Updated Milestones for the Proposal

"High-Performance Cornerstone Technologies for the National Virtual Observatory"

Prepared By:

T. Prince (Principal Investigator), G. B. Berriman (Project Manager), D. Curkendall, J. Good, L. Husman, J. Jacob, and R. Williams (Co-Investigators)

Please Direct Questions To:

G. B. Berriman

gbb@ipac.caltech.edu
☎ (626) 397-9548; FAX (626) 397-7354

Submitted to NASA, November 16, 2001

First Revision: December 26, 2001;
Second Revision: January 14, 2002.

Project Goals: MONTAGE, an image-mosaic service for the NVO.

MONTAGE: Mosaics On the TerAGrid Express

This project will deploy a portable, compute-intensive service that will deliver *science-grade custom* mosaics on demand, with requests made through existing portals. *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. It will be applied in the first instance to the 2MASS, SDSS and DPOSS image data sets. The software will be portable and publicly available: while it will run operationally on the emerging Teragrid cluster, it will available for download for running on local clusters and individual workstations.

The due dates given here assume a starting date of February 1, 2002

A) Software engineering plan completed:

- Publish the Project Management plan on a project web page. The plan will include the following:
    - Science use cases and detailed requirements, approved by the Customer Review Board (CRB). The CRB will be specifically asked to provide guidance on the spatial scales and pixel sampling frequencies over which science grade mosaics should be served.
    - Technical Approach
    - Software Engineering plan
    - Management Approach
    - Quality Assurance
    - Configuration Management

- Deliver a Risk Management strategy that will be available internally to the project and the sponsor.

Due Date: May 31, 2002

E) Code baseline delivered:

Describe the physical problem that is addressed by the baseline code. The baseline code will include the ability to remove backgrounds through application of a low pass filter.

YourSky Mosaic Service

- Baseline the performance of the *yourSky* image mosaic service, which offers custom access to mosaics of user-selected size, coordinate system, and projection, with bilinear sampling. The performance will be baselined on the JPL PowerOnyx. It is equipped with 8 x 200 MHz MIPS R10K processors and running at a clock speed of 194 MHz. Performance and scaling curves will be published on the project web page. Documented source code will be delivered to the project web page.

***Background Removal Service*:**

- Baseline a general background removal service that provides background removal decoupled from the mosaic service. Perform a global fit to the 2MASS and DPOSS data with a high-pass filter that adjusts images to a common sky mode but preserves background flux. Write background parameters to a meta-database. Publish baseline performance to fit the images in those parts of the sky publicly released by 2MASS and DPOSS. Documented source code will be delivered to the project web page.

All software referred to above will be documented and made publicly available through the project web site.

Publish, on the project web page, initial system requirements and design documents.

Due Date: June 30, 2002

H) Design Policy for Interoperability and Community Delivery Agreed On

Declare that this project will use Virtual Sky and OASIS as its primary clients. The API for other clients to run the mosaic engine will be published on the project web page. Declare that the operational version of the mosaic engine, *yourSky*, will run on the NPACI *Teragrid*, and that the project will serve mosaics from three surveys, 2MASS, DPOSS and SDSS. Under the guidance of the Customer Release Board, we will establish policies for the public release of the service and for naming customers who will be permitted access to the service on a shared-risk basis before public release. Declare working relationships with collaborators at NPACI (SDSC), 2MASS, DPOSS and SDSS.

The Design Policy for Interoperability and Community Delivery will be published on the project web site.

Publish on the Project Web page updated requirements and design documents and initial test plans.

Due Date: July 31, 2002

B) First Annual Report delivered to project web site

Deliver report to the project web page.

Due Date: August 31, 2002

F) First Code Improvement demonstrated

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 *YourSky* web form.

YourSky Mosaic Engine –

- Ensure conservation of energy in mosaics
- Handle image rotations in all WCS projections
- Metric: The following metrics apply to science grade mosaics; their precise values will be established through the guidance of the CRB:
  - 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.
  - 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).

- ***Apply Background Removal Parameters that support background subtraction models:***
  - Common sky model that preserves total flux
  - Preserve point sources only
  - Preserves feature on a scale that allows science analysis (we anticipate 1 to 5 degrees, as noted above).

Documented source code made publicly available via the project web site.

Due Date: February 28, 2003

I) Interoperability Prototype

- The improved *YourSky* 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.

- The improved *YourSky* code delivered in Milestone F) and running on the Teragrid will be interoperable with the *OASIS* and *VirtualSky* 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 visualize the images.

Due Date: July 30, 2003

C) Second Annual Report

Deliver second annual report to project web site.

Due Date: August 30, 2003

G) Second Code Improvement

Code Improvements

- 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.

- 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.

Documented source code will be made publicly available via the project web site.

Due Date: February 28, 2004

J) Full interoperability

- 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 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.

- Publish on the project web site updated requirements & design docs, and updated test plan and test reports, and a draft Users' Guide.

Due Date: August 15, 2004

K) Customer Delivery

Community delivery of the fully operational service. Deployment will be announced at a major national astronomical conference, where we will:

- Demonstrate custom mosaics from 2MASS, DPOSS and SDSS with sustained throughput of 30 square degrees (e.g. thirty 1 degree x 1 degree mosaics, 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 and all input data previously staged on these disks.
- Present scaling curves specifying the variation in compute time with mosaic size and number of user requests
- Deliver a Users' Guide that will include descriptions of how to access the service, known caveats and limitations, and a thorough description of the impact of background removal on the scientific integrity of the images.

Documented source code will be made publicly available via the project web site.

Publish on the project web page final versions of requirements, design, test plan and test results. Deliver to the project web page an updated Users' Guide and a Maintenance Guide.

Due Date: Jan 10, 2005

D) Final Report

Delivery of final report to web site.

Due Date: February 15, 2005