

We present a working prototype system query system which demonstrates howVO data-models may be used as the underlying basis for a VO-wide query. This system, NX ("New XML query"), overlies, and extends the W3C standard for query on XML documents, XQuery (we refer to this extended version as "VOQuery"). The NX system is designed to allow query on data described in terms of VO-wide data models which are built up from the VO Quantity (ref). We have extrapolated a "VO Catalog" model for the purposes of this demo, and can use this model to query data which are stored in one or more SQL data bases using essentially nothing more than VOQuery (our demo model will include. Our demo is implemented on top of a test data-base holding SuperMacho data from NOAO.

In this poster we will describe the contents of the NX system, how it may be deployed, and future efforts for development. All NX software is GPL licensed and freely available for use and/or modification.

P1.2.1 Assembly and Classification of Spectral Energy Distributions – A new VO Web Service

Hans-Martin Adorf, GAVO, Max-Planck-Institut fuer extraterrestrische Physik, Florian Kerber, ST-ECF, European Southern Observatory, Gerard Lemson, GAVO, Max-Planck-Institut fuer extraterrestrische Physik, Alberto Micol, ST-ECF, European Southern Observatory, Roberto Mignani, European Southern Observatory, Thomas Rauch, Universitaet Tuebingen

We describe a new Web service which operates in two steps. As a first step, a list of observed spectral energy distributions (SEDs) is assembled. The assembly process is initiated with an input list of sky-positions that are used for querying multiple distributed catalogues covering different wavelength intervals. The sources returned from the different catalogues are spatially matched using a probabilistic method similar to XMatch.

As a second step, this list of SEDs may be submitted to a classifier that uses a library of theoretical SED's as templates for classification. For each observed SED the three best-matching theoretical SEDs are identified.

A science case has been selected for testing the capabilities of the Web service described.

P1.2.2 FROG: Time Series Analysis for the Web Service Era

Alasdair Allan, University of Exeter

The FROG application is part of the next generation Starlink software work and released under the GNU Public License (GPL). Written in Java, it has been designed for the Web and Grid Service era as an extensible, pluggable, tool for time series analysis and display. With an integrated SOAP server the packages functionality is exposed to the user for use in their own code, and to be

used remotely over the Grid, as part of the Virtual Observatory (VO).

P1.2.3 The Chandra Bibliography Database: Bibliographic Statistics

Sarah Blecksmith, SAO, John Bright, SAO , Arnold Rots, SAO, Sherry Winkelman, SAO, Paul Green, SAO, Mihoko Yukita, SAO

We have started to generate and catalog bibliographic statistics for the Chandra mission. As a result, the databases associated with the Chandra bibliography continue to expand in scope and purpose. Publication data, including number of citations, are collected for each Chandra proposal and stored in the bibliography database. Scientific metrics based on these data are computed monthly and stored in the metrics database for trending analysis of publication statistics. The design of the system and some results for illustration are presented. In addition, we have discovered some unexpected operational uses for the bibliographic database which will also be discussed.

This work is supported by NASA contract NAS8-03060 (CXC).

P1.2.4 MIGALE: millstones and roadmap

Igor Chilingarian, Sternberg Astronomical Inst. of Moscow State U., Philippe Prugniel, CRAL Observatoire de Lyon, Hector Flores, GEPI Observatoire de Paris, Jean Guibert, GEPI Observatoire de Paris, Regis Haigron, GEPI Observatoire de Paris, Isabelle Jegouzo, GEPI Observatoire de Paris, Frederic Royer, GEPI Observatoire de Paris, Francoise Tajahmady, GEPI Observatoire de Paris, Gilles Theureau, GEPI Observatoire de Paris, Jacques Vetois, GEPI Observatoire de Paris

MIGALE (Multiparametric Virtual Instrument to Study the Evolution of Galaxies) is a collaboration between two laboratories in France: CRAL (Lyon) and GEPI (Paris), Sofia observatory (Bulgaria) and Moscow State University

(Russia). The project is now supported by Programme National Galaxies (France) and two national virtual observatories, VO-France and RVO. The observational projects on two large telescopes (VLT and Russian 6-m) proposed by MIGALE are being carried out.

The best known resource provided and maintained by MIGALE is HyperLeda - the database containing homogenized information for millions of galaxies.

MIGALE distributes the Pleinpot open source software platform that can be used to easily build astronomical databases and online tools and has vocation to conform to the VO standards in the nearest future. PostgreSQL is used as DBMS.

In 2004 three large Pleinpot powered databases related to the studies of the evolution of galaxies were opened to the public:

*GIRAFFE archive, http://dbgiraf.obspm.fr/ArchiveGiraffe/
*HiGi, http://klun.obs-nancay.fr/
*ASPID, http://alcor.sao.ru/db/aspid/

We built the prototypes of the tools that would constitute the proposed virtual instrumentation in 2005-2007:

*DisGal3D to make deconvolution of 3D spectral data
*PEGASE.HR to build high resolution synthetic spectra of galaxies
*SPIKeR to determine the internal kinematics and stellar content of galaxies from integrated light spectra

We plan to release three databases supposed to become major sources of scientific data for galaxies:

*DisGal: 3D data for distant galaxies
*Fabry-Perot: 3D data from Fabry-Perot interferometers all over the world
*ASPID-GE: reduced 3D data from MPFS (6m telescope)

P1.2.5 Telescope Observation Markup Language (TOML) Definition and Use

Shaun de Witt, Joint Astronomy Centre, USA, Tim Jenness, Joint Astronomy Centre, USA, Frossie Economou, Joint Astronomy Centre, USA, Martin Folger, ATC, UK

An increasing number of observatories now use Flexible Scheduling to ensure they are used in the most efficient manner possible. This necessitates the development of Observation Preparation tools, with the output from these being delivered to observatories before the start of an observing semester. In many cases, this information is transferred in XML.

In this poster, we present the XML definition developed for JCMT and UKIRT to facilitate this transfer of information. We also show how we use this definition to validate observing programs and update constraints on-the-fly without the need to release new versions of our Observation Preparation Tool.

P1.2.6 The CIAO Website Unveiled

Elizabeth Galle, Harvard-Smithsonian Center for Astrophysics, Doug Burke, Harvard-Smithsonian Center for Astrophysics, Chris Stawarz, Harvard-Smithsonian Center for Astrophysics, Antonella Fruscione, Harvard-Smithsonian Center for Astrophysics

The Chandra Interactive Analysis of Observations website (CIAO, http://cxc.harvard.edu/ciao/) is the primary resource available to users of the CIAO software package. Several hundred pages of content are written in XML, allowing for a baseline set of text from which many types of documentation may be created. Development and production HTML versions of the site are generated from the XML via conversion scripts and XSL stylesheets.

We present an overview of the back-end of the CIAO website, including custom markup tags, stylesheets, and CSS. Design decisions, browser support, and content management are also discussed.

The success of the project led to the use of this system in maintaining five public websites at the CXC.

This work was supported by the Chandra X-ray Center under NASA grant NAS8-03060.

P1.2.7 VOStat: a distributed statistical toolkit for the Virtual Observatory

Matthew J. Graham, Caltech, S. G. Djorgovski, Caltech, A. Mahabal, Caltech, Roy Williams, Caltech, G. Jogesh Babu, Penn State, Eric D. Feigelson, Penn State, Robert Nichol, Portsmouth, Dan Vanden Berk, Carnegie Mellon, Larry Wasserman, Carnegie Mellon

The nature of astronomical data is changing: data volumes are following Moore's law with a doubling every 18 months and data sets consisting of a billion data vectors in a 100-dimensional parameter space are becoming commonplace. Sophisticated statistical techniques are crucial to fully and efficiently exploit these and maximize the scientific return. A long-standing limitation, however, on the range and capability of such analyses has been the paucity of non-proprietary software.

VOStatistics is the result of a cross-disciplinary collaboration between astronomers and statisticians to meet these challenges; it is a prototype knowledge-based statistical toolkit implemented within the VO paradigm for the entire astronomical community. VOStatistics consists of an easily extensible distributed web services-based framework transparently accessed via an open-source client GUI (available from our website: http://www.vostat.org).

A few exploratory science applications are presented to demonstrate the functionality currently offered by VOStat. This includes multi-resolutional k-dimensional trees for clustering and outlier detection and more traditional

techniques, drawn from the open-source R statistics package, such as principal component and survival analyses.

This work is supported in part by the NSF grant DMS-0101360.

P1.2.8 The ESO Telescope Bibliography Web Interface -- Linking Publications and Observations

Uta Grothkopf, ESO, Nausicaa Delmotte, ESO, Norbert Rainer, ESO

We present a web interface that links scientific papers based on ESO observations with underlying observing proposals and archival data. It offers search options for bibliographic details of individual papers as well as for queries by observatory site, telescope, instrument and program ID. It is based on WDB, a Web interface to SQL Databases written in Perl. This toolset provides an easy mechanism for crossing databases, by allowing data retrieved to be converted into hyper-text links. Thus, the ESO Telescope Bibliography Web Interface is now linked to the ESO Observing Programs Interface of the ESO/ST-ECF Science Archive Facility. The reverse link has also been implemented. The ESO Telescope Bibliography can also be queried through a filter implemented at the ADS Abstract Service, with active links to the ESO Archive. These services are prerequisites to the astronomical Virtual Observatory as they allow the user to keep track of the entire lifetime of a scientific proposal: from scheduling, to observations and publications.

P1.2.9 WBUCS: A web simulator for deep galaxy fields

Daniel Magee, UCO/Lick Observatory, Rychard Bouwens, UCO/Lick Observatory, Garth Illingworth, UCO/Lick Observatroy

Today's deep high resolution multiwavelength surveys contain a wealth of information about galaxies at different epochs. Fully exploiting this information to trace out galaxy assembly requires the ability to archive galaxy samples and resimulate these samples at different

redshifts. On this poster, we describe some of the tools developed for these ends as well as a simulator soon to be available on the web for public use. Based upon a pixel-by-pixel modeling of object SEDs and their selection volumes, this simulator provides users with the ability to make realistic multicolor simulations of galaxy fields from galaxy samples at all redshifts: from z~0 samples selected from the Sloan Digital Sky Survey to z~1-6 samples selected from the Great Observatories Origins Deep Survey. Users only need specify the pass bands, noise, and PSFs, or equivalently the exposure times on well-known instruments like HST or ground-based telescopes. As such, this simulator provides the community with a real world virtual observatory, useful both at the proposal phase and for making comparisons with observations in hand. The engine for the simulator is the well-known BUCS library developed for galaxy evolution studies, while the web interface is similar to simple web based forms used in popular image archives.

P1.2.10 Configurable format conversion tool from XML to FITS with binary extensions

Keiichi Matsuzaki, Hajime Baba, Akira Miura, Koushirou Shuutou

We are now developing configurable format conversion tool from XML format to FITS format with binary extensions. Structure of input data of this tool is specified by an XML schema language -- RELAX NG -- and that of output is specified by template FITS file. Since RELAX NG is general purpose language and this tool accepts mapping of arbitral data item in XML data structure, we can widely use this tool in data processing in the astronomy.

This tool gives one easy solution to connect ‘modern’ XML technologies with ‘traditional’ type FITS file processing. The step of conversion can be minimized one. This tool avoids designer/developer of each application programming validation and transformation since this tool can be fully configured by the schema language and the mapping definitions. Compared with general purpose computer language, roles of the schema language and the mapping definition are limited. When we use these definitions, we can shorten period of development keeping reliability of the system.

Initial implementation of this tool will be available within this year. We are now planning to use this tool in the process of telemetry status of Astro-F and Solar-B satellites which ISAS will launch in 2005-2006. In these applications, specifications of input are created by transformation of existing database of telemetry definition which also minimizes effort for development.

P1.2.11 Automatic Electronic Transfer of Astronomical Data and Metadata

Geoffrey Melnychuk, CADC, Adrian Damian, CADC, Sharon Goliath, CADC, Ling Shao, CADC, Severin Gaudet, CADC, Norm Hill, CADC, David Bohlender, CADC

The Canadian Astronomy Data Centre (CADC) has developed a system for the electronic transfer of telescope observations and metadata from the observatory to the facility's archive site. The system is currently in use at the Gemini North and South and CFHT observatories. Each step in the process between the time data become available at the telescope until they become available to an archive user is automated. This includes data verification, file transfer, ingestion of catalogue metadata, and the insertion of the file into the CADC's archive tracking system. In this poster, we present the benefits of automatic transfer, provide details of our simple, easily-modified design and implementation of the system, and discuss future developments such as the automatic distribution of data to PI's that this system makes possible.

P1.2.12 Spreading DIRT with Web Services

Marc W. Pound, U. Maryland, Mark G. Wolfire, U. Maryland, N.S. Amarnath, U. Maryland, Raymond L. Plante, U. Illinois/NCSA

Over the past 5 years, we have created and developed the Web Infrared Tool Shed (WITS), a collection of web pages and programs (toolboxes) for the analysis of astronomical objects that emit primarily in the wavelength range 2 mum to 1 mm. WITS provides the astronomical community with a uniform set of models over a large parameter space of source properties. One of the toolboxes is the Dust InfraRed Toolbox (DIRT), a comprehensive Java applet for modeling astrophysical process in circumstellar dust shells around young and evolved stars.

The DIRT database currently holds about 600,000 model spectral energy distributions. To facilitate greater flexibility in retrieval and visualization of these models, we aim to provide access to the model database through a variety of WebServices. We describe here a simple service that retrieves a requested model and returns it as a VOTable. This is the first step towards interoperability with a Theory VO. We also describe the classes of Web Service that a theory-based VO should provide.

P1.2.13 XML Web Services at CDS

André Schaaff, CDS, Jean-François Nicolas, CDS

Some of the CDS services are provided through SOAP services since October 2002. New services are added depending on the need of users and they are now all compliant to the IVOA Support Interfaces 0.2 proposal.

To increase the availability, CDS SOAP services are mirrored at NASA ADS since May 2004.

WSMonitor, a light online tool designed to visualize statistics about Web Services (logs for each mirror, logs

for each Web Service interface …) is the most recent development.

On the server side, the logs are generated via a LogHandler (Axis is used at CDS but it may be done easily with .Net).

On the client side, the user interface is based on SVG (clickable cheese, histogram …).

This tool is not dedicated to soap services and can be adapted quickly to another use cases (also used at CDS to visualize statistics about a cluster).

The CDS XML Web Services portal is now a part of the CDS Developer's corner (http://cdsweb.u-strasbg.fr/devcorner.gml) where VO developers can find information about how to download and implement parsers, web services …

P1.2.14 The AAS Professional-Amateur Electronic Registry

Jeff Stoner, Astronomical League, Janet Stevens, Astronomical League

The AAS's Working Group on Professional-Amateur Collaboration (WGPAC) was established to create an online registry to allow amateur and professional astronomers interested in collaboration to connect.

Amateur astronomers will enter approximate geographic location, skills, contact information, and available observing equipment.

Professional astronomers (and profession astronomy groups) will enter what types of observations and assistance they need, and contact information.

The registry can be reached via the WGPAC home page at http://www.aas.org/wgpac/

P1.2.15 A New SIMBAD Software and System - Simbad 4

Marc Wenger, CDS France, Anais Oberto, CDS France, Jean-Philippe Lejal, Universite de Nancy France, Sylvaine Jaehn, Univ. Techno. de Belfort/Montbelliard France, Olivier Dellicour, Haute Ecole Rennequin Sualem Belgium, Benoit Baranne, Univ. Techno. de Belfort/Montbelliard France, Julien Deprez, Haute Ecole Rennequin Sualem Belgium

The current version of Simbad is running since 1990. It was designed for command line queries and evolved to client/server mode. In the same time, its contents increased from 700,000 objects to more than 3,000,000, and the number of queries exploded from a few hundred to an average of 15,000 daily.

Fifteen years later, it was time to design a new version taking in account the evolution of technologies in languages, DBMS, networking, and software architecture.

The poster will present the architectural design of the new Simbad system. The choices made for the development language (JAVA), the DBMS(Postgresql) will be analyzed and justified. The integration of Simbad4 in the Virtual Observatory functionalities will be emphasized.

This new Simbad system will be launched mid 2005.

P1.2.16 RADAR: A Fast, Scaleable and Distributable Archive Inventory Service

Anzhen Zhang, Caltech, Tom Jarrett, Caltech, Anastasia Alexov, Caltech, Bruce Berriman, Caltech, John Good, Caltech, Mihseh Kong, Caltech, Naveed Tahir-Kheli, Caltech, Serge Monkewitz, Caltech

The NASA/IPAC Infrared Science Archive (IRSA) has recently deployed the Recursive Archive Digest and Reference (RADAR) service, which returns an inventory of IRSA's holdings in response to a spatial query, and offers one-click download of data and links to IRSA's data access services. RADAR also supports inventories and data access from remote archives; the current implementation supports access to The Multimission Archive at STScI (MAST) Spectral and Image Scrapbook and NED Basic Data. When complete, RADAR will persist the results of multiple queries in "data collections," and will provide tools that will allow users to augment collections, remove data from them, modify search criteria, resubmit jobs and check job status. RADAR is supported by an evolution of IRSA's component based architecture. It utilizes a fast estimation service and runs under Request Management Environment (ROME) funded by NVO.

P1.3.1 An Astrocomp-Grid based parallel tool for the analysis of light curve sequences of magnetically active close binaries

U. Becciani, INAF - Osservatorio Astrofisico di Catania, A. F. Lanza, INAF - Osservatorio Astrofisico di Catania, A. Costa, INAF - Osservatorio Astrofisico di Catania, M. Rodono`, Dip. di Fisica e Astronomia, Università di Catania

We present some preliminary results of the analysis of long-term light curve sequences of magnetically active close binaries by means of a numerical parallel code made available through the Astrocomp web portal. The code looks for the best values of the photometric parameters and evaluates their confidence intervals for eclipsing binaries with cool spots on their surface. Specifically, cool spots produce distortions of the shape of the light curves that may cause systematic errors in the determination of the luminosity ratio, the fractionary radii of the component stars and the inclination of the orbital plane. Our method of analysis reduces the effects of spots on the determination of such parameters by simultaneously fitting a long-term sequence of data during which the spots' coverage and distribution show remarkable changes thus allowing us to correct for their perturbation (cf. Rodon\`o, Lanza \& Becciani 2001, A\&A 371, 174).

The large amount of computational work required by our approach is managed by means of a parallel code based on MPI. The load balance of the computation is achieved by estimating the time required to analyse a given light curve and distributing subsets of light curves to the available processors in order to get similar overall computation times.

The code has been made available through a web-based user-friendly interface called Astrocomp. It allows a registered remote user to run the parallel code on a set of high performance computing resources.

Astrocomp is a project developed by the INAF-Astrophysical Observatory of Catania, University of Roma La Sapienza and ENEA (http://www.astrocomp.it).

P1.3.2 AstroGrid and the Registry: Enabling Resource Discovery

Kevin Benson, Mullard Space Science Laboratory, UCL, London, UK, Brian McIlwrath, CLRC/Rutherford Appleton Laboratory, Didcot, UK, Elizabeth Auden, Mullard Space Science Laboratory,UCL, London, UK, Dave Morris, Institute of Astronomy, Cambridge University, Cambridge, UK

The UK Virtual Observatory (VO) project AstroGrid (see http://www.astrogrid.org and related talks at this meeting) began in 2001 and is nearing the successful completion of its first release in December 2004.

AstroGrid increases scientific research possibilities by enabling access to distributed astronomical data and information resources. AstroGrid provides a standardised framework for publishing, co-operative analysis and visualisation of data supported by the following components:

* Data discovery and resource directory - the 'Registry'
* The virtual data repository.
* A set of integrated applications for seamless data manipulation,
* A workflow capability to sequence several applications working on a
data set, and
* Heterogeneous data access and manipulation accessories.

This poster covers the back-end 'Registry' component of AstroGrid, which provides the research astronomer with the capability to discover data and application resources relevant to their specific problem. We demonstrate the standards required for multiple data resources to interoperate via AstroGrid and how other components discover resources through the Registry.

The AstroGrid registry is constructed using the eXist XML database, and is queriable using XQuery. Each resource available through AstroGrid is described by means of standard resource descriptors, implemented via a standard XML schema. AstroGrid implements the International Virtual Observatory Alliance's (http://www.ivoa.ne t) agreed VOResource standard. We describe how, using the IVOA standard, we are able to 'harvest' registry information from other VO projects, thus allowing astronomers access to resources located for example in the USA, Europe, Japan and so forth.

P1.3.3 Integration of an MPP system and astrophysical applications in the INFN-GRID

Alessandro Costa, INAF Astrophysical Observatory of Catania (Italy), Ugo Becciani ,INAF Astrophysical Observatory of Catania (Italy), Antonio Calanducci, INAF Astrophysical Observatory of Catania (Italy)

We present the middleware modification made to integrate an IBM-SP parallel computer into the INFN-GRID (the INFN Production Grid for Scientific Applications) and the results of the applications runs made onto the IBM-SP to test its operation within the grid. The IBM-SP system is an 8-processors 1.1 GHz using AIX5.2 operating system.

Its hardware architecture poses a major challenge to the integration into the grid network because it does not support the LCFGng (Local ConFiGuration system Next Generation) facilities.

In order to obtain the goal without the facilities of the LCFGngserver (RPM based) we have properly tuned and compiled the middleware in the IBM-SP, implementing the Grid Services and schedulers for the job execution and monitoring. The testing phase was successfully passed by running a set of MPI jobs through the grid into the IBM-SP by means of the User Interface.

Specifically the tests was made by using MARA, a public code for the analysis of light curve sequences, that is made accessible through the Astrocomp portal, a web based interface for astrophysical parallel codes. The integration in the INFN-GRID of the IBM-SP did not require to stop the production of the system.

It can be regarded as a paradigmatic case for the integration of machines using different operating systems.

This work is carried out in the framework of the DRACO project (Datagrid for Italian Research in Astrophysics and Coordination with the Virtual Observatory) http://wwwas.oat.ts.astro.it/draco.

P1.3.4 The AstroGrid Common Execution Architecture

Paul Harrison, Jodrell Bank Observatory, Noel Winstanley, Jodrell Bank Observatory, John Taylor, Royal Observatory Edinburgh

The UK Virtual Observatory (VO) project AstroGrid (see http://www.astrogrid.org and related talks at this meeting) began in 2001 and is nearing the successful completion of its first release in December 2004.

This paper describes AstroGrid's Common Execution Architecture (CEA). This is an attempt to create a reasonably small set of interfaces and schema to model how to execute a typical astronomical application within the VO. In this context an application can be any process that consumes or produces data, so in existing terminology this could include

* A legacy unix command line application
* A database query
* A web service

The CEA has been designed primarily to work within a web services framework, with the parameter passing mechanism layered on top of this so that the web interface for all applications is described by a single constant piece of WSDL - the differences between applications are expressed by the registry entries for each application. Within AstroGrid we have created pluggable components that can wrap legacy command-line applications, HTTP GET/POST applications and databases as CEA compliant web services, which when combined with the Astrogrid Workflow component make distributed processing within the VO a reality.

P1.3.5 A Java Thick Client User Interface for Grid Processing.

Ted Hesselroth, Spitzer Science Center

A user interface (CAPRI) which is configurable at runtime has been developed which allows application features to be maintained and upgraded on a central server, available to users without the need for reinstalling software.

The user interface is specified by an XML file accessed through a URL and parsed by the open-source SWIX library, which returns a completely laid-out container with the application's controls.

A set of generic model-view-controller-actions classes are also instantiated based on parsing of the input XML file. Hierarchical relationships present in the XML file are reflected in membership relationships among the classes. An event-driven architecture with a central event handler allows for convenient extensibility.

Client/server software is based on the Java Web Services package. SAAJ message passing based on SOAP is the transaction medium between the client and server, allowing input and output files to be exchanged.

The server has access to data and computing resources and brokers the requested computation. Secure shell is supported as an option for transfers between the server and computing grid. Sun Grid Engine software is used to manage the cluster of processing nodes, which can be diskless workstations booted via PXE from the grid engine manager system.

This application has been deployed at the Spitzer Science Center to allow rapid interactive processing of science data.

P1.3.6 The Publisher's AstroGrid Library

Martin Hill, ROE, Noel Winstanly, JBO

AstroGrid is developing a Publishers AstroGrid Library to help data owners publish their data to the Virtual Observatory as painlessly as possible. PAL requires only configuration for common data formats (such as RDBMSs and FITS collections), but can also be extended by publishers to handle special data forms. It is designed not only to be compatible with the VO standards (NVO cone-search, SkyNode, ADQL, SIAP) but also to provide extra features that are not yet AVO standards. We describe both how to use it and how it works.

P1.3.7 EdSkyQuery-G: a data federation Grid architecture for astronomy

Brian Hills, edikt, the University of Edinburgh , Aileen Campbell, edikt, the University of Edinburgh, Bob Mann, Royal Observatory of Edinburgh, Tom Sugden, EPCC, the University of Edinburgh, Martin Hill, Royal Observatory of Edinburgh, Rob Baxter, EPCC and edikt, the University of Edinburgh

The identification of observations of particular celestial objects in multiple, large (often multi Tera-byte) heterogeneous databases distributed around the world lies at the heart of the Virtual Observatory concept. The EdSkyQuery-G project aims to provide a scalable architecture for doing this through extending the ideas prototyped in the SkyQuery .NET web service developed at Johns Hopkins University. EdSkyQuery-G is being developed with future integration with AstroGrid in mind, and it will inform the specification of the International Virtual Observatory Alliance's OpenSkyQuery data integration standard.

The EdSkyQuery-G architecture extends ideas developed by SkyQuery and the Global Grid Forum’s Data Access and Integration Services working group. EdSkyQuery-G builds on

top of two distinct data access middleware developed at the University of Edinburgh: Eldas, a suite of J2EE data access services developed by the Edikt project; and OGSA-DAI, the tomcat/axis database access services developed by EPCC and released as part of the Globus Toolkit. EdSkyQuery-G uses Eldas and OGSA-DAI interchangeably, building a true loosely-coupled distributed data federation architecture based on emerging web and Grid service standards.

In this poster we present the architecture and current status of EdSkyQuery-G, including results from the initial prototype version.

P1.3.8 Grist: Grid Data Mining for Astronomy

Joseph C. Jacob, Jet Propulsion Laboratory, California Institute of Technology, Roy Williams, California Institute of Technology, Jogesh Babu, The Pennsylvania State University, S. George Djorgovski, California Institute of Technology, Matthew J. Graham, California Institute of Technology, Daniel S. Katz, Jet Propulsion Laboratory, California Institute of Technology, Ashish Mahabal, California Institute of Technology, Craig D. Miller, Jet Propulsion Laboratory, California Institute of Technology, Robert Nichol, ICG, University of Portsmouth, PO1 2EG, UK, Daniel E. Vanden Berk, The Pennsylvania State University, Harshpreet Walia, Jet Propulsion Laboratory, California Institute of Technology

The Grist project is developing a novel grid-technology based system as a prototype research environment for astronomy with massive and complex datasets. When complete, this knowledge extraction system will consist of a library of distributed grid services controlled by a workflow system, compliant with standards emerging from the grid computing, web services, and virtual observatory communities. The science drivers for Grist include finding high redshift quasars, studying peculiar variable objects, search for transients in real-time, and the fitting of SDSS QSO spectra to measure black hole masses. Grist services will also be part of a compelling vehicle for outreach as a

component of the “hyperatlas” project to serve high-resolution multi-wavelength imagery over the internet. In support of these science objectives, the Grist framework will provide the enabling fabric to tie together distributed grid services in the areas of data access, federation, mining, source extraction, image mosaicking, catalog federation, data subsetting, statistics (histograms, kernel density estimation, and R language utilities exposed by VOStatistics services), and visualization. Interactive deployment and control of these distributed services will be provided from an intuitive, graphical desktop workflow manager. The new grid services paradigm explored in Grist will pave the way for a new era of distributed astronomy, with tremendous flexibility that allows software components to be deployed as services that are: (i) controlled and maintained by the authors; (ii) close to the data source for efficiency; or (iii) controlled by the end users so they have control over policies and level of service.

P1.3.9 The AstroGrid Portal and Workflow

Jeff Lusted, University of Leicester, Phil Nicolson, University of Leicester, Patricio F Ortiz, University of Leicester, Roy Platon, Rutherford Appleton Labs, Peter Shillan, Royal Observatory Edinburgh, John D Taylor, Royal Observatory Edinburgh, Noel Winstanley, Jodrell Bank

AstroGrid (see http://www.astrogrid.org), a UK eScience project with collaborating groups drawn from the major UK data archive centres, is creating the UK's first virtual observatory. We describe here the 'portal' and 'workflow' components of the AstroGrid system. Related presentations at this meeting cover the AstroGrid deployment, the international context, and technical details of other system software components.

The 'portal' component forms the browser-based user-interface for Astrogrid. The interface is constructed to allow research astronomers a rich and usable interface to the functionality provided by AstroGrid.

The following capabilities are provided, with full details provided for:

(1) Login and registration with a Community

(2) Formation of queries to be submitted against astronomical datasets

(3) Design of workflows consisting of multiple steps, where a step can consist of a query or some other registered astronomical tool

(4) Submission of workflows where individual steps can be executed at different data centers

(5) Interrogation of jobs (submitted workflows) to see progress and overall status.

(6) Browsing of the virtual file space supported by AstroGrid.

The design of queries and workflows depends upon being able to browse metadata in the AstroGrid registry, which contains metadata covering both astronomical datasets and tools. The registry is described in a related paper (Benson et al, these proceedings).

The virtual file space (MySpace in AstroGrid parlance) contains files owned by a particular user. These can be inputs and outputs from workflows. The files are browsable.

We note how Astrogrid is using International Virtual Observatory Alliance (http://www.ivoa.net) agreed standards.

P1.3.10 The VST data reduction: an application for GRID infrastructure

Mikhail Pavlov, INAF-O.A.Capodimonte, Juan M. Alcalà , INAF-O.A.Capodimonte, Aniello Grado, INAF-O.A.Capodimonte, Enrico Cascone, INAF-O.A.Capodimonte, Giulio Capasso, INAF-O.A.Capodimonte, Valeria Manna, INAF-O.A.Capodimonte, Fabio Pasian, INAF-O.A.Trieste, Leopoldo Benacchio, INAF-O.A.Padova

The VLT Survey Telescope (VST) is a cooperation between ESO, INAF-OAC-Napoli and the European OmegaCam Consortium for the realization, construction and operation of a wide-field imaging facility. The telescope will have an aperture of 2.6m. OmegaCam is a mosaic camera with 32 CCDs (2k x 4k chips, 15mu pixels). The corrected field of view of the system will be 1 square degree with a plate scale of 0.21 arcsec/pix.

The Sloan photometric system (u', g', r', i',z) will be adopted, but the Johnson B and V filters will be also available. It is foreseen that VST will start operations by the end 2005 at Paranal, in Chile.

The VST+OmegaCam system implies an enormous data flow that calls for adequate planning, archiving, scientific analysis and support facilities: the expected VST data flow will be about 150 Gbyte of raw data per night.

The data reduction will be performed using parallel processing.

The strategies for the processing of the VST images will be based on the guidelines of the Astronomical Wide-Field Imaging System for Europe(ASTRO-WISE).

The Italian project Enabling platforms for high performance computational GRIDS oriented to scalable virtual organizations started in 2002. The work package 10 (WP10) of this project is devoted to Astrophysical applications.

Under the framework of this WP, the INAF-OAC-Napoli is studying the possibility to use the GRID infrastructure for the VST data reduction.

This contribution presents strategies for the gridization of the VST data reduction pipeline.

P1.3.11 Access Control in AstroGrid Software

Guy Rixon, Institute of Astronomy, University of Cambridge. David Morris, Institute of Astronomy, University of Cambridge, Kevin Benson, Mullard Space Science Laboratory

Many parts of the Virtual Observatory (VO) require access control. These include the resource registry and the storage services "MySpace"(by AstroGrid) and "MyDB" (by JHU) in which the contents of the system are set at run time via the public interfaces of the VO; improper access could cripple the system. The VO needs an access-control system that prevents both accidental misuse and malicious damage. The system must be strong enough to resist an outright attack without impeding normal scientific usage.

AstroGrid (see http://www.astrogrid.org and related talks at this meeting) is working towards a system consisting in registration of users in astronomical communities (e.g. university departments, satellite-mission teams); federation of communities and services to provide a single-sign-on authentication-system; the ability to assign access to groups of users, the group membership being managed by the community; a flexible expression of access rights in authorization registers; the ability to divide authorization issues between the service provider and the community service.

We describe the security arrangements achieved in the AstroGrid-1 infrastructure: community services for user registration; single sign-on based on one-use passwords; authorization policy for MySpace. We outline the planned evolution for the AstroGrid-2 products: authentication based on public-key cryptography and the distributed authorization-system.

P1.3.12 Integrated Access to Solar Data using EGSO

Isabelle F. Scholl, International Space University, Robert D. Bentley, University College London, André Csillaghy, University of Applied Sciences Aargau

The European Grid of Solar Observations (EGSO) is a grid testbed funded by the European Commission's Fifth Framework Programme under its Information Society Technologies (IST) thematic priority. The project started in 2002 and is designed to provide enhanced access to solar and related data around the world.

EGSO is working closely with other virtual observatory (VO) projects in the solar physics and related domains. This includes the US Virtual Solar Observatory (VSO) and the Collaborative Sun-Earth Connector (CoSEC) - both funded by NASA. More recently, we have been working with the Virtual Space Plasma Observatory (VSPO) and are also in contact with the Virtual Heliospheric Observatory (VHO). Through Living With a Star (LWS) and the new IAU Working Group on "International Data Access" (Solar and Heliospheric), the VOs are studying ways of ensuring interoperability from the "sun to dirt".

The EGSO grid is composed of two main components, roles to set up the grid and, catalogs and registries to allow roles to answer users' queries. Catalogs are made of lists of observations, events and features (a new service provided by EGSO). Registries are built from these catalogs and organized in order to enhance search capabilities. In this paper we will mainly focus on the catalogs and registries model and how it serves the user interface.

EGSO R4 is now being Beta-tested by users and anyone interested in helping with this should contact one of the authors. More information about EGSO can be found under http://www.egso.org .

P1.3.13 Prototypes of a Computational Grid for the Planck Satellite

Giuliano Taffoni, INAF-O.A.Trieste, Giancarlo De Gasperis, Università Tor Vergata Roma, Davide Maino, Università Statale di Milano, Andrea Zacchei, INAF-O.A.Trieste, Giuliano Castelli, INAF-O.A.Trieste, Claudio Vuerli, INAF-O.A.Trieste, Fabio Pasian, INAF-O.A.Trieste, Salim Ansari, ESA/ESTEC, Jan Tauber, ESA/ESTEC, Torsten Ensslin, MPA Garching, Roberto Barbera, INFN-Catania

PlanckGrid is a project aimed at assessing the possibility of developing the pipelines setup for processing Planck data on the Grid infrastructure. The amount of data collected by the satellite during its sky surveys requires an extremely high computational power both for reduction and analysis. For this reason a Grid environment represents an interesting layout to be considered when processing those data.

The already contributed software that simulates the mission and generates the final products taking as an input a simulated sky, has been partially ported to the Grid infrastructure and collected in pipelines or workflows by different applications.

This project coordinates two main initiatives: the ESA and INAF-OATs joint collaboration and the INAF-OATs – GILDA activity. ESA and INAF-OATs set up a test Grid based on Globus Toolkit 2 and on the GridAssist workflow manager whose results are expected by the end of September 2004. Planck simulation pipelines have been also successfully ported on GILDA, a test-bed Grid infrastructure setup to host test-bed applications that at a later stage will be proposed as test-bed for EGEE . We have successfully simulated one year of sky survey for the whole set of radiometers of the LFI instrument, reduced the data and generated temperature maps. Our successful tests demonstrate that porting Planck pipelines on the Grid is possible and will bring great benefits in processing data at the two Planck DPCs. This paper presents the work already done in gridifying such partial pipelines together with the remarkable results obtained from those activities.

P1.3.14 WebCom-G: Implementing an astronomical data analysis pipeline on a Grid-type infrastructure

Seathrun O Tuairisg, John Cunniffe, Michael Browne, Andy Shearer, John Morrison

The recent upsurge in astrophysical research applications of grid technologies, coupled with the increase in temporal and spatial sky-coverage by dedicated all-sky surveys and on-line data archives, have afforded us the opportunity to develop an automated image reduction and analysis pipeline. Written using Python and Pyraf, the Python implementation of the IRAF package, this has been tailored to act on data from a number of different astronomical instruments. By exploiting inherent parallelisms within the pipeline, we have augmented this project with the ability to be run over a network of computers. Of particular interest to us is an investigation into the latency penalties in running the pipeline within a cluster and between two clusters. We have used a condensed graph programming model, the Grid middle-ware solution WebCom-G, to realize Grid-implementation. We describe how a re-organisation of such an astronomical image analysis structure can improve operational efficiency and show how such a paradigm can be extended to other applications of image processing. It is intended to use this project as a testbed for eventually running our image processing applications over a grid network of computers, with a view toward possible implementation as part of a virtual observatory infrastructure.

P1.3.15 Astronomical database related applications in the Grid.it project

Alessandra Volpato, INAF-O.A.Padova, Giuliano Taffoni, INAF-O.A.Trieste, Serena Pastore, INAF-O.A.Padova, Claudio Vuerli, INAF-O.A.Trieste, Andrea Baruffolo, INAF-O.A.Padova, Riccardo Smareglia, INAF-O.A.Trieste, Giuliano Castelli, INAF-O.A.Trieste, Fabio Pasian, INAF-O.A.Trieste, Leopoldo Benacchio, INAF-O.A.Padova

The Astronomical Observatories of Trieste and Padova are involved in the Grid.it project, a multidisciplinary research project, funded by the Italian Ministry for Education, University and Research, aiming at 'Enabling platforms for high-performance computational grids oriented to scalable virtual organizations'.

The workpackage of this project named Grid Applications for Astrophysics focuses on exploring the use of grid technologies for the development of astrophysical applications. Within this workpackage the main goal of the OAPd group is to study the portability to the Grid of an existing system for the consultation of large astronomical catalogues, currently serving on the net the Second Guide Star Catalog (GSC-II). In the same framework, the Astronomical Observatory of Trieste (OAT) Technology Group is dealing with the similar problem of integrating in the Grid the archive of observational data from the Italian Galileo National Telescope (TNG) and, as a further step, to provide processing for the data retrieved through grid-enabled pipelines.

Since the grid infrastructure provided by the Italian National Institute for Nuclear Physics (INFN) is based on GT-2, an adequate model for DBMSes is not available. For this reason, INFN and INAF institutes are collaborating to design and implement a suitable architecture to integrate DBMSes in the existing grid infrastructure. In the meanwhile INAF OATs and OAPd have developed temporary solutions to circumvent the problem, by adopting two complementary approaches: one based on a client-server paradigm and the other on Web-Services architecture. In this paper we report about these activities and the preliminary results obtained so far.

P1.3.16 Grid-related activity in progress at INAF

Claudio Vuerli, INAF-O.A.Trieste, Leopoldo Benacchio, INAF-O.A.Padova, Fabio Pasian, INAF-O.A.Trieste, Juan Alcalà, INAF-O.A.Capodimonte, Andrea Baruffolo, INAF-O.A.Padova, Ugo Becciani, INAF-O.A.Catania, Enrico Cascone, INAF-O.A.Capodimonte, Adriano Fontana, INAF-O.A.Roma, Giuseppe Longo, Università di Napoli

INAF, the Italian National Institute of Astrophysics, is currently involved in several Grid-related projects, at both the national and international levels.

Grid.it is a multi-disciplinary national project with the purpose of setting up and testing a national Grid infrastructure for the Italian research community. Astrophysics contributes with two applications. Their goal is to access astrophysical archives and catalogues (namely the pilot archive of the Telescopio Nazionale Galileo and the GSC-II) and to gridify the pipelines working on retrieved data (e.g. ASTRO-WISE). Work has been done to make the Grid middleware able to access and publish DBMS entities on the Grid.

DRACO is a national project extending Grid.it. Five INAF institutes plus two Universities are involved in this project, aimed at using Grid technology for processing and mining huge amounts of data. Test-bed applications are contributed, while some nodes are validating dedicated technological solutions. DRACO is also the project allowing the Italian community to participate in the Virtual Observatory; several DRACO applications are VO-enabled already.

Several initiatives are currently in progress to gridify pipelines for the Planck project. A grid-aware pipeline working on simulated Planck data will be officially proposed to the EGEE project. In the meantime Planck pipelines are being tested on Gilda, a test-bed Grid infrastructure setup to host applications to be later ported on EGEE. A joint collaboration with ESA has also been started for this purpose. Test pipelines using the Planck simulation software have been successfully tested on the Grid infrastructure in use for Grid.it and Gilda.

Tuesday Poster Session (P2.x.x)

P2.1.1 matplotlib - a portable Python plotting package

Paul Barrett, STScI, John Hunter, Univ. of Chicago, Todd Miller, STScI, Jin-Chung Hsu, STScI, Perry Greenfield, STScI

matplotlib is a portable 2D plotting and imaging package aimed primarily at scientific, engineering, and financial data. matplotlib can be used interactively from the Python shell, called from python scripts, or embedded in a GUI application (GTK, Wx, Tk, Windows).Many popular hardcopy outputs are supported including jpg, png, postscript and SVG. Features include multiple figures and axes, interactive navigation, many predefined line styles and symbols, images, antialiasing, alpha blending, date and financial plots, W3C compliant font management and freetype2 support, legends and tables, pseudocolor plots, mathematical text and more. It works with both numarray and Numeric. The goals of the package, basic architecture, current features (illustrated with examples), and planned enhancements will be described.

P2.1.2 Mosaicing with interferometric radio telescopes: Efficient algorithm for imaging and image plane corrections.

S. Bhatnagar, NRAO Socorro, K. Golap, NRAO Socorro, T.J. Cornwell, NRAO Socorro

High resolution imaging of objects larger than the primary beam of the antennas in an interferometric radio telescope require the mosaic imaging mode. Mosicing is expected to be widely used with upcoming new telescopes like the ALMA and the EVLA. Development of algorithms to solve for the errors which limit the dynamic range with this mode of imaging is scientifically important.

Since the emission spans many primary beams, primary beam and the pointing offsets related image plane errors adversely affect, and currently limit the dynamic range of

the mosaiced images. Efficiency of algorithms to solve for such antenna based errors using the measured data itself, depends on the accuracy of the image reconstruction algorithm. Such algorithms as well as the image reconstruction algorithms are both computationally expensive. The dominant cost of such an algorithm comes from the computation of the derivative of the objective function with respect to the pointing offsets. Residual visibility computation also require the image model. This in turn requires re-sampling (gridding) of the observed visibilities from all the mosaic pointing observations on a regular grid for Fourier inversion. The dominant cost of imaging, particularly for large number of pointings, strongly depends on the cost of gridding the visibilities on a regular grid.

In this paper, we present a relatively efficient visibility gridding algorithm and the results from an algorithm for solving for the antenna based pointing offsets. Implications on the imaging dynamic range and computational costs for mosaicing will also be discussed.

P2.1.3 Software for detection of Optical Transients in observations with Rapid Wide-Field Camera

Anton Biryukov, SAI, SAO, Gregory Beskin, SAO, Sergey Bondar, RIPI., Kevin Hurley, SSL., Eugeny Ivanov, RIPI, Sergey Karpov, SAO, Elena Katkova, RIPI., Alexei Pozanenko, IKI., Ivan Zolotukhin, SAI, SAO

A software complex for detection and investigation of optical transients (OTs) in observations with rapid wide-field camera has been developed. Data stream is a sample of digitized TV-CCD frames with size of 1360x1024 pixels, exposure time of 0.13 sec and frame frequency of 7.5 Hz. Software performs the following tasks: real time data transfer to the LAN; accumulation of initial data (total volume up to 0.5 Tb per night); real time data reduction - detection and classification of OTs (stationary and moving), determination of their equatorial coordinates, magnitudes and parameters of trajectory, their possible identification with known objects and transmission of information about OTs (Alerts) to the local and global networks.

It takes 0.4 seconds (3 successive frames) to conclude about detection of OT and to calculate of detected object parameters.

OT detection algorithm is based on the pixel-to-pixel comparison of current frame with frame averaged over 10-100 previous ones. All observed objects initially considered as motional. It allow to separate really motion and stationary transients in the field. For each frame the Awaiting Area (AA) is found for every observed Object on basis of its trajectory parameters. In this area algorithm suppose to detect appropriate object on current frame. The Object is considered as stationary (real) OT, if after 2 frames, its position of first detection lies inside current AA.

The faintest detectable object has 11.5 magnitude (close to V-band). During 150 nights of observations from autumn 2003 till summer 2004 in average 100 satellites and 200 meteors were registered per night.

P2.1.4 The BUCS Library for Cloning Complete Galaxy Samples

Rychard Bouwens, UC Santa Cruz, Garth Illingworth, UC Santa Cruz, Dan Magee, UC Santa Cruz

Interpreting today's high-resolution multiwavelength data on galaxies can be a challenging task. Real galaxies have complex morphologies, which depend upon wavelength, and can be found on images with a wide variety of beam smearing and signal-to-noise. To properly measure the evolution of galaxy structure and morphology across cosmic time, these properties need to be treated. It is to this end that we have developed the BUCS software suite. With this package, it is possible to select complete samples of galaxies off an image set and then project these objects to higher redshift accounting for pixel k-corrections, cosmic surface brightness dimming, and image specific noise and PSF. The end result of these simulations are no-evolution image sets that can be directly compared against the observations. This suite is now available to do science and as an illustration of its capabilities we describe its use in interpreting z~6-8 galaxies (i and z-dropouts) from the

Hubble Ultra Deep Field. Work is also under way to port these capabilities to the web. The purpose of this talk is to introduce this effort to the astronomical community as a whole.

P2.1.5 Extending Sherpa with S-Lang

Douglas Burke, Harvard-Smithsonian Center for Astrophysics, Aneta Siemiginowska, Harvard-Smithsonian Center for Astrophysics, Stephen Doe, Harvard-Smithsonian Center for Astrophysics

The modeling and fitting program Sherpa is one of the cornerstones of the CIAO X-ray astronomy analysis software package. It allows users to fit their one- and two-dimensional data sets - such as spectra, radial profiles, or images - in an interactive environment. The addition of the S-Lang interpreted language (http://www.s-lang.org/) in the CIAO 3.0 release offers the user the capability to not only control Sherpa from a scripting language, but also to customize and extend Sherpa's capabilities by using the S-Lang interfaces to CIAO's infrastructure (libraries). In this talk we will describe a number of extensions to Sherpa written in S-Lang, including: plot customizations; alternative user interfaces to Sherpa's commands; extra plots for comparing fits and image data; and interactive data manipulation. These extensions require no compilation of code by the user and are available for download from the CIAO web pages (http://cxc.harvard.edu/ciao/).

P2.1.6 A visualisation system for large remote astronomical images

Anil Chandra

The Remote Visualisation System (RVS) is a server-side distributed software system developed at the Australia Telescope National Facility. The RVS addresses the need for quick visualisation and analysis of large astronomical images stored in remote locations. It is a system in which the majority of the processing is done on the RVS Server, freeing the client from most of the complex processing tasks. One of the results of the Virtual Observatory is the collection of different data archives that are now available for public use. These archives hold data of various sizes and in the future, some archives may want to provide images that are gigabytes in size. The RVS allows users to visualise and interact with images in remote locations like these archives regardless of the size of the data.

Users of the RVS will not download images to their computers for viewing and analysis. The RVS Server performs all downloading and processing of images and sends the resulting rendered image to the user. A high bandwidth connection between the RVS Server and the data archive will result in a much faster transfer of the image than if the users downloaded the image to their local computers.

The RVS Server offers its services via SOAP interface, which allows custom clients to be built to make use of the services. Clients can vary from web browsers and applets to fully standalone applications written in any language on any platform.

P2.1.7 Extending and Customising DS9

De Clarke, UCO/Lick Observatory, Steve Allen, UCO/Lick Observatory, Bill Joye, Harvard CFA

UCO/Lick Observatory adopted DS9 as the standard real-time and quick-look display for DEIMOS (commissioned Summer 2002) and subsequent instruments. Lick software staff made several modifications and extensions to DS9, taking advantage of its open architecture and the ease of adding Tcl/Tk code to the DS9 core at runtime. We worked in close collaboration with DS9 author Bill Joye on both core modifications and external 'plug-ins'. Our poster describes features added to DS9 (in the core or externally), and strategies for integrating DS9 with real time image acquisition. The Lick software team invites other DS9 users to form a community and discuss code-sharing, APIs and other DS9 issues.

P2.1.8 Radio Frequency Interference excision without a reference horn

Tim Cornwell, NRAO, Rick Perley, NRAO, Kumar Golap, NRAO. Sanjay Bhatnagar, NRAO

Radio Frequency Interference is an increasingly important threat to radio astronomy, particularly as new instruments with wide frequency ranges and high sensitivity come into operation. RFI mitigation can be performed pre and post correlation of the incoming signals. We describe a post-correlation approach in which closure relations obeyed by the RFI are used in the excision process. Similar approaches use a reference horn to isolate the intefering signals. Instead our approach uses the observations themselves to characterize the interference. Although computationally expensive, our approach is robust and has the potential to allow observing even in the presence of dense, low level RFI.

P2.1.9 Two-Dimensional Fitting of Brightness Profiles in Galaxy Images with Evolution Strategies

Juan Carlos Gomez, INAOE, Olac Fuentes, INAOE, Ivanio Puerari, INAOE

In this work we present a two-dimensional fitting algorithm to model brightness profiles of spatially well resolved elliptical and spiral galaxies from an image. The algorithm is based on Evolution Strategies (ES), an optimization technique that have been used in other astronomical problems with good results. The problem can be seen as an optimization problem because we need to minimize the difference between the image and the model. Following this idea, we use ES instead a traditional algorithms such as Simplex, Quasi-Newton or Levenberg-Marquardt, because of traditional algorithms have problems with local minima, and in most cases when dealing with real images they can not converge to a solution.

For the fitting we used two models: the de Vaucoleurs profile and an exponential disk. The model is constructed with generalized ellipses that fit brightness profile of

the image following the equations given by the two models used. In the end, we have an artificial image that represents the light distribution in the real image. Results presented here show that ES is a well suited method to work with two-dimensional fitting in complex problems, when traditional algorithms fail.

P2.1.10 Porting the Starlink Software Collection to GNU Autoconf

Norman Gray, Starlink, Tim Jenness, Joint Astronomy Centre, Alasdair Allan, Starlink, David Berry, Starlink, Malcolm Currie, Starlink, Peter Draper, Starlink, Mark Taylor, Starlink, Brad Cavanagh, Joint Astronomy Centre

The Starlink Classic Software Collection (USSC) currently runs on three different unix platforms and contains approximately 130 separate software items, totaling over 6 million lines of code using a mix of Fortran, C, Tcl and Perl. The proliferation of requests for ports to new operating systems (including multiple variants of Linux), in conjunction with a decrease in the level of support for the classic software collection, has lead to a decision to modify the build system from the current collection of makefiles with hard-wired OS configurations to a scheme involving feature-discovery via GNU Autoconf.

As a result of this work, we have already ported the USSC to Mac OSX and Cygwin. This poster will present the issues involved in a substantial reorganization of a large legacy code base, including the difficulties in extending the autoconf system to properly support Fortran.

P2.1.11 Requirements for a Future Astronomical Data Analysis Environment

Preben Grosbol, ESO, et al.

Most of the systems currently used to analyze astronomical data were designed and implemented more than a decade ago. Although they still are very useful for analysis, one often would like a better interface to newer concepts like

archives, Virtual Observatories and GRID. Further, incompatibilities between most of the current systems with respect to control language and semantics make it cumbersome to intermix applications from different origins.

An OPTICON Network, funded by EU FP6, started this year to discuss high-level needs for an astronomical data analysis environment which could provide a flexible access to both legacy applications and new astronomical resources. The aim is to establish requirements and basic design recommendations for such an environment. This paper presents the current state of this work.

P2.1.12 Multidrizzle: Automated Image Combination and Cosmic-Ray Identification Software

Robert Jedrzejewski, STScI, Warren Hack, STScI, Christopher Hanley, STScI, Ivo Busko, STScI, Anton Koekemoer, STScI

Multidrizzle began as a collection of FORTRAN code and Image Reduction and Analysis Facility (IRAF) scripts for the combination of Hubble Space Telescope imaging data. Through the use of Python and modules such as NUMARRAY, PYFITS, and TRAITS, we have developed a robust application, in less than one year, that is the primary tool in the Advanced Camera for Surveys data reduction pipeline for image combination and cosmic-ray cleaning. We describe the motivations for moving the existing Multidrizzle code to Python, highlight the new data analysis tools that were developed in support of Multidrizzle, and illustrate how easily Multidrizzle can be expanded to support new instruments and data reduction algorithms. In addition, the removal of IRAF dependencies allows Multidrizzle to run under Microsoft Windows and avoids users having to install IRAF if their only need is to combine HST images. It has also significantly reduced the I/O load by avoiding the heavy use of intermediate files passed from one IRAF task to another.

P2.1.13 Harmonic development of an arbitrary function of the Moon/Sun/planets coordinates to Poisson series

Sergey Kudryavtsev, Sternberg Astronomical Institute of Moscow State University

A new algorithm of spectral analysis of an arbitrary function of the Moon/Sun/planets coordinates tabulated over a long period of time is proposed. Expansion of the function to a Poisson series is directly made where the amplitudes and arguments of the series' terms are high-degree time polynomials (as opposed to the classical Fourier analysis where the terms' amplitudes are constants and the arguments are linear functions of time). This approach leads to an essential improvement in accuracy of the harmonic development of the tabulated function over a long-term interval and reduction of the approximating series' length.

To test the algorithm, we calculated the Earth-Moon distance on every day within [1000BC-5000AD] by using the ELP2000-85 analytical theory of lunar motion and then made the spectral analysis of the tabulated values with help of the new algorithm. All coefficients of the Poisson series composing the ELP2000-85 theory have been successfully found; the maximum deviation of the lunar distance values calculated by our series from those given by the exact ELP2000-85 model does not exceed 1.5 centimeters over the whole interval of six thousand years.

The work is supported in part by grant 02-02-16887 from the Russian Foundation for Basic Research.

P2.1.14 AXedrizzle - Spectral 2D Resampling using Drizzle

Martin Kuemmel, Space Telescope - European Coordinating Facility, Jeremy Walsh, Space Telescope - European Coordinating Facility, Soeren Larsen, Space Telescope - European Coordinating Facility, Richard Hook, Space Telescope - European Coordinating Facility

The aXe spectral extraction software was designed to extract spectra from grism images such as those taken with the Advanced Camera for Surveys on HST. Aided by an object catalogue and a configuration file, aXe extracts and calibrates the spectra in an unsupervised way. The usual

way to combine spectra from several dithered grism images is to coadd the individual 1D spectra extracted for each object to create a deep 1D spectrum. This procedure rebins the data twice, and a complex weighting scheme is required to propagate pixel defects and cosmic ray affected pixels through the reduction process. In order to mitigate these drawbacks we have developed aXedrizzle, in which a deep 2D spectrum of each object is formed by co-adding all the 2D spectra in the individual images. For the 2D resampling of the individual 2D spectra we employ the stsdas.drizzle software. The transformation coefficients for drizzle are computed such that the combined drizzled image resembles an ideal long slit spectrum with the dispersion direction parallel to the x-axis and cross-dispersion direction parallel to the y-axis. The final 1D extraction of the spectra can be done with standard aXe tasks, or any other longslit extraction program.

In this contribution the new reduction scheme is introduced, and its advantages are discussed extensively. Results from Hubble Ultra Deep Field ACS/HRC Parallels program reduced with aXedrizzle are presented to demonstrate the feasibility of this first implementation of the drizzle code to combine spectral data. Possible applications to ground-based multi-object spectroscopy are discussed.

P2.1.15 A survey of software for the manipulation and visualisation of 3-D sub-mm spectral line datasets

Jamie Leech, Joint Astronomy Centre, Tim Jenness, Joint Astronomy Centre, Brad Cavanagh, Joint Astronomy Centre

3-dimensional spectral line datasets are the end products of observations from single-dish telescopes and interferometers using widely differing instrument technologies over a broad wavelength range. Consequently, a range of software packages exist for the data reduction, analysis and visualisation of such 3-D data sets. These packages vary greatly in scope, generality, age, maturity and the availability of support and further development effort. Instruments soon to be commissioned at the JCMT, such as the 350 GHz heterodyne focal plane array (HARP) and

the corresponding ACSIS digital auto-correlation spectrometer will present new challenges for the manipulation of large 3D spectral cubes. This poster presentation will survey a range of relevant software packages, with particular emphasis on identifying packages suitable for the handling of submillimetre data sets generated by focal plane arrays such as HARP/ACSIS on the JCMT. The conclusions will be broadly relevant to future software development and re-use for a wide variety of astronomical instruments and wavelength regimes.

P2.1.16 Generalized algorithm of entropic deconvolution

S.F. Likhachev, V. A. Ladygin, I.A. Guirin, Astro Space Center of P.N. Lebedev Institute

A new deconvolution algorithm is proposed for VLBI image restoration. This algorithm is a modification of the method of maximum entropy (MME) with the capability of selecting a choice of parameters of the target function that must be maximized. The algorithm provides a scaled invariance of the solution of a given system of transcendental equations with restrictions, and allows the calculation of the correct value of RMS for errors. It also provides a fast convergence for a given norm and the geometric progression rate of convergence. This developed algorithm is non-linear and provides the non-negativity of the inensity value. Moreover, it does not require prior knowledge of dispersion and does not contain Lagrangian factors. In comparison with traditional MME algorithms, the developed algorithm is equally suitable for both extended and core-jet types of radio sources. In addition, generalization of the algorithm for multi-frequency imaging deconvolution is also possible. This approach is very important for future VLBI projects such as EVLA and Space VLBI (Radioastron). The experimental tests of the algorithm are presented. The algorithm is implemented in the ASL for Windows software project.

P2.1.17 Using FITSCUT to Create Color Images

William Jon McCann, The Johns Hopkins University

FITSCUT is a standalone program, written in C, and licensed under the GPL that was developed to create color images on-the-fly for the Advanced Camera for Surveys Science Data Archive WWW interface.

FITSCUT has many powerful features, including:
* image resizing and zooming
* image alignment using WCS
* intensity auto-scaling using a histogram
* automatically generating a WCS compass
* intensity scaling using: asinh, histogram equalization, log, square root
* creating output in FITS, JPEG, or PNG format

I will describe the design and demonstrate its use. I will also explain how it can be used to create a powerful science archive using only Free Software tools.

P2.1.18 WAX: A High Performance Spatial Auto-Correlation Application

Serge Monkewitz, Infrared Processing and Analysis Center, Caltech, Sherry Wheelock, Infrared Processing and Analysis Center, Caltech

We describe the algorithms employed by WAX, a spatial auto-correlation application written in C and C++ which allows for both rapid grouping of multi-epoch apparitions as well as customizable statistical analysis of generated groups. The grouping algorithm, dubbed the swiss cheese algorithm, is designed to handle diverse input databases ranging from the 2MASS working point source database (an all sky database with relatively little coverage depth) to the 2MASS working calibration source database (a database with sparse but very deep coverage). WAX retrieves apparitions and stores groups directly from and to a DBMS, generating optimized C structures and ESQL/C code based on user defined retrieval and output columns. Furthermore, WAX allows generated groups to be spatially indexed via the HTM

scheme and provides fast coverage queries for points and small circular areas on the sky. Finally, WAX operates on a declination based sky subdivision, allowing multiple instances to be run simultaneously and independently, further speeding the process of merging apparitions from very large databases. The Two Micron All Sky Survey will use WAX to create merged apparition catalogs from their working point and calibration source databases, linking generated groups to sources in the already publicly available all-sky catalogs. For a given 2MASS source, this will allow astronomers to examine the properties of many related (and as yet unpublished) 2MASS extractions, and further extends the scientific value of the 2MASS data sets.

P2.1.19 Image smoothing and segmentation based on Shrira-Pesenson equation

M. Pesenson, M. Moshir, D. Makovoz, D. Frayer, D. Henderson

We introduce a new to image processing field nonlinear partial differential equation (PDE) intended for image enhancement. The equation has traveling wave solutions approximating jumps, thus leading to detection of sharp boundaries in images. Based on singular solutions of PDEs, this work opens new possibilities for noise reduction and segmentation of images. The new approach is compared with the classical scale-space theory of image processing which is based on nonlinear diffusion equations.

P2.1.20 SALSA:an interactive tool for parallel analysis and visualization of Astronomical data sets

Thomas Quinn, University of Washington, Graeme Lufkin, University of Washington, Greg Stinson, University of Washington, Filippo Gioachin, University of Illinois, Orion Lawler, University of Illinois, Laxmikant Kale, University of Illionois

Both astrophysical simulations and large astronomical surveys are producing catalogs containing of order one

billion objects. Analyzing such data sets is difficult with a serial package, and custom coding of parallel software is tedious.

To meet these challenges, we have developed Salsa, a parallel, interactive analysis tool for point-like data such as particles in an N-body simulation or object catalogs from a sky survey. The user runs a graphical client application on their desktop, which communicates with the server running on a parallel machine. The client controls the display of the simulation, which is rendered by the server. The client can enter high level code, like Python, that is executed on the server, implementing new functionality and providing programmatic control.

The client is currently written in Java and the server is written in the CHARM++ parallel language for maximum portability. The client/server pair of Salsa is currently used for active research in N-body simulations, and its use on Sloan Digital Sky Survey data will be demonstrated. The server can read particle or catalog data, render different types of visualizations, create and manipulate groups of objects, and generate statistics on these groups. Information about the object data and control over groups is exposed to the embedded Python interface. The client manipulates the simulation view, defines groups, and sends user-written code to the server.

P2.1.21 A Data Analysis Package for Bolometer Ground Testing

Bernhard Schulz, Caltech/IPAC, Lijun Zhang, Caltech/IPAC, Hien Nguyen, Caltech/JPL, Ken Ganga, Caltech/IPAC, Warren Holmes, Caltech/JPL

ESA's Herschel space observatory, to be launched in 2007, will be sensitive to Far Infrared wavelengths beyond 60 microns. The longer wavelength interval between 200 and 670 micron will be covered by SPIRE, a combination of broadband camera and Fourier transform spectrometer. SPIRE will use exclusively spiderweb bolometers as detectors, which are manufactured and tested at JPL. We describe a data analysis package developed at the NASA Herschel Science Center at IPAC in support of the testing activity. The package consists of a widget based viewer, allowing immediate

display and limited processing of the 193 recorded data channels in the lab. For the subsequent analysis a suite of subroutines is provided that are combined via IDL scripts. The software is used to provide fast, consistent, and automatic derivation of detector noise, optimum bias, thermal time constant, load curve modeling, and more for arrays consisting of 19 to 139 detectors and additional dark, resistor, and thermistor pixels.

P2.1.22 Improved Reduction Algorithm for ISOPHOT-S Chopped Observations

Bernhard Schulz, Caltech/IPAC

The astronomical data gathered by the extrinsic photoconductors on-board the ISO satellite were greatly affected by the high energy radiation environment in space. In addition so called transients with long time constants make automated pipeline processing of the data a challenge. ISOPHOT-S is the high sensitivity spectrometer of the ISOPHOT instrument, featuring two 64 element Si:Ga arrays for the wavelength range of 2.5 to 11.6 micron. It collected a wealth of almost 300 observations of extragalactic nuclei at a spectral resolution of ~100. Most of those observations were performed using the chopped observing mode, justifying a dedicated effort to revise and improve the automatic processing techniques in order to derive a homogeneously reduced dataset with realistic uncertainties. The improvements compared to the original standard pipeline processing include techniques like ramp subdivision, smoothed sigma kappa deglitching, spike filtering and accounting for a non-Gaussian signal distribution due to the glitch residuum.

P2.1.23 Taming The Measurement Equation With MeqTrees

O.M. Smirnov, ASTRON, J.E. Noordam, ASTRON

A Measurement Equation is used to predict values of the data measured with a particular instrument, e.g. a radio telescope. It is a combined model of the instrument and the observed object(s). One way of implementing an arbitrary M.E. is by means of 'MeqTrees', which can also be used to

solve for arbitrary subsets of its parameters. The poster presents the general structure and functionality of MeqTrees.

P2.1.24 Monte Carlo Image Analysis in Radio Interferometry

R. V. Urvashi, NRAO/NMT, T.J. Cornwell, NRAO

Image analysis, such as component fitting of radio interferometric images has traditionally been based on likelihood techniques applied to the deconvolved images. The analysis usually ignores the uncertainty arising from the process of deconvolution. Ideally one would estimate the properties of components representing the entire emission present in the raw, dirty image. In practice, this is not feasible given the large dimensionality of the parameter space. We present an intermediate approach in which a Bayesian image analysis is performed to fit components to sub-regions of the dirty image, taking full account of the point spread function. Prior to the fitting of emission inside a given region, the emission outside of the region is removed from influence by subtracting a previously deconvolved image. Our method produces samples of the posterior distributions for the number and parameters of elliptical gaussian components within the region of interest. We compare the performance of this approach to the standard methods.

P2.1.25 Making the Most of Missing Values: Object Clustering with Partial Data in Astronomy

Kiri Wagstaff, Jet Propulsion Laboratory, California Institute of Technology, Victoria G. Laidler, Computer Sciences Corporation, Space Telescope Science Institute

Modern classification and clustering techniques typically identify a set of useful features to define a parameter space. Clustering methods then group the objects in that parameter space, with a goal such as automatically distinguishing between stars and galaxies based on magnitude, color, or shape parameters. However, some of the

selected features may be unavailable for some objects. For example, shape may not be well-defined for objects close to the detection limit, and objects of extreme color maybe unobservable at some wavelengths.

The usual methods for handling data with missing values, such as data imputation (estimating the missing values) or marginalization (deleting all items with missing values), rely on the assumption that missing values occur by random chance. While this is a reasonable assumption in other disciplines, the fact that a value is missing in an astronomical catalog often carries important information about the object. We demonstrate a clustering analysis algorithm, KSC, that a) uses all observed values and b) does not discard the partially-observed objects. KSC uses soft constraints defined by the fully observed objects to assist in the grouping of objects with missing values. We present an analysis of objects taken from the Sloan Digital Sky Survey to demonstrate how imputing the values can be misleading and why the KSC approach can produce more appropriate results.

P2.1.26 JDBC Driver for the AIPS++ Table System

Jason Ye, Wes Young, Boyd Waters

Integrating Java with AIPS++ can provide many advantages that cannot be realized with AIPS++ alone. Beyond simplifying architecture and code, use of Java in astronomical processing is promising because of its standardized nature, widely available tool packages and its exceptional GUI rendering abilities. We have implemented a Java Database Connectivity (JDBC) driver for the AIPS++ table system. This allows us to use the Table Query Language (TaQL) to query and manipulate the database from Java and provides a standard interface between the AIPS++ table system and future Java applications.

P2.1.27 The Visible Light Magnetograph at the Big Bear Solar Observatory: hardware and software, first results

Sergiy Shumko, BBSO, Alexandra Tritschler, BBSO, Carsten Denker, NJIT, John Varsik, BBSO, William Marquette, BBSO, Vladimir Abramenko, BBSO, Philip Goode, BBSO.

The Visible-light Imaging Magnetograph (VIM) is a magnetograph system for observations in the wavelength range from 550 nm to 700 nm. The instrument was designed for high spatial and high temporal observations of the solar photosphere and chromosphere. VIM utilizes the remodelled Coude-feed of the 65 cm vacuum telescope at the Big Bear Solar Observatory (BBSO). In this paper we present current status of the instrument. We provide description of hardware and software developed to control the VIM. We present first results and images obtained with the VIM.

P2.2.1 Using Perl in Basic Science and Calibration Pipelines for Spitzer Infrared Array Camera Data

Heidi Brandenburg, IPAC/SSC/Caltech, Russ Laher, IPAC/SSC/Caltech, Patrick , IPAC/SSC/Caltech, Mehrdad Moshir, IPAC/SSC/Caltech, Jason Surace, IPAC/SSC/Caltech

Object oriented Perl language pipelines generate calibration products and basic calibrated data from raw images taken by the Infrared Array Camera (IRAC) onboard NASA's Spitzer Space Telescope. The pipelines gather input data and control files, initiate database interactions, and manage data flow through C, C++ and Fortran component programs. The compiled component programs perform instrumental signature correction, calibration, and data characterization.

Core pipeline functionality is provided by two compact Perl object hierarchies - one for pipelines and another for images. The objects allowed flexible and agile response to change during Spitzer's first year of operations, offsetting the cost of utilizing interpreted Perl in production data processing.

P2.2.2 A photometric comparison of the Eclipse and ORAC-DR pipelines for ISAAC

Malcolm Currie, Starlink, RAL.

Photometric fidelity is important even for data-reduction pipelines, where it may be critical to assess sky conditions during observations, and for automated reductions of Virtual Observatory raw data.

Here I compare pipeline products from three ISAAC infrared-imaging datasets reduced with the ESO Eclipse package, and the equivalent Starlink pipeline written in ORAC-DR. Two of these datasets are single observations of crowded fields with and without nebulosity; and one is sparse, but it permits comparison of repeated observations of stars spread

over two nights. I also report on whether a self-flat using the sky, or a twilight flat gives better results.

P2.2.3 Design of the SCUBA-2 Quick-Look Display and Data Reduction Pipeline

Andy Gibb, University of British Columbia, Tim Jenness, Joint Astronomy Centre, Douglas Scott, University of British Columbia, Frossie Economou, Joint Astronomy Centre, Dennis Kelly, UKATC, Wayne Holland, UKATC

SCUBA-2, scheduled for delivery in early 2006, will be the largest submillimetre bolometer array ever built. Data from this instrument will be stored at a rate of 200~Hz generating approximately 0.5~TB per 16 hour night; unheard of for a submillimetre bolometer array. This paper will present the overall design of the quick-look display system and the pipeline, and discuss some of the unique algorithmic challenges that will have to be resolved.

P2.2.4 MIPS Germanium Pipeline Organization and Implementation

David Henderson, Spitzer Science Center/IPAC, David Frayer, Spitzer Science Center/IPAC, Ted Hesselroth, Spitzer Science Center/IPAC, Meyer Pesenson, Spitzer Science Center/IPAC

The MIPS Germanium data reduction pipelines present challenges to remove a wide variety of detector artifacts and still operate efficiently in a loosely coupled multiprocessor environment. The system scheduling architecture is designed to sequentially execute four stages of pipelines. Each pipeline stage is built around Perl scripts that can invoke Fortran/C/C++ modules or Informix database stored procedures. All inter-pipeline communication is via the database.

The pipeline stages are elimination of nonlinear and radiation artifacts in the flux measurement, calibration of the fluxes with both onboard and stellar calibration

sources, applying post-facto pointing information, and assembling individual exposures into mosaics.

P2.2.5 Caltrans Keeps the Spitzer Pipelines Moving

Wen Lee, Russ Laher, John Fowler, Mehrdad Moshir

The computer pipelines used to process digital astronomical images from NASA's Spitzer Space Telescope require various input calibration-data files for characterizing the attributes and behaviors of the onboard focal-plane-arrays and their detector pixels, such as operability, dark-current offset, linearity, on-uniformity, muxbleed, droop, and point-response functions. The telescope has three science instruments, each with three or four spectral-band-pass channels, depending on the instrument. Moreover, each instrument has various operating modes (e.g., full array or sub-array in one case) and parameters (e.g., integration time).

Calibration data are needed by pipelines for generating both science products (production pipelines) and higher-level calibration products (calibration pipelines).The calibration files, which arecreated in various formats either

---

P2.2.6 VISTA Data Flow System: Pipeline Processing for WFCAM and VISTA

James R. Lewis, Cambridge Astronomy Survey Unit, M.J. Irwin, Cambridge Astronomy Survey Unit, S.T. Hodgkin, Cambridge Astronomy Survey Unit, Peter S. Bunclark, Cambridge Astronomy Survey Unit, D.W. Evans, Cambridge Astronomy Survey Unit, R.G. McMahon, Cambridge Astronomy Survey Unit

The UKIRT Wide Field Camera (WFCAM) on Mauna Kea and the VISTA IR mosaic camera at ESO, Paranal, with respectively 4 Rockwell 2k x 2k and 16 Raytheon 2k x 2k IR arrays on 4m-class telescopes, represent an enormous leap in deep IR survey capability. With combined nightly data-rates of typically 1TB, automated pipeline processing and data management requirements are paramount. Pipeline processing

of IR data is far more technically challenging than for optical data. IR detectors are inherently more unstable, while the sky emission is over 100 times brighter than most objects of interest, and varies in a complex spatial and temporal manner. In this presentation we describe the pipeline architecture being developed to deal with the IR imaging data from WFCAM and VISTA, and discuss the primary issues involved in an end-to-end system capable of: robustly removing instrument and night sky signatures; monitoring data quality and system integrity; providing astrometric and photometric calibration; and generating photon noise-limited images and astronomical catalogues.

P2.2.7 Time Domain Explorations With Digital Sky Surveys

A. Mahabal (Caltech), M. J. Graham (Caltech), S. G. Djorgovski (Caltech), R. Williams (Caltech), P. Kollipara (Caltech), E. Krause (Caltech), B. Granett (Caltech), C. Baltay (Yale University), D. Rabinowitz (Yale University), A. Rengstorf (NCSA/UIUC), R. Brunner (NCSA/UIUC), M. Bogosavljevic (Caltech), A. Bauer (Yale University), P. Andrews (Yale University), N. Ellman (Yale University), S. Duffau (Yale University), J. Snyder (Yale University), N. Morgan (Yale University), J. Musser (Indiana University), S. Mufson (Indiana University), M. Gebhard (Indiana University)

One of the new frontiers of astronomical research is the exploration of time variability on the sky at different wavelengths and flux levels. We have carried out a pilot project using DPOSS data to study strong variables and transients, and are now extending it to the new Palomar-QUEST synoptic sky survey. We report on our early findings and outline the methodology to be implemented in preparation for a real-time transient detection pipeline. In addition to large numbers of known types of highly variable sources (e.g., SNe, CVs, OVV QSOs, etc.), we expect to find numerous transients whose nature may be established by a rapid follow-up. Whereas we will make all detected variables publicly available through the web, we anticipate that email alerts would be issued in the real time for a subset of events deemed to be the most interesting. This real-time process entails many

challenges, in an effort to maintain a high completeness while keeping the contamination low. We will utilize distributed Grid services developed by the GRIST project, and implement a variety of advanced statistical and machine learning techniques.

P2.2.8 Spectral Extraction in the Spitzer IRS Pipeline

Bob Narron, Spitzer – Caltech, Jim Ingalls, Spitzer – Caltech, Phil Appleton, Spitzer – Caltech, Clare Waterson, Spitzer – Caltech, Jing Li, Spitzer – Caltech, Ted Hesselroth, Spitzer – Caltech, Iffat Khan, Spitzer – Caltech, Fan Fang, Spitzer - Caltech

For all Spitzer IRS science frames, a standard spectral extraction product is created by the pipeline. Creation of these products is done using the following three steps: (1) location, finding the center of the source, (2) integration, integrating a rectangle at each wavelength, and (3) calibration, applying the final calibration and tuning factors to each order. The extraction product is an ASCII table which, for each wavelength, gives a flux, uncertainty, and bad pixel flags.

The standard extraction product is produced by the pipeline in a completely automated way. However, if a user wants to control the details of the extraction process there is an interactive package called "SPICE”. This package provides the ability to run the standard pipeline modules, but with user specified processing parameters.

P2.2.10 Planck/LFI Pipeline - The Demonstration Model

F. Pasian, INAF-Osservatorio Astronomico di Trieste, D. Maino, Universita' degli Studi di Milano, dip. Di Fisica, A. Zacchei, INAF-Osservatorio Astronomico di Triest, C. Baccigalupi, SISSA-International School for Advanced Studies, Trieste on behalf of the DPC development Team.

LFI is one of the two instruments installed on board Planck, the M3 mission of ESA's Horizon 2000+ programme.

The Demonstration Model (DM) is the second version of the LFI DPC pipeline, built on top of the BBM (BreadBoard Model), which was released in July 2002. Testing the BBM allowed the DPC to refine the user requirements and the architectural design for the DM.

The DM will be integrated at DPC using the IDIS Process Coordinator as pipeline environment and FITS files as standard data format; in the future, the pipeline will be interfaced to the IDIS DMC (Data Management Component), a data structure common to the whole Planck project.

The DM goal is to demonstrate that the current pipeline is able to handle the foreseen LFI data flow and to perform an end-to-end processing of the data, from telemetry to the production of the scientific results. The DM development was mainly concentrated on understanding and removing systematic effects; at the moment the pipeline is able to detect and remove all systematic effects which are understood and modeled by the Planck simulation pipeline, such as thermal fluctuations, 1/f noise, sidelobes effects, beam distortion, etc.

The Demonstration Model design and integration scheme will be described.

P2.2.11 FLIPPER: a FLexIble PiPEline fRamework

Francesco Pierfederici, NOAO

FLIPPER is a lightweight pipeline framework able to handle blackboard-based pipelines. Its ease of configuration and operation and the fact the it is extremely lightweight make FLIPPER the ideal solution for both small and large pipeline systems.

Here are some of the main features:

- Lightweight
  * Only one polling process per pipeline.
  * Pollers use coroutines and continuations (wherever possible) to reduce overhead associated with PThreads.

- Easy to install
  * FLIPPER depends on Python 2.2+ only. No other third-party software is needed.
  * FLIPPER runs on both UNIX (including Mac OS X) and Windows machines.
  * No compilation needed.

- Easy to use:
  * Pipelines are fully described in XML files (one file per pipeline).
  * Uses the same basic concepts as OPUS (triggers, blackboard, modules etc.).

- Flexible:
  * Pipeline modules are activated when trigger conditions are met.
  * The most popular trigger types are supported (time, file and event/blackboard based).
  * Pipeline modules can define pre and post-processing actions.
  * Pipeline modules can be written in any language.

FLIPPER is currently being used for the NOAO Mosaic Pipeline and has been chosen as pipeline framework for the NOAO NEWFIRM pipeline system. The present paper describes the architecture, features and design choices of FLIPPER in detail.

P2.2.12 Wavelength calibration in physical model based calibration pipelines

M. Fiorentino, M. Rosa

For the case of the Space Telescope Imaging Spectrograph (STIS on HST) we show how the wavelength calibration of complex 2D-spectrograms such as echelle and long slit spectra can be achieved in a fully self-consistent manner by a physical model based pipeline. The ST-ECF is currently upgrading the STIS data calibration pipeline in a step-by-step manner. Gradually the CalibrationEnhanced CE-calstis pipeline will be able to generate the geometric transformations necessary during wavelength calibration and spectral extraction from a physical (ray trace) model of the STIS optics. In parallel, the same model based approach is at the core of a "calibration-data-pipeline" which will provide the supporting reference data from a self-consistent automatic analysis of calibration exposures, a process that currently in the classical pipeline needs to be done painstakingly by hand.

We emphasize that this concept and even the actual implementation (code) can, should, be re-used for other instruments as well. Our method makes heavy use of the "Simulated Annealing" technique to optimize the many-parameter configuration data for such a physical model based description of instruments.

P2.3.1 Analyzing the Cas A Megasecond in Less than a Megasecond

John E. Davis, M.I.T., John C. Houck, M.I.T., Glenn E. Allen, M.I.T., Michael D. Stage, M.I.T.

This spring the Chandra X-ray Observatory concluded its megasecond observation of the Cassiopeia A supernova remnant using its Advanced CCD Imaging Spectrometer (ACIS).The nine long exposures that made up the megasecond observation produced a total of more than 300 million events on the ACIS-7 CCD, giving an average of about 300 events per pixel, with brighter parts of the remnant contributing as many as 10000 events per pixel. We took this ``spectrum in every pixel'' observation quite literally by extracting the spectrum in nearly every pixel and then fit the spectra on an adaptive spatial grid to produce high resolution maps of line emission, Doppler velocities, and plasma temperatures. Although this required extracting and fitting a quarter of a million spectra, we were able to produce the high resolution maps in a matter of days. This task was facilitated by the rapid scripting afforded by the S-Lang interpreter as embedded in the Interactive Spectral Interpretation System (ISIS) application that is distributed as part of the Chandra Data Analysis System (CIAO). In this paper, we describe how we used the interpreter and its modules, and present some of the data structures and algorithms employed for processing the data from the megasecond observation.

P2.3.2 Linearization of Spitzer Infrared Spectrometer Data Via Minimization of Chi-Square With Correlated Errors

John Fowler, California Institute of Technology

The Spitzer Infrared Spectrometer data are taken via read-without-reset measurements to obtain multiple samples forming a photometric "ramp" for each pixel in an echellogram. Each ramp is linearized via a quadratic model. After linearization, a quality-assurance test is performed to determine how linear each pixel's ramp has become. This

is accomplished by fitting a straight line to the ramp via chi-square minimization. The goodness of fit is of primary importance, since this determines whether the inevitable deviations from linearity are statistically significant given the estimated photometric noise. Because the latter is dominated by photon noise which is summed up the ramp, the chi-square parameter used to measure goodness of fit must include the effects of correlated errors. This paper describes the construction of the full error covariance matrix and its use in the chi-square minimization.

P2.3.3 Post-pipeline processing of Deep Spitzer data for the GOODS Legacy Project

David Grumm, STScI, Stefano Casertano, STScI, Mark Dickinson, NOAO, Sherie Holfeltz, STScI

The GOODS Legacy Project has acquired the deepest images planned to be taken by Spitzer from 3.6 to 24 micron, with effective exposure times of up to 100 hours for IRAC. Optimal analysis of these data requires an effective calibration noise better than a part in 10,000. Starting from the high-quality calibrated frames delivered by the Spitzer Science Center pipeline, we employ a variety of techniques to validate and improve the calibration quality and obtain combined deep images that fulfill the requirements imposed by the GOODS science goals. Post-pipeline processing includes correction for instrumental artifacts, improved alignment, verification of the flat field quality via self-calibration, and image combination using drizzle-related techniques. We present a quantitative analysis of the noise and PSF properties of the final images thus obtained, and indicate the improvements obtained with each of the techniques employed.

P2.3.4 Towards the Optimal Association of Hubble ACS Images.

Jonas Haase, ST-ECF/ESO, Daniel Durand, CADC, Richard Hook, ST-ECF/ESO, Alberto Micol, ST-ECF/ESA

The ACS Associations project is aiming to provide combinations of exposures from the Hubble Space Telescope Advanced Camera for Surveys (ACS) similar to those already produced for the older WFPC2 camera.

The goal is to associate exposures of a region in the sky with as few restrictions as possible, e.g. regardless of time, science goal or principal investigator of the original observational programmes.

Tests on an initial prototype for the ACS Associations are being carried out in order identify methods that will lead to the most interesting science product.

Several ways of matching or cross-correlating the science frames, to find their relative shifts and rotations and then correct the astrometric information in the image header world coordinate system (WCS), are being tested against each other and the most stable and effective way of stacking the frames is being sought.

At this stage there is also a good opportunity to revisit the implicit assumptions carried over from the WFPC2 Associations and find ways of breaking up the paradigms to allow for more flexible and potentially more useful associations. The association discovering algorithm has been made more complex and the possibility of offering user-defined on-the-fly associations with association strategies matching a user science goal is also being investigated.

P2.3.5 Concurrent science and wavecal COS data

Philip Hodge, STScI

Wavelength calibration for the Cosmic Origins Spectrograph (COS) is achieved by taking an exposure (a wavecal) of an internal emission-line lamp and comparing the observed location with the expected location. In order to correct for possible mechanism drifts during exposures which can be more than an hour long, a new observing mode has been proposed whereby the line lamp would be turned on briefly several times during a science exposure. This is practical because the science and wavecal spectra fall on different parts of the detector, and because COS data will normally be taken in time-tag mode, where the time and pixel coordinates are recorded for each detected photon. This paper describes the calibration of such data.

P2.3.6 Real time response DB system of MAXI ground software

Mitsuhiro Kohama, RIKEN,, Hitoshi Negoro, NIHON U., Naoyuki Kuramata, JAXA, Hiroshi Tomida, JAXA, Haruyoshi Katayama, JAXA, Tatehiro Mihara, RIKEN, Shiro Ueno, JAXA, Masaru Matsuoka, JAXA, Nobuyuki Kawai, T.I.Tech, Atsumasa Yoshida, Aoyama-Gakuin U., Emi Miyata, Osaka U., Hiroshi Tunemi, Oksaka U., and MAXI mission team

Monitor of All-sky X-ray Image (MAXI) is the first astrophysical payload which will be mounted on the Japanese Experiment Module (JEM) Exposed Facility of International Space Station (ISS) in 2008.

The continuous two method down links will enable us to alert the astronomers in all over the world to the appearance of X-ray transients, novae, bursts, flares etc. We have been developing the real time quick response DB system to realize this alert system. The prototype model had been produced, and we have evaluated its performance.

P2.3.7 Mission in a Nutshell: The Gaia Parameter Database

Uwe Lammers, European Space Agency, Jos de Bruijne, European Space Agency

Gaia, ESA's ambitious space astrometry mission targeted for launch around mid-2010, is currently in the advanced study/definition phase. Around 280 European scientists and a number of industrial groups are working on various preparatory mission aspects ranging from the payload design to the detailed feasibility study of the on-ground data processing. In this situation the Gaia Parameter Database has been implemented to address the need for a common, consistent, complete, and up-to-date repository of mission parameter data in support of all scientific and technical activities on a project-wide level. The system is designed as a Web-based database application with a single entry point and supports both online and offline access via standard browsers and related tools. Scalar and multidimensional data are organized in a hierarchical manner, allowing to uniquely express interrelationships among them. Elemental data is stored in a relational database system and reconstructed by access software as an in-memory XML tree which then allows the seamless rendering into various output formats (e.g.\ HTML, Java/C++ source code, \LaTeX, etc.) via XSLT. Controlled releases of the database contents are made at irregular intervals and put under CVS control transparent access to which is offered likewise through the system's central entry point. The system is fully generic and could easily be adopted to other projects with similar needs.

P2.3.8 Global covariance analysis of the SIM astrometric grid

Valeri Makarov, MSC & JPL

The all-sky, high accuracy astrometric grid of ~1300 red giants will provide the basis for astrometric analysis of target stars and extragalactic objects observed with the Space Interferometry Mission (SIM) facility. Due to the differential character of interferometric measurements with

SIM, the grid data will be correlated globally, the more strongly the closer the stars on the sky. Those correlations will have far-reaching significance for astrophysical projects using ensembles of stars, e.g., determination of the distance to the LMC, and the rotation curve of the Galaxy. The ongoing grid simulations with the newly developed

---

P2.3.9 Processing of 24 Micron Data at the Spitzer Science Center: the First Year of Flight Data

F. Masci, SSC Caltech, J. Fowler, SSC Caltech, R. Laher, SSC Caltech, F. Fang, SSC Caltech, D. Makovoz, SSC Caltech, W.-P. Lee, SSC Caltech, S. Stolovy, SSC Caltech, D. Padgett, SSC Caltech, M. Moshir, SSC Caltech

We present an overview of the pipeline flow and reduction steps involved in the processing of imaging data acquired with the 24 micron array on board the Spitzer Space Telescope. This is one of three arrays in the Multiband Imaging Photometer for Spitzer (MIPS) instrument. It provides 5.3 x 5.3 arcminute images at a scale of ~2.5 arcseconds per pixel corresponding to a sampling of the point spread function which is a factor of ~1.4 above Nyquist. A scan-mirror allows dithering of images on the array without the overhead of moving and stabilizing the spacecraft. It also provides an efficient means of mapping large areas of sky. Prior to distribution of data to observers, raw images undergo several stages of automated processing at Caltech's Spitzer Science Center (SSC). These include the removal of instrumental artifacts (some of which are scan-mirror dependent) to create a Basic Calibrated Data product (BCD), pointing reconstruction and refinement thereof and the mosaicking of BCDs pertaining to an observational request

P2.3.10 The Impact of the ACIS Readout Streak and Pileup on Chandra Source Detection

Michael McCollough, Harvard-Smithsonian Center for Astrophysics, Arnold Rots, Harvard-Smithsonian Center for Astrophysics

As part of the development of the Chandra Level 3 (L3) data products (Chandra Source catalog) the impacts of the ACIS readout streak and pileup have been examined. A method has been developed which allows us to determine, for a given observation, which columns of the ACIS CCDs are impacted by readout streaks and may result in the detection of false sources. A discussion is given of how to identify and determine the characteristics of real sources located in the readout streak. The implications of this for the creation of L3 products is discussed. Additionally various methods of removing the readout streak from the image are examined.

We also examine the possibility of using the readout streak to help gage impact of pileup on the spectrum of the strong sources. The use of the readout streak to help calibrate the amount of pileup is explored.

This work is supported by NASA contract NAS8-03060 (CXC).

P2.3.11 Hubble Space Telescope – Science Instruments Trending

Manfred P. Miebach, Space Telescope Science Institute

With the decision to cancel the Servicing Mission 4 (SM 4) for the Hubble Space Telescope using the Space Shuttle, engineers are now focusing on preserving the existing hardware on HST to the maximum extend until (hopefully) during a robotic mission critically needed spare units (e.g. flight batteries, gyroscopes) will be exchanged to continue successful science.

The upgrade of the Hubble Space Telescope (HST) ground system was completed in December 1999 providing the Flight

Operations Teams and the Systems Engineers a variety of powerful engineering tools to monitor and trend all subsystems on-board the spacecraft

One subsystem of the Ground System is the Space Telescope Engineering Data Store, the design of which is based on modern Data Warehouse technology. The process of populating the Data Warehouse with HST historical telemetry data for the entire lifetime of the spacecraft has been completed, providing access to HST engineering data since the launch of HST in April 1990.

This paper will provide hands-on experience from an end user perspective using the Data Warehouse as an HST engineering telemetry archive for trending the Science Instruments of the Hubble Space Telescope. Engineering Teams are using HST telemetry extensively for

Spacecraft Anomaly resolutions

Science Instrument trending

Improve Instrument operational efficiency with the overall idea to extend the life of the hardware and maximize science output of the observatory.

Keywords: Telemetry Processing, Telemetry archiving, Spacecraft Trending, Long-term trending, WEB-based telemetry access system, Data Warehousing, Hubble Space Telescope

---

P2.3.12 Modern Statistical Methods for GLAST Event Analysis

Robin D Morris, USRA-RIACS, Johann Cohen-Tanugi, Stanford Linear Accelerator Center

We describe the ongoing development of an event analysis algorithm for the Large Area Telescope (LAT) instrument on the Gamma-ray Large Area Space Telescope (GLAST).We show how it is possible to construct an algorithm that incorporates accurately the physics of the detector, both in terms of the processes that produce secondary particles and photons and most importantly, in terms of the multiple scattering processes of the charged particles. The posterior distribution is modeled as a mixture, where the discrete modes of the mixture correspond to different physical processes that can cause the same detector

response. We use sequential importance sampling ("particle filters") to determine the distribution over the components of the mixture, that is, the possible configurations in terms of secondary processes and particles, and Markov chain Monte Carlo to explore the details of the distribution in each configuration. This allows us to estimate the mixture distribution (over configurations), where each mixture component is non-gaussian due to the multiple scattering and the detector geometry. The result is an algorithm that estimates the full distribution over the energy, azimuth and elevation of the incident photons, and aims to extract as much information as possible from the detector response.

P2.3.13 Filtering of signal dependent noise applied to MIPS (70mn and 24mn).

M. Moshir , M. Pesenson , D. Frayer, D. Henderson

Linear filtering and nonlinear median filtering are not useful for signal-dependent noise removal, due to the nonlinear coupling of signal and noise. Therefore decoupling of the noise from signal is highly desirable which is done by applying a nonlinear transformation. The transformation is followed by a linear filter and the inverse nonlinear transformation. This procedure is used for the reduction of signal-dependent noise.

P2.3.14 Use of the GENIE software for Automated Detection of Features From MSX/Spitzer Images

L. Rottler, S. Carey, S. Brumby

At IPAC there is great interest in finding or developing software to do automated detection of features, particularly extended features, from astronomical images. For point like objects having a definable PSF or extended objects with regular geometries this is a reasonably straightforward process. Many robust algorithms to do this have been developed and are in common use throughout the astronomical community. However for extended objects with irregular geometries, such as nebulae in absorption near the Galactic plane or extended emission nebulae, there is no general robust technique to automatically detect and identify these features.

To address this issue we use a software package, GENIE: a hybrid evolutionary genetic algorithm-based system developed at the Los Alamos National Laboratory. It is designed to evolve, generate, and implement image processing algorithms to detect and classify diverse kinds of features embedded in hyperspectral images. In this paper, we demonstrate it's applicability to the problem of automatic detection and identification of irregularly shaped features in multi-wavelength MSX and Spitzer images and discuss the robustness of the algorithms generated by GENIE to do this for several classes of objects.

P2.3.15 The COSMOS 2-Degree HST Survey: Data Handling and Mining

Patrick Shopbell, Caltech, Nicholas Scoville, Caltech, Bruce Berriman, IPAC

COSMOS is a cycle-12/13 HST Treasury program that is using the ACS camera to uniformly survey a two-square-degree region of the sky. COSMOS is by far the largest contiguous area ever observed with HST, and as such will provide a critical data point in the two-dimensional phase space comprised of survey depth versus area. The wide spatial coverage of the COSMOS survey will enable us to map galaxy formation, the dark matter distribution, and the evolution of large-scale structure, up to and including the largest, most massive structures.

The COSMOS ACS survey, which is currently ~50% complete, is only one part of the project, which also involves multiwavelength observations of the field, including ground- and space-based radio, submillimeter, IR, optical, UV, and X-ray coverage. Spectra of some 50,000 individual objects will also be obtained to further elucidate the three-dimensional structures in the field. These datasets, while not particularly large by today's standards, are quite diverse, leading us to develop novel ways of presenting the data to users. In this paper, we present an overview of the techniques that we have developed so far and the directions in which we are progressing to provide adequate visualization and data mining tools to users of the COSMOS datasets.

P2.3.16 JAVA based Astronomical Software - The HERSCHEL/PACS Common Software System (PCSS) as example

Ekkehard Wieprecht, Max Planck Institut fuer extraterrestrische Physik, Garching/Germany, R. Huygen,Institute of Astronomy K.U. Leuven, Leuven/Belgium, B. Vandenbussche,Institute of Astronomy K.U. Leuven, Leuven/Belgium, W. de Meester,Institute of Astronomy K.U. Leuven, Leuven/Belgium, J. Brumfit,ESA - Astrophysics Division, Noordwijk/Netherlands, J.J. Mathieu,ESA - Astrophysics Division, Noordwijk/Netherlands, N. de Candussio,ESA - Astrophysics Division, Noordwijk/Netherlands, S. Ott,ESA - Astrophysics Division, Noordwijk/Netherlands, H. Siddiqui,ESA - Astrophysics Division, Noordwijk/Netherlands, A. Contursi,Max Planck Institut fuer extraterrestrische Physik, Garching/Germany, S. Osterhage,Max Planck Institut fuer extraterrestrische Physik, Garching/Germany, M. Wetzstein,Max Planck Institut fuer extraterrestrische Physik, Garching/Germany

ESA's Herschel Space Observatory to be launched in 2007, is the first space observatory covering the full far-infrared and submillimetre wavelength range (60 -670 microns).

The Photodetector Array Camera & Spectrometer (PACS) is one of the three science instruments. It employs two Ge:Ga photoconductor arrays and two bolometer arrays to form imaging line spectroscopy and imaging photometry in the 60 - 210 micron wavelength band.

The HERSCHEL ground segment is based on a common, object oriented database system - the Herschel Common Science System (HCSS). It is implemented using JAVA technology and written in a common effort by the HERSCHEL Science Center and the three instrument teams.

The PACS Common Software System (PCSS) is based on the HERSCHEL Common Software System (HCSS). It was designed for a smooth transition between the different phases of the project. Instrument engineers use PCSS during instrument tests, calibration specialists for instrument characterisation on ground and in orbit, and finally the observer shall use it for data reduction.

The design shall allow also to use the same components for automatic processing, quick look analysis and interactive processing.

We give insides of design aspects of such a highly complex system and present user experiences with PCSS, gathered during first instrument tests of PACS.

We will also discuss the idea of a software system used for all phases of the mission against other possible software development strategies.

Is JAVA suitable as base language for astronomical/interactive systems?

Wednesday Poster Session (P3.x.x)

P3.1.1 FITS World Coordinates for DEIMOS Slitmasks

Steven L. Allen, UCO/Lick Observatory

The DEep Imaging Multi-Object Spectrograph (DEIMOS) installed at the Keck-II telescope began producing direct images and multi-slit spectra in 2002. Pixel data from the CCDs are stored in multi-extension FITS (MEF) files compatible with the NOAO/IRAF mosaic. Each IMAGE extension also contains keywords for multiple World Coordinate Systems that conform to the recently adopted FITS WCS standards.

Initially the WCS keywords in DEIMOS images only describe coordinate systems related to the pixel layout of the mosaic detector. There have been no WCS keywords for finding coordinates on the metal of the multi-object slit masks or for finding coordinates on the sky.

We created a calibration slitmask with a regular grid of holes, loaded it into DEIMOS, and observed at different rotation angles. We report on the success of the reduction of these observations that establish the WCS from CCD pixels to the milling coordinate system on the metal of the slitmask. We show how the ds9 image viewer is able to use the slitmask WCS to provide an undistorted view of a DEIMOS slitmask. We also take an initial look at how the non-linearities in this WCS might be handled by the algorithms described in the current draft of WCS Paper IV on Distortions in FITS coordinates. Finally, we consider the remaining steps that will be needed for EIMOS FITS files to contain an astrometrically precise celestial coordinate system.

P3.1.2 The Use of CANbus in CARMA

Andrew Beard, Caltech/OVRO, Stephen L. Scott, Caltech/OVRO

The CARMA project is using the Controller Area Network (CAN) communications bus, an industrial bus with many applications in the automotive sector. CAN is used between specialized monitor/control devices, implemented with a Philips XAC microcontroller, and compact PCI crates running Linux.Central to the system design is the partitioning of the 29 bit extended CAN message Id into several CARMA specific identifiers. These identifiers form the basis for a communication protocol that is used in monitor and control. Libraries are available for the microcontroller for communications over the CANbus and for common support functions. For the Linux host, the CARMA CANbus library consists of a C++ framework which provides a software abstraction layer for CAN devices while hiding much of the low-level communication details. The framework is flexible enough to be interfaced to a variety of monitor/control systems either directly via multiple inheritance (e.g. CORBA/IDL defined interfaces) or indirectly through delegation.

We will discuss details of our protocol, the libraries, several of the applications, and the limitations to the model. The applications include the CARMA monitor and control system, a CAN-over-Tcp/Ip server with remote clients in LabView and VisualBasic, and a firmware downloader.

P3.1.3 Architecture of the eSTAR/WFCAM Transient Object Detection Agent

Brad Cavanagh, Joint Astronomy Centre, Alasdair Allan, University of Exeter, Tim Jenness, Joint Astronomy Centre, Frossie Economou, Joint Astronomy Centre, Paul Hirst, Joint Astronomy Centre, Tim Naylor, University of Exeter, Andy Adamson, Joint Astronomy Centre

The eSTAR Project and the Joint Astronomy Centre are developing an agent for automated transient and moving object detection, identification, and alerts. The agent is being developed for the United Kingdom Infrared Telescope's Wide Field Camera, and will use output from a real-time data reduction pipeline to cross-correlate results with known objects, taken from pre-existing survey databases and a running database of objects previously observed with WFCAM. The agent will be able to identify transient and moving objects and alert astronomers depending on their specific interests. This poster will describe the architecture of the agent and its integration with the eSTAR Project, systems at the JAC, and external resources.

P3.1.4 IRAF Package for GNIRS Data Reduction

Andrew Cooke, CTIO, Bernadette Rodgers, Gemini Observatory

In August 2003 NOAO and Gemini began working together to improve the IRAF based data reduction packages used at the two telescopes. Here we report on the GNIRS (Gemini Near Infra-Red Spectrograph) Gemini/IRAF package - one of the first fruits of the collaboration.

The package helps the user generate high quality, calibrated spectra from the raw observations. It supports longslit, IFU (Integral Field Unit) and cross-dispersed observing modes - the processing of different data types is handled automatically, with the user following a similar reduction procedure in all cases. It can also be used as

part of an automated pipeline and is generic enough to support other instruments with a similar data format.

This work is based on the earlier NIRI (Near Infra-Red Imager) spectroscopy routines, written by Joe Jensen (Gemini). All the tasks have been rewritten and, although they remain CL scripts, they now call compiled tasks (e.g. gemextn, gemlog, gemarith) to implement common functionality. This gives cleaner, more reliable code and an improved user response.

In our poster we illustrate the flow of data through the package, outlining how GNIRS observations are reduced.

The package is currently (Aug 2004) being tested. It will be released with the next version of the Gemini software (target Sep 2004).

P3.1.5 An Overview of the Mount Stromlo Observatory Telescope Control System

M.A. Jarnyk, G.R. Hovey

A new telescope control system, MSOTCS, is being written to support the remote control and automation functionality of the Siding Spring 2.3m telescope. Key features of MSOTCS include: the use of a proprietary telescope pointing/astrometry kernel called TCSpk; QNX6 as the operating system; the use of a central shared database for all astrometric, configuration and control system data; and the use of a publish and subscribe database for the telescope data required by multiple remote clients. MSOTCS takes the form of an embedded controller which communicates with external user interfaces via RPC.

P3.1.6 N-body simulations with GRAPE-6a and MD-GRAPE2 acceleration hardware

Vicki Johnson, Interconnect Technologies Corporation, Alper Ates, Pomona College

The first generation NBodyLab infrastructure for astrophysical nbody calculations was demonstrated at ADASS 2002. The authors present new capabilities, applications and simulation results using two types of GRAPE cards (the newly available GRAPE-6a and the older MD-GRAPE2).

The GRAPE cards are designed by researchers in Japan for low-cost hardware acceleration of nbody calculations in particle simulations. Applications using GRAPE cards in large arrays have won many Gordon Bell supercomputing prizes in the past decade, and new generations are currently being developed.

The experience of the authors working with these new technologies in a open software framework, accessible via the web, will benefit the astronomical community.

P3.1.7 Optimizing Observing Sequence Design for both Periodic and Non-periodic Phenomena: A Bayesian Approach

Mark D. Johnston, Jet Propulsion Laboratory, Russell Knight, Jet Propulsion Laboratory

The problem of designing observing sequences to detect and characterize periodic phenomena occurs regularly in astronomical investigations. Examples of current interest include Cepheid variable searches in external galaxies (with Hubble Space Telescope) to determine the extragalactic distance scale, and future high accuracy astrometric observations of nearby stars with SIM, the Space Interferometry Mission, to search for planetary companions. Various sampling strategies have been proposed to obtain good phase coverage over an interesting range of periods. Recently, Loredo and Chernoff have proposed the use of "Bayesian adaptive exploration", a model-based Bayesian method that exploits observations made to date to

determine the best future observation times according to a maximum information criterion. While this method makes the best possible use of any results already obtained, it does not address the "bootstrap" problem of scheduling in advance of any data collection. It also is highly compute-intensive, which is a especially problematic when an integrated observing schedule for hundreds of targets is required, taking into account all of the various other constraints and preferences that come into play. In this paper we report on our progress on addressing these issues. We have developed approximate expressions for the uniformity of frequency coverage that can be used when scheduling to assess candidate sample times. We describe the results obtained using these estimators, and compare them with detailed simulations. We describe our progress and plans for integrating optimizing criteria for both periodic and non-periodic observations into a single observation sequence.

P3.1.8 The Remote Control System for the ASTE Telescope

Takeshi Kamazaki, Univ. of Tokyo, Hajime Ezawa, National Astronomical Observatory of Japan, Nobuyuki Yamaguchi, National Astronomical Observatory of Japan, Kenichi Tatematsu, National Astronomical Observatory of Japan, Nario Kuno, Nobeyama Radio Observatory, Kiyohiko Yanagisawa, Fujitsu, Jun Maekawa, Fujitsu, Osamu Horigome, Fujitsu

The Atacama Submillimeter Telescope Experiment (ASTE) is a project to operate a 10-m submillimeter telescope in the high altitude site (4,800 m) at Atacama desert in northern Chile.

The key to successful telescope operation under the severe environment of the observing site is to realize a stable remote control system. The remote control system for ASTE consists of a newly developed operating software capable for remote observation and a satellite network facility (56-64 kbps), which connects the telescope site to the outer world including the operation base in San Pedro de Atacama(altitude 2,400 m) or institutes in Japan. The control software was developed based on the existing COSMOS3 system, which has been used for the 45-m telescope

and the Millimeter Array (NMA) of the Nobeyama Radio Observatory (NRO) in Japan. This new system for ASTE employs the following features:

- Remote operation through narrow-band network connection (56-64 kbps).
- Robust system to survive with sudden network connection failure.
- Traffic shaping to guaranty necessary communication for telescope control.
- Accommodate multiple remote operating sites, including privilege control and easy addition of remote sites.

We have succeeded in the first remote observation from the operation base in San Pedro de Atacama in July 2002, followed by the first successful remote observation from Japan in October 2003.

P3.1.9 The ALMA Real Time Monitor and Control Bus

Jeff Kern, NRAO, Rodrigo Amestica, NRAO

The Atacama Large Millimeter Array (ALMA) requires high precision real time control in a distributed system. To meet the unique requirements of ALMA a special purpose monitor and control bus, based on the industry standard CAN bus, has been developed. We will discuss the design and performance of the ALMA Monitor & Control Bus as well as its software implementation in a Linux based real-time operating system.

P3.1.10 On the use of IDL for instrument control.

Steve Mazuk, The Aerospace Corporation, Catherine Venturini, The Aerospace Corporation

The Aerospace Corporation has developed a near-infrared and visible spectrograph that is used for astronomical observations at Lick observatory’s 3 meter telescope. This paper describes the instrument control and data handling system, which employs the Interactive Data Language (IDL) for both the user interface and instrument control. The system employs IDL in a client-server design to control all

aspects of data acquisition, and has been operational for several years. The use of IDL has simplified the system design and allowed for extensive modifications.

P3.1.11 The CARMA Database System

David M. Mehringer, U Illinois/NCSA, Raymond L. Plante, U Illinois/NCSA, Stephen L. Scott, Caltech/OVRO, N. S. Amarnath, U Maryland, Athol Kemball, U Illinois/NCSA, Harold Ravlin, U Illinois/NCSA

The Combined Array for Research in Millimeter-wave Astronomy (CARMA) will merge the Owens Valley Radio Observatory (OVRO) millimeter array and the Berkeley-Illinois-Maryland Association (BIMA) millimeter array and will add the new Sunyaev-Zeldovich Array (SZA) to form a powerful telescope located in California which will be fully operational in 2005.

CARMA will consist of 23 heterogeneous antennas as well as numerous other hardware and software subsystems.

The CARMA Database System (CDBS) will store the monitor data from all the subsystems as well as astronomical data and logging information. As a result, the CDBS must fulfill a number of challenging requirements, including:

* It must ingest of order 35,000 monitor point samples every half-second, which is equivalent to about 1 TB/week.
* Because of the high ingest rate, its storage overhead must be minimal.
* It must handle and quickly process a wide variety of complex queries.
* It must store data for the entire lifetime of the instrument (a few decades) since users must be able to access data for any point in the telescope's past. The CDBS must also operate for several years after the telescope has stopped operating to allow archival research.

After evaluating three freely available RDBMSs, we have chosen MySQL as our backend. The last requirement above has led us to develop a mostly RDBMS-neutral API using ODBC. This strategy will allow us to easily swap in another ODBC-compliant RDBMS backend in the future should the need arise.

This presentation will discuss some of the details of this complex system.

P3.1.12 Database of Instrumental Characteristics in a Real Observatory

Alberto Micol, ESA/ESO Space Telescope European Coordinating Facility, Diego Sforna, ESA/ESO Space Telescope European Coordinating Facility

The build of a central repository of the various characteristics of the science instruments contingent of HST and ESO Observatories is here illustrated.

This repository, along with smart user interfaces, is going to be a fundamental tool for many and equally important reasons. If properly constructed, and by taking into account that the users of such tool are not necessarily knowledgeable of all the internal intricacies of the Observatory Data Flow System, this tool can be used by the various groups within the Observatory operations (from the User Support to the Quality Control, down to the Archive), and by the general public.

While discussing the various problems a data provider has to solve on the way, an illustration of the first implementation will be shown.

Indications on how this tool will be beneficial to the efficiency of the Observatory operations will be given. Last but not least, applications in the Virtual Observatory context will be suggested especially with reference to a parallel presentation by D. Sforna entitled 'A metadata layer to enable VO-compliant access to the ESO/ST-ECF Archive'.

P3.1.13 The EVLA Software System Overall Design

Tom Morgan, NRAO, Socorro, NM, Kevin Ryan, NRAO, Socorro, NM, Ken Sowinski, NRAO, Socorro, NM, Boyd Waters, NRAO, Socorro, NM

We review the overall design of the software system being constructed to operate the Expanded Very Large Array (EVLA). The scope of the system spans proposal preparation and submission to the NRAO through archiving of final output data products in forms suitable for export to data reduction packages and for data mining. Many VLA functions currently being performed by hand will be replaced by automated subsystems. Other software-based procedures as well as monitor and control systems and data archives are being redesigned to handle enhanced correlator configurations and higher output data volumes, and high bandwidth receiver, sampler and data transfer technologies being incorporated into the antennas. Each of the major software subsystems will be addressed with respect to inputs, outputs, functionality and connectivity to data bases and other system components.

P3.1.14 Building a General Purpose Beowulf Cluster for Astrophysics Research

Matthew Phelps, Harvard - Smithsonian Center for Astrophysics

The CfA is a large, multi-discipline astrophysics research facility run jointly by the Smithsonian Institution and Harvard University. The challenges of designing and deploying a high performance, Linux based, Beowulf cluster for use by many departments and projects are covered. Considerations include hardware, infrastructure (space, cooling, networking, etc.), and software (particularly scheduling systems).

P3.1.15 A new Data Flow System for the Nordic Optical Telescope

Saskia Prins, Nordic Optical Telescope, Peter M. Sorensen, Nordic Optical Telescope

The Nordic Optical Telescope (NOT) is a 2.6 m telescope for observations at optical and near-infrared wavelengths. It is operated jointly by Denmark, Finland, Iceland, Norway, and Sweden on La Palma, Canary Islands.

To remain scientifically interesting for the Nordic community in the 8-10m era, the NOT focuses its resources on flexible instrumentation and near real-time data assessment to accommodate quick-response observations. To this end we are currently implementing a new data flow system which includes new FITS formats and headers, quality control tools and science data reduction pipelines, metadata databases and several mostly low-level upgrades to the observing system, i.e., new TCS, SOAP for subsystem access, RAID storage array, gigabit switch.

P3.1.16 The CARMA Correlator

Kevin P. Rauch, Univ. MD, Rick Hobbs, Caltech, David W. Hawkins, Caltech

We describe the software development effort for the first-light correlator of the Combined Array for Research in Millimeter-wave Astronomy (CARMA), a merger of the Owens Valley Radio Observatory (OVRO) and Berkeley-Illinois-Maryland Association (BIMA) millmeter arrays expected to reach first-light in 2005. The digital hardware relies on programmable logic devices (FPGAs), signal processors (DSPs), and Linux host computers acting in concert to produce baseline visibility data. We summarize the hardware configuration, the distribution of computational tasks among components, and describe the tools and techniques used to program these devices.

P3.1.17 PESO - The Python Based Control System of Ond\v{r}ejov 2m Telescop

Petr Skoda, Astronomical Institute, Jan Fuchs, Astronomical Institute, Jaroslav Honsa, Astronomical Institute

Python has been gaining a good reputation and respectability in many areas of software development. We have chosen Python after getting the new CCD detector for coud\'e spectrograph of Ond\v{r}ejov observatory 2m telescope. The VersArray detector from Roper Scientific came only with the close source library PVCAM of low-level camera control functions for Linux, so we had to write the whole astronomical data acquisition system from the scratch and integrate it with the current spectrograph and telescope control systems.

The final result of our effort, PESO (Python Exposure System for Ond\v{r}ejov) is a highly comfortable GUI based environment allowing the observer to change the spectrograph configuration, chose the detector acquisition mode, selecting the exposure parameters and monitoring the exposure progress.

All the relevant information from control computers is written into FITS header by the PyFITS module and the acquired CCD frame is immediately displayed in SAO DS9 window using the XPA calls.

The GTK based front end design was drawn in the Glade visual development tool, giving the shape and position of all widgets in single XML file, which is used in Python by simple call of PyGlade module.

We describe our experience with design and implementation of PESO, stressing the easiness of quick change of GUI together with the capability of separated testing of every module using the Python debugger IPython.

P3.1.18 The robotic Liverpool Telescope data flow model

Robert. J. Smith, Liverpool John Moores University on behalf of the LT Operations Group, Liverpool John Moores University

The Liverpool telescope is a 2m fully robotic telescope own by Liverpool John Moores University (UK) and operating at Observatorio Roque de los Muchachos in the Canary Islands. It is entirely autonomous and is obtaining data for a wide variety of science programmes without any nightly supervision, dynamically scheduling the observations according to both weather conditions and certain `fairness' criteria. In order to take advantage of the telescope's capabilities; flexibility, efficiency and very rapid response to targets of opportunity; an integrated procedure of data handling had to be developed. For example, imaging data of photometric fields are reduced whilst the telescope acquires the next field, providing almost real-time data quality assessments, which can be fed back into the scheduling decisions. We will describe the full data flow model from phase II observation specification, though data acquisition, automated pipeline reduction and distribution to the appropriate astronomer using secure transfer protocols.

P3.1.19 Managing ESO's Operations Data Flow systems

Dieter Suchar, European Southern Observatory, Benoît Pirenne, European Southern Observatory

The Systems Operations Support Group at ESO is responsible for the management of ESO's entire data flow from the proposal submission (Phase I) through to archive data distribution. Only three FTEs are handling over 100 operations critical Linux and Sun servers.

Our systems administration is not only focused on the standardization and automated installations of our systems. We are mainly focusing on predictable systems operation. Our challenges include:

- A storage-centric computer system: Mainly disks, but also other storage media are the core of the setup: CPUs, RAM and interfaces are simply there to support them.

- Single computer systems have been replaced with grid blades and clustered systems, multiplying the amount of machines to maintain.

- The density of power consuming equipment has dramatically increased and the existing facilities had to undergo extensive electrical upgrades. Air-conditioning had to be replaced due to the corresponding growth in heat dissipation.

The challenge is not anymore to smoothly install and manage operational servers, but to fulfill the growing storage needs of the astronomical user community in time, while simultaneously adapting the infrastructure. This contribution will describe how we are addressing those challenges.

P3.1.20 Gbit/s I/O at the European VLBI data processor - clustered silicon and long fiber

Harro Verkouter, JIVE, Huib Jan van Langevelde, JIVE, Friso Olnon, JIVE, Arpad Szomoru, JIVE, Steve Parsley, JIVE, Mike Garrett, JIVE

In the European VLBI Network, a number of high-data-rate radio-astronomy R&D projects are taking shape. We deal with Gbits/s on both on the input-side and the output side of the EVN MkIV VLBI correlator at JIVE.

The EVN MkIV correlator is capable of handling 16Gbit/s input data rate and a maximum output data rate of 1.28Gbit/s (160MByte/s).

At the input side of the correlator various developments are well underway. Hard disk recorded VLBI observations are supported. Another project deals with enabling real-time imaging VLBI, which was demonstrated to work on April 28th 2004.

Work is in progress to enable capturing the full output rate of the MkIV correlator. This will yield high time/frequency resolution data, which will enable more scientific output from VLBI datasets as the field-of-view will be enhanced by a factor of 100 or more.

In order to match the expected data rate and volume, a combined storage/compute cluster was installed. The hardware is tuned for processing large volumes in a streaming fashion: InfiniBand is used as cluster interconnect.

Research is going on in the field of distributed processing. New algorithms will be developed to enable parallel processing of VLBI data. This effort is done in conjunction with other radio-astronomical institutes in an EU funded effort.

An archive with access methods like the Astrophysical Virtual Observatory is developed when we are able to generate fine-tuned (in time/frequency resolution and/or desired field-of-view) datasets from the very large raw datasets. Raw datasets typically will be on the order of 4TB for a canonical VLBI experiment.

P3.1.21 Telescope Automation and Remote Observing Software (TAROS)

Greg Wilson, Adam Czezowski, Gary Hovey, Mark Jarnyk, Jon Nielsen , Bill Roberts, Kim Sebo, Dione Smith, Annino Vaccarella,Peter Young

TAROS is a system that will allow for the Australian National University telescopes at Siding Spring Observatory to be operated in fully automated, remote interactive or local interactive observing modes. The TAROS system is operated by a Java front-end GUI and we have employed the use of several Java technologies - such as Java Message Service (JMS) for communication between the telescope and the remote observer, Java Native Interface to integrate existing data acquisition software written in C++ (CICADA) with new Java programs and the JSky collection of Java GUI components for parts of the remote observer client. In this poster the design and implementation of the TAROS system is described.

P3.1.22 Optical Camera with high temporal resolution to search for transients in the wide field

Ivan Zolotukhin, Sternberg Astronomical Institute, Moscow, Russia, Gregory Beskin, Special Astrophysical Observatory, Nizhnij Arkhyz,, Karachai-Cherkessia, Russia, Anton Biryukov, Sternberg Astronomical Institute, Moscow, Russia, Sergey Bondar, Research Institute for Precision Instruments, Moscow, Russia, Kevin Hurley, Space Science Laboratory, University of California, Berkeley, Evgeny Ivanov, Research Institute for Precision Instruments, Moscow, Russia, Sergey Karpov, Special Astrophysical Observatory, Nizhnij Arkhyz, Karachai-Cherkessia, Russia, Elena Katkova, Research Institute for Precision Instruments, Moscow, Russia, Alexey Pozanenko, Space Research Institute, Moscow, Russia

One of the poorly understood aspects of gamma-ray bursts (GRBs) is prompt optical emission. For its successful registration one have to observe independently from space borne gamma-ray telescopes and use optical instruments with wide field of view.

To realize this approach wide field optical camera with high temporal resolution has been developed. Main objective (15 cm diameter, F/1.2) of the camera projects 21x16 degrees area onto image intensifier photocathode (gain - 150, scaling factor - 0.22). Special optics transfers image from intesifier's output to the TV-CCD camera with frame frequency of 7.5 Hz (0.13 sec exposure time).

Observational data (17 Mb/sec transfer rate) is transmitted to the local PC which broadcasts it through the LAN to the storage computer equipped with RAID and to the PC for real-time processing. The chosen network configuration allows to keep raw data obtained during 8 hours observing set with 0.13 s resolution (volume up to 0.5 Tb) and detect on the fly both stationary and moving OTs up to V < 11.5. System provides classification of detected objects, magnitude and coordinates with 35 accuracy in 0.4 sec (3 frames).

The camera operated in test mode from May to November, 2003; since November 2003 the camera monitors HETE-2 WXM field of view. Expected numbers of GRBs hitting in field of view of the camera which will be registered simultaneously with WXM (HETE-2) and BAT (SWIFT) are 1-2 and 4 per year, correspondingly.

P3.2.1 INTEGRAL reprocessed data and public data archive

Mathias Beck for the INTEGRAL Science Data Centre

This contribution will describe the format and benefits of the refined INTEGRAL data format. We will show how to access the data and the reduction software and give examples how to run the INTEGRAL specific software as well as general tools relevant in high energy astrophysics.

In addition, we will give information on the access to the data archive during the first three months after the availability of public data on July 19, 2004.

P3.2.2 The Design of the W. M. Keck Observatory Archive

G. Bruce Berriman, Michelson Science Center and Infrared Processing and Analysis Center, Thomas Bida, Lowell Observatory, David Ciardi, Michelson Science Center, Albert Conrad, W. M. Keck Observatory, Anastasia Laity, Infrared Processing and Analysis Center, Jeffrey Mader, W.M. Keck Observatory, Naveed Tahir-Kheli, Infrared Processing and Analysis Center, Hien Tranh, W. M. Keck Observatory

The Michelson Science Center and the W. M. Keck Observatory are building an archive that will serve data obtained at the Keck Observatory. The archive has begun operations and is ingesting level 0 (uncalibrated) observations made with the recently upgraded High Resolution Echelle Spectrometer (HIRES); these observations will be publicly accessible after expiration of a proprietary period. Observatory staff have begun using the archived data to determine the long-term performance of the HIRES instrument. The archive is housed at the Michelson Science Center (MSC)and employs a modular design with the following components:1. Data Evaluation and Preparation: images from the telescope are evaluated and native FITS headers are converted to metadata that will support archiving; 2. Trans Pacific Data Transfer: metadata are sent daily by email and ingested into the archive in a highly fault tolerant fashion, and

FITS images are written to DVD's and sent to IPAC each week; 3. Science Information System: inherited from the NASA/IPAC Infrared Science Archive, it provides all the functionality needed to support database queries and processing of requests; and a web-based 4. User Interface, a thin layer above the information system that accepts user requests and returns results. The design offers two major cost saving benefits: it overcomes the geographical separation between the telescope and the archive and enables development at Keck and at MSC to proceed independently; and it permits direct inheritance of the IRSA architecture.

P3.2.3 Software Architecture of the Spitzer Space Telescope Uplink/Archive

Joe Chavez, Annie Hoac, Tatiana Goldina, Xiuqin Wu

The Spitzer Science Center (SSC) provides a set of science user tools to support planning and archive access via the Internet. We will present the software architecture and design principles that underlie the Uplink/Archive subsystem of the SSC. Included in the discussion will be a review of the original Uplink architecture as presented in P1-59 ADASS 1999 and the evolutionary changes for the current deployment. The Archive subsystem is based on the same set of core components used in the Uplink subsystem but is based on Web services technology to allow open access to the Archive. Web services technology provides a basis for searching the archive and retrieving data products.

P3.2.4 Data Processing Discovery Agents in Gemini Science Archive

Adrian Damian, Canadian Astronomy Data Centre, Norm Hill, CADC, Severin Gaudet, CADC, David Bohlender, CADC, Sharon Goliath, CADC, Geoffrey Melnychuk, CADC, Colin Aspin, GEMINI

Gemini Science Archive (GSA) is a new science archive developed by the Canadian Astronomy Data Center (CADC) to

provide the scientific community with tools for effective on-line access to data collected by the Gemini telescopes. The first release of the GSA basic archive (aka Phase 1) is scheduled for fall 2004, with other subsequent releases containing advanced capabilities to follow after.

Discovery agents (DA) represent one of the key design concepts of the advanced GSA archive. A discovery agent is a software component designed to search the archive catalog in order to discover patterns of input data and perform relevant processing on the selected data. Typically, a DA uses a specified criteria to scan the catalog for associations and process the selected data in order to extract previously unknown information .

DAs are at the core of the data mining functionality associated with the GSA archive. They play an important role in achieving the GSA's goal of implementing an effective archive that boosts scientific productivity and ensures that maximum value was extracted from the expensive-to-obtain observational data. A number of DAs are going to be implemented in the future releases of GSA, part of the advanced capabilities features. They will perform tasks such as: generation of previews and other data products, identification of data associations etc. The paper provides detailed description of the DA framework in GSA.

P3.2.5 The D4A Digitiser

Jean-Pierre De Cuyper, Royal Obs. Belgium Lars Winter, Hamburg, Germany

The D4A (Digital Access to Aerial- and Astro-photographic Archives) project aims to acquire the necessary know-how, hardware and software to digitise the astro-photographic collections of the Royal Observatory of Belgium and the aerial-photographic collections of the National Geographic Institute and the Royal Museum of Central Africa in collaboration with AGFA-Gevaert. The D4A digitiser under construction consists of a granite based Aerotech ABL 3600 open frame air bearing XY positioning system, with custom build automatic plate holder assembly suited for mounting glass plates and film sheets up to 350mmx350mm. Extended with an automatic film roll transport system and a plate tray handling and storage assembly with plate rotator. The optical subsystem consists of a cooled CMOS camera from Vector International (12bit ADC) mounted to a Schneider Xenoplan telecentric 1:1 lens and a back light illumination system. The light source is a very bright LED (lifetime min. 50000h) computer controlled by a precision power supply for adjusting the exposure of each individual subimage.

The D4A digitiser is intended to measure astronomical (up to 350mmx350mm) as well as aerial (up to 240mmx204mm) photographs at a rate of at least 6 plates per hour. Special software is needed for handling the 18 GByte/h data rate with online processing and storage of all images.

The goal is to provide astrometrically and photometrically calibrated digital images with overlayed identified stars. The extracted information will be stored in a database. The digital images will be compressed lossless in order to reduce the storage size and time.

P3.2.6 FITS keyword database at the European Southern Observatory

Adam Dobrzycki, Nausicaa Delmotte, Jens Knudstrup, Benoît Pirenne, Nathalie Rossat, Andreas Wicenec, Stefano Zampieri

The European Southern Observatory (ESO) is one of the largest astronomical observatories in the world. It manages numerous telescopes, which use various types of instruments and readout detectors. The data flow process at ESO's observatories involves several steps: telescope setup, data acquisition, pipeline processing, quality control, archivisation, distribution of data to the users. Each of those steps provides information which is stored in FITS header keywords of the data products.

At present, definitions of those keywords are provided in "Data Dictionaries; there is one Dictionary for each context, i.e. instrument, telescope system, observatory, etc. Because of the complexity of the ESO system, a large number of dictionaries (presently approaching a hundred) need to be maintained. Also, the format of the dictionaries does not allow for storage of information that can be very helpful, e.g. whether keyword is mandatory or optional, keyword dependencies, allowed ranges, etc. This information is critical for automated back-end data quality control, an essential step in data flow from the observatory as large as ESO.

The above considerations clearly show a need for a centralised keyword database including both the traditional dictionary definitions and the supporting information. We present the design of such a database, to be implemented at ESO.

P3.2.7 Overview of the Gemini Science Archive

Severin Gaudet, Canadian Astronomy Data Centre, Adrian Damian, CADC, Norm Hill, CADC, David Bohlender, CADC, Sharon Goliath, CADC, Geoffrey Melnychuk, CADC, Colin Aspin, GEMIN

Gemini Science Archive (GSA) is a new science archive developed by the Canadian Astronomy Data Center (CADC) to provide the scientific community with tools for effective on-line access to data collected by the Gemini telescopes. The first release of the GSA basic archive (aka Phase 1) is scheduled for early fall 2004, with other subsequent releases containing advanced capabilities to follow after.

Benefiting from the knowledge and experience of other telescope archives built and operated by CADC, GSA was also designed to be one of the first scientifically effective archive of ground-based operations. As such, from the very beginning GSA was considered an integrated part of the planning, observation, calibration, data reduction and data distribution process that occur at Gemini. This fundamental shift in data archiving, allowed for the implementation of some novel approaches to the data handling and processing:

1. support for data produced directly by telescopes' instruments as well as other related data (weather, logs etc)
2. database schema to separately store meta-data and catalogs
3. easy to edit data dictionary to specify data formats, ranges, transformations etc

4. rapid availability of data due to electronic transfer of data from the Gemini telescopes
5. generation of derived data for easy access by the user or for compatibility with other telescopes

The paper addresses all these aspects of GSA design and also lays down new features planed for future releases.

P3.2.8 Meta-Data Management in the Gemini Science Archive

Sharon Goliath, CADC, Adrian Damian, CADC, Severin Gaudet, CADC, Norm Hill, CADC, David Bohlender, CADC

The Gemini Science Archive (GSA) is a new science archive developed by the Canadian Astronomy Data Center (CADC) to provide the scientific community with tools for effective on-line access to data collected by the Gemini telescopes.

CADC has developed a data dictionary (DD) for the GSA to meet the goal of providing dependable and consistent archive data. This DD defines the rules that govern data manipulation on obtaining meta-data from a variety of data sources. The DD is specified in XML, and so provides flexibility while reducing source code maintenance. The paper will discuss the rules that may be specified for the data sources and individual meta-data attributes, the role of the data dictionary in populating the GSA, verification of the consistency of the dictionary itself, and possible future improvements in the data dictionary.

P3.2.9 The Architecture of the NASA/IPAC Infrared Science Archive (IRSA)

John C. Good, Infrared Processing and Analysis Center, Caltech, Anastasia Alexov, Infrared Processing and Analysis Center, Caltech, G. Bruce Berriman, Infrared Processing and Analysis Center, Caltech, Nian-Ming Chiu, Infrared Processing and Analysis Center, Caltech, Thomas H. Jarrett, Infrared Processing and Analysis Center, Caltech, Mih-seh Kong, Infrared

Processing and Analysis Center, Caltech, Anastasia Clower Laity, Infrared Processing and Analysis Center, Caltech, Serge Michel Monkewitz, Infrared Processing and Analysis Center, Caltech, Naveed D. Tahir-Kheli, Infrared Processing and Analysis Center, Caltech, Saille Warner-Norton, Infrared Processing and Analysis Center, Caltech, Angela Zhang, Infrared Processing and Analysis Center, Caltech

IRSA employs a component-based architecture based on a set of portable, standalone data access and processing programs. New archive interfaces are built by calling these modules through a standard command-control protocol. The inherent re-use of this approach makes the system highly extensible and dramatically simplifies supporting a wide range of projects.

For instance, the IRSA tools are also being used for the WM Keck Observatory archive, interfaces to the Spitzer Space Observatory archive and support for the COSMOS Hubble Treasury Program. The architecture seamlessly supports National Virtual Observatory protocols via output mode switches in the data access modules.

All this requires a remarkably small number of components. The external interfaces are: Gator, for general catalog queries; Atlas, for spatial queries of mixed collections of tables and image sets; and an interface currently under development to support complex multi-table database queries. These in turn utilize the data access modules described above including: ISISQL; for DBMS access; and Montage, for general image reprojection and mosaicking.

Two signed JAVA applets are used for presentation and data fusion across IRSA and external archives (another form of interoperability):Oasis, for display of images and tables; and QtPlot(an extension of Berkeley's PtPlot) for general XY plotting.

The architecture is rounded out by two integration tools: RADAR, which provides general region statistics and can organize the execution of all the above services for a complete view of a region of the sky; and ROME, a general request management system developed as part of the NVO.

P3.2.10 FITS Image Cutout

Vivek Haridas, JHU, William O'Mullane, JHU. Alex Szalay, JHU, Tamas Budavari, JHU, Maria Nieto-Santisteban, JHU, Samuel Carliles, JHU, Aniruddha Thakar, JHU

The Sloan Digital Sky Survey is on its way to build a very large map of the universe. We look at implementing web services that help in getting image cutouts from the FITS files in the SDSS archive. These cutouts would be provided in the traditional FITS format as well as in the gray scale jpeg image format.

Some amount of complexities exists in creating cutouts in both the formats. The process of getting cutouts in the FITS format would require World Coordinate System values to be recalculated for the newly created FITS files according to the positioning of the cutouts in the original archive images. Also, there is a possibility of the cutouts not existing completely in any single FITS file in the archive.

The jpeg image cutouts provided are actually a gray scale visual rendering of the FITS Image cutouts. In addition to the above complexities, sky noise calibration, arcsinh scaling and gamma corrections are performed on the image data.

P3.2.11 Mining the Astronomical Archives for the "Tadpole" Galaxy

T.H. Jarrett, IPAC/Caltech, Ismael Perez Fournon, IAC, Gordon Stacey, Cornell U, Donovan Domingue, IPAC/Caltech

We present results from an investigation of the peculiar "Tadpole Galaxy" (UGC 10214).This work is motivated by the Spitzer Wide-area Infrared Extragalactic Survey (SWIRE)observations of the ELAIS N1 region, which includes UGC10214and its local environment. The Spitzer imaging data comprise the four mid-infrared channels of IRAC and the 24-micron MIPS observations. To supplement these mid-infrared data, we have

(1) carried out a ground-based program to image the Tadpole at high resolution in the optical and near-infrared windows, and
(2) systematically searched through publicly available astronomical mission archives. Our presentation includes the data mining process used to collect and organize the complex and disparate data sets that come from mission/observatory archives, notably the IRSA (infrared), MAST (HST), HEASARC (high energy), NED &LEDA (extragalactic) and CDS/VizieR (general) archives. We spotlight the use of three virtual observatory tools: IRSA's inventory and data exploration services RADAR and ATLAS, and the NVO data inventory service.

P3.2.12 Astronomical Computing within Data Archives

Adrian Pope, Alex Szalay, Jim Gray

As astronomical datasets continue to grow in size we begin to face serious issues with transporting data from archives to computing resources used to do calculations with the data. We have begun to attack this problem by performing as many calculations as possible within the archive. One example is the use of a SQL Server archive with Hierarchical Triangular Mesh (HTM) spatial indexing software to pixelize the low redshift Sloan Digital Sky Survey (SDSS) Main Galaxy Sample (MGS) before performing Large Scale Structure (LSS) analyses. The resulting counts-in-cells data is an order of magnitude smaller than the input catalog, dramatically reducing the amount of data transported to external programs. We also present a system for analytically calculating the spatial relationships between spherical polygons within an archive that has been used to analyze the complex geometry of SDSS survey footprint and masks, a necessity for preparing samples of data suitable for detailed LSS studies. The software was written (by Alex Szalay and Jim Gray) in the high level Transact-SQL language, with the use of the HTM library. Finally we describe the creation and utilization of a Monte Carlo realization of the SDSS for studying the effects of calibration errors and masking on LSS studies. The realization contains hundreds of millions of random points and was created using a cluster of SQL Server machines.

P3.2.13 The WFCAM Science Archive and other WFAU work

Mike Read, Institute for Astronomy University of Edinburgh, Nigel Hambly, Institute for Astronomy University of Edinburgh, Eckhard Sutorius, Institute for Astronomy University of Edinburgh, Bob Mann, Institute for Astronomy University of Edinburgh, Ian Bond, Institute for Astronomy University of Edinburgh

The Wide Field Astronomy Unit (WFAU) of the Institute for Astronomy, University of Edinburgh will host the data taken with WFCAM, the new wide-field infrared imager on UKIRT. This, the WFCAM Science Archive, is the final stage in the data flow from telescope to user.

Starting in 2005 WFCAM will carry out five main surveys over a 7 year period. Observations will typically generate 100Gb of data per night. The goal of the WFCAM Science Archive is to allow fast and flexible science exploitation of this large volume of data.

This drives the design of the archive which is described in detail. Areas covered include hardware choices, curation tasks and user access.

Archive scalability is also a key issue as the WFCAM Science Archive forms the first phase VISTA Data Flow System (VDFS) science archive. WFAU's involvement with the VDFS and other developments are briefly discussed.

P3.2.14 ESO Archive Services

Nathalie Rossat, Nausicaa Delmotte, Nathalie Fourniol, Benoît Pirenne, Adam Dobrzycki

The ESO/ST-ECF Science Archive is a joint collaboration of the European Southern Observatory (ESO) and the Hubble Space Telescope - European Coordinating Facility (ST-ECF). The archive provides access to data from both the Hubble Space Telescope (HST) and the European Southern Observatory (ESO).

The ESO Science Archive offers a number of on-line information services relevant to ESO users and to the community at large. In this paper, we will go through the different steps that enable a user to retrieve raw and processed data, from the query form to the actual processing. The concept of "Header On the Fly (Hotf)" that "repairs" possibly wrong data headers with meta-data from a database is also presented, together with the fact that calibrated data from ESO instruments will shortly be made available.

The emphasis in this poster will be on how the meta data describing each ESO observation is kept in a relational database in instrument specific tables (IST), and how this information is used to correct data upon distribution. The three avenues for making ESO science products available to the VO -and linked to their respective raw frames- will also be explored.

Keywords: Databases, Query Form, Information Retrieval, Header On The Fly, FITS, Calibrated Science Products

---

P3.2.15 Recent developments of the ISO Data Archive

Alberto Salama, ESA, ISO Data Centre Team, ESA

ESA's Infrared Space Observatory (ISO) performed 30,000 scientific observations, covering all areas of astronomy, as the world's first true infrared observatory. Two spectrometers, a camera and an imaging photo-polarimeter jointly covered wavelengths from 2.5 to 240 microns. Launched in 1995, ISO was operational until May 1998. All data had been re-processed with the end-of-mission calibration to populate the first homogeneous ISO Data Archive, which opened to the community in December 1998.

Through the ensuing four years of the Post-operations Phase, ESA's ISO Data Centre developed and refined the ISO Data Archive to offer the ISO data to the worldwide astronomical community, and together with the several National Data Centres, worked to fill the archive with the best systematically processed and calibrated data products which could be achieved for the huge ISO database. These products allow users to select from the archive data sets of interest for deeper study with interactive analysis tools.

During ISO's Active Archive Phase, which runs until December 2006, the ISO Data Centre continues to work with active National Data Centres to bring the archive into its final shape. In this talk we will review the major upgrades of the ISO Data Archive accomplished in this phase: (i) the incorporation of Highly Processed Data Products, the result of dedicated projects focused on cleaning the pipeline products from residual instrumental artifacts, (ii) the detailed characterisation of each observation via a Data Quality Report and (iii) the integration within the Virtual Observatories.

P3.2.16 The NOAO Data Cache Initiative - Building a Distributed Online Datastore

Rob Seaman, NOAO Data Products Program, Irene Barg, NOAO DPP, Nelson Saavedra, NOAO DPP, Chris Smith, NOAO DPP, Nelson Zarate, NOAO DPP

The Data Cache Initiative (DCI) of the NOAO Data Products Program is a prototype Data Transport System for NOAO and affiliate facilities. DCI provides pre-tested solutions for conveying data from our large suite of instrumentation to a central online datastore. The heart of DCI is an extension of the Save-the-Bits safe store, running for more than a decade. The iSTB server has been simplified by the removal of STB's media handling functionality, and iSTB has been enhanced to remediate each incoming header with information from a database of NOAO instrumentation and an interface to the NOAO proposal database.

DCI was commissioned with data from telescopes on Kitt Peak and Cerro Tololo. Data at a particular site flow into a mountain wide data cache. The iSTB server runs on the KP and CT data cache machines, receiving connections from a few dozen supported instrument configurations. As each image in the queue is processed, keywords are added to each header such as proposal ID and principle investigator. A directory location is generated for the datastore given the telescope, date, etc. Checksums are computed and the file is compressed.

Each file is transferred from the mountain to the corresponding datastore at the Tucson or La Serena data centers using an rsync-based queue adopted from NCSA. From each data center, the files are transported to the other NOAO data center and also to NCSA for off-site storage using the Storage Resource Broker (SRB) of the San Diego Supercomputer Center. Thus we have three copies of each file on spinning disks or near-online. Major institutional users will be given access to the datastores.

P3.2.17 FITS Query Language

Eric Sessoms, NRAO

FITS Query Language (FQL) imposes a relational database model on collections of FITS files. The chief advantages of FQL are openness and ease-of-use. Openness is achieved through an ODBC (Open DataBase Connectivity) interface that makes FQL available to a wide variety of clients, including many applications that do not have available or easy-to-use FITS readers. Ease-of-use is achieved through an SQL (Structured Query Language) implementation that simplifies the correlation of data from multiple sources.

We show how the declarative nature of FQL can replace procedural data access code, allowing correct applications to be produced more quickly, and we discuss the role FQL plays in data analysis for the Green Bank Telescope (GBT).

P3.2.18 The SIP Convention for Representing Distortion in FITS Image Headers

David Shupe, Spitzer Science Center, Caltech, Mehrdad Moshir, Spitzer Science Center, Caltech, Richard Hook, Space Telescope European Coordinating Facility, STScI, Jing Li, Spitzer Science Center, Caltech, David Makovoz, Spitzer Science Center, Caltech, Robert Narron, Spitzer Science Center, Caltech

The SIP convention provides a means for representing non-linear geometric distortion as polynomial coefficients in FITS header keywords. This scheme has been incorporated into the imaging products of the Spitzer Science Center (SSC) and it is under consideration for wider use. The SIP keywords are interpreted by several standard analysis tools, including WCSTools, SAOImage, DS9, GAIA/Starlink AST, IDLAstro, and drizzle, as well as by the SSC's "Mopex" software. We describe the motivation for the SIP keywords, and provide definitions of how they are used together with standard FITS WCS keywords to relate pixel coordinates to the sky. Several examples illustrate the application of this scheme to actual distorted data.

P3.2.19 New Features for VO-Enabled Data Intensive Science with the SDSS Data Release 3

Ani Thakar, JHU, Alex Szalay, JHU, Jim Gray, Microsoft Research, William O'Mullane, JHU, Tamas Budavari, JHU, Maria Nieto-Santisteban, JHU, George Fekete, JHU, Nolan Li, JHU, Robert Lupton, Princeton

With the third data release (DR3), the SDSS Catalog Archive Server (CAS) serves up 3 Terabytes of catalog data for nearly 200 million celestial objects via the SkyServer web pages at skyserver.sdss.org.Several changes and new features have been added that facilitate data-intensive science with the data and advance VO standards and technologies.

1. The CasJobs batch query system unveiled at ADASS XIII now allows stored procedures and functions in the user's MyDB. This enables users to bring their program to the data rather than the other way around, which is far more efficient for data-intensive applications. We replicate user account details between installations to allow distributed CasJobs configurations. This form of replication raises issues that are crucial to the VO community. We are also attempting to secure the system using X509 certificates through WS-Security. This ties in with the distributed storage effort within the VO.

2. New color images in the Visual Tools based on the ASINH stretch that are deeper and more feature-rich. We briefly discuss the algorithm.

3. The traffic logging system is now quite detailed and extensive, allowing many useful and interesting statistics to be gathered. The Support Services specification proposed

for IVOA includes a logging section modeled on this approach.

4. A much faster HTM (spatial index) library with an upgraded interface.

5. An enhanced object crossid facility allowing users to run their own SQL query on the matching objects.

We briefly describe these features and list future enhancements anticipated with SQLServer Yukon and SDSS-DR4.

P3.2.20 SOLARNET – The federation of the Italian solar data archives

Cosimo Antonio Volpicelli, INAF-Osservatorio Astronomico Torino, Corrado Perna, INAF-Osservatorio Astronomico Monte Porzio – Roma, Alberto Cora , INAF-Osservatorio Astronomico Torino

We describe the implementation of the national project SOLARNET(SOLar ARchivesNETwork) aimed at federating the heterogeneous Italian solar archives into a VO framework like a single integrated database and providing users with tools to search for specific data sets and retrieve them. It interoperates using the SOAP XML Web Services exposed by each single site and managed via a unified Portal. Different user interfaces allow search of all participating data archives by different input parameters. The Portal provides three main functionalities. The first one, the Registration service allow at any new data archive or service to join the federation filling a simple web form, the second perform distributed astronomical queries on all nodes involved and collects the returned data in different formats, the third one will let you navigate interactively among the metadata documentation for all the data stored in the remote databases showing you what each exposed table or column means.

Currently the SOLARNET federation includes 5 data archives: SOLAR (SOho Long-term ARchive) European mirror of the SOHO NASA/GSFCarchive at Torino observatory, TSRS telescope archive SOLRA (SOLar Radio Archive) at Trieste observatory,

archive of RISE/PSPT (Precision Solar Photometric Telescope)network of Rome, Mauna Loa and Sacramento Peak observatories, DISCO (Data Interface to the Sun at Capodimonte Observatory) at Naples and Catania Solar Data Archive at Catania observatory .

SOLARNET implementation is using Sun's Java Web Services Developers Pack v1.3 with Apache Axis and JAXB to bind XML schema to java classes.

P3.2.21 A Conceptual Domain Model for the NOAO Science ArchivE

Phillip Warner, NOAO, Rafael Hiriart, NOAO, Frank Valdes, NOAO, Tod Lauer, NOAO, Sean Points, NOAO

The NOAO Science Archive (NSA) team is working on a design for the next generation science archive. The goal is to expand the current NSA data holdings by providing access to all raw data produced by observatories that are under the direction of NOAO and partner organizations, and to NOAO Pipeline-processed data products. Access will be restricted during the proprietary period (nominally 18 months), after which the data will be publicly available to all other entities, such as NSA users and the VO community.

A domain model for the NSA has been developed using classes and concepts from a document produced by Valdes (2000), which defines classes to describe the elements of an astronomical observation. The NSA domain model incorporates and generalizes these classes to provide a generic model for data acquisition. Other classes representing data products derived by pipelines, e.g., calibration, reduction, and analysis pipelines, are also included. Concepts modeled by the generalization of these classes include collections of elements involved in data acquisition, collections of element states (where relevant), high-level data classification, and internal and external resources.

P3.2.22 A Case Study of Applying Object-Relational Persistence in Astronomy Data Archiving

Shilong Stanley Yao(1), Rafael Hiriart(2), Irene Barg(1), Phillip Warner(1), Dave Gasson(1)

(1) {yao, ibarg, pwarner, dgasson}@noao.edu Data Products Program, NOAO/AURA, 950 North Cherry Ave., Tucson, AZ 85719
(2) rhiriart@ctio.noao.edu NOAO/CTIO, Casilla 603, La Serena, Chile

Astronomy data archiving manipulates large volumes of semantic-rich data. Meta-data, as the portal to the massive data products, is of special interests, because it contains numerous science parameters describing a large number of domain objects with complex structures and interactions. An effective modeling of meta-data is essential to the archive design. NOAO Science Archive (NSA) team is developing a comprehensive domain model to reach this goal. Java and an object model derived from the domain model well address the application layer of the archive system. However, since RDBMS is the best proved technology for data management, the challenge is the paradigm mismatch between the object and the relational models.

Transparent object-relational mapping (ORM) persistence is a successful solution to this challenge. In the data modeling and persistence implementation of NSA, we are using Hibernate, a well-accepted ORM tool, to connect the object model in the application layer and the relational model in the database layer. Thus, the database is isolated from the Java application. The application queries directly on objects using a DBMS-independent object-oriented query API, which frees the application developers from the low level JDBC and SQL so that they can focus on the domain logic. We present the detailed design of the NSA release 3 data model and object-relational persistence, including mapping, retrieving and caching. Persistence layer optimization and performance tuning will be analyzed. The system is being built on J2EE framework, so the integration of Hibernate into the EJB container and the transaction management are also explored.

Author Index

Sunday October 24, 2004		
10:00 – 13:00 And 13:00 – 16:00	Huntington Gardens and Museum Tour (1 hour guided, 1 ½ hour self-guided)	Huntington Gardens and Museum
16:00 – 21:00	Registration	Cordova
16:30 – 18:00	FITS (BoF.1)	Maple
19:00 – 21:00	Reception	Piazza

Monday October 25, 2004		
8:30 – 9:00	*Plenary Session* Dr. George Helou, IPAC Executive Director	Justine
9:00 – 10:15	*Imaging Algorithms (O1)*	Justine
10:30 – 11:00	*VOSpec: A tool for handling Virtual Observatory compliant Spectra (FM.1)*	Aspen
11:15 – 12:15	*Ground-based Observatories I (O2)*	Justine
12:15 – 13:45	Lunch	
12:30 - 13:45	*Focus Demonstration*	Aspen
13:45 – 15:15	*Great Space Observatories (O3)*	Justine
16:00 – 17:30	*Future Astronomical Data Analysis Environments* (BoF.2)	Maple
	Sky Indexation, Pixelization, and the VO (BoF.3)	Cypress
18:00 – 21:15	JPL Tour, Museum, Multimission Image Processing Laboratory	JPL

Tuesday October 26, 2004		
8:30 – 9:00	*Plenary Session* Dr. David Baltimore, President Caltech	Justine
9:00 – 10:30	*Great Space Observatories I (O4)*	Justine
10:30 – 11:00	*Focus Demonstration*	Aspen
11:15 – 12:15	*The Virtual Observatory: Handling the Data (O5)*	Justine
12:15 – 13:45	Lunch	
12:30 – 13:00	*Focus Demonstration)*	Aspen
13:45 – 15:15	*The Virtual Observatory: Grid (O6)*	Justine
15:30 – 16:00	*Spitzer Pride – Science User Tools (FT.4)*	Aspen
16:00 – 17:15	*The Virtual Observatory: Mining and Analysis (O7)*	Justine
18:30 – 21:30	Banquet	FrontRunner

Wednesday October 227, 2004		
8:30 – 9:00	*Plenary Session* Dr. Charles Elachi, Director JPL	Justine
9:00 – 10:30	*Ground-based Observatories II (O8)*	Justine
10:30 – 11:00	*Focus Demonstrations*	Aspen
11:15 – 12:15	*Detection Algorithms (O9)*	Justine
12:15 – 13:45	Lunch	
12:30 – 13:00	*Focus Demonstration)*	Aspen
14:15 – 16:00	*More Algorithms (O10)*	Justine
1545 – 1600	Closing Remarks	Justine
16:00	Conference Closed	