An International Virtual-Data Grid Laboratory
for Data Intensive Science

Paul Avery, University of Florida
Ian Foster, University of Chicago
Rob Gardner, Indiana University
Harvey Newman, California Institute of Technology
Alexander Szalay, Johns Hopkins University

Submitted to the 2001 NSF Information and Technology Research Program
Proposal #0107441
A. Introduction

We propose to establish and utilize an international Virtual-Data Grid Laboratory (iVDGL) of unprecedented scale and scope, comprising resources in the U.S., Europe, and other world regions. Our goal in establishing this laboratory is to drive the development, and transition to global production use, of Petabyte-scale virtual data applications required by frontier computationally oriented science. In so doing, we seize the opportunity presented by a convergence of rapid advances in networking, computer technology, Grid infrastructure, and application sciences, as well as substantial investments in data-intensive science now underway in Europe, U.S., and Asia. We expect experiments conducted in this unique international laboratory to influence the future of scientific investigation in a wide range of disciplines, including astronomy, astrophysics, high energy physics, and bioinformatics. A significant additional benefit of this facility is that it will empower a set of universities who normally have no access to top tier facilities, hence bringing the methods and results of international scientific enterprises to a diverse, world-wide audience.

Data Grid technologies embody entirely new approaches to the analysis of large data collections, in which the resources of an entire scientific community are brought to bear on the analysis and discovery process, and data products are made available to all community members, regardless of location. Large projects such as the NSF-funded GriPhyN and the European Union DataGrid project are developing the basic technologies required to create working data grids. What is missing is (1) the deployment, evaluation and optimization of these technologies on a production scale, and (2) the integration of these technologies into production applications. These two missing pieces are hindering progress in the basic IT research underlying data grids, and are thereby slowing adoption of data grid technology by the scientific community. In this project we aim to overcome these obstacles to progress.

The following figure illustrates the structure and scope of the proposed virtual laboratory. Laboratory users will include international science experiments such as LIGO, SDSS, ATLAS, CMS, and NVO (National Virtual Observatory), as well as outreach activities and Grid technology research efforts. The laboratory itself will be created by deploying a carefully crafted data grid technology base across a set of resource sites, each of which provides substantial computing and storage capability managed by iVDGL software. The 20+ resource centers, of varying sizes, will include U.S. sites put in place specifically for the laboratory, sites contributed by the European Union and potentially other international collaborators, existing experiment facilities, and facilities placed at outreach institutions. In addition, a Grid Operations Center will provide the essential operational element required to ensure functionality and to reduce operational overhead on resource centers.

Specific tasks to be undertaken in this project include the following. (1) Construct the international virtual laboratory, including development of new techniques for low-overhead operation of a large, internationally distributed facility; (2) adapt experiment-specific analysis tools to exploit iVDGL features, including the Virtual Data Toolkit of the GriPhyN project; (3) conduct ongoing and comprehensive evaluations of both data grid technologies and the experiment’s analysis systems in the iVDGL, studying performance at all levels from network to application in a coordinated fashion, and (4) based on these evaluations, formulate system models that can be used to guide the design and optimization of Data Grid systems and applications. The experience gained with information systems of this size and complexity, providing transparent access to massive distributed data collections, will be applicable to large-scale data-intensive problems in many fields of science, engineering, and eventually industry and commerce. Such systems will be needed in the coming decades as a central element of our information-based society.

It is our intention to promote learning and inclusion both via the integration of minority institutions into the fabric of the iVDGL (by placing resource centers at those institutions to facilitate participation in the project) and via the integration of a diverse set of U.S. universities into the scientific program of participating physics and computer science experiments (by expanding the list of institutions participating in Grid projects). The project will be conducted by an exceptional team of leading application scientists and computer scientists and will be incorporated within the management structure of GriPhyN to tightly integrate the application and research efforts.
B. Data-Intensive Science and Data Grids

A new generation of frontier physics experiments is coming on line whose success depends on the ability to acquire, analyze, and disseminate vast quantities of data. The experiments that are the focus of this proposal are the CMS\textsuperscript{3} and ATLAS\textsuperscript{5} experiments at the LHC (Large Hadron Collider) at CERN, Geneva, LIGO\textsuperscript{6,7,8} (Laser Interferometer Gravitational-wave Observatory) and SDSS\textsuperscript{9} (Sloan Digital Sky Survey). For the next two decades, the LHC will probe the TeV frontier of particle energies to search for new states and fundamental interactions of matter. LIGO will detect and analyze, over a similar span of time, nature's most energetic events sending gravitational waves across the cosmos. SDSS is the first of several planned surveys that will systematically scan the sky to provide the most comprehensive catalog of astronomical data ever recorded. The federation of these catalogs will form the basis for the planned National Virtual Observatory\textsuperscript{10} (NVO).

Exploring the scientific wealth of these experiments presents new problems in data access, processing and distribution, and collaboration across networks. The LHC experiments, for example, will accumulate data volumes of hundreds of petabytes of raw, derived and simulated data. Finding, processing and separating out the rare “signals” in the data will be hard to manage and computationally demanding\textsuperscript{11}. The intrinsic complexity of the problem requires tighter integration than is possible by scaling up present-day solutions using Moore’s-law technological improvements. As a result, in previous work we defined a four level\textsuperscript{12} Grid structure: Tier 1, a US regional center for ATLAS, CMS, LIGO or SDSS, located at a laboratory or university; Tier 2, University Regional Centers (URCs); Tier 3, University Research Groups (URGs); and Tier 4, an access layer made up of PCs, workstations and other devices. As in GriPhyN, we note the need to access and manage virtual data, which may not physically exist except as a specification for how it is to be calculated or fetched. Petascale Virtual-data Grids (PVDGs), which manage this additional complexity, offer a new degree of transparency in the delivery of both data-handling and processing resources to end-user applications.

We plan in this project to implement a PVDG hierarchy that will enhance the discovery potential of the experiments while providing invaluable real-world experience with Grid systems. Our project will (1) exploit and extend the synergy between the Computer Science teams and the physics teams in the GriPhyN project by integrating Grid tools with scientific applications, (2) forge new connections with major European Grid efforts (and eventually Grid efforts in all world regions), (3) create with them a large-scale research environment in the iVDGL that will be of immense value to researchers in many disciplines and yield unique insights, and (4) exploit this laboratory to build Grid-enabled data analysis systems and support-tools, conduct large scale tests to prove the systems, and produce scientific results. We expect the iVDGL to provide the basis for an international Grid infrastructure that would be used by researchers around the world.

C. Development and Deployment of an International VDGL

PVDG concepts have been recognized as central to scientific progress in a wide range of disciplines. Simulation studies have demonstrated the feasibility of the basic concept and projects such as GriPhyN are developing essential technologies and toolkits. The next step is to create facilities to enable large-scale experimentation. The history of studies of nonlinear systems of this kind makes it clear that experimentation at scale is required for correct insights into the key factors controlling the system behavior, and the derivation of effective strategies for system operation that combine high resource utilization levels with acceptable response times. For PVDGs, “at scale” embraces issues of geographical distribution, ownership distribution, size of user population, performance, partitioning and fragmentation of requests, processing and storage capacity, and heterogeneity of demands.

The iVDGL is designed to provide these facilities. This proposed international research laboratory, to be developed jointly with the UK\textsuperscript{13} and EU\textsuperscript{14} and potentially others, will couple resources and researchers worldwide, providing a unique experimental environment that will allow the at-scale experiments required for progress in both application and computer science. Application scientists will use the iVDGL to perform Tera- and Peta-scale data analyses, both as part of carefully designed technical experiments (“data challenges”) and as part of their regular operations (“productions”); computer scientists will provide tools to application scientists for use during these experiments, and will also have access to the iVDGL for their own purposes. To maintain fidelity to the multi-level Grid hierarchy described in Section B, we will implement several URCs (Tier 2) and several smaller URGs (Tier 3), all at US universities, with some sites funded by this proposal and some by other means. We will fund through this proposal several URG sites at institutions historically underrepresented in large research projects, exploiting the Grid’s potential to utilize intellectual capital in diverse locations and extending research benefits to a much wider pool of students.
The iVDGL will ramp up over time to approximately 20 “core” sites with 20-30 “part-time” centers, each having computing, storage, and network resources consistent with its role in the hierarchy, plus a distributed Grid Operations Center. The core sites, all at URC universities in the US and Europe, and funded by their respective national agencies, will comprise the iVDGL foundation, organizing activities for the different experiments and occasionally putting together large-scale exercises utilizing a large fraction of the total sites, from national laboratories to small clusters. The iVDGL sites would be constructed as follows: URC sites (Tier 2) funded by this proposal and other sources (~12, at US universities); URG sites (Tier 3) funded by this proposal and other sources (~10, at US universities); laboratory sites operated by US agencies (Fermilab, Brookhaven, Argonne, Caltech); and European Data Grid tested sites (~15, at CERN and in UK, France, Germany, Italy, perhaps elsewhere). In addition, we have already had encouraging discussions with other international participants in Russia, Japan (KEK), Canada, and South America (AMPATH), and Pakistan, who we anticipate contributing resources as well, hence allowing the iVDGL to grow eventually to some 60 or more sites worldwide. These sites will be connected via a variety of national networks and international links.

The creation of a coherent and flexible experimental facility of 20-60 sites will require careful, staged deployment, configuration and management. It must be possible to manage the entire distributed facility as a single unit, loading new software and detecting and diagnosing faults remotely—and, whenever possible, centrally. It must be possible to reconfigure resources dynamically to meet the needs of different disciplines and experiments. At the same time, individual sites and experiments will require some autonomy, particularly when providing cost sharing on equipment. We propose to address these issues in two ways. First, working in partnership with the NCSA Alliance “Gbox” project (charged with producing standard software for Grid-enabled clusters), we will develop specifications and implementations for standard local and remote management and monitoring software, with the goal of making it straightforward for a site cluster to participate in iVDGL. Second, we will take the first steps towards the creation of a “GGOC,” a Global Grid Operations Center charged with operating the iVDGL as a NOC manages a network. The outcome of both activities should be of considerable interest to other scientific disciplines.

In addition to the management software, a major component of the laboratory will be the definition and dissemination of a suite of standard Data Grid middleware services and tools. This middleware layer will draw from software developed within the NSF GriPhyN project and the European Data Grid initiative, as well as other leading commercial and academic packages. The iVDGL will also identify limitations in the state-of-the-art management and middleware software that we will be deploying and using in evaluation studies. For many of these components, the iVDGL will represent the largest operational configuration ever tried, so we expect to learn many useful lessons from these experiments. We expect the system to evolve substantially over the five-year period, as limitations are observed and corrected. We also observe that deployment across the iVDGL should prove attractive to other developers of advanced software packages, for this reason.

D. Integration of Grid Tools with Scientific Applications

Application scientists will use the iVDGL to perform large-scale data analyses, both as part of carefully designed experiments (“data challenges”) and in some cases as part of their regular operations; computer scientists will provide tools to application scientists during these experiments, and will have access to iVDGL for their own purposes. The major activities of the experiments will be to integrate virtual data concepts into their core software base and then to perform large-scale experiments on the iVDGL. These experiments will involve extremely demanding computations on terascale data (for SDSS and LIGO) and petascale data (for ATLAS and CMS) that will be quite unprecedented in their scope.

LHC: To search for the Higgs particles thought to be responsible for the origin of mass, Supersymmetry and other new interactions and symmetries of nature, physicists need to seamlessly and efficiently access data and computing resources across international scales. Thus ATLAS and CMS are both developing experiment-specific, or “core” application software, as well as common middleware tools, that must work in this global environment. For both experiments, the iVDGL provides a realistic, wide-area distributed environment in which their Grid-integrated software environments can be tested at scales needed for major production exercises from 2002-2005. ATLAS is developing the ATHENA analysis framework and an object database management system (ODBMS). CMS is developing several major subsystems, notably the Object Reconstruction for CMS Analysis (ORCA), the Interactive Graphical User Analysis (IGUANA) environment, and the CMS Analysis and Reconstruction Framework (CARF). Persistent objects in CMS software are handled by CARF using an ODBMS which supports the concepts of virtual (location-independent and medium-independent) data access. For both experiments, the ODBMS and software
framework needs to be supported by the GriPhyN Virtual Data Toolkit to successfully provide transparent data access and delivery, and virtual data products across >150 sites. Monitoring tools providing feedback to users to guide their requests for data and their expectations as a function of data volume and data storage location are needed. Efforts are underway to integrate the Grid tools into core software, to help produce the Grid tools, and to monitor, measure and simulate the Grid systems and derive strategies for efficient data handling and workflow management.

LIGO/LSC: The software analysis environment being developed by the LIGO Scientific Collaboration follows a design that enables the software to run on a large number of platforms operating under several unix operating systems. The core software being developed by LIGO Laboratory is based upon an object oriented (C++) layered design with APIs that communicate over a dedicated distributed system of processors to perform massive and parallel computations on clustered computer networks. There will be at least four and possibly more instantiations of this analysis environment operating autonomously at sites across the LIGO WAN. The baseline design was not scoped to enable interprocess communications between geographically isolated resources within the LIGO Laboratory and the collaboration. The goal to the proposed LIGO/LSC effort will be to Grid-enable this software environment, using iVDGL to extend the currently restricted functionality in four major ways. (1) LIGO core software API components need to be integrated with the GriPhyN Virtual Data Toolkit to enable Grid-based access to the LIGO databases; (2) LIGO search algorithm software must be ported to the iVDGL in order to use its distributed computing capabilities; (3) The GriPhyN Virtual Data Toolkit will be used to replicate LIGO data across the iVDGL; (4) LIGO will work with its international partners in Europe (GEO600 in UK/Germany, Virgo in Italy/France) to establish a network-based analysis capability based on the iVDGL which includes sharing of data across the iVDGL.

SDSS will have two initial sites for the iVDGL, with different functionalities. The Fermilab node will create large amounts of Virtual Data through reprocessing the 2.5 Terapixels of SDSS imaging data. We will quantify the shearing of galaxy images by gravitational lensing due to the effects of the ubiquitous dark matter. Data will be reprocessed on demand from regions with multiple exposures, exploring temporal variations, discovering transient objects like distant supernovae. The JHU node will consist of the parallel catalog database of the project (SX), and perform Virtual Data computations consisting of advanced statistical tools, measuring spatial clustering and their dependence on galaxy parameters. These analyses will lead to new insights into the galaxy formation process: are galaxy types determined by “nature or nurture”, and measure the fundamental parameters of the Universe. The algorithms scale typically as N^2 to N^3 with the number of objects. With 10^8 objects in the SDSS catalog, they represent a substantial computational challenge. In order to accomplish this goal, we need to (a) convert the database to operate within the grid environment, (b) create a grid-enabled version of the advanced statistical tools (c) integrate the two in the iVDGL environment. A third site would be added later, to implement a virtual data service based on the SDSS data for the whole astronomical community, and provide educational content for the wide public, accessible through the National Virtual Observatory. The NVO, and its global counterpart is seen as a major customer of our VDG technology. Early access for the astronomy community to the iVDGL resources will accelerate the development of Virtual Observatories over the whole world.

E. Resources

E.1. Leverage from Existing Projects, Activities, Resources and Expertise

The iVDGL is possible and desirable because a rich set of projects is far advanced in the process of developing next-generation “Grid” technologies. It is these technologies that will make it possible for the discipline science and IT research communities to profit quickly and decisively from the proposed international research facility. iVDGL principals are deeply involved in these various projects and so are well placed to coordinate closely with them. Relevant IT activities such as the Condor, Globus, and Storage Resource Broker projects are pioneering the technologies required for security, resource management, data access, and so forth in Grid environments. These projects are already partnering with the discipline science projects to prototype various Data Grid components and applications, although a lack of suitable large-scale experimental facilities has made serious evaluation difficult.

US physicists in this project have also done much collaborative work with computer scientists on distributed databases, starting with the GIOD (Globally Interconnected Object Databases) project in 1997, followed by the PPDG (Particle Physics Data Grid, DoE) and ALDAP (Accessing Large Data archives in Astronomy and Particle physics, NSF) projects, and most recently the GriPhyN project. Major network resources that will be used for advanced testing of Grid concepts include the instrumented NTON Caltech-SLAC research link at OC-48 going to OC-192, and the US-CERN transatlantic link at OC-3 (2000) going to OC-12 by 2002. Ongoing semi-annual large-
scale data production and distribution exercises by CMS (20 TB and 1000 CPUs in 2001) will be used to test and drive further developments, based on early work with such tools as the Grid Data Management Prototype (GDMP) developed by Caltech and the European DataGrid project. These efforts are collectively developing key enabling Data Grid technologies and seeking to apply these technologies to their disciplines.

Finally, the physicists and astronomers in this project are the software and computing leaders of their respective experiments, and have access to considerable computing infrastructures at Fermilab, Brookhaven and Argonne laboratories. SDSC’s presence here and in GriPhyN will allow larger facilities to be exploited for suitable tests.

E.2. The Need for a Large, Integrated Project

We believe that the large scale of the activities undertaken here, the scientific importance and broad relevance of the scientific experiments, and the strong synergy among the four physics experiments and between physics and computer science goals, together justify the submission of this proposal to the ITR program as a Large Project. The international scale of the iVDGL alone requires a US component that is commensurate with the multi-million dollar funded projects in Europe and the UK and that has the leverage to expand the collaboration to partners in other world regions. Only a collaborative effort such as that proposed here can provide the opportunities for integration, experimentation and evaluation at scale that will ensure long-term relevance. A large integrated project also provides tremendous education and outreach possibilities and can reach a broader audience than multiple small projects.

F. Merits of Proposed Work

Leveraging Grid Research, Tools and Technology: We will be applying the fruits of Computational Grid research to enable scientific discovery for several leading experiments and by extension to other scientific domains. We expect major new discoveries both in particle- and astrophysics from the use of the iVDGL environment.

University Partnerships: This proposal will impact university research in several areas. First, universities make up two of the four levels in the worldwide IT infrastructure of our experiments, making them a critical component in the computational hierarchy and giving universities a powerful voice in large-scale scientific experiments. Second, the universities in this program will be leading the effort to integrate grid tools with their scientific applications, greatly increasing the scientific potential of the experiments. Finally, the ability to provide seamless access to globally distributed data collections removes the enormous barrier faced by researchers requiring access to raw data from distant sites and empowers smaller institutions with limited budgets to contribute intellectually to the analysis.

Global Partnerships: Global partnerships are present at two different levels in this proposal. First, all our experiments involve far-flung international collaborations that make their successful operation important to a global community. The integration through this proposal of the international computational infrastructures of these global experiments will have a tremendous impact on the science they produce. Global partnerships are also present in the proposed iVDGL. By incorporating resources from the US, UK and European Union, the iVDGL will not only provide a large-scale laboratory for grid research and tests by our experiments and other scientific communities, but will promote the development of a common Grid infrastructure that would later be extended to other world regions.

Promoting Teaching and Learning: We are committing significant budgetary resources to a broad and novel program that will integrate new communities into our scientific research while providing education and outreach consistent with the large size of our project. The main ideas are outlined here and will be discussed in more detail in the full proposal. Our outreach program has several thrusts: (1) provide hardware to universities historically underrepresented in scientific research projects to allow them to participate in our multi-level iVDGL; (2) utilization of full-time E/O faculty to be hired at Florida and University of Texas, Brownsville and the resources of NCSA/NPACI EOT program; (3) course development at many of the participating institutions; (4) creation of professionally designed web projects to promote teaching and learning by “direct discovery” methods that allow students and the general public to interact with “safe” datasets and simulation environments that we will provide. We have begun talks with ThinkQuest on obtaining initial grants to develop student-oriented web projects based on the application sciences and grid technology. SDSS also has an ongoing collaboration with Microsoft to provide a website, modeled after the Terraserver to provide color images of the sky, tied to an object catalog and utilizing a special interface to support educational activities. The website will also provide content for additional educational layers.
G. Management

G.1. Integration and Coordination Strategy; Management Plan

Although the iVDGL will use and integrate software and middleware from a variety of sources, it will work particularly closely with the GriPhyN ITR project and the four experiments that participate in GriPhyN. As described above, it will use software from the GriPhyN project and physics applications to develop and demonstrate the operation of the iVDGL. For this reason, management of the iVDGL project will be closely integrated with GriPhyN and with the experimental projects and with the funding agencies for the experiments, NSF, DOE, and NASA. We plan the following organization:

1. Leverage of GriPhyN Management: Resource allocations and policy questions will be managed by the existing GriPhyN organization. As described in the GriPhyN Management plan, the organization consists of two directors (the GriPhyN PIs), a full-time Project Coordinator, and three Boards for Project Coordination, Collaboration matters, and External Advice. The Project Coordination Board, on which the management of the four experiments is represented, advises the Directors and Project Coordinator closely on goals, budgets, and schedules, ensures good integration with the development and, eventually, the production needs of the experiments.

2. Deputy Project Coordinator for the iVDGL: In order to keep the Information Technology research of GriPhyN and its deployment and integration in the iVDGL closely coordinated, the Project Coordinator will have overall responsibility for day-to-day work in both areas. It is expected, however, that the Project Coordinator will take primary responsibility for the GriPhyN IT tasks, for which the position was originally planned, while a new full-time position of Deputy Project Coordinator for the iVDGL will be created to run the challenging new task of creating and operating the Laboratory described in this proposal.

3. Facilities organizations of the four experiments: Each of the developmental regional centers to be created by this project will have a dual role, first as part of the national and international iVDGL described above, and second as a developmental part of the facilities of one of the four participating experiments. Each center will be associated with one of the experiments. The local operators of each center will be part of the facilities organization of the associated experiment, who will manage it for the iVDGL project. These operators will have the responsibility of operating the center to meet the needs of both GriPhyN and of the particular experiment. In order to permit the necessary integration of data grid capabilities with each experiment’s application software, a substantial part of the operation of each center will be focused on integration and interoperation with other facilities of the particular experiment, with operations assigned by the experiment. The remainder of the center operation will be devoted to the extended, unified iVDGL in which all centers work as one Laboratory, with operations assigned by the iVDGL project. Any conflict between these modes of operation will be resolved by the GriPhyN/iVDGL Directors and Project Coordinator.

G.2. Coordination with International Efforts

The four experimental projects are international in scope and will need the production Virtual Data Grids that are based on the research proposed here to function seamlessly across international and interregional boundaries. Thus it is important to seek early opportunities to extend our Laboratory to include the developmental facilities that are emerging in other regions. We are in close touch with the DataGrid project of the European Union and a number of related national projects. As we write this preproposal, we expect to be able to carry out joint experiments of interconnecting laboratories with the DataGrid project itself, the UK DataGrid project and probably other national projects. These projects are presently discussing among themselves the appropriate coordination structure to ensure compatibility and interoperability, and to plan the operational aspects of joint operational tests and experiments.

G.3. Resource Allocations

A substantial fraction of the resources requested in this proposal will be devoted to the developmental regional centers that together will comprise the iVDGL. As noted above, these new centers will need to be integrated with the associated experiment organizations as well as with the iVDGL project. Since each of the experiments is itself composed of many US institutions (in addition to international partners), it is expected that each experiment will have a transparent process for selecting the institutions to operate its regional centers. Each of the experiments is currently developing the system for this selection, taking into account which institutions are interested in the assignment, opportunities for leveraging local infrastructure, other potential cost sharing, and the capabilities of the personnel at each site. All selections will be made in consultation with the funding agencies to ensure coordination.
REFERENCES:


3 Using tools such as the MONARC Simulation System (I. Legrand et al.). See http://www.cern.ch/MONARC/.


8 The LIGO Experiment, http://www.ligo.caltech.edu/.


12 In the case of the LHC, where the datasets start in the multi-petabyte range, we have extended the concept to a five level system. The site of the particle accelerator (CERN), where the data is first acquired and subjected to first-round processing, is designated as a Tier0 center.

13 Statement of support from UK PPARC is quoted in the supporting document on International Collaboration.

14 Statement from Fabrizio Gagliardi, Project Director of the European Union DataGrid Project is quoted in the supporting document on International Collaboration.

15 The GGOC also will cover monitoring and performance tracking, and problem resolution both for the Grid system software and for the experiments' major Grid-enabled applications, analogous to a large-scale application service provider. Grid operation will involve coordinated operations of the GGOC with Regional Grid Operations Centers (RGOC) in a manner compatible with the tiered Grid structure used by the experiments.


19 These subsystems are now entering the third major cycle of development, the “fully functional” stage.

20 This refers to initial interactive implementations. Once the information from these interactive tools is extracted and understood, automated agent-based tools may take over some of these manual functions, especially in large-scale production environments.
21 GEO600 gravitational wave experiment home page, http://www.geo600.uni-hannover.de/.


31 SDSS Skyserver website, for a demonstration see http://dart.pha.jhu.edu/sdss/