The NSF National Virtual Observatory
TeraGrid Utilization Proposal to NRAC
Multi-year, Large Research Collaboration

Roy **Williams**, PI *California Institute of Technology*
Andrew **Connolly**, coI *University of Pittsburgh*
Jeffrey **Gardner**, coI *Pittsburgh Supercomputing Center*

Senior Personnel

Bruce **Berriman** *IPAC, California Institute of Technology*
Tamas **Budavari** *Johns Hopkins University*
Ewa **Deelman** *ISI, University of Southern California*
Michael **Feldmann** *California Institute of Technology*
Jim **Gray** *Microsoft Corporation*
George **Kremenek** *San Diego Supercomputer Center*
Christopher **Miller** *Carnegie Mellon University*
Reagan **Moore** *San Diego Supercomputer Center*
Robert **Nichol** *Carnegie Mellon University*
Alex **Szalay** *Johns Hopkins University*
Dan **Vanden Berk** *University of Pittsburgh*

Abstract

Astronomy faces a data avalanche. Breakthroughs in telescope, detector, and computer technology have allowed astronomical surveys to produce terabytes of images and catalogs. The NSF National Virtual Observatory (NVO) is a multiyear effort to build tools, services, registries, protocols, and standards that can extract the full knowledge content of these massive, multi-frequency data sets. We propose here to initiate the process whereby the computational resources of the TeraGrid can be combined with the NVO to enable astronomers to explore and analyze these new data sets in order to understand the physical processes that drive the formation and evolution of our universe. In this initial component of a multiyear effort we propose to use the TeraGrid to (1) expose massive data to massive computing through NVO protocols, (2) run representative applications to explore that data, (3) foster new projects in astronomy that use NVO services and TeraGrid resources, and (4) encourage new ways to use supercomputing facilities for science.

1 Introduction

The next decade will witness the completion of several new and massive surveys of the Universe. These surveys will span the range of the electromagnetic spectrum from the X-rays (the ROSAT, Chandra, and XMM satellites) through the optical and ultraviolet (the SDSS, GALEX, LSST surveys) to the measurements of the cosmic microwave background in the submillimeter and radio (the WMAP and PLANCK satellites). Individually, each of these huge data sources will lead to major advances in our understanding of the processes that drive the formation and evolution of the Universe. However, it is only when these datasets are combined - by collating data from several different surveys or matching simulations to observations - that their full scientific potential will finally be realized; the scientific returns from the total will far exceed those from any one individual component. The recognition of this by the astronomical community has led to a major new initiative in astrophysics: the National Virtual Observatory (NVO; http//www.us-vo.org).

The development of NVO, an NSF sponsored Information Technology Research project, together with related efforts ongoing world-wide (through the International Virtual Observatory Alliance) has the potential to revolutionize scientific research in astrophysics. The NVO is designed to support the federation of astronomical sky surveys; seamlessly interlacing data across the full electromagnetic spectrum. It will provide tools to explore within and extract data from these massive, multi-frequency surveys. Such an unprecedented access will provide an astronomer with the ability to determine the detailed correlations within these data, to understand the physical processes that give rise to these correlations and to potentially identify new classes of astrophysical phenomena.

1.1 TeraGrid for both Testbed and Applications

Over the last two years the NVO has developed the infrastructure to deliver data through standard service types, thus providing access at a higher semantic level than files and record sets. These “smart pipes” deliver source catalogs, images, and spectra in response to requests that are framed as sky regions, wavelength regions, and other criteria. The implementations of the standard services are registered in a new distributed registry structure that will enable publication and discovery of these access services as well as their use in applications.

Because of these standard access methods, federation of astronomical databases is greatly simplified. It is now possible to build archives and tools that are interoperable through these standards. In the following, we propose a three-fold approach to provide a coherent approach to bringing the NVO to the Grid:

- The creation of an NVO Testbed, where astronomical data is copied to TeraGrid resources and data archives are made NVO compliant.
- The execution of high-performance applications that use NVO services and tools to create astronomical knowledge.
- The deployment high performance applications as services to the wider astronomical community.

The applications identified within this proposal cover a broad range of scientific questions and methodologies. They encompass algorithms that comprise large numbers of small, independent work units as well as tightly coupled parallel applications that can utilize the massive parallelism of each TeraGrid resource. We believe these applications are representative of the majority of future NVO grid services. As such they will not only result in answers to fundamental science questions, but they will also serve as an important diagnostic capability for understanding and optimizing the interface between the Grid and the NVO. All applications discussed here are full-fledged and have been developed and tested in smaller grid environments and are ready to be integrated into the TeraGrid framework. Each will yield new and valuable scientific results that would not have been possible without TeraGrid resources.

Our initial TeraGrid proposal requests an allocation of 257,000 Service Units to enable the following research and development:

- Replication of sky surveys onto the TeraGrid for high-speed access
- Porting of NVO services onto the TeraGrid to enable large scale data analysis
- Execution of 6 astronomy data processing applications: 3 data pipeline applications for intermediate processing of survey data and 3 data mining applications for extracting knowledge from surveys

To enable these applications we require a storage space that will house the roughly 20 million images from 5 separate sky surveys, and is needed for analyses that process the entire sky and/or compare objects in multiple sky surveys.

2 NVO-TeraGrid Testbed

Reagan Moore and George Kremenek, San Diego Supercomputer Center

2.1 Data Replication and the NVO

Many current sky surveys are available through NVO protocols from their curators, but the bandwidth may be quite low. For example, the Sloan Digital Sky Survey (~15TB when complete) could be downloaded from Fermilab in 30 days. Many of the NVO grid applications described in this proposal require faster data access rates to remain compute and not I/O bound. To provide an order of magnitude increase in data access, the surveys and their NVO access services must be replicated on higher performance storage systems. Based upon the access rates that have already been demonstrated, it will be possible to analyze an entire 10TB survey in 3 hours once replicated onto the TeraGrid storage. This will enable analyses that process an entire sky survey to a derivative product, to mine the entire archive for patterns, or to federate multiple such surveys.

Tests of the performance of the TeraGrid infrastructure have been tracked, with the following verified results:

- Sustained data transfer rates between SDSC and NCSA have been demonstrated at 250MB/sec over the TeraGrid Network using the Storage Resource Broker. A 10TB collection can, therefore, be moved between TeraGrid sites in 12 hours.
- Data has moved at rates up to 5GB/sec between the SAN and the TeraGrid Linux cluster at SDSC. With this rate a 10TB collection can be accessed in 34 minutes.
- Replication of the 2-Micron All Sky Survey (10TB) and the Digital Palomar Observatory Sky Survey (3TB) onto the SAN at SDSC through use of the Storage Resource Broker. Initial releases of the collections have been successfully installed on the TeraGrid.

The initial component of this research program is to provide the resources necessary to accomplish the subsequent TeraGrid applications. We propose to replicate additional surveys onto TeraGrid Resources. These include the Sloan Digital Sky Survey (~15TB), the United States Naval Observatory B collection (~8 TB), and the MACHO survey (~8TB). Each of these surveys is either wide angle (e.g. 5000 square degrees from the SDSS) or full sky or multitemporal. They encompass static catalogs, catalogs of moving sources from proper motion surveys (USNO-B), and time dependent synoptic surveys (MACHO). The ability to combine data from the listed surveys will advance the conduct of science by enabling large-scale federated analyses of the entire sky.

Replication resource requirements in aggregate are:

- 50 TB of data storage. The total number of images is on the order of 20 million.
- Archival copy of the data on tape, a total of 100TB of tape media. We will store two copies of the images, one on the SAM-FS system at SDSC, and a second on the HPSS system at either SDSC or Caltech. The two copies are required for disaster recovery, and should be implemented on geographically remote sites.
- Disk space to support analysis of the data, up to 20TB at a time for comparisons between surveys.
- Database support for collections.

2.2 Interfacing to NVO Data Services

NVO has developed standard services for accessing images, catalogs, and spectra. These include the Simple Image Access Protocol (SIAP), the OpenSkyNode, and the Simple Spectral Access Protocol (SSAP). Much of the NVO project is directed to the propagation of these standards to astronomical data archives, and we request some TeraGrid resources for building and deploying these services as conduits for the data in the NVO-TeraGrid Testbed.

The data that feeds the projects below will be obtained through these NVO-compliant services with *Atlasmaker*, *Montage* and the *Resolved Star Formation in Galaxies* application using the Image Access protocol (SIAP), the *Quasar Spectra* application using the Spectral Access protocol (SSAP), and the *N-Point Correlation* application using OpenSkyNode.

We request 2000 Service Units to port additional NVO services onto the NVO-TeraGrid testbed to support the data analysis, and image manipulation.

3 Astronomical Grid Applications

Initially, six astronomy applications will be executed under this proposal. The first three are pipeline applications that perform data-intensive intermediate processing on NVO databases. The outputs of these applications would be stored as part of the TeraGrid data grid. The remaining three applications are data mining and simulation tasks designed to analyze and interpret databases that comprise the NVO. Throughout the next section we outline the scientific drivers for each of these components, their algorithms and grid implementations and the resources required to execute them successfully on the TeraGrid.

3.1 Atlasmaker: A Grid-based Implementation of the Hyperatlas

Roy Williams and Michael Feldmann, California Institute of Technology

The Atlasmaker project is using Grid technology, in combination with NVO interoperability, to 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. The Atlasmaker software does resampling and mosaicing of image collections, and is well-suited to operate with the Hyperatlas standard (Williams 2003a). 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).

Atlasmaker is related to Montage (section 3.3): both are doing image mosaicing. However, the latter is focused on a request portal for single mosaics of particular regions of the sky, while Atlasmaker will create large standard atlases to be mined.

3.1.1 Science

We expect these atlases to be a new paradigm for knowledge extraction in astronomy, as well as a magnificent 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.

Often the combination of multiple datasets reveals more knowledge than is present in the components. An optical survey may show clusters of galaxies, and an x-ray survey may show emitting sources. The combination, however, reveals that the x-rays come from the space between the galaxies of the clusters, and we infer the hot intracluster gas that was not obvious in either waveband. This phenomenon is known as *data federation*. Extracting knowledge from federated catalogs is well understood; the next logical step is large-scale federated imagery. This means that images are resampled to a common pixel plane, and catalogs extracted from the joint pixel space.

The set of images being federated might represent the sky at different wavelengths, times, resolutions, polarizations, etc. Images may have been derived from combining and computing with other images. In many cases there is knowledge in the federated imagery that is not present in the federation of catalogs derived from the same original images. It is our contention that the value of this federation can outshine the loss of data quality in the resampling; indeed we believe that image federation opens a new data-centric window on the Universe.

Atlasmaker is a suite of software (Williams 2003a, b) that builds mosaics from NVO-compliant image sets—i.e. images that are exposed through the VO Simple Image Access Protocol. Atlasmaker scavenges a grid for enough computational resources to build terabyte-size atlases, which are then available for federation and data-mining. Atlasmaker is available on the open-source model, and is based on a Grid architecture of relocatable programs and web services. There is a choice of two trusted codes for the resampling kernel: Montage from NASA IPAC/JPL/Caltech and Swarp from the French Terapix project. Each offers different advantages in terms of quality and speed.

When a user requests a given data product, the manager will check for its existence in the cache that stores already computed products and, if it is found, return it the user. If not found the data product will be computed. The names of surveys are resolved by the NVO registry to get required metadata, including current URLs for the NVO services that can provide external image data.

3.1.2 Implementation: Service Based Grid Scavenger

Atlasmaker is designed to actively find and exploit Grid resources that it can use. It is a self-contained package that can be wrapped into a tar file that self-deploys in a robust way. This package is combined with a collection of service requests (which survey to render on which atlas pages) and thrown to a machine for computation. We hope to use TeraGrid hotpage services (when available) to decide the destination automatically.

The Atlasmaker code is written in Python, a highly portable and powerful scripting language, with a central core of pixel manipulation written in C (Montage and Swarp), with MPI parallelism as well as multi-threading. The applications obtain flexibility through the use of services that may reside anywhere on the net; the services are replicated for fault tolerance. Data is stored and retrieved from a distributed file system based on SRB (Storage Resource Broker), a central pillar of the TeraGrid data system. Currently we have implementations for the TeraGrid queue manager, PBS, and also for the Condor environment. The code has run at Caltech, SDSC, and NCSA, using data services at SDSC, Caltech, and Johns Hopkins.

3.1.3 Resource Estimates

We propose to build atlases from some of the major sky surveys. We will initially focus on surveys from 2MASS (10 TB), SDSS (0.5 TB currently released, 15 TB eventually), and DPOSS (3 TB). Resource requirements are estimated for the 2MASS data as this drives the requirements. For a single bandpass (out of 3), and a single atlas page of 5.5 degrees square (out of 1734 pages), we estimate the data transfer and CPU requirements. For a single atlas page we require 2100 input files of 2MB in size (a total of 4.2 GB). A single Itanium CPU can process a single file in 7.5 seconds using Swarp or 140 seconds using Montage. It would, therefore, require 4.6 CPU hours per bandpass per atlas page. This amount of data can be transferred from internally within the Grid in 16.8 seconds. The transfer requirements are, therefore, well suited to the TeraGrid. To process the full 2MASS survey would require 24000 service units. We have asked for about double this (50000 SU) to include multiple surveys. We also require storage space for the resulting projections. A 1-arcsecond projection across the eleven bandpasses of 2MASS, SDSS, and DPOSS requires 16 TBS of storage, or 32 TBs of tape media including replication.

3.2 Resolving star formation in galaxies

Andrew Connolly, University of Pittsburgh
Alex Szalay and Tamas Budavari, Johns Hopkins University

3.2.1 Science

Given a multi-spectral image of the sky or even just multi-spectral cutouts around individual galaxies the resolved nature of these images provides a unique opportunity to study the physical processes that drive the formation and evolution of galaxies within our universe. Previous work has shown that the environment in which a galaxy resides can significantly affect the star formation activity within the galaxy itself. The overall or integrated star formation rate of a galaxy appears to be suppressed in galaxies residing in overdense regions such as the central cores of clusters (Gomez et al 2003, Lewis et al 2002, and Hogg et al 2003). Within these galaxies in overdense regions the star formation appears to be more centrally condensed than for field galaxies (Rose et al 2001).

A number of mechanisms for causing this circumnuclear activity (as opposed to the more general star formation in HII regions within the disks of galaxies) have been suggested. These include tidal stripping of the gas from the outskirts of a galaxy as it impacts the intracluster medium, merging of galaxy subcomponents and galaxy harassment (Moore et al 1996). To differentiate between these multiple scenarios we need to determine how star formation is distributed **within** galaxies not just the total luminosity of a galaxy (e.g. in its simplest case ram pressure stripping will result in a suppression of star formation in the outer disk of a galaxy). Such a study is, however, limited due to the relative small sample of galaxies available in current studies of the internal distribution of star formation (e.g. the Rose study comprised less than 30 galaxies). We require several orders of magnitude increase in the number of galaxies to get a representative sample of galaxies; one that includes galaxies with a wide range of luminosities and spectral types as well as galaxies residing in over and underdense regions.

We can gain this order of magnitude increase in sample size if we consider the multi-spectral imaging data available through the NVO. It is well known that, from multicolor photometric data, we can estimate the redshift (photometric), age, metallicity, dust and star formation rate of galaxies by fitting spectral energy distributions (SEDs; Bruzual and Charlot 1993) to their integrated fluxes. This SED fitting technique can be extended to consider age, metallicity and star formation as a function of position within a galaxy (i.e. a measure of the resolved star formation). This can be achieved either on a pixel basis (Conti et al 2003) or for groupings of pixels with similar colors (Budavari et al. 2003). The choice of how we subdivide the data is simply dependent on the signal-to-noise of the image and the number of resolution elements.

3.2.2 Implementation

Our implementation of the SED fitting procedure is built upon a multi-resolutional grid (standard hill climbing search techniques are not well suited to this task due to the presence of multiple minima within the likelihood space and the need to characterize these local minima not just find a simple global minimum for the fit). Local minima within the four dimensional space are initially identified in the lowest resolution mode of the grid. The grid is then progressively refined to increase the resolution of the likelihood space around each of these local minima. This enables an accurate measure of the minima together with the form of the likelihood space around these minima (i.e. the likelihood surface is not typically symmetric about the minima). Applying this technique we characterize, for each pixel, the full likelihood space at low resolution with increased resolution around our regions of interest. To minimize the size of the output from this analysis we marginalize the likelihoods for each parameter (star formation, age, metallicity and dust) and characterize these one-dimensional likelihood functions as a sum of Gaussians. In this we way, for each pixel, we typically extract approximately 24 parameters (typically we can characterize the one dimensional likelihood functions with only two Gaussians). All software to date has been written in ANSI C and C++. The techniques have been implemented and run under the Condor and PBS environments on small and large clusters of CPUs.

3.2.3 Resource requirements

We propose applying the fitting techniques described above to a representative sample of galaxies selected from the SDSS (approximately 500,000 resolved galaxy images). The galaxies in this sample have a median redshift of z=0.1 and a typical size of 50x50 pixels (corresponding to a spatial resolution of 0.5 kpc per pixel). In the low-resolution mode (i.e. the lowest resolution of the grid used to characterize the global likelihood function) it takes approximately 180s on a 2 GHz machine to analyze a 50x50 pixel image (the typical image resolution of the SDSS at these bright magnitudes). The higher resolution grid minimizations and the marginalization of the likelihood functions increase the overall system time by approximately 20%. Therefore, to characterize the 500,000 images within the SDSS, we require approximately 30,000 CPU hours. To store the outputs of this analysis we require approximately 200-400 GB of disk space (dependent on how we store the final marginalized likelihoods).

3.3 Montage

Bruce Berriman, IPAC, California Institute of Technology,
Ewa Deelman, ISI, University of Southern Californi, and many others

3.3.1 Science

Montage acts as a virtual telescope providing on-demand processing of astronomical images to deliver science grade mosaics of any projection or spatial sampling. Such an application has broad applicability in astronomy (see Montage 2002 for science applications; see also Atlasmaker this proposal). The goal is to create an output image which is as close as possible to that which would have been created if the sky had been observed with an instrument with the same resolution as the output image. When fully deployed, Montage will provide the computational tools for computing science grade full sky images, and for producing multi-wavelength astrophysical images. Examples are the image products expected to be delivered by the Spitzer Space Telescope (formerly SIRTF) Legacy program. These products will consist of infrared images measured with Spitzer, supplemented with multi-wavelength images from the ground for maximal scientific return. DPOSS, SDSS and 2MASS provide the broad wavelength base for these images.

The computational problem implied by this goal is that the flux in the input pixels must be redistributed to the output pixels in proportion to the overlap between the input and output pixels. The key to a general solution lies in the recognition that the preservation of astrometric (positional) and calibration fidelity of the input images requires that input and output pixels must be projected on to the celestial sphere, with the overlaps calculated by classical spherical trigonometric algorithms for polygon overlaps. Since this would be done using Grid resources, the parallelization inherent in the architecture can be exploited to the maximum. All re-projection jobs can be added to a pool of tasks and performed by as many processors as are available. The same is true of the other list driven processes inherent within Montage.

Montage is related to the Atlasmaker project (section 3.1); however the latter is focused on building standard data products (multiwavelength atlases) rather than a portal to service requests (i.e. individual regions of the sky).

3.3.2 Implementation

Our goal is to enable efficient execution of important astronomy applications on the TeraGrid. In some cases, the execution of applications will be performed using the Pegasus software, which stands for Planning for Execution in Grids (PEGasus). So far, we have been able to run the Montage application on a Condor pool and have performed a series of runs on the TeraGrid nodes at SDSC and NCSA.

We have developed a prototype architecture for on-demand production of image mosaics on the TeraGrid. A web service located at JPL creates an abstract workflow description of a mosaic request made to Montage by calling the DAG-building script. This workflow description, written in XML, is submitted to Pegasus. Pegasus, developed at ISI as part of the GriPhyN project, is able to map and schedule application workflows expressed in an abstract fashion, i.e. without reference to particular data and resources.

The Pegasus Concrete Planner queries the Globus Replica Location Service (RLS) to find the location of the input files. These are specified by their logical filenames in the DAG. RLS returns a list of physical locations for the files. The information about the available data is then used to optimize the concrete workflow from the point of view of Virtual Data, by reusing existing data products. Pegasus queries the Transformation Catalog to determine if the components are available in the execution environment and to identify their locations. The Transformation Catalog performs the mapping between a logical component name and the location of the corresponding executables on specific compute resources. Transfer nodes are added for any of these files that need to be staged in, so that each component and its input files are at the same physical location. If the input files are replicated at several locations, the concrete planner currently picks the source location at random. Finally transfer nodes and registration nodes, which publish the resulting data products in the RLS, are added if the user requested that all the data be published and sent to a particular storage location.

So far, Pegasus has been used to execute applications from various scientific domains, such as high-energy physics, astronomy and gravitational wave physics. The proposed work would extent Pegasus to execute applications on a new platform—the TeraGrid and on a greater scale.

3.3.3 Resource Requirements

With over 90% of processing time required to generate a mosaic consumed by resampling and redistribution of flux, a single 2MASS image of 1024 x 512 pixels requires 100 seconds of processing on a 2.0 GHz P4 box. Serial processing of a full sky survey of 4,000,000 2MASS images would take over 12 compute-years. Speed-up is easily achieved by taking advantage of the parallelization inherent in the Montage design. The reprocessing and resampling of individual images can be performed on as many processors as are available, and we have taken advantage of this in running Montage on the TeraGrid. We will compute image mosaics 90x90 degrees at the center of our galaxy and in the Orion-Taurus region. These will take 14000 compute hours in total. For compute resources and storage we request 20,000 service units and 0.75TB storage to support the goals of the next phase of our work.

3.4 Fitting Quasar Spectra

Dan Vanden Berk, University of Pittsburgh
Sebastian Jester, Fermilab
Robert Nichol, Carnegie-Mellon University

3.4.1 Science

Model fitting to observed spectra is a common problem in science, particularly so in astronomy. In the case of quasars and other powerful active galaxies, the spectrum—the flux density as a function of photon wavelength or frequency—contains information about the energy source, the radiation physics, and the geometric structure of the objects. Of particular interest is the mass of the central object, almost certainly a supermassive black hole, which can be derived once model parameters have been determined from the spectrum. A primary scientific goal of this program is to measure the distribution of black hole masses throughout cosmic history, using over 50,000 quasar spectra available from the Sloan Digital Sky Survey component in the NVO.

The spectroscopic models for quasars consist of 5 basic components: a continuum power-law, a Balmer continuum contribution (free-bound emission of Hydrogen), emission lines from non-Iron transitions, complex emission from overlapping Iron transitions, and emission from the host galaxy. The widths of the non-Iron emission lines, combined with the continuum luminosity, give a reliable measure of the central black hole mass. In addition, the measurements allow numerous other investigations to be carried out, ranging from the physics of accretion phenomena, to the possible determination of cosmological parameters.

The primary problem is that the parameter space is enormous. A single spectrum may be modeled with well over 100 parameters, mainly because of the large number of emission lines. The values of the parameters can often range over an order of magnitude. In addition, the various components are not independent, which necessitates a search over the entire parameter space simultaneously. While an interactive gradient-based fitting program may be adequate for a small handful of spectra, we need a reliable automatic routine that can be applied to our database of over 50,000 spectra.

3.4.2 Implementation

Traditional spectral modeling routines usually attempt to simplify the problem by making assumptions about the continuum and the Iron emission. For example, a power-law "continuum" is usually "fit" by assuming certain spectral regions are free of emission line contamination, and then the function is subtracted before continuing. However, it is well known (e.g. Vanden Berk et al. 2001) that there are virtually no regions free of emission line contamination—the continuum and emission must be accounted for simultaneously. In addition, a single universal Iron emission template is usually subtracted from the spectra—a technique which has proven to be greatly inadequate. Finally, the problem of contamination of the quasar spectrum by light from the host galaxy is almost never addressed.

We have developed a genetic-algorithm (GA) based routine to find near optimal solutions to the model fits of quasar spectra. The routine employs the standard genetic operators of crossover, mutation, and elitism, as well as "strategy parameters" which vary the search space as optimal solutions are approached. To describe the steps briefly, an initial population of random individuals (solutions) is generated, and the goodness of fit (determined by χ^2 values in this case) of each individual is evaluated. Members of the first generation are selected at random, but weighted by fitness, for "breeding" and "mutation" in order to form the next generation of solutions. The process is iterated until either an acceptable fitness or maximum number of iterations is reached.

The use of GAs for spectral fitting is not new (e.g. McIntosh et al. 1995), but they are not yet widely employed in astronomy. Our algorithm uses a number of novel features to improve the performance of standard GAs. For example, we have included "strategy parameters" which determine the probabilistic parameter search radius during mutations. The strategy parameters either shrink or expand depending on whether the best solutions change little or greatly from generation to generation. We have also used spectral principal component analysis (Yip et al. 2003) to reduce the complexity of the host galaxy component to a small number of eigenvectors.

3.4.3 Resource Requirements

Our task is to apply our GA algorithm to the spectra of over 50,000 quasars from the Sloan Digital Sky Survey component of the NVO. The routine generates near optimal solutions on a single-processor 2GHz Pentium4 machine in approximately 1 hour for a single spectrum. The request is therefore for 50,000 service units. Our strategy for the grid application is to process each spectrum independently on a separate node, simultaneously using as many nodes as are available. This is in a sense a fiducial Grid project, wherein large numbers of bite-sized computations may be asynchronously scheduled via Condor, PBS or the TeraGrid queue manager on heterogeneous TeraGrid platforms.

3.5 N-Point Correlation Functions of Galaxies

Robert Nichol, Carnegie Mellon University
Jeffrey Gardner, Pittsburgh Supercomputing Center
Andrew Connolly and Joerg Colberg, University of Pittsburgh
Jim Gray, Microsoft Corporation

3.5.1 Science

A fundamental ambition of any scientific analysis is to quantify any clustering within the data. This is particularly true in astrophysics as gravity naturally leads to a highly clustered universe. Cosmologists have chosen to characterize such clustering using n-point correlation functions (see Peebles 1980) which are a non-parametric measure of the higher order moments of a distribution of points. For example, the 2-point function of galaxies in the Universe quantifies the excess (or deficit) of pairs of galaxies as a function of their separations, compared to a Poisson distribution. Likewise, the 3-point function is the excess (or deficit) of triplets (triangles), and so on up the hierarchy of combinations (4-point, 7-point, *etc.*). If the structure in the Universe originated from Gaussian initial conditions predicted by Inflation, then all the higher order correlation functions (3-point and above) would only depend upon the 2-point function, the lowest order correlation function. Any other set of initial conditions, or non-linear physics, would result in a non-trivial set of higher order correlation functions and therefore, it is imperative that we accurately measure these functions to test our fundamental cosmological theories and assumptions. Precision measurements of the higher-order correlations of galaxies are now possible due to the availability of high quality data from the SDSS survey.

Through a multi-disciplinary research team of astrophysicists, computer scientists and statisticians (the PiCA group), we developed a fast algorithm using kD-trees to speed up the calculation. The key is to avoid an explicit search through the data set for all configurations of data points. Our innovation is to use multiple kD-trees together (*i.e.*, using n kD-trees for the n—point function) to represent the data and searches for configurations that satisfy a matching criteria. For the 2-point function on 10^6 SDSS galaxies, we can achieve a 10^4-fold speed-up over the naive quadratic algorithm and a 25-fold speed-up over other tree-based algorithms. The 3-point and 4-point results show even larger speed-ups.

A further refinement to the standard n-point function is the projected n-point function which, although more costly, is better suited to astronomical surveys. In galaxy surveys, for example, the 2D position of the galaxies on the sky is known far more accurately than their distance from us. Consequently, it is better to cast the correlation function in terms of the projected separation of galaxies on the sky and their distance along the line of sight (as opposed to using their x-y-z positions as in the standard case). We still build the kD-tree in x-y-z coordinates—this still turns out to be more efficient than having a tree in projected space—but we calculate the projected separations while traversing the tree. Like standard 3-point, we are the only group to have implemented projected 3-point for more than 60,000 galaxies.

3.5.2 Implementation

We have recently taken our concept further and re-engineered our kD-tree methods into a sophisticated and highly optimized parallel implementation for both standard and projected n-point. Our application is both scalable and portable. In order to minimize communication latency, we employ a data caching algorithm that predicts future off-processor communication and attempts to preload the required data. This algorithm is so successful that even on 128 PEs, we preload 99.99% of off-processor data before it is needed. In order to achieve maximum performance and portability, all parallel communication is encapsulated in a Machine Dependent Layer (MDL), a small library of functions that is written for each parallel architecture in order to take advantage of the strength of each one. To date, there are MPI, PVM, POSIX Threads, Cray SHMEM, and Quadrics/Elan MPI versions of MDL. All I/O is done through the XDR library, ensuring binary compatibility between all architectures. In the future, we will also include XML headers. Consequently, in a Grid environment we achieve maximum portability between heterogeneous Grid resources while simultaneously capitalizing on the attributes of each one.

As part of our development, we are creating a webservice interface to the n-point code that we will make available through the NVO. Our webservice include the option of running n-point analyses on local or distant machines and will interface with TeraGrid scheduling software.

3.5.3 Resource Requirements

We have presently applied our n-point correlation function codes to computing the standard 3-point function for 130,000 galaxies from the SDSS DR1 database. Our current benchmarks for standard 3-point require 40 cpu-hours on a 1.3 GHz Itanium2 per 100 triangle configurations. These initial measurements, while state of the art compared to earlier works, are far from ideal. Quantifying the form of the correlation function will require the next generation of 3-point SDSS measurements which requires significant computer resources: for this we turn to TeraGrid. Specifically, we propose to scale-up the standard 3-point calculations in three ways:

- The SDSS dataset has expanded and, during the time period of this TeraGrid proposal, there will be 500,000 galaxies with 3D position information; a factor of four increase in sample size.

- We need to increase the number of triangle configurations explored from 100 to 1000. To date we are forced to use quite course binning (where a bin represents a single triangle configuration) to cover fully the of possible configuration space. This results in a loss of resolution and we have potentially "averaged-over" subtle features in the 3-point e.g. the predicted peak in the 3-point at 2Mpc (see Takada & Jain 2003).
- The error on the 3-point function needs to be calculated from an ensemble of fake universes (that have the same shape as the real data). These fake Universes are now easy to construct either using N-body simulations or fast non-linear approximations. We would like to use a minimum of 10 fake universes to measure the error on our 3-point.

Taken together, these advances will require approximately 60000 SUs, or over 1000 times the computation of our current results. We propose to use the parallel capability of our code to fully exploit the unique architecture of the TeraGrid.

We have also calculated the projected 3-point function for the current SDSS DR1 database. The 130,000 galaxies requires approximately 400 node hours on a 1.3GHz Itanium2. In standard n-point, the tree can be traversed separately for each triangle configuration, allowing the problem to be divided up into smaller independent work units, where each unit is a single triangle configuration. However, in the projected case it is more efficient to traverse the tree only once and compute all triangle configurations at each tree node. Thus, calculating the projected 3-point function for the SDSS is a single, tightly-coupled massively parallel job. Our goal is to calculate the projected 3-point function for the new SDSS dataset of 500,000 galaxies, which given our current benchmarks will require 25,000 node-hours on the on the PSC TCS.

We have written a version of our MDL that takes advantages of the capabilities of the Elan library of the PSC TCS machine, and contributes greatly to the scalability of our code. Consequently, we hope to run the projected 3-point on upwards of 1000 PEs on the TCS. For the standard 3-point calculations, we anticipate using the majority of our node-hours on the TCS as well, since the most computational time will be spent on the minority of time consuming triangle configurations that need to be done in parallel.

Our long-term goal for future proposals is to replicate these calculations for the higher order correlation functions, e.g. the 4-point function, as well as to include higher-order cross-correlations with other datasets, e.g. the 3-point cross-correlation between SDSS & WMAP. The substantial computational and interconnect requirements of these ambitious future projects will make TeraGrid platforms the ideal, if not the only, resources by which to realize the full scientific potential of these datasets.

3.6 The Cosmic Microwave Background Grid

Christopher Miller and Robert Nichol, Carnegie Mellon University

3.6.1 Science

Since the discovery of fluctuations in the Cosmic Microwave Background (CMB) in 1992 (Smoot et al. 1992), a goal in observational cosmology has been to use well-understood models of the CMB to constrain the physical parameters that govern the evolution of our Universe. The physical model requires the numerical integration of equations of motion for a perfect fluid (the Universe is theorized to be a perfect fluid at the time the CMB was created). While a single model may only require 10 to 100 seconds to compute, the real computational challenge comes from the need to calculate over 10^6 models to fully sample the cosmological parameter space!

The need for the large number of models arises because the cosmological models occupy a high dimensional space that is degenerate over many of the parameters. Ideally, one would construct a grid of models for all variations of the chosen parameters. A minimization technique would then be applied to the data and the best model parameters and 1σ confidence intervals reported. At approximately one minute per model a grid of six parameters can, however, be extremely time consuming. Two procedures have, therefore, been developed to overcome these numerical issues: adaptive grids and Markov Chain Monte

Carlo (MCMC). While these approaches are efficient computationally neither is entirely satisfactory. The adaptiveness of a grid is typically subjective to a specific project while the MCMC technique does not map the full parameter space and does not fare well in high dimensional degenerate spaces such as the CMB models.

3.6.2 Implementation

Given that the determination of the basic cosmological parameters that define the geometry of our universe is a fundamental goal of modern cosmology it is essential that we be able to accurately and robustly characterize not only the optimal solution but also the uncertainties on this fit. We have, therefore, developed a new non-parametric technique to analyze CMB data that can fully characterize the likelihood space when fitting cosmological parameters (Miller et al. 2002). Our method relies on multiple realizations of a candidate universe, and so is complementary to the MCMC procedure. To implement this technique to determine the optimal fit (and uncertainties) to the CMB power spectrum as measured by the WMAP satellite we require a grid of 10^6 CMB models that we can analyze and then make available to the cosmological community. From a scientific viewpoint, our technique provides an important frequentist (or multiple realization) perspective on the current CMB data, which to date has been analyzed using Bayesian (i.e. non-frequentist) techniques. While the two approaches are expected to agree in the limits of large datasets and well-informed priors, neither the datasets nor the priors of today are sufficient. Thus, the scientific merit of a new approach is clear.

In addition, the toolbox required for this research is well established (and published). For instance, the code we use to generate the CMB models is a standard in the field (CMBFAST—Seljak and Zaldarriaga 1996) and the non-parametric method we have developed has been used in more limited analyses of older CMB data. Our ability to apply our technique to the new CMB data from the Wilkinson Microwave Anisotropy Probe (WMAP) hinges on large-node computing resources such as TeraGrid.

3.6.3 Resource Requirements

We need to calculate 10^6 models where a single model runs an average of one minute on a 2 GHz Xeon processor. For each model, we will also calculate a goodness-of-fit based on our non-parametric technique (Miller et al. 2002). Altogether, we require 20000 compute hours to accomplish our goal. On completion, we will make our grid of models available to the community through a fast database server. We expect the availability of these models to play an important role in NVO science for the coming decade, as new and higher quality CMB data become available (through WMAP and later Planck).

4 Summary of Resources Requested

The bulk of the Service Units requested will be used by the 6 astronomy data processing applications we have described, and a further 2000 SUs will be needed to port NVO services. These are detailed in the table below.

Application	Service Units
Atlasmaker	50,000
Resolving Star Formation in Galaxies	30,000
Montage	20,000
Fitting Quasar Spectra	50,000
N-Point Correlation Functions of Galaxies	85,000
Cosmic Microwave Background Grid	20,000
Porting NVO Services	2,000

TOTAL:	**257,000**

We will also require 50 TB of archival storage for 20 million images from the 5 principle sky surveys, 100TB of online storage for these images and replicas, plus an additional 20TB of online storage to support the data analysis. For the hyperatlas catalog, 16TBs of storage is required. The total archival storage space that is needed is 116TBs.

5 Project Management and Future Plans

5.1 TeraGrid Resource Management

The testbed and applications described above represent the initial components of a multiyear effort to bring the NVO onto the TeraGrid. These applications were selected due to their mature state of development, their utility in optimizing how the TeraGrid might interact with a broad range of NVO applications and their ability to produce fundamental scientific returns in a short amount of time. Each application is representative of the sort of work that the NVO consortium is encouraging in the area of high-performance computing.

Coordination and management of the resources awarded under this grant will be undertaken by the co-Is of this proposal (Williams, Connolly and Gardner) with the oversight of the NVO Executive Committee. This will entail ensuring the individual science are met together with the dissemination to the astronomical community of the knowledge gained through working at the interface of the TeraGrid and the NVO. Over the course of this multi-year, large research collaboration, we expect that the nature of the grid services available through the NVO will evolve with new projects arising that can effectively use the TeraGrid (see below for an outline of our next generation of NVO-TeraGrid applications). Allocation of resources for new applications will be coordinated by the coIs under the direction of the NVO Executive Committee. The committee will review proposed new applications to determine their suitability as NVO-TeraGrid components and their resource requirements (drawing upon the expertise of independent, external reviewers as necessary). The development of accepted new applications will be assisted through the use of expedited TeraGrid accounts.

5.2 Future Development of the NVO-TeraGrid Applications

Over the course of this proposal new grid based applications will be implemented that build upon the services described in this proposal. While this process will clearly evolve with the development of the NVO we note here a number of components that are under active development and that we expect to be included in subsequent requests for TeraGrid resources. These applications include ones that will increase the interactivity and visibility of grid services in the community as well as new analysis and data mining routines.

5.2.1 On-Demand Services

Montage and the QSO spectral fitting components will be extended to provide an on-demand computational engine for the astronomical community. This will enable an outside user to request finding charts or images of regions of the sky at any of the frequencies available through the NVO TeraGrid data services or to upload spectra for analysis by the QSO spectral fitting grid service. These applications will be presented as a web-based portal, and we would like to put TeraGrid resources behind these. An authorized TeraGrid user from this project would take responsibility for the service, guaranteeing that anonymous requests are sufficiently small, and that larger requests are directed to obtain a personal allocation and log in with a certificate.

5.2.2 Image Federation

Montage and Atlasmaker are currently designed to provide a series of multi-spectral images of regions of the sky with a common image resolution. Current use cases for these applications focus on providing finding charts or images of regions of the sky that have been observed by a new telescope or at a new wavelength (i.e. to determine the properties of astronomical sources that have already been detected). Building upon these services we can, however, optimally combine these multi-spectral images to detect sources simultaneously at all frequencies (Szalay, Connolly, Szokoly 1999). This multispectral image detection algorithm has the potential to detect sources that would be missed by an analysis of any one of the individual frequencies (e.g. sources that are just below a detection threshold in any one image may be very significant detections if we coadd all images) and possibly to identify new classes of astrophysical phenomena.

5.2.3 Data Mining Services

Development of data mining grid services will be driven by the new generation of astronomical surveys that are sampling the temporal domain (e.g. SuperMacho, Quest, Pan-STARRS and the LSST). The goals of these surveys are to identify variable and moving sources (from supernovae to near-Earth asteroids) in almost real-time. The opening of the time domain provides many challenges for which the TeraGrid, combined with the NVO, will be a unique resource. Source detection on data sets that comprise 8TB per night where we require an almost real-time turnaround (to identify supernovae or asteroids with sufficient time for followup observations at other telescope facilities) constitutes a substantial computational and storage challenge.

5.2.4 Theory-Observation Interface

Our near-term goals have focused on the analysis of observational data. The longer term goal of this proposal, and the NVO as a whole, is to interface observations and theory. The development of grid services for interacting with large observational data sets naturally extends to the data produced from large N-boby simulations of the universe. Our expectation is that simulated data sets will evolve into synthetic observations that will be analyzed through the NVO grid services. To accomplish this the NVO-TeraGrid data services will be expanded to include simulated as well as the observational data. With standard NVO access protocols grid services developed will be applicable to both simulated and observational data which will facilitate direct comparisons between the properties of the observed and simulated universe.

6 Project Team Qualifications

To accomplish both the near and long term goals of this program to bring the NVO onto the TeraGrid we have compiled a collaboration of astrophysicists and computer scientists who share the vision that the TeraGrid provides a unique opportunity to address some of the most intractable problems in astrophysics. Our team contains a broad expertise in high performance computing ranging from the development of algorithms for analyzing astrophysical data to the question of optimal storage of massive data sets and the integration of databases with data mining applications.

We expect substantial scientific return through the fusion of astronomical science and TeraGrid. Two of the core applications have already been ported and implemented on the TeraGrid with the remaining applications all running on smaller compute grids.

Research to be undertaken through this proposal is fully funded under a number of NSF awards. These include four NSF Information Technology Research awards; the **National Virtual Observatory**, **Algorithms for Cosmology**, **Grist** (Grid Services for Data Analysis in Astronomy), and **PiCA** (Pittsburgh Computational AstroStatistics).

7 References

- Brunner, R., Djorgovski, S.G., Szalay, A., *Virtual Observatories of the Future*, Astronomical Society of the Pacific Conference Series, Vol. 225, pp. 257-264, June 2000.
- Budavari, T., Szalay, A., Connolly, A., Csabai, I., 2003, AJ, in preparation
- Conti, A., Connolly, A., Hopkins, A., Budavari, T., Csabai,I., Szalay, A., Schmidt, S., Petrovich, N., 2003, AJ, 126, 2330.
- Deelman, E., Blythe, J., Gil, Y., Kesselman, C., *Workflow Management in GriPhyN*, in Grid Resource Management, J. Nabrzyski, J. Schopf, and J. Weglarz editors, Kluwer, 2003.
- Deelman, E., *et al*., *Mapping Abstract Complex Workflows onto Grid Environments*, Journal of Grid Computing, Vol.1, no. 1, 2003, pp. 25-39.
- Deelman, E., *et al*., *Grid-Based Galaxy Morphology Analysis for the National Virtual Observatory*, Proceedings of SC 2003.
- Gomez, P., et al, 2003, ApJ, 548, 210
- Hogg et al, 2003, ApJ in press
- Jing, Y.P., & Boener, G., 2003, Astroph. J., submitted, astro-ph/0311585
- Lewis, I. et al. 2002, MNRAS, 334, 673
- Miller, Nichol, Genovese, & Wasserman, 2002, Astroph. J., 565, L67
- Miller, C.J., Nichol, R.C., Genovese, C., & Wasserman, L., 2002, Astroph. J., 565, L67
- Montage 2002, Science Use Cases, `http://montage.ipac.caltech.edu/docs/usecases.html`
- Peebles, P.J.E., 1980, *The Large-Scale Structure of the Universe*, Princeton, NJ: Princeton UP.
- Rajasekar, A., Wan, M., Moore, R., Kremenek, G., Guptil, T. *Data Grids, Collections, and Grid Bricks*, Proc. 20th IEEE Symposium on Mass Storage Systems and Eleventh Goddard Conference on Mass Storage Systems and Technologies, San Diego, 2003.
- Rajasekar, A., Wan, M., Moore, R., Jagatheesan, A., Kremenek, G., *Real Experiences with Data Grids - Case studies in using the SRB*", Int. Symp.High-Performance Computer Architecture, Kyushu, Japan, 2002.
- Rajasekar, A., Moore, R., Ludäscher, B., Zaslavsky, I., *The Grid Adventures: SDSC's Storage Resource Broker and Web Services in Digital Library Applications*:, 4th Russian Conference on Digital Libraries, Dubna, Russia, 2002.
- Rose, J.A., Gaba, A.E., Caldwell, N., & Chaboyer, B. 2001, AJ, 121, 793
- Seljak, U., & Zaldarriaga, M., 1996, ApJ, 469, 437
- Smoot, G. F. et al., 1992, Astroph. J., 396, L1
- Szalay, A. S., Connolly, A. J., Szokoly, G. P., *Simultaneous Multicolor Detection of Faint Galaxies in the Hubble Deep* Field, AJ, 117, 68, also astroph/9811086.
- Takada, M., & Jain, B., 2003, Mon. Not. Royal Ast. Soc., 344, 857
- Williams, R. D., Djorgovski, S. G., Feldmann, M. T., and Jacob, J. C., 2003a, *Hyperatlas, A New Framework for Image Federation*, `http://arxiv.org/abs/astro-ph/0312195`
- Williams, R. D., Djorgovski, S. G., Feldmann, M. T., and Jacob, J. C., 2003b, *Atlasmaker, A Grid-based Implementation of the Hyperatlas*, `http://arxiv.org/abs/astro-ph/0312196`
- Williams, R. D., Feldmann, M. T., 2003c, *Atlasmaker, Software to build Atlases of Federated Imagery*, 2003, `http://www.cacr.caltech.edu/projects/nvo/atlasmaker/`
- Williams, R. D., et. al., *Multi-Wavelength Image Space: Another Grid-Enabled Science*, Concurrency & Computation, vol 15 (2003), pp 539-549.
- Williams R. D., *Grids and the Virtual Observatory*, pp 837-858 in *Grid Computing: Making The Global Infrastructure*, eds. F. Berman, A. J.G. Hey, and G. C. Fox, Wiley, 2003.