Montage: Architecture and Applications of an Astronomical Image Mosaic Engine

http://montage.ipac.caltech.edu/

Bruce Berriman, John Good and Anastasia Laity *IPAC*

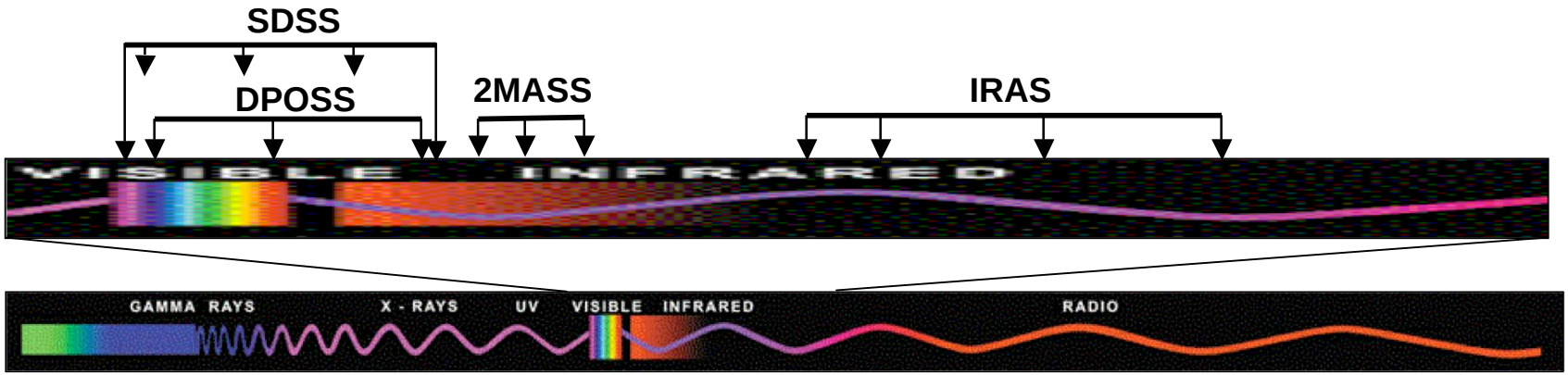
Joseph C. Jacob and Daniel S. Katz *JPL*

Ewa Deelman, Gurmeet Singh and Mei-Hui Su *ISI*

Thomas Prince *Caltech*

Science Drivers for Montage

- Imaging surveys are addressing fundamental questions in astronomy
- Montage overcomes limitations in maximizing scientific return from these image datasets
  - Many astrophysical structures are extended and structure on the sky contaminated by variable background radiation from instrument or sky
  - Survey results are published in widely varying coordinates, map projections, sizes and spatial resolutions => limits study of structures and sources over multiple wavelengths
- Montage delivers science grade image mosaics
  - Preserves astrometry and flux of input images
  - Delivers mosaic according to users specifications of projection, coordinates, spatial sampling, mosaic size, image rotation
  - Rectifies background radiation to a common level



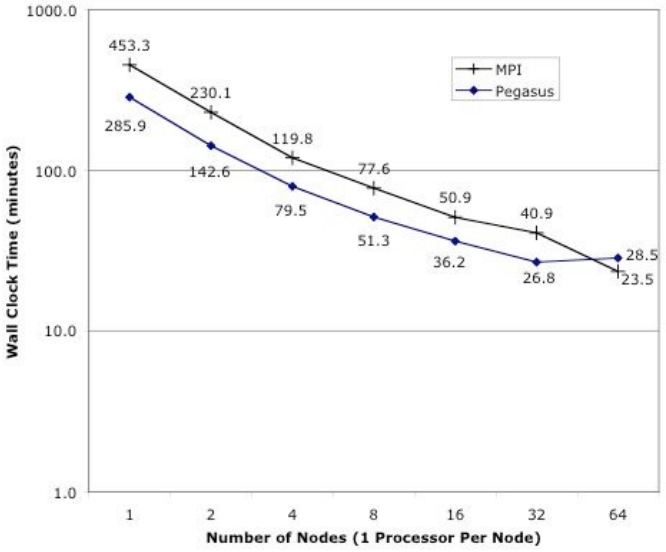
Montage Design Drivers

- Major design drivers
  - Support generation of data products by incorporating Montage into pipelines & processing environments
  - Support variety of hardware architectures used by astronomers: single processor; a cluster of multiple processors with a shared file system; multiple clusters, each with a shared file system; grid processing
  - Support on-request image mosaic requests submitted through a web interface
- Montage Architecture
  - Stand-alone "core" processing modules for performing steps in developing image mosaic
    - Reprojection of input images; rectification of background radiation to a common level; co-addition of rectified, reprojected images to form mosaic
    - Utilities for managing mosaics - tiling, creating browse products, …
    - Processing flow controlled by calls to simple executives
    - Code written in ANSI C for performance and portability

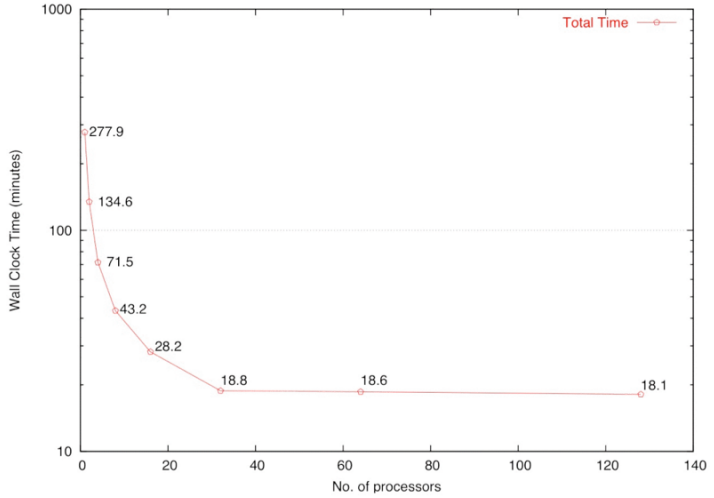
Design Drivers *cont*

- Montage Architecture - Parallel Technologies
  - Message Passing Interface (MPI)
    - Best performance
    - Requires a set of processors with a shared file system
  - Grid Technology
    - Very good performance
    - Built-in fault tolerance
    - Can use multiple sets of processors

Wall Clock Times to Build 6 x 6 Deg mosaic for MPI and Grid



6 x 6 degree 2MASS Mosaic, NCSA TeraGrid Pegasus Portal



Applications of Montage - I

- Generation of Science Data Products
  - Visible and Infrared Survey Telescope for Astronomy (VISTA) (under evaluation) NEW
  - Spitzer Space Telescope Legacy Projects
    - SWIRE - trace the evolution of dusty, star-forming galaxies, evolved stellar populations, and active galactic nuclei to redshifts of z ~ 3
    - GLIMPSE and MIPSGAL - infrared surveys of galactic plane to study global history of star formation and energetics of interstellar medium NEW
  - COSMOS Hubble Treasury Program
    - Multi-wavelength studies of the large scale distribution of matter in the Universe
  - The INT/WFC Photometric H-alpha Survey (IPHAS) of the Northern Galactic Plane
    - Deep survey in the red (Sloan R and I bands) NEW
  - National Virtual Observatory
    - All-Sky 2MASS mosaic

Applications of Montage -II

- Incorporation into Science Processing and Data Access Environments
  - Spitzer Space Telescope Outreach
    - Generating as educational products distributed through the *Cool Cosmos* website
    - Leverages Montage's capabilities of generating mosaics in uncommon projections, often the best one for E/PO products
  - IRSA (NASA's InfraRed Science Archive)

NEW

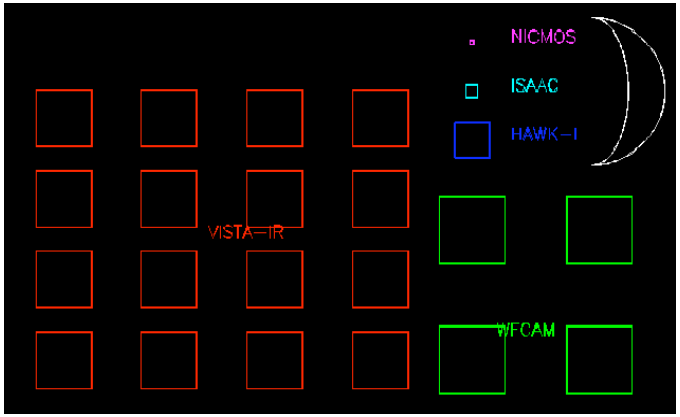
    - Supports services for accessing all-sky image surveys
    - Includes creation of browse products, image cutout services
  - NSF National Virtual Observatory (NVO) *Atlasmake*r project
    - Hyperatlas grid, a standard set of image parameters for displaying image data sets originally in different projections, coordinates, . . .
    - Leveraging Montage to deliver tools that will generate map images showing spatial coverage of astronomical data sets
  - UK Astrogrid Virtual Observatory
    - Montage adopted as a "key application" for end users to generate custom mosaics

Applications - VISTA

Visible and Infrared Survey Telescope for Astronomy (2007)

- 4-m class imaging survey telescope in Chile
- 67-million pixel camera (1.6 deg FoV)

Field of View of the 16 2048 x 2048 arrays



Evaluation of Montage

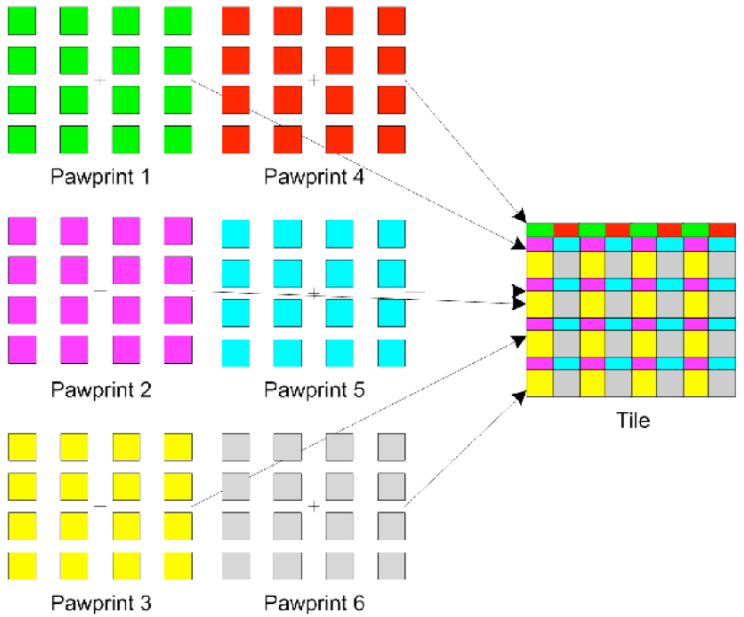
•Six stepped exposures fill a "Tile" with at least two exposures

•Automated pipeline processes each "pawprint",

• Montage is under evaluation as part of post-pipeline backend to stitch the pawprints together to create the tile

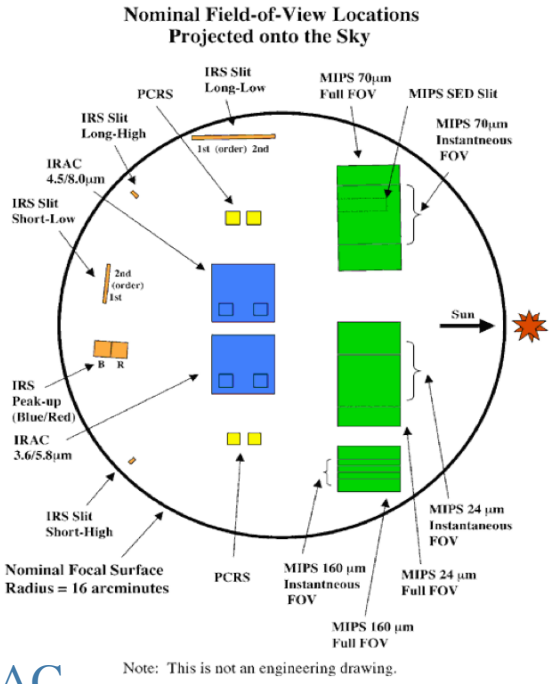
- Compute the image geometry to high precision and co-add the individual images
- Compensate for variability in sky conditions
- Reprojection of supporting observations for quality assurance

Observing Mode



GLIMPSE and MIPSGAL: Complementary Surveys of the Galactic Plane

Survey	Instrument	Bands (μm)	Field-of-View (arcmin)
GLIMPSE	IRAC	3.5, 4.5, 5.8, 8.0	5.2. x 5.2
MIPSGAL	MIPS	24	5.4 x 5.4
		70	5.25 x 2.6
		160	0.5 x 0.5



Generate "multi-wavelength" images of the Galactic as a resource for studying star formation - process IRAC, MIPS, Midcourse Space Experiment (8.3, 12.1, 14.6 and 21.3 μm) and 2MASS (1.2, 1.6 and 2.2 μm)

Three-color section of the Galactic Plane measured by the IRAC

Applications -2MASS All Sky Mosaic

Why Build An All-Sky Mosaic?

- Value-added 2MASS product combines overlapping images and includes deep survey data
- Demonstrate applicability of computing at scale to astronomical image data sets
  - pathfinder for future missions
- Demonstrate management of data sets at scale
- Shakedown TeraGrid hardware at SDSC

Input : 4,121,440 files, 2 MB in size (32-bits)

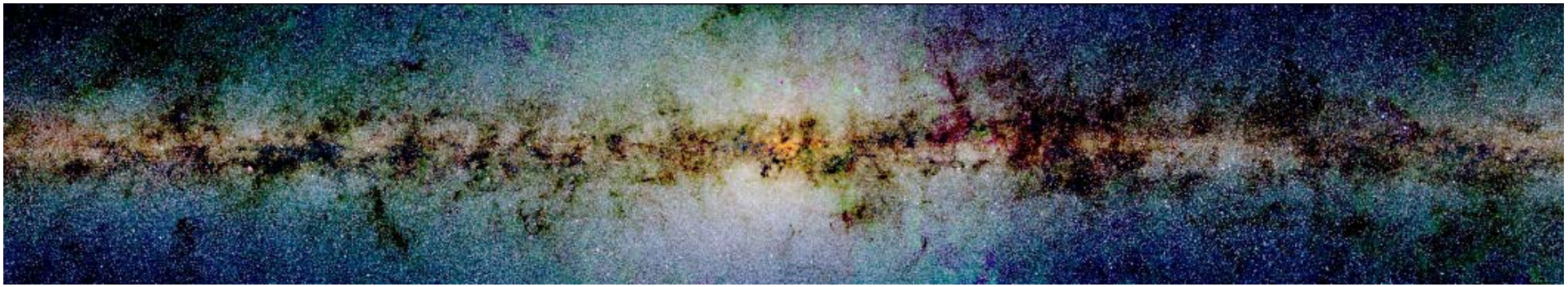
Output

- 1734 x 4 GB plates 6 deg on a side for 3 bands (64 bits)
  - Each plate tiled into a 12x12 array of 26-MB files
- The total size is about 20 TB in 750,000 tiles.

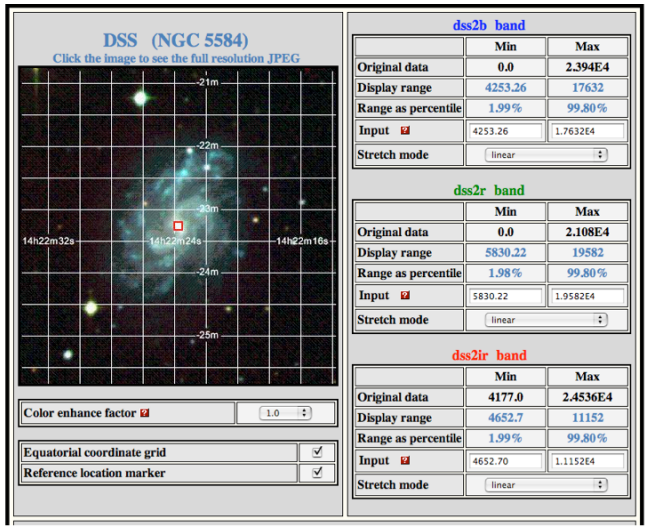
Processing Statistics

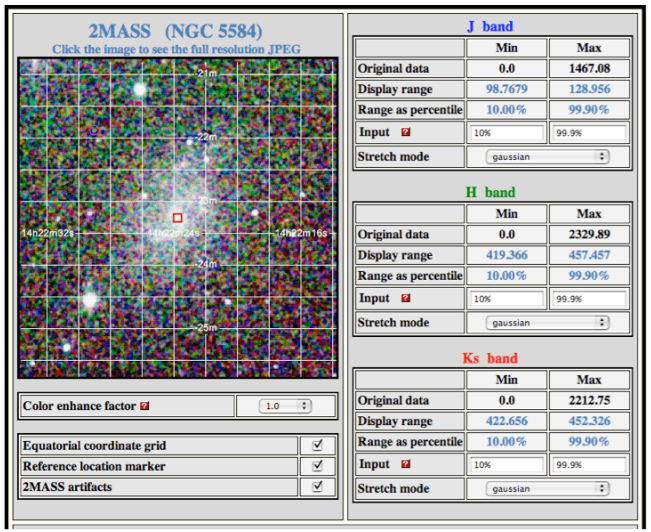
- MPI Python module managed the process
- 13,000 CPU hours
- Files on mosaic boundaries opened multiple times
  - 14 TB in 6,275,494 files opened

Below: Section of 3-color 2MASS mosaic along the Galactic plane



Applications - All Sky Image Access





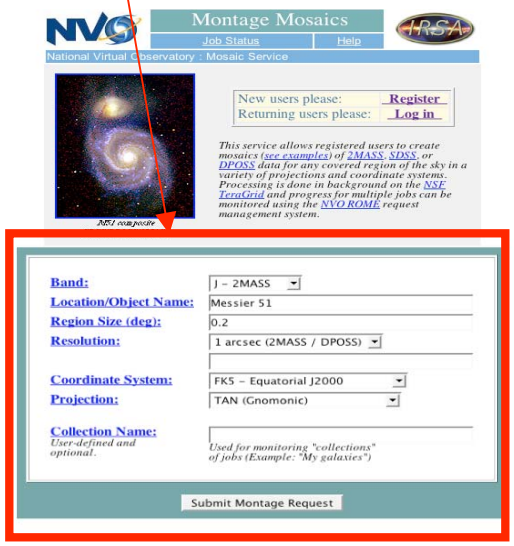
http://irsa.ipac.caltech.edu/applications/FinderChart

- IRSA's Finder Chart Service
  - Creates on-demand 3-color images of three major sky surveys - SDSS, 2MASS, DSS
- Quick Look Quality Assurance of fields astronomers are observing
  - Are moving objects real or artificial?
  - Are objects of unusual color inherent in the object or the field it is in?

Left: 3-color images of the the Galaxy NGC 5584 in the 2MASS and DSS surveys

An On-Request Compute Service

Enter mosaic request



Monitor job status

Request ID	Collection	Status	Message	Data	Result page	Submit Time
82	bruce	COMPLETED	Mosaic complete	Data Available	Result page	Fri Dec 16 13:58:57 PST 2005
83	bruce	COMPLETED	Mosaic complete	Data Available	Result page	Fri Dec 16 15:18:26 PST 2005
84	bruce	ERROR	Invalid password.			Fri Dec 16 15:19:08 PST 2005
85	bruce	COMPLETED	Mosaic complete	Data Available	Result page	Fri Dec 16 15:51:29 PST 2005
86	bruce	ERROR	Invalid password.			Fri Dec 16 15:51:53 PST 2005

View messages

Messages from application: Grid Montage (process ID 19902)

Grid Montage run started.

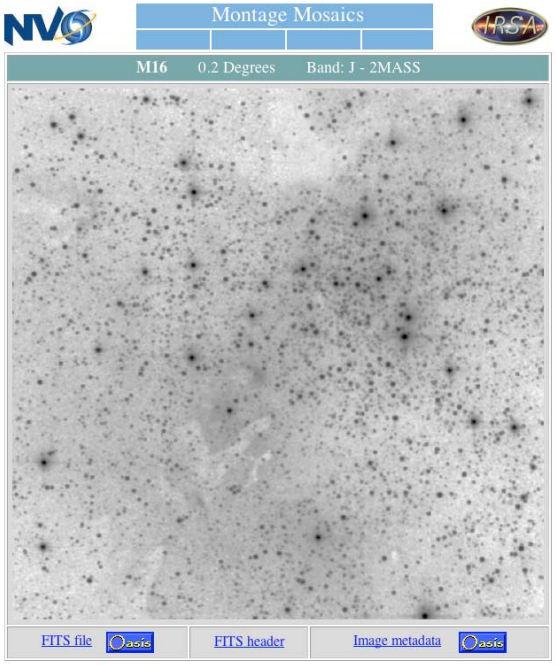
Header generated successfully

GLOBUS submission successful (status: https://tg-login1.ncsa.teragrid.org:39031/1861/1134770339/
15 images in overlap region
Retrieved 15 images from archive (15 sec elapsed)
Reprojected 10 images so far (88 sec elapsed)
Reprojected 15 images (0 failed, 0 did not overlap region) (125 sec elapsed)
29 overlap regions
Overlap analysis done for 10 images so far (133 sec elapsed)
Overlap analysis done for 20 images so far (139 sec elapsed)
Background correction overlap analysis complete (147 sec elapsed)
Background corrected 10 images so far (151 sec elapsed)
Images background corrected (154 sec elapsed)
Final mosaic created (155 sec elapsed)
JPEG of final mosaic generated (156 sec elapsed)
Data copied to permanent store (157 sec elapsed)
Mosaic complete (168 sec elapsed)

Results: here

Mosaic complete.

View and download mosaic

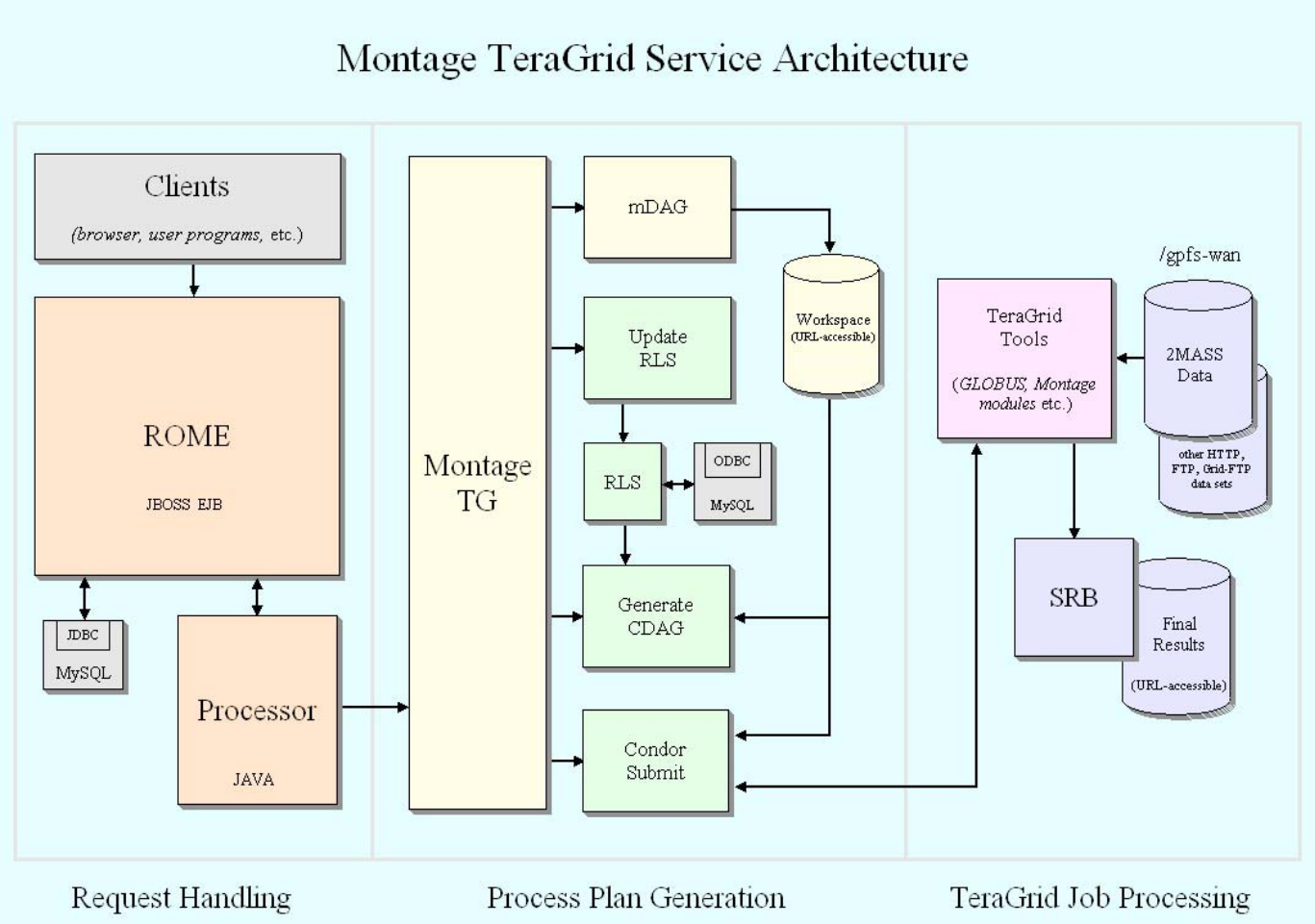


This mosaic took 157 seconds.

For more information on using Oasis for data display and analysis, click here

Processing history

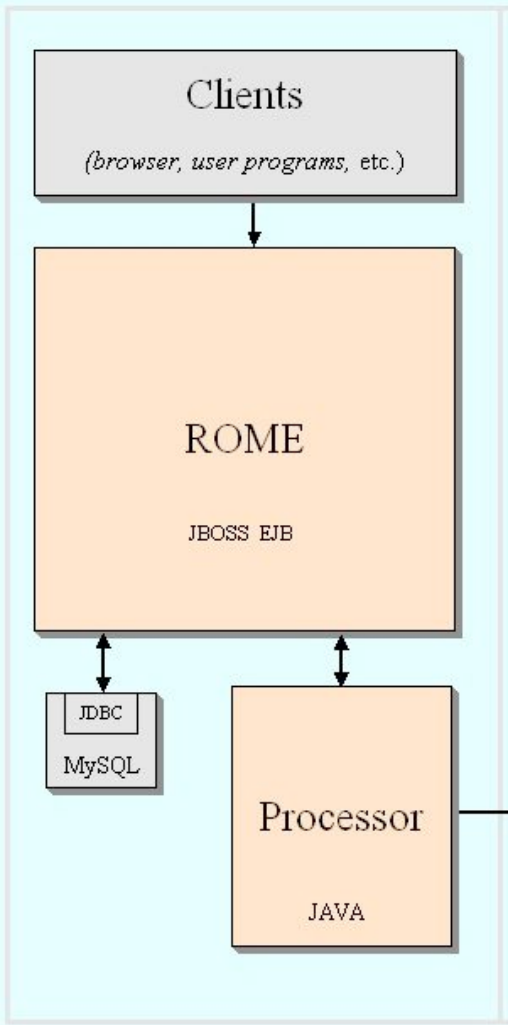
On-Request Service Architecture



IPAC TeraGrid

Request Handling

- Request Management Environment (ROME).
  - Developed at IPAC for the NVO
  - Enterprise Java Beans (EJBs)
  - Accept processing requests from users (via servelets)
  - Manage processing queues
    - Ensure that resources distribute processing among users
  - Handles monitoring and user notification

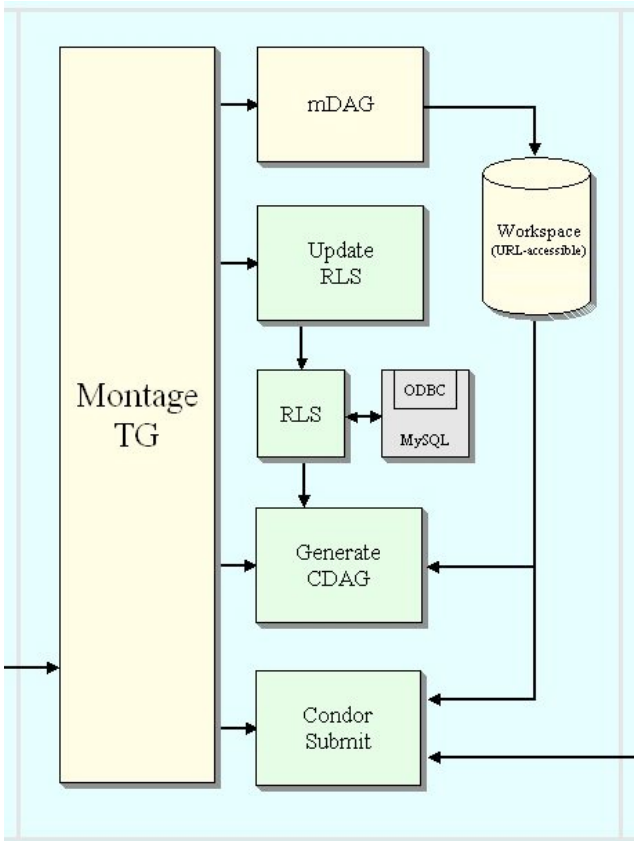


Request Handling

Process Plan Generation

- Montage application implemented via a Condor DAG (Directed Acyclic Graph)
  - Generate an abstract DAG
  - Process input parameters, set up storage space
  - Query resources to discover images needed
  - Pegasus processes abstract DAG into a Condor-specific DAG targeting TeraGrid resources.
    - Pegasus in turn utilizes the Globus Replication Location Service (RLS) to find physical location of file resources

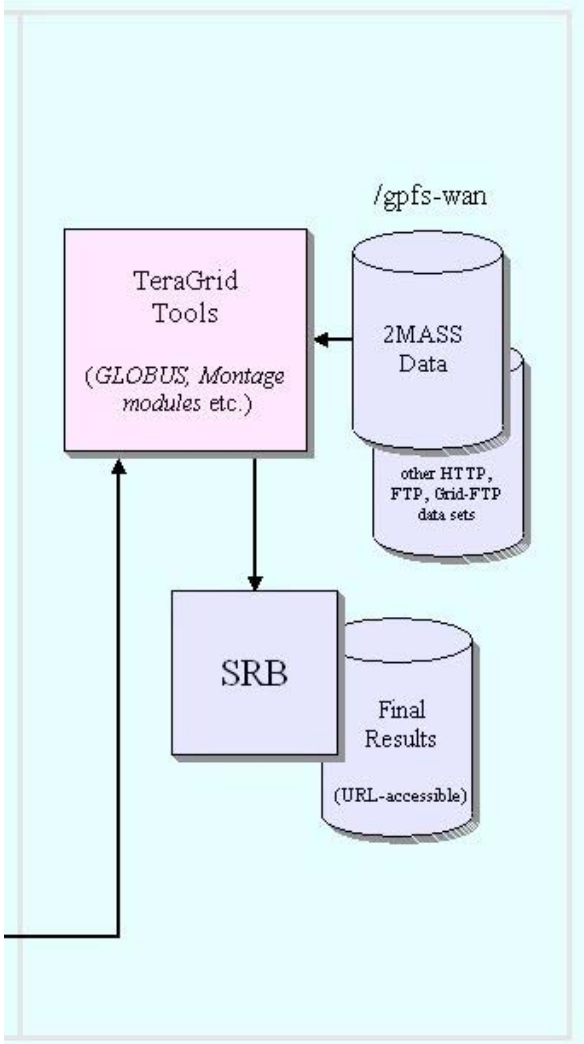
Custom to Montage



Process Plan Generation

TeraGrid Processing

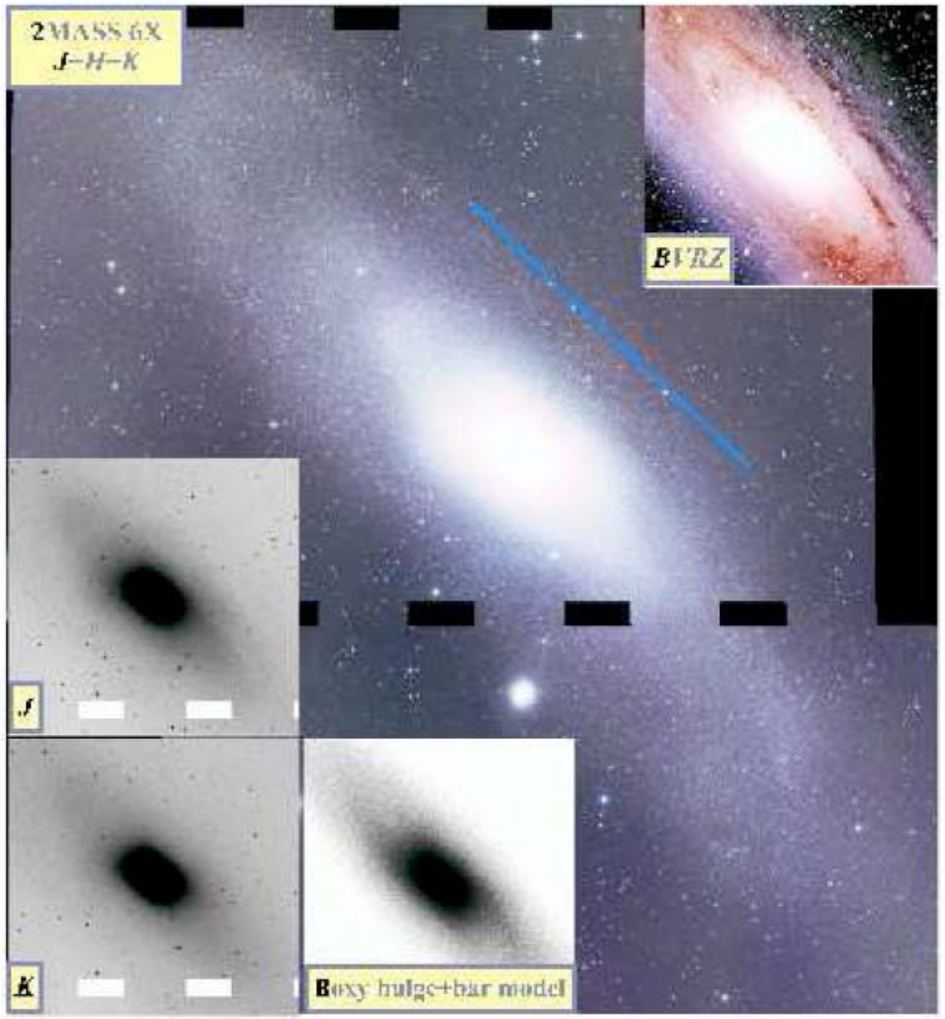
- TeraGrid tools. This environment processes the mosaic
  - Marshals compute resources and intermediate file storage space,
  - Handles errors by rescheduling subtasks, reports results back through the submitting machine to ROME and thence to the user.
- Uses copy of the 2MASS data stored at SDSC in a distributed file system (/gpfs-wan) as input
- Places results in a URL-accessible location within SDSC's Storage Resource Broker (SRB) system.
- Flexible design - backend can be, e.g., a Condor pool
  - IPAC installing 20-machine cluster of Dell Dual Core 2-processor 3.2 GHz Xeon compute servers at IPAC.



TeraGrid Job Processing

Science Result - Verification of a Bar in the Spiral Galaxy M31

- Beaton et al. Ap J Lett in press
- Montage used to generate mosaic of M31 galaxy
- Background rectification brings out “boxy bulge” - signature of bar



Summary

- Use of Montage in astronomy is growing
  - Generation of science products
  - Performing scientific research
  - Performing “pathfinder” studies in demonstrating processing of images at scale
- On-request service under evaluation by astronomers and will be publicly deployed in Fall 2006
- To learn more, visit the project web site at *http://montage.ipac.caltech.edu/*

Right: 3-color mosaic of the star formation region NGC 6357
“War and Peace Nebula”

Combines MSX and 2MASS data sets; 2.5 degree square mosaic - 9,000 pixels on a side