

# Montage

**Montage: An Astronomical Image Mosaic Service  
for the National Virtual Observatory**  
<http://montage.ipac.caltech.edu>

**Cooperative Agreement Number NCC 5-626**

**Second Annual Report  
(Period September 1 2002 – August 31 2003)**



## **Objective:**

The Montage project will deploy a portable, compute-intensive service that will deliver science-grade custom mosaics on demand. Science-grade in this context requires that terrestrial and instrumental features are removed from images in a way that can be described quantitatively; custom refers to user-specified parameters of projection, coordinates, size, rotation and spatial sampling. Montage leverages the image mosaic algorithms already deployed in the *yourSky* image mosaic server.

Montage will generate mosaics from input files that comply with the Flexible Image Transport System (FITS) standard and contain images whose projections comply with the World Coordinate System (WCS) standards. In operations, Montage will be deployed on the emerging Distributed Terascale Facility (hereafter, TeraGrid), where it will process requests for 2Micron All Sky Sky Survey (2MASS), Sloan Digital Sky Survey (SDSS) and Digital Palomar Observatory Sky Survey (DPOSS) image mosaics; the requests will be made through existing astronomy World Wide Web portals.

Montage's performance goal is to sustain throughput of 30 square degrees (e.g. thirty 1 degree x 1 degree mosaics, or one 5.4 degrees x 5.4 degrees mosaic, etc.) per minute on a 1024x400MHz R12K Processor Origin 3000 or machine equivalent with sustained bandwidth to disk of 160 MB/sec.

## **Approach:**

Deep, wide area imaging surveys have assumed fundamental importance in astronomy. They are being used to address such fundamental questions as the structure and organization of galaxies in space and the dynamic history of our galaxy. One of the most powerful probes of the structure and evolution of astrophysical sources is their behavior with wavelength. However, this power has yet to be fully realized in the analysis of astrophysical images because survey results are published in widely varying coordinates, map projections, sizes and spatial resolutions. Moreover, the spatial extent of many astrophysical sources is much greater than that of individual images. Astronomy therefore has need for a general image mosaic engine that will deliver image mosaics of arbitrary size in any common coordinate system or map projection and at any spatial sampling. Montage aims to provide this service.

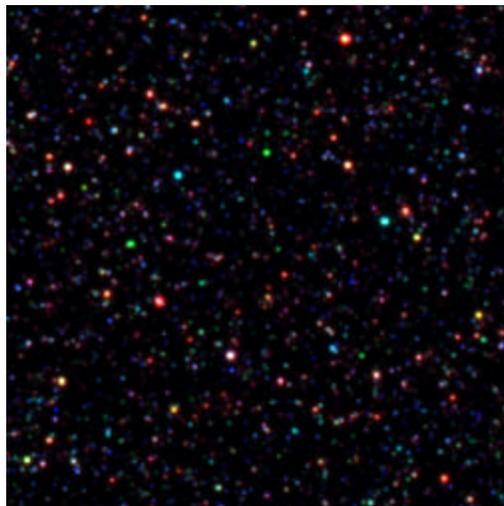
The key to our technical approach is to develop a flexible framework that will support many custom user cases and processing needs. Cases include compute and time intensive mosaics covering wide areas of the sky, small mosaics generated on a desktop as part of a small research project or observation planning program, and mosaics generated as a standard science product as part of a processing pipeline.

## Scientific Accomplishments

### **Data Validation and Processing Support for Spitzer Space Telescope (formerly SIRTf) Legacy Teams.**

The Galactic Legacy Infrared Mid-Plane Survey Extraordinaire (GLIMPSE), Spitzer Wire Area Infrared Experiment (SWIRE), and “From Molecular Cores to Planet-forming Disks” (c2d) teams are actively using Montage to support data simulation, mission planning, quality assurance and pipeline development.

In support of Quality Assurance programs, GLIMPSE and SWIRE have made particular use of Montage as a reprojection engine to co-register images measured at different wavelengths, spatial samplings, coordinate systems and projections. SWIRE has combined images from 2MASS in the J, H and K bands with images from the Spitzer Infrared Array Camera (IRAC), and GLIMPSE has combined images from 2MASS at J, H and K with Midcourse Space Experiment (MSX) images at 8  $\mu\text{m}$ .; see figures 1 and 2.



**Figure 1: A synthetic image of the Lockman Hole measured by the Infrared Array Camera aboard Spitzer. It has been generated by mosaicking 2MASS images with simulated IRAC images. All the sources in the mosaic are distant galaxies.**



**Figure 2: A 2MASS and MSX Mosaic of G305.3, with a 1x1 degree field of view centered on 305.3 +0.2. This is a 3-color combination using 2MASS J (blue) and H (green) bands, as well as MSX (red). This is a science validation image generated by the GLIMPSE project.**

### **Atasmaker**

The Atasmaker project is using Montage as base code. It is a Grid technology project that, when used in combination with NVO interoperability, will create new knowledge resources in astronomy. The product is a multi-faceted, multi-dimensional, scientifically trusted image atlas of the sky, made by federating many different surveys at different wavelengths, times, resolutions, polarizations, etc. Atasmaker performs resampling and mosaicking of image collections and is well suited to operate with a proposed “hyperatlas” standard. Requests can be satisfied via on-demand computations or by accessing a data cache. Computed data is stored in a distributed virtual file system, such as the Storage Resource Broker (SRB). We expect these atlases to be a new and powerful paradigm for knowledge extraction in astronomy, as well as a way to build educational resources. The system is being incorporated into the data analysis pipeline of the Palomar-Quest synoptic survey, and is being used to generate all-sky atlases from the 2MASS, SDSS, and DPOSS surveys for joint object detection.

### **Background Images for Astronomical Query Services.**

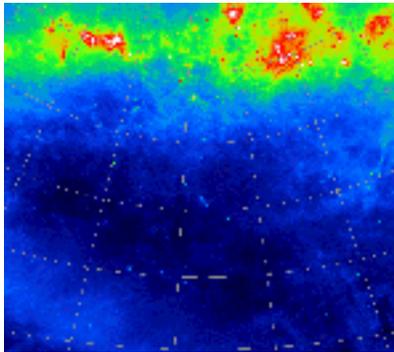
The NASA/PAC Infrared Science Archive (IRSA) is actively using Montage to generate images that will support visualization of locations of sources and image footprints that satisfy spatial queries to its scientific data holdings. Images in existing image collections are inadequate for this purpose because they rarely coincide with the locations and radii of queries.

## NVO Middleware Demonstration Project

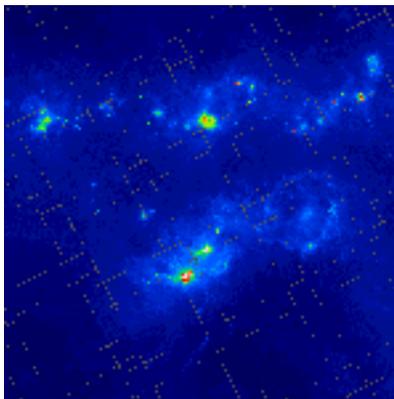
As part of its commitment to the National Virtual Observatory (NVO), IPAC is developing the Request Object Management Environment (ROME). It is a simple, portable request management environment that can work in conjunction with existing browsers, http services and custom clients to support reliable execution of long-lived jobs. It communicates status information to the clients.

IPAC hosted a demonstration of how this system will work in operations. The demonstration used Montage as a compute-intensive service. Users placed orders for IRAS image mosaics through a web page, and they were processed by Montage, running on a server “behind” ROME.

Figures 3, 4 & 5 show images generated by this demonstration project.



**Figure 3: IRAS Mosaic of Andromeda and the Galactic Plane.** This image is an full resolution (1.5 arcminute pixels) mosaic of the 60-micron IRAS image (ISSA) data. It covers a 60-degree square region of the sky centered on the Andromeda Galaxy (Messier 031) and was generated in Galactic coordinates (the Galactic plane runs across the top of the image). The image is actually 2400 pixels square and has been resampled down here by a factor of three.



**Figure 4: IRAS Mosaic of Orion and the Galactic Plane.** This region is just south of the Galactic plane in a direction almost exactly opposite the Galactic center. Since in

that direction we are nestled up against the inside edge of a spiral arm, the various clouds of dust, gas and associated star-forming regions are spread out more than in other directions (where we see them at a greater distance).

The plane of the galaxy runs across the top of the image, and the active area below that coincides with the visible constellation of Orion (his head is in the general vicinity of the ring of bright emission on the right and his belt runs through the very bright area on the left).

The image, comprised of 31 images covering an area 45 degrees across, took 9072 seconds to process on a 300 MHz Sun UltraSPARC-III.

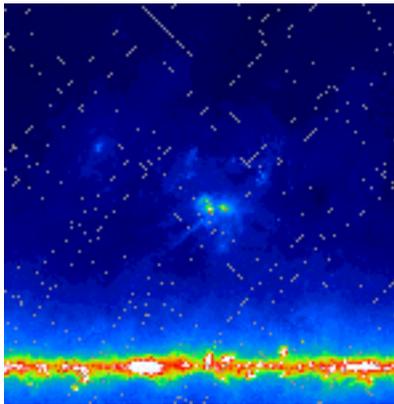


Figure 5: IRAS Mosaic of Rho Ophiuchus, 45 degrees wide in Galactic coordinates. The Galactic center is bottom center of the image. Processing of the 32 original images took 9677 seconds on a 300 MHz Sun UltraSPARC-III.

## Technology Accomplishments.

### *Progress Towards Milestones*

Table 1 summarizes the progress towards milestones during this reporting period.

Table 1: Montage Milestones for Period September 1 2002 – August 30 2003

Milestones Satisfied	
Milestone	Deliverables
F	Develop Science Grade Mosaics that conserve energy and support background removal, with metrics specified through the guidance of the CRB and scientifically validated under its auspices. Access to this service will be through a modification of the existing <i>yourSky</i> web form.  <i>YourSky Mosaic Engine</i>

	<ul style="list-style-type: none"> <li>* Ensure conservation of energy in mosaics</li> <li>* Handle image rotations in all WCS projections</li> <li>* Metric: The following metrics apply to science grade mosaics; their precise values will be established through the guidance of the CRB:</li>   <li>* Reduction in the average deviation from the measured energy per unit area (we anticipate roughly 50%) when constructing mosaics in at least 10 WCS projections with any image rotation.</li> <li>* Spatial scale of mosaics and spatial re-sampling of pixels that allow science analysis (we anticipate 1 to 5 degrees spatial scale; and full, 1/2, 1/4, and 1/8 resolutions).</li>   <li>* <b><i>Apply Background Removal Parameters that support background subtraction models:</i></b></li>   <li>* Common sky model that preserves total flux</li> <li>* Preserve point sources only</li> <li>* Preserves feature on a scale that allows science analysis (we anticipate 1 to 5 degrees, as noted above).</li>   <p>Documented source code made publicly available via the project web site.</p> </ul>
<b>Milestones In Progress</b>	
<b>Milestone</b>	<b>Deliverables</b>
<b>I</b>	<ul style="list-style-type: none"> <li>* The improved <i>YourSky</i> code delivered in Milestone F) will run on the Teragrid Linux cluster. Performance comparison between the PowerOnyx and the Teragrid will be published on the web page.</li> <li>* The improved <i>YourSky</i> code delivered in Milestone F) and running on the Teragrid will be interoperable with the <i>OASIS</i> and <i>VirtualSky</i> clients, in that users place an order for a custom mosaic through these clients, receive notification of the completion of the request, and are able to visualize the images.</li> </ul>

Delivery of Milestone F was delayed to allow development of a complete Users Guide, and perform thorough testing and validation to assure the preservation of astrometric and calibration fidelity of the input data.

***Technical Accomplishments***

Version 1.7.1 of Montage has been publicly released, and is available through the Montage website at <http://montage.ipac.caltech.edu/docs/download.html>. The Montage distribution consists of 20 modules that contain 7560 lines of code. The distribution also includes all supporting libraries, build instructions, a build test, and validation test data sets. The system has been rigorously tested using 2MASS images for ten of the most

common projections supported by the World Coordinate Systems (WCS). It was subject to 2595 test cases, which yielded 119 defect reports. Altogether, 116 of them were closed, and the impact of the remaining three were described in the documentation.

A complete Users Guide is available at <http://montage.ipac.caltech.edu/docs>. The guide includes the following:

### Montage Design

- Description of Components
- Montage Algorithms
- Detailed Design Document (PDF)
- Extending Montage to Larger or Different Problems
- Science Use Cases

### How to Run Montage

- Montage Tutorial: How to Build a mosaic of m101
- Supported WCS Projections
- Montage Header Templates
- API
- Debug Levels
- Caveats
- Montage Performance
- Troubleshooting Montage

### Software Installation

- Downloading Montage
- System Requirements
- Building Montage

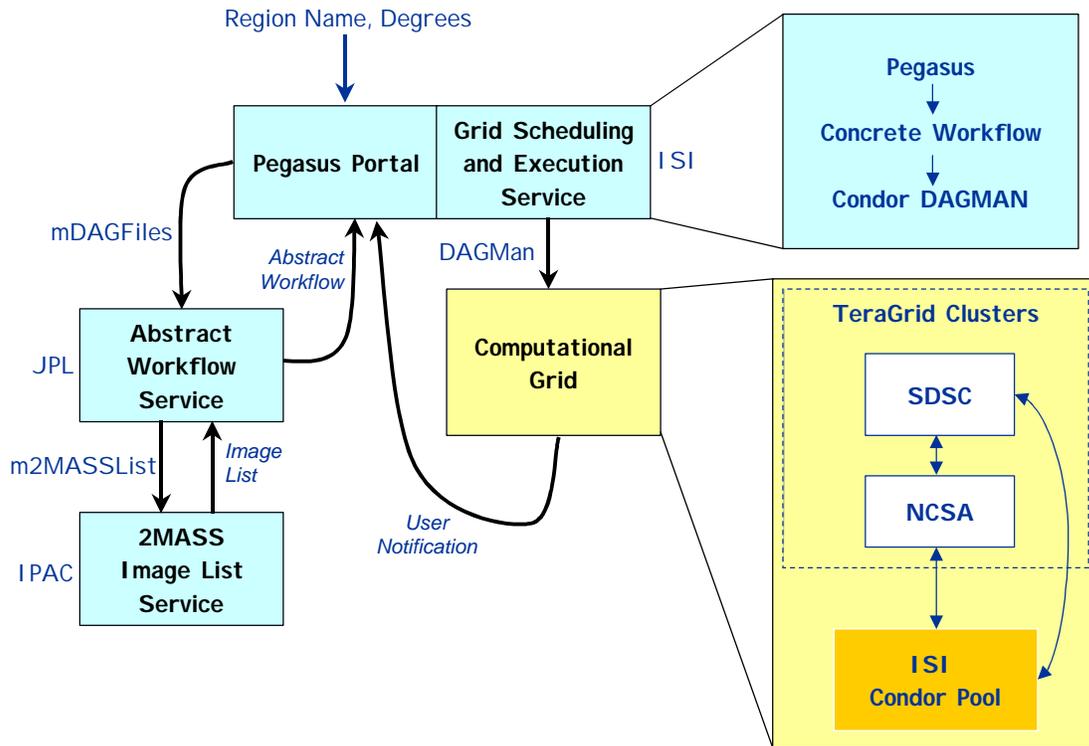
### Montage Test Suite

- Test Suite Overview
- Montage Build Test
- Montage System Tests
- Photometric and Calibration Accuracy
- Third Party Validation by SWIRE Team

### Reference Materials

- Montage Calltrees

To satisfy Milestone I, we have, in collaboration with the staff at Information Sciences Institute, USC, begun to develop a distributed architecture that will accept requests for mosaics from a web page, process the request on the Teragrid, and return the mosaic to the user. Figure 6 captures a preliminary architecture that identifies the required components.



**Figure 6: Preliminary distributed architecture for running Montage on the Teragrid.**

### Status/Plans

The work plan for September 1 2003 through August 31 2004 calls for us to meet the milestones in Table 2.

### Table: Summary work plan for Sept 1 2003 - Aug 31 2004

Milestone	Deliverables
<b>I</b> (in progress)	<ul style="list-style-type: none"> <li>* The improved <i>YourSky</i> code delivered in Milestone F) will run on the Teragrid Linux cluster. Performance comparison between the PowerOnyx and the Teragrid will be published on the web page.</li> <li>* The improved <i>YourSky</i> code delivered in Milestone F) and running on the Teragrid will be interoperable with the <i>OASIS</i> and <i>VirtualSky</i> clients, in that users place an order for a custom mosaic through these clients, receive notification of the completion of the request, and are able to visualize the images.</li> </ul>
<b>C</b>	Second Annual Report:  Deliver second annual report to project web site.
<b>G</b>	Second Code Improvement:  Code Improvements  <ul style="list-style-type: none"> <li>* The improved YourSky code per milestone I) will run on the Teragrid. The achievable computational speed-up will depend on the performance of the Teragrid as deployed. We propose two performance metrics: A target computation speedup that ignores I/O time and a target overall speedup that includes both computation and I/O times. We will achieve a target performance that is equivalent to a computation speedup of 64 and an overall speedup, including I/O, of 32, for a 5 degree x 5 degree 2MASS mosaic (which will be the most computation intensive dataset) on a 128x1GHz (128 GFlops) target machine with a sustained bandwidth to disk of 160 MB/sec.</li> <li>* Cache results locally for commonly requested regions. Develop cache of at least 2 TB. Metric: Demonstrate speed-up when cached mosaic is requested. Publish speed-up figures.</li> </ul> Documented source code will be made publicly available via the project web site.
<b>J</b>	Full Interoperability:  <ul style="list-style-type: none"> <li>* Demonstrate that the compute engine accepts requests from the OASIS and VirtualSky clients for mosaics from the 2MASS, DPOSS and SDSS surveys, processes the request (includes accessing cached images as necessary), notifies the user regarding the status and availability of a mosaic, which can be visualized by the user. Visualization includes full user control of the image in real-time: pan/zoom, cropping, scaling, resampling, color table, stretch, and histogram equalization.</li> <li>* Publish on the project web site updated requirements &amp; design docs, and updated test plan and test reports, and a draft Users' Guide.</li> </ul>

## *Outreach*

The Montage project does not support its own outreach program, but the E/PO team at IPAC uses the software actively. Dr. Robert Hurt has contributed the following description of how Montage is used by this team:

“Montage has proven to be a useful tool for education and public outreach by facilitating the quick creation of large mosaics of interesting regions of sky with relatively little effort. Many of the datasets at IPAC are composed of image tiles covering large areas of sky broken into very small pieces. Large mosaics around visually complex regions of sky make it much easier to stimulate interest in this data. Such large mosaics are typically very labor-intensive using standard astronomical tools, so few of them tend to be done.

The powerful flexibility of Montage makes it a great asset for generating large images that would not be practical given typical EPO time and resource constraints. We have used Montage to make a number of extensive mosaics from the 2MASS and IRAS datasets. Both of these are very quick procedures. For 2MASS, useful tools exist to quickly extract the image datafiles from online archives, and then stitch them together easily and consistently. We have created a number of large mosaics from IRAS data using the simple ROME interface to this dataset. This interface shows the real power of Montage to operate behind the scenes for a user who only wants to worry about overall image parameters and the end product.

Other powerful applications of Montage for our EPO efforts will include visualizing full-sky datasets. The manipulation of such huge datasets can be highly resource-intensive and cumbersome and generally requires custom code. However, we will be soon using Montage to quickly and easily reproject all-sky datasets into formats that can easily be turned into immersive viewers, backdrops for realistic 3D animations, and even maps/globes that can be distributed on-line”

## *Demonstration at SC 2003*

At SC03, Montage will have a presence in three booths in the exhibit hall. We plan to build a lightbox display of Montage featuring the image in Figure 7 (below) to show in the NASA booth, and to hand out a flyer about Montage to interested visitors. We will also present the middle third of an hour-long talk in the NPACI booth on the high-performance computing and data handling needs of astronomy, particularly Montage (J. Jacob, D. Katz). San Diego Supercomputer Center (SDSC) will present the first third of the overall presentation (R. Moore), and Caltech’s Center for Advanced Computing Research the last third (R. Williams). Our partners at Information Sciences Institute, USC will help us conduct live demos of an early version of the Montage Teragrid portal in the Argonne National Laboratory (ANL) booth.

## **Publications:**

“Montage: An On-Demand Image Mosaic Service for the NVO”. G. B. Berriman, J. C. Good, D. W. Kurkendall, J. C. Jacob, D. S. Katz, T. Prince & R. Williams.

“Astronomical Data Analysis Software and Systems – XII” (ASP Conf Series 295), 343.

**Abstract:** Montage will deliver a generalized toolkit for generating on-demand, science-grade custom astronomical image mosaics. “Science-grade” in this context requires that terrestrial and

instrumental features are removed from images in a way that can be described quantitatively. "Custom" refers to user-specified parameters of projection, coordinates, size, rotation and spatial sampling, and whether the drizzle algorithm should be invoked. The greatest value of Montage will be its ability to analyze images at multiple wavelengths, by delivering them on a common projection, coordinate system and spatial sampling and thereby allowing analysis as if they were part of the same multi-wavelength image. Montage will be deployed as a compute-intensive service through existing portals. It will be integrated into the emerging NVO architecture, and run operationally on the Teragrid, where it will process the 2MASS, DPOSS and SDSS image collections. The software will also be portable and publicly available.

"An Architecture for Access to A Compute Intensive Image Mosaic Service in the NVO"  
G. Bruce Berriman, David Curkendall, John Good, Joseph Jacob, Daniel S. Katz, Mihseh Kong, Serge Monkewitz, Reagan Moore, Thomas Prince, Roy Williams. "Virtual Observatories", A. Szalay, ed. (SPIE Conference 4686), 91

**Abstract:** The National Virtual Observatory (NVO) will provide on-demand access to data collections, data fusion services and compute intensive applications. The paper describes the development of a framework that will support two key aspects of these objectives: a compute engine that will deliver custom image mosaics, and a "request management system," based on an e-business applications server, for job processing, including monitoring, failover and status reporting. We will develop this request management system to support a diverse range of astronomical requests, including services scaled to operate on the emerging computational grid infrastructure. Data requests will be made through existing portals to demonstrate the system: the NASA/IPAC Extragalactic Database (NED), the On-Line Archive Science Information Services (OASIS) at the NASA/IPAC Infrared Science Archive (IRSA); the Virtual Sky service at Caltech's Center for Advanced Computing Research (CACR), and the *yourSky* mosaic server at the Jet Propulsion Laboratory (JPL).

"Multi-Wavelength Image Space: Another Grid-Enabled Science"  
Roy Williams, Bruce Berriman, Ewa Deelman, John Good, Joseph Jacob, Carl Kesselman, Carol Lonsdale, Seb Oliver, Tom Prince. *Concurrency & Computation*, vol 15 (2003), pp 539-549 (Refereed)

**Abstract:** We describe a new Grid-enabled branch of astronomy: multi-wavelength images. To see sky images in the same pixel space, they must be projected to that space, a compute-intensive process. There is thus a virtual data space induced that is defined by an image and the applied projection. This virtual data can be created and replicated with Planners and Replica catalog technology developed under the GriPhyN project. We plan to deploy our system (MONTAGE) on the US Teragrid. Grid computing is also needed for ingesting data -- computing background correction on each image -- which forms a separate virtual data space. Multi-wavelength images can be used for pushing source detection and statistics by an order of magnitude from current techniques; for optimization of multi-wavelength image registration for detection and characterization of extended sources; and for detection of new classes of essentially multi-wavelength astronomical phenomena. The paper discusses both the grid architecture and the scientific goals.

## **Presentations:**

"Montage: An On-Demand Image Mosaic Service for the NVO". G. B. Berriman, paper presented at "Astronomical Data Analysis Software and Systems – XII". Baltimore, MD.

"Progress report on Montage." G. B. Berriman. Presentations given to the National Virtual Observatory Management Team Meetings, December 2002 (Chicago), April 2003 (Pasadena)

“Montage: An Image Mosaic Service for the NVO”. G. B. Berriman. Presentation given to the IPAC Users Committee (April 2003).

"Data Access and Visualization Using Clusters and other Parallel Computers," D. S. Katz, A. Bergou, G. B. Berriman, G. Block, J. Collier, D. Curkendall, J. Good, L. Husman, J. Jacob, A. Laity, P. P. Li, C. Miller, L. Plesea, T. Prince, H. Siegel, and R. Williams, “On the Use of Commodity Clusters for Large-Scale Scientific Applications”, Tysons Corner, Virginia, July 2003

### ***Presentations at Supercomputers***

We had a handout and slideshow at the NASA booth, and demos in the NPACI and CACR booths.

Slides and handouts can be downloaded from  
<http://montage.ipac.caltech.edu/documents.html>.

## **Point of Contact**

Dr. G. Bruce Berriman, 100-22 Caltech, Pasadena, CA 91125  
Montage Project Manager  
Phone (626) 395-1817; FAX (626) 397-7354  
Email: [gbb@ipac.caltech.edu](mailto:gbb@ipac.caltech.edu)

## **Media References**

ESDC News, Summer 2003. “Supercomputer Serves Up Giant Slice of Milky Way”. See <http://esdcd-news.gsfc.nasa.gov/2003.Summer/03.ct-2mass.html>

Envision, Volume 19, Number 1. January-March 2003. See <http://www.npaci.edu/envision/v19.1/>  
Figure 7 was the cover art for this issue of Envision.



**Figure 7: The Galactic Plane Near the War and Peace Nebula. A three-color mosaic derived from images in the Second Incremental Data Release of the 2 Micron All Sky Survey (2MASS). This picture shows dust clouds and nebulosity in the plane of our Galaxy, and it combines images at three near-infrared bands. The mosaic contains 12,000 individual pixels on a side. There are 347 individual images in each of three bands: J (1.25  $\mu\text{m}$ ), H (1.65  $\mu\text{m}$ ), and K (2.2  $\mu\text{m}$ ). This image has been reduced to roughly 1/24 of full resolution.**

### **Patents**

Not applicable

### **New Technology Reports**

A New Technology Report (NTR) was submitted for the Montage first public release (version 1.7.1): "Montage: An Astronomical Image Mosaic Service for the National Virtual Observatory". The Montage Team credited on this NTR include (in alphabetical order) Attila Bergou (see Graduate Students/Post-docs Section below), Bruce Berriman, Anastasia Laity, John Good, Joseph Jacob, Daniel Katz, Thomas Prince, and Roy Williams.

### **Graduate Students/Post-docs**

Mr. Attila Bergou completed his undergraduate education at Carnegie Mellon University, graduating with a B.S. in computer science and a B.S. in physics in May of 2002. He then was a graduate co-op student at Jet Propulsion Laboratory for one year working with the Montage Team as well as the Quantum Computer Technologies Group. Now he is on academic leave to pursue his PhD in Physics at Cornell University and expects to finish in the spring of 2008.