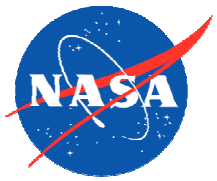
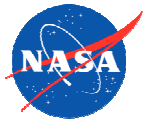
The Montage Architecture for Grid-Enabled Science Processing of Large, Distributed Datasets

Joseph Jacob, Daniel Katz, and Thomas Prince (JPL)
Bruce Berriman, John Good, and Anastasia Laity (IPAC)
Ewa Deelman, Gurmeet Singh, and Mei-Hui Su (ISI)

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



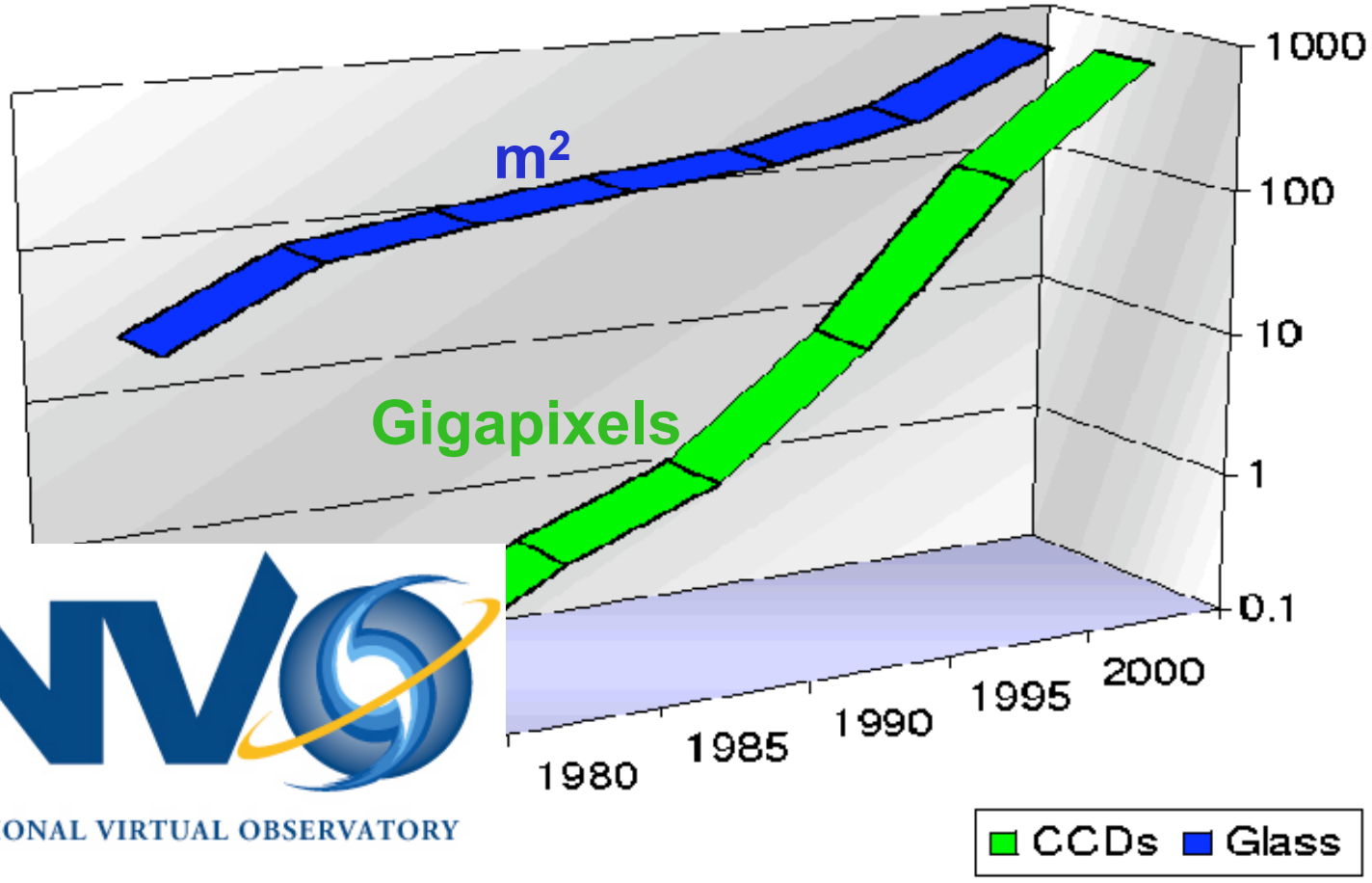


Earth/Planetary/Space Selected Shared Traits

- Large, distributed datasets
- Image mosaics a necessity
- Need for Grid Computing

The Data Avalanche!

Growth in Aperture & Focal Plane
Of Institutionally Managed Observatories



NATIONAL VIRTUAL OBSERVATORY

Selected Image Archives

IRAS	1 GB	1 arcmin	All Sky	4 Infrared Bands
DPOSS	4 TB	1 arcsec	All Northern Sky	1 Near-IR, 2 Visible Bands
2MASS	10 TB	1 arcsec	All Sky	3 Near-Infrared Bands
SDSS-DR2	5 TB	0.4 arcsec	3,324 square degrees (16% of Northern Sky)	1 Near-IR, 4 Visible Bands

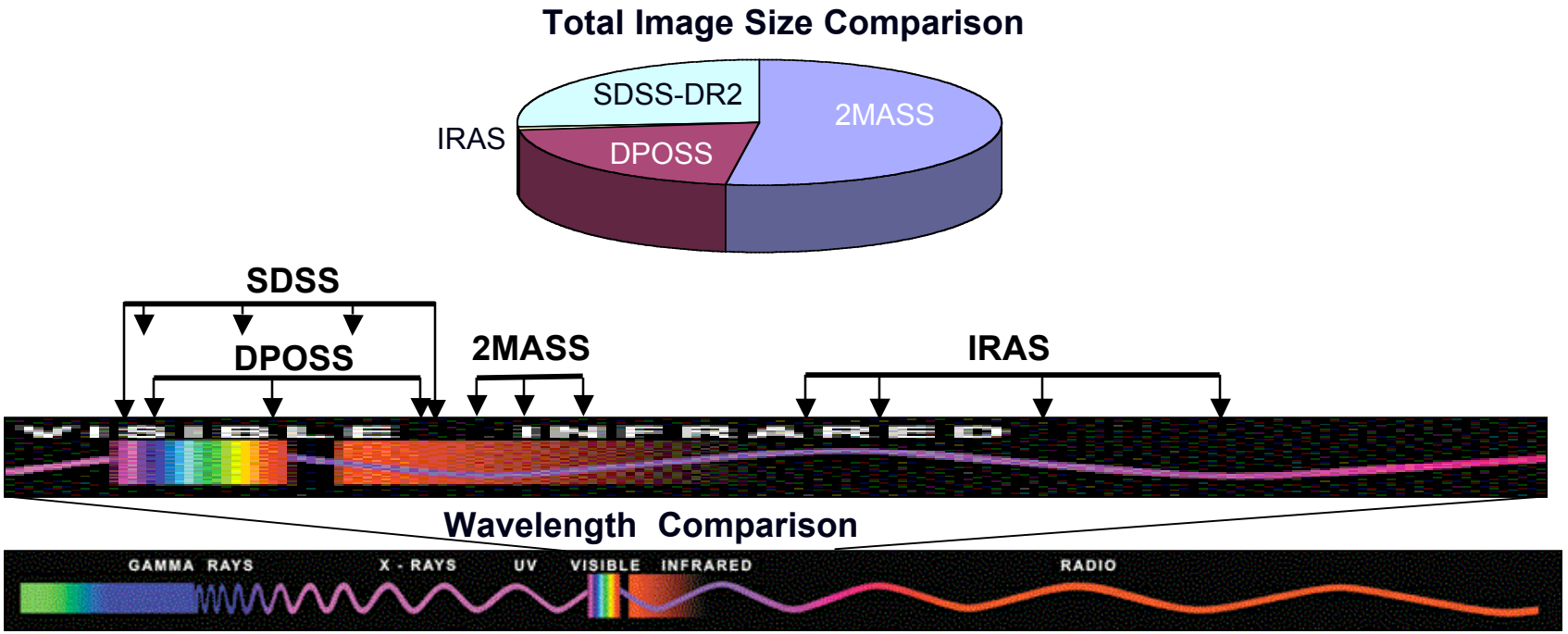
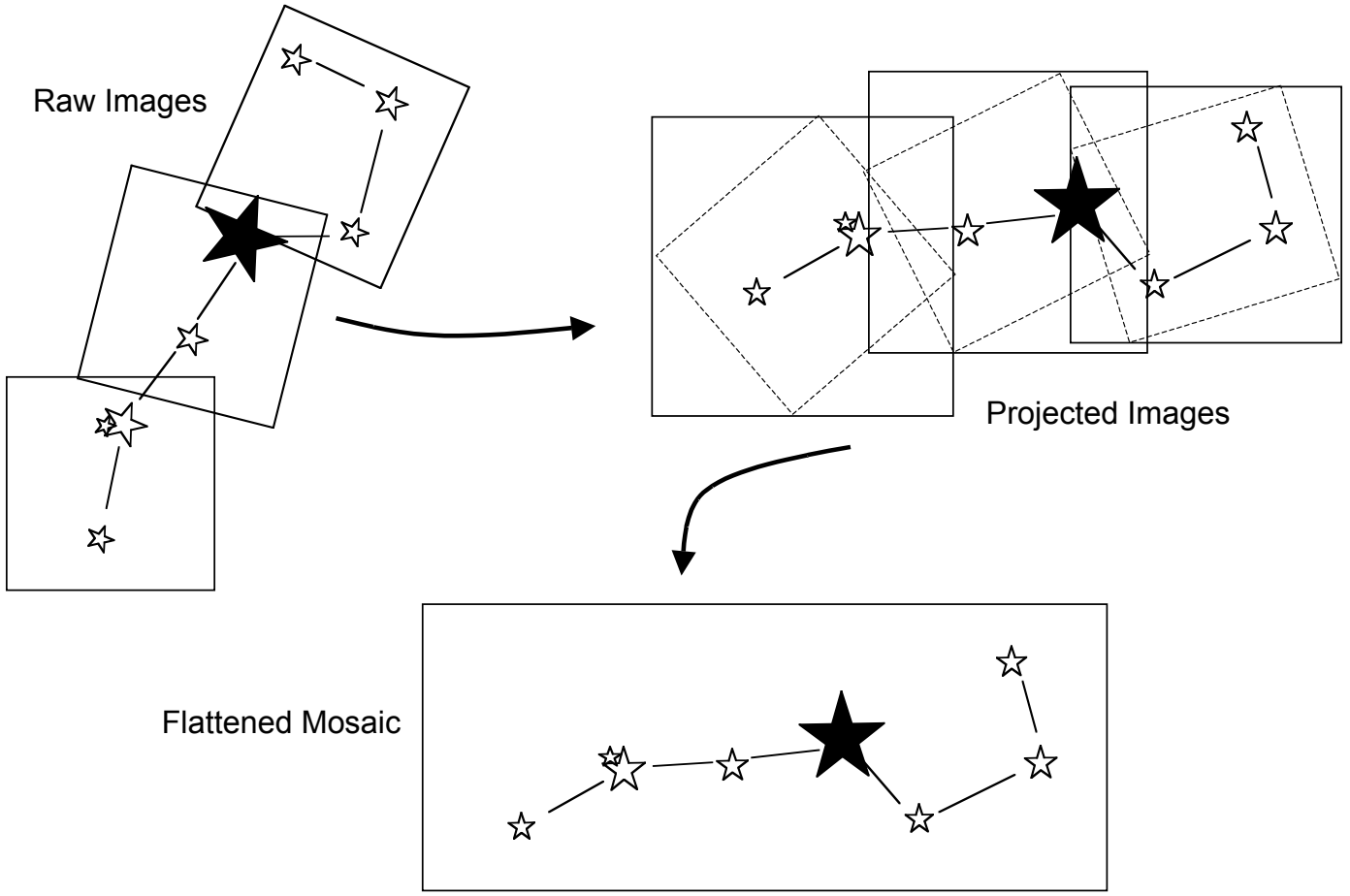


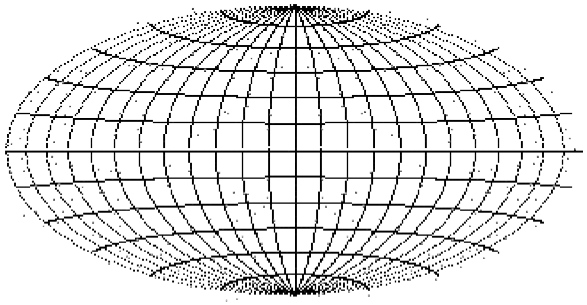
Image Reprojection and Mosaicking

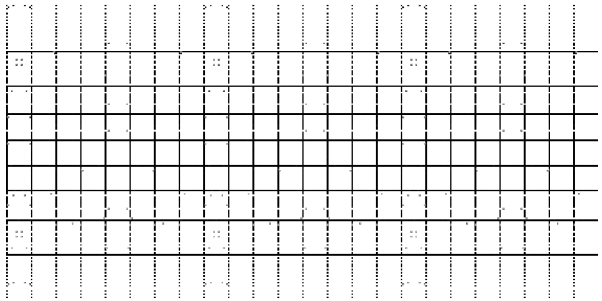
FITS format encapsulates the image data with keyword-value pairs that describe the image and specify how to map pixels to the sky

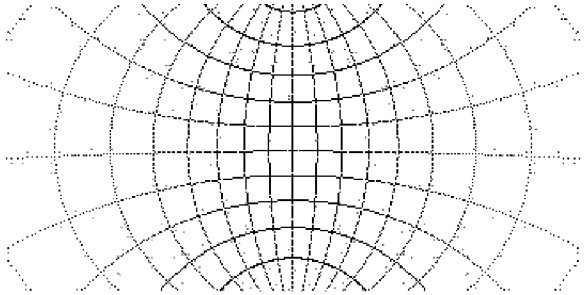


World Coordinate System (WCS)

- Coordinate Systems: Rotation of axes in 3-D space; e.g., Galactic, Ecliptic, J2000 Equatorial, B1950 Equatorial, etc.
- WCS projections: How coordinates map to each axis; e.g., TAN (Tangent plane), CAR (Cartesian), etc.

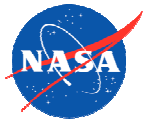






Science Drivers for Mosaics

- Many important astrophysics questions involve studying regions that are at least a few degrees across.
  - Need high, uniform spatial resolution
  - BUT cameras give high resolution or wide area but not both => need mosaics
  - required for research and planning
- Mosaics can reveal new structures & open new lines of research
- Star formation regions, clusters of galaxies must be studied on much larger scales to reveal structure and dynamics
- Mosaicking multiple surveys to the same grid – **image federation** – required to effectively search for faint, unusual objects, transients, or unknown objects with unusual spectrum.

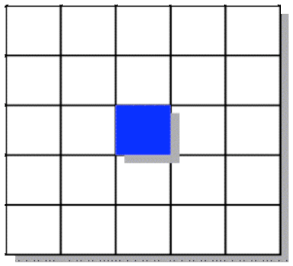


What is Montage?

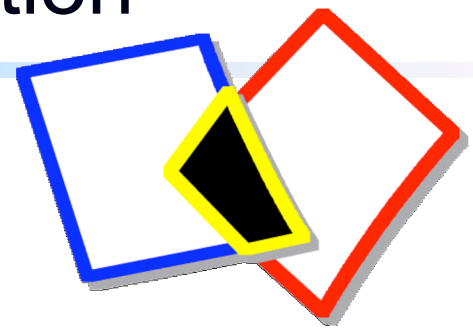
- Delivers custom, science grade image mosaics
  - User specifies projection, coordinates, spatial sampling, mosaic size, image rotation
  - Preserve astrometry & flux
  - Background modeled and matched across images
- Modular “toolbox” design
  - Loosely-coupled engines for Image Reprojection, Background Matching, Co-addition
    - Control testing and maintenance costs
    - Flexibility; e.g custom background algorithm; use as a reprojection and co-registration engine
  - Implemented in ANSI C for portability
- Public service will be deployed on the *TeraGrid*
  - Order mosaics through web portal

Montage_v1.x Reprojection

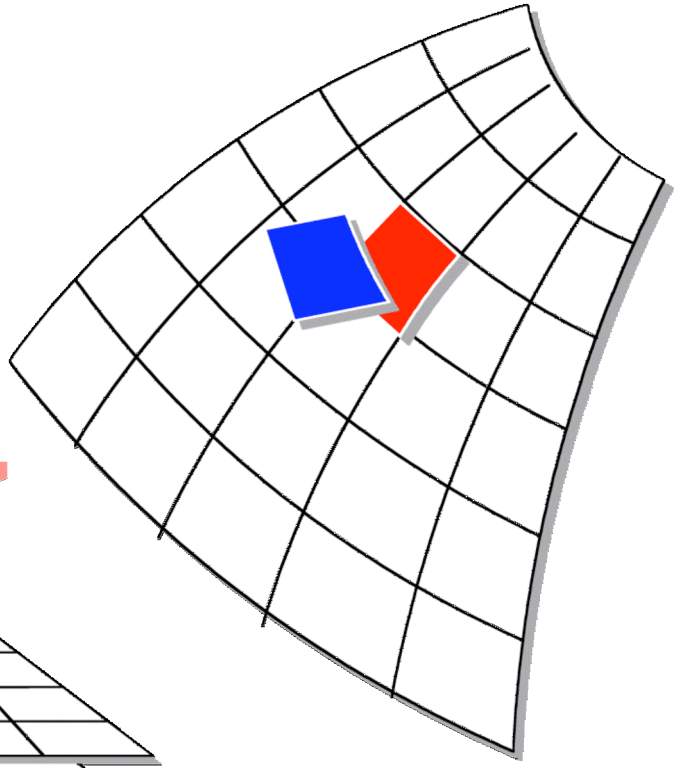
Arbitrary Input Image



Central to the algorithm is accurate calculation of the area of spherical polygon intersection between two pixels (assumes great circle segments are adequate between pixel vertices)



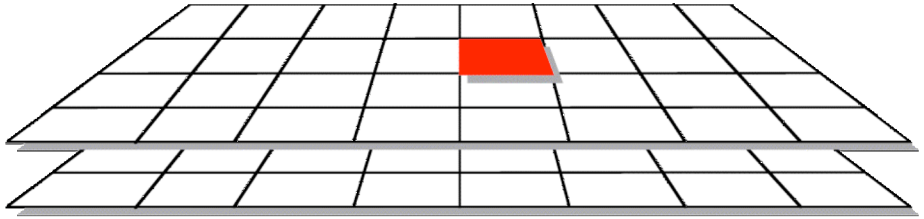
Input pixels projected on celestial sphere



```
SIMPLE =                    T /
 BITPIX=                  -64 /
 NAXIS  =                    2 /
 NAXIS1=                  3000 /
 NAXIS2=                  3000 /
 CDELT1=                -   3.333333E-4 /
 CDELT2=                -   3.333333E-4 /
 CRPIX1=                1500.5 /
 CRPIX2=                1500.5 /
 CTYPE1='RA---TAN'
 CTYPE2='DEC--TAN'
 CRVAL1=               265.91334 /
 CRVAL2=               -29.35778 /
 CROTA2=                   0. /
END
```

FITS header defines output projection

Output pixels projected on celestial sphere



Reprojected Image

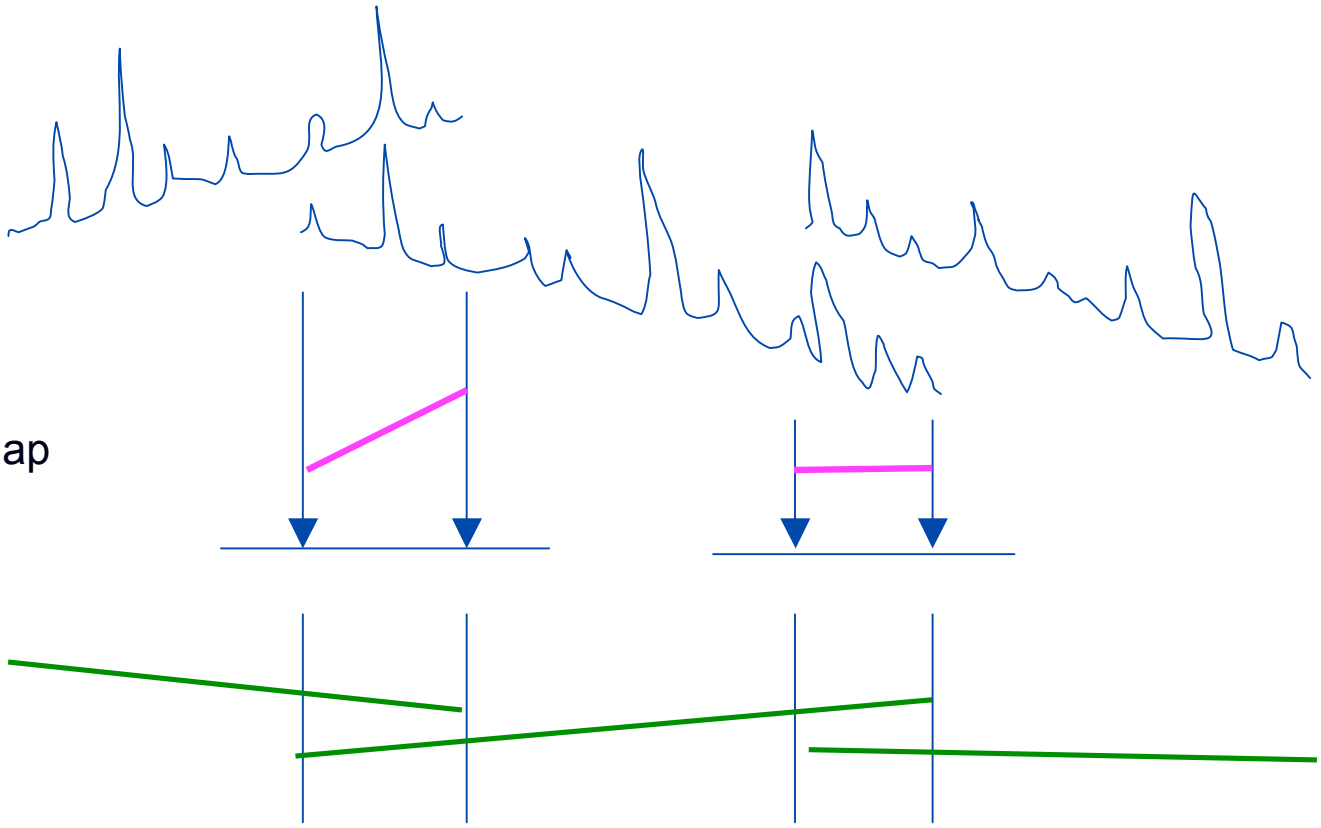
Montage Background Matching

Project pixels to output projection

Fit ramps to overlap differences

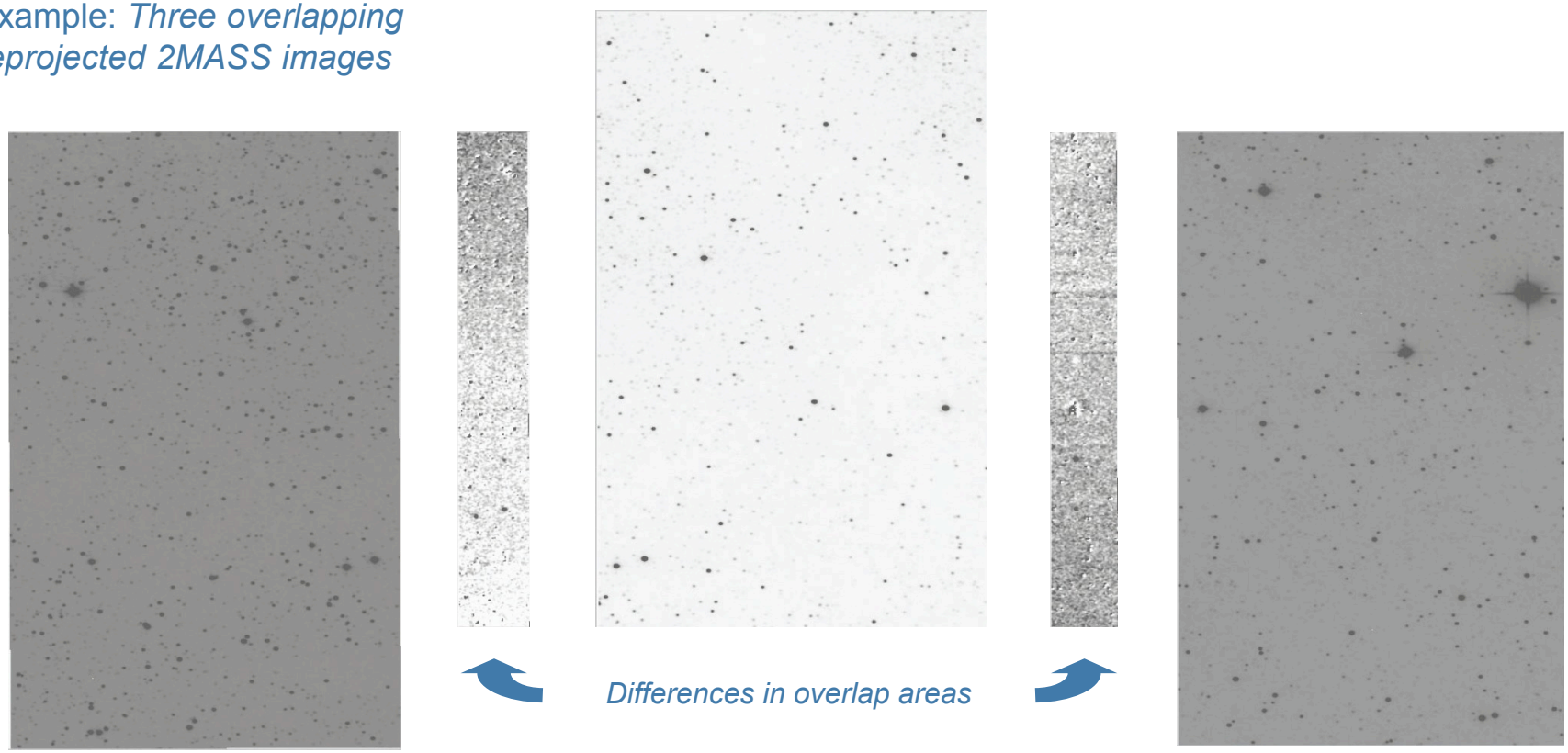
Find correction ramp for each image.

Subtract from Pixel values



Montage: Background Correction Procedure

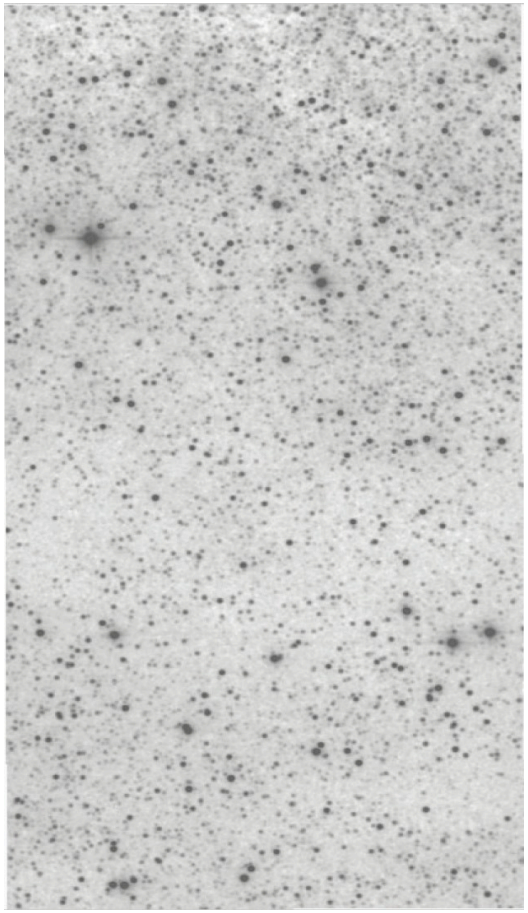
Example: *Three overlapping reprojected 2MASS images*

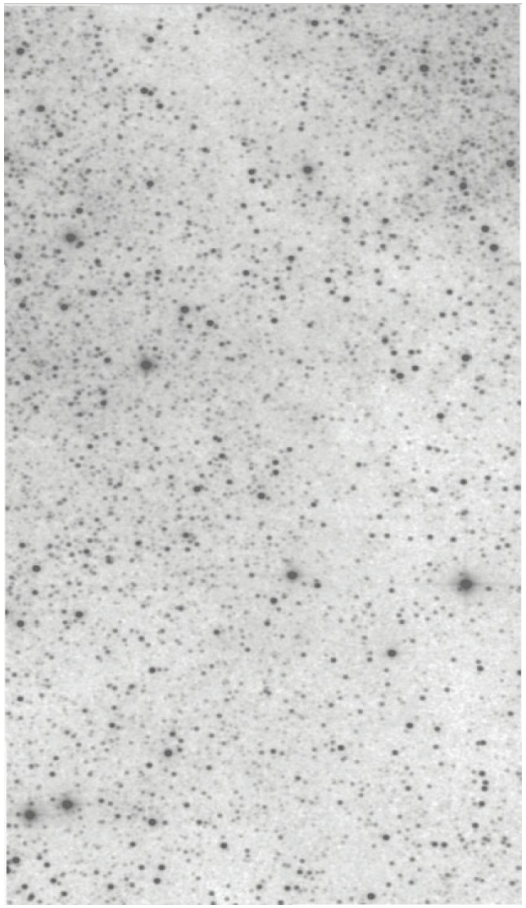


Differences in overlap areas

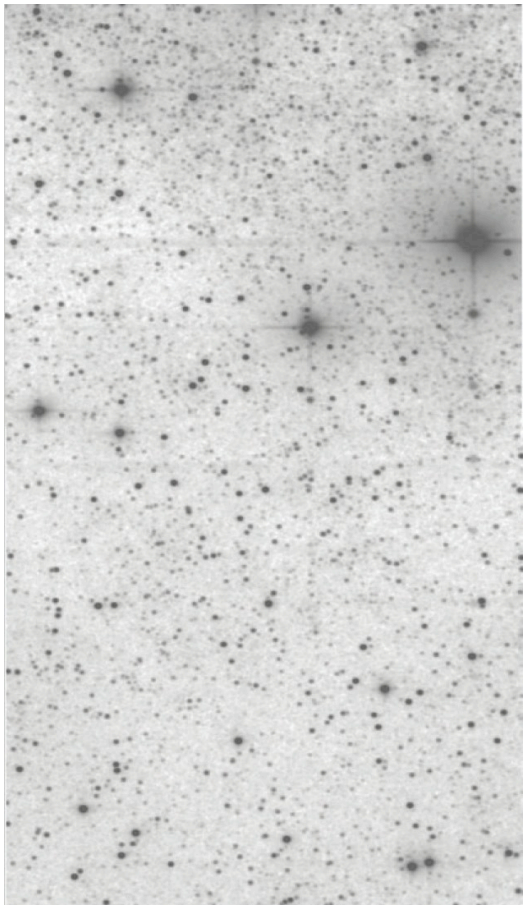
- A correction is calculated for each image based on all the differences between it and its neighbors (an approximation to a least squares fit to the difference data with brightness outlier pixels excluded). The correction is currently a plane but could be a higher order surface.
- This is done for all images, then half the correction determined is applied (to a parameter database; equivalent numerically to applying it to the images.
- The process is iterated until step difference for all images becomes small.

Montage: Background Correction Results





Reprojected Background-Corrected Images



First Public Release of Montage

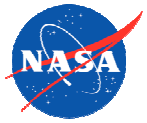
Serial Processing of Images - Montage_v1.7

- Available for download via a clickwrap license issued by Caltech at *http://montage.ipac.caltech.edu*
  - User's Guide
- Emphasizes accuracy in photometry and astrometry
  - Images processed serially
  - Reprojection performed on surface of sphere
- BUT generality at expense of speed
- AND mosaic size limited to available memory

- Performance
  - Mosaic of 54 2MASS images, 1 deg x 1 deg
  - Pentium-4 2.26 GHz, 1 GB RAM

Reprojection	5500 s
Background Modeling	55 s
Rectification	28 s
Co-addition	11 s

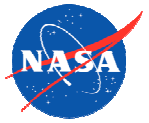

Rho Ophiuchi
324 2MASS images in each band => 972 images
On a 1 GHz Sun, mosaicking takes about 15 hours



Montage_v1.7 Computations

- Building a mosaic from N 1024 x 512 pixel 2MASS images on a single processor 1.4 GHz Linux box takes roughly (N x 100) seconds (with the reprojection algorithm used in Montage_v1.7)
- 98-99% of this time is in the reprojection, which can be perfectly parallelized (this doesn't embarrass us)

Dataset	# of images	Size of each image	Sky coverage	Total number of pixels (x 10^{12})	Storage size (TB)	Processing time for all data in 1.4 GHz IA32 processor hours (x 1,000)
2MASS	~ 4 million	~ 17' x 8.5' at 1"	~ 100%	~ 2.1	~ 8	~ 111
DPOSS	~ 2,600	~ 6.6° x 6.6° at 1"	~ 50%	~ 1.4	~ 3	~ 74
SDSS (DR1)	~ 50,000	~ 13.6' x 9' at 0.4"	~ 25%	~ 1.2	~ 2.4	~ 65

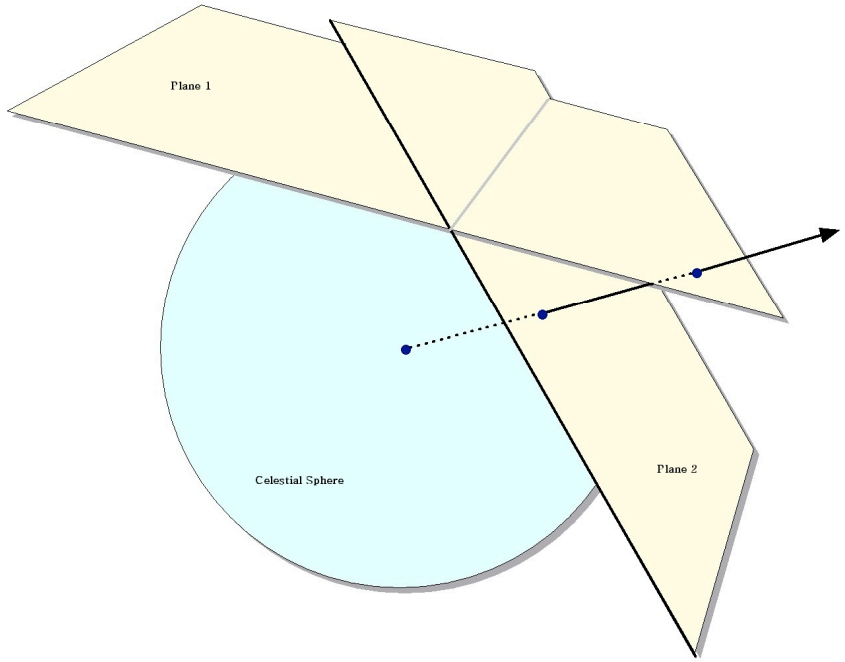


Improvements in Upcoming Montage Second Release - Montage_v2.x

- Algorithmic Improvements
  - Fast plane to plane projection
  - Improved memory efficiency
- Montage TeraGrid Portal

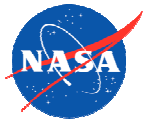
Custom Reprojection Algorithms

- Transform directly from input pixel to output pixels
  - Approach developed by Spitzer for tangent plane projections
  - Augment with “distorted” gnomonic projections
    - Pixel locations distorted by small distance relative to image projection plane
  - **Performance improvement in reprojection by x 30**



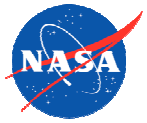
AND Co-addition no longer limited by memory - output images read into memory one line at a time, co-added and written to disk

=> 30% performance degradation acceptable



Montage: The Grid Years

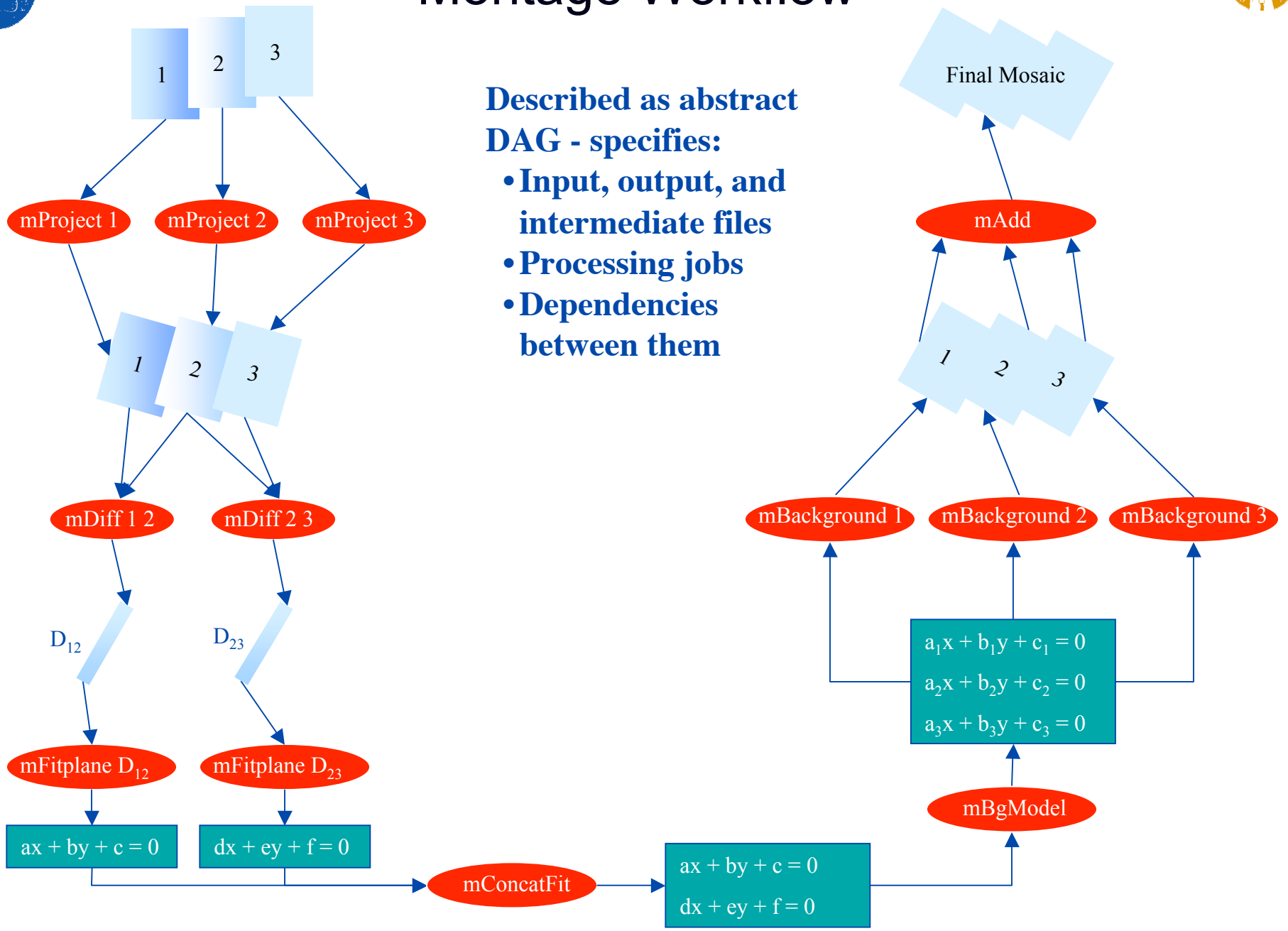
- Exploit parallelization inherent in Montage modular design
  - Grid is an abstraction - array of processors, grid of clusters, …
  - Montage has modular design - run on any environment
- Prototype architecture for ordering a mosaic through a web portal
  - Request processed on a computing grid
  - Prototype portal uses the Distributed Terascale Facility (TeraGrid)
    - This is one instance of how Montage could run on a grid
  - A NASA CNIS task at JPL is using Montage for large scale mosaicking on the Information Power Grid



Montage: The Grid Years (cont.)

- Prototype version of a methodology for running on any “grid environment”
  - Many parts of the process can be parallelized
  - Build a script to enable parallelization
    - Called a *Directed Acyclic Graph* (DAG)
    - Describes flow of data and processing
      - Describes which data are needed by which part of the job
      - Describes what is to be run and when
    - Standard tools can execute a DAG

Montage Workflow

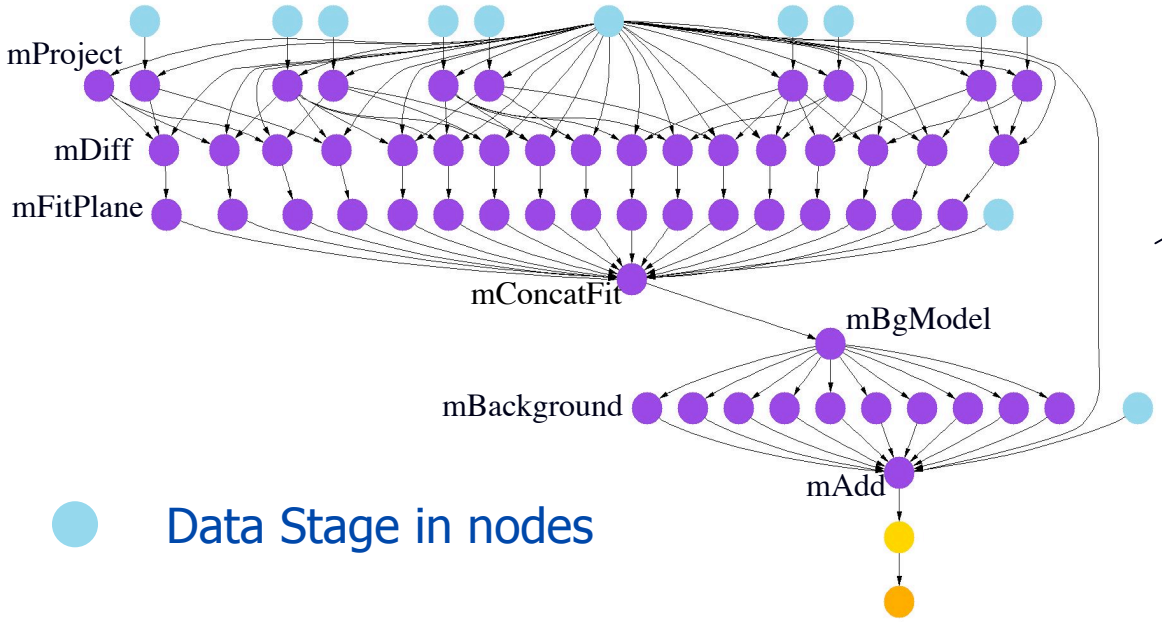


Described as abstract DAG - specifies:

- **Input, output, and intermediate files**
- **Processing jobs**
- **Dependencies between them**

Montage on the Grid

Example DAG for 10 input files



- Data Stage in nodes
- Montage compute nodes
- Data stage out nodes
- Registration nodes

Maps an abstract workflow to an executable form

Pegasus

Grid Information Systems

Information about available resources, data location

Condor DAGMan

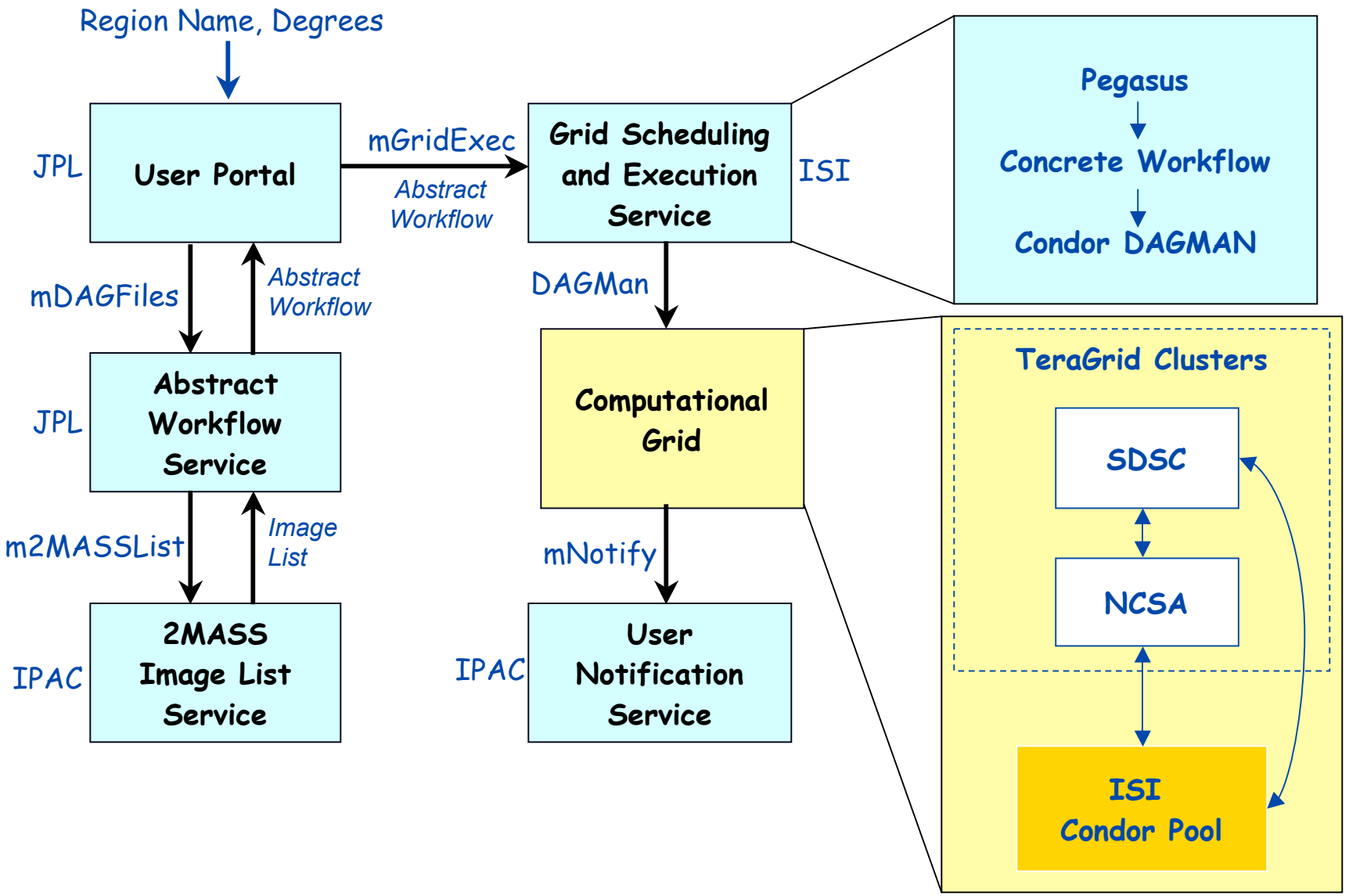
Executes the workflow

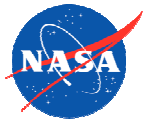
MyProxy

Grid

User's grid credentials

Montage Grid Prototype





TeraGrid Performance

Job	# Jobs	Avg Run Time (s)
mAdd	1	94
mBackground	180	2.64
mBgModel	1	180
mConcatFit	1	9
mDiff	482	2.89
mFitplane	483	2.55
mProject	180	131
Data Transfer In	183	5-30
Data Transfer Out	1	1080

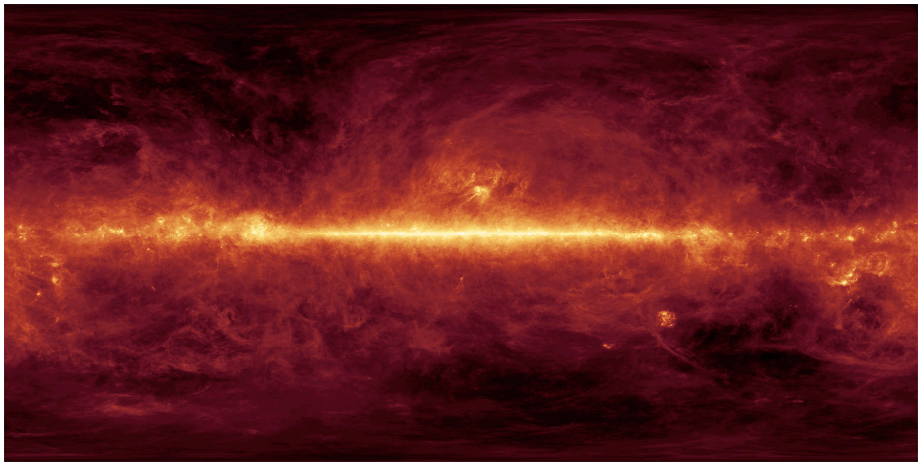
2 deg x 2 deg 2MASS mosaic of M16

Workflow Run Time: 107 min (1515 jobs)

Exposes highest degree of parallelism

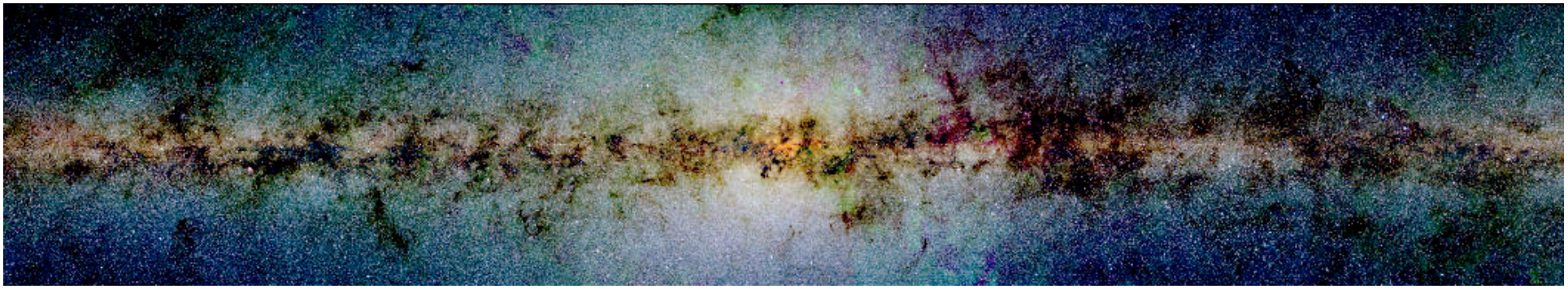
Overhead in scheduling lots of small jobs => Reduce overheads by aggregating nodes

Sample Montage Mosaics



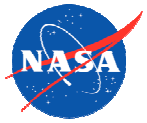
100 µm sky; aggregation of COBE and IRAS maps (Schlegel, Finkbeiner and Davis, 1998)

- 360 x 180 degrees; CAR projection



2MASS 3-color mosaic of galactic plane

- 44 x 8 degrees; 36.5 GB per band; CAR projection
- 158,400 x 28,800 pixels; covers 0.8% of the sky
- 4 hours wall clock time on cluster of 4 x 1.4-GHz Linux boxes



Summary

- Earth and Space science both have huge, complex, distributed datasets
  - Need image mosaics to make sense of it all
  - Grid computing a natural fit
- Montage is a custom astronomical image mosaicking service that emphasizes astrometric and photometric accuracy
- First public release, Montage_v1.7.1, available for download at the Montage website
- Montage_v2.x includes algorithmic enhancements for fast reprojection, and emphasizes grid computing
- A prototype Montage service has been deployed on the TeraGrid; ties together distributed services at JPL, Caltech IPAC, and ISI
- Montage is also being used on the Information Power Grid for large scale (all sky) mosaics
- Montage website: http://montage.ipac.caltech.edu/