April 25, 2001

An International Virtual-Data Grid Laboratory for Data Intensive Science

Submitted to the 2001 NSF Information and Technology Research Program
Proposal #0122557

Bruce Allen	University of Wisconsin, Milwaukee
Paul Avery[*]	University of Florida
Keith Baker	Hampton University
Rich Baker	Brookhaven National Lab
Lothar Bauerdick	Fermi National Accelerator Laboratory
James Branson	University of California, San Diego
Julian Bunn	California Institute of Technology
Manuela Campanelli	University of Texas, Brownsville
L. Samuel Finn	The Pennsylvania State University
Ian Fisk	University of California, San Diego
Ian Foster[†]	University of Chicago
Rob Gardner[†]	Indiana University
Alan George	University of Florida
John Huth	Harvard University
Stephen Kent	Fermi National Accelerator Laboratory
Carl Kesselman[‡]	University of Southern California
Scott Koranda	University of Wisconsin, Milwaukee and NCSA
Albert Lazzarini‡	California Institute of Technology
Miron Livny‡	University of Wisconsin
Reagan Moore	San Diego Supercomputing Center
Richard Mount	Stanford Linear Accelerator Center
Harvey Newman[†]	California Institute of Technology
Vivian O'Dell	Fermi National Accelerator Laboratory
Tim Olson	Salish Kootenai University
Ruth Pordes	Fermi National Accelerator Laboratory
Lawrence Price‡	Argonne National Lab
Jenny Schopf	Northwestern University
Jim Shank	Boston University
Alexander Szalay[†]	The Johns Hopkins University
Valerie Taylor	Northwestern University
Torre Wenaus	Brookhaven National Lab
Roy Williams	California Institute of Technology

[*] Principal Investigator
[†] co-Principal Investigator
[‡] Major author

Table of Contents

B. Project Summary

We propose to establish and utilize an international Virtual-Data Grid Laboratory (iVDGL) of unprecedented scale and scope, comprising heterogeneous computing and storage resources in the U.S., Europe—and ultimately other regions—linked by high-speed networks, and operated as a single system for the purposes of interdisciplinary experimentation in Grid-enabled data-intensive scientific computing.

Our goal in establishing this laboratory is to drive the development, and transition to every day 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, information technology, Data Grid software tools, and application sciences, as well as substantial investments in data-intensive science now underway in the U.S., Europe, and Asia. We expect experiments conducted in this unique international laboratory to influence the future of scientific investigation by bringing into practice new modes of transparent access to information in a wide range of disciplines, including high-energy and nuclear physics, gravitational wave research, astronomy, astrophysics, earth observations, and bioinformatics. iVDGL experiments will also provide computer scientists developing data grid technology with invaluable experience and insight, therefore influencing the future of data grids themselves. A significant additional benefit of this facility is that it will empower a set of universities who normally have little access to top tier facilities and state of the art software systems, 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 interdisciplinary efforts such as the NSF-funded GriPhyN and European Union (EU) DataGrid projects are engaged in the research and development of 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 the development of large-scale data-grid applications application design methodologies, thereby slowing the transition of data grid technology from proof of concept to full adoption by the scientific community. In this project we aim to establish a laboratory that will enable us to overcome these obstacles to progress.

Laboratory users will include international scientific collaborations such as the Laser Interferometer Gravitational-wave Observatory (LIGO), the ATLAS and CMS detectors at the Large Hadron Collider (LHC) at CERN, the Sloan Digital Sky Survey (SDSS), and the proposed National Virtual Observatory (NVO); application groups affiliated with the NSF PACIs and EU projects; outreach activities; and Grid technology research efforts. The laboratory itself will be created by deploying a carefully crafted data grid technology base across an international set of sites, each of which provides substantial computing and storage capability accessible via iVDGL software. The 20+ sites, of varying sizes, will include U.S. sites put in place specifically for the laboratory; sites contributed by EU, Japanese, Australian, and potentially other international collaborators; existing facilities that are owned and managed by the scientific collaborations; and facilities placed at outreach institutions. These sites will be connected by national and transoceanic networks ranging in speed from hundreds of Megabits/s to tens of Gigabit/s. An international Grid Operations Center (iGOC) will provide the essential management and coordination elements required to ensure overall functionality and to reduce operational overhead on resource centers.

Specific tasks to be undertaken in this project include the following. (1) Construct the international laboratory, including development of new techniques for low-overhead operation of a large, internationally distributed facility; (2) adapt current data grid applications and other large-scale production data analysis applications that can benefit from Data Grid technology to exploit iVDGL features; (3) conduct ongoing and comprehensive evaluations of both data grid technologies and the Data Grid applications in the iVDGL, using various (including agent-based) software information gathering and dissemination systems to study 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, and at a later stage to guide the operation of the iVDGL itself. The experience gained with information systems of this size and complexity, providing transparent managed access to massive distributed data collections, will be applicable to large-scale data-intensive problems in a wide spectrum of scientific and engineering disciplines, and eventually in industry and commerce. Such systems will be needed in the coming decades as a central element of our information-based society.

C. Project Description

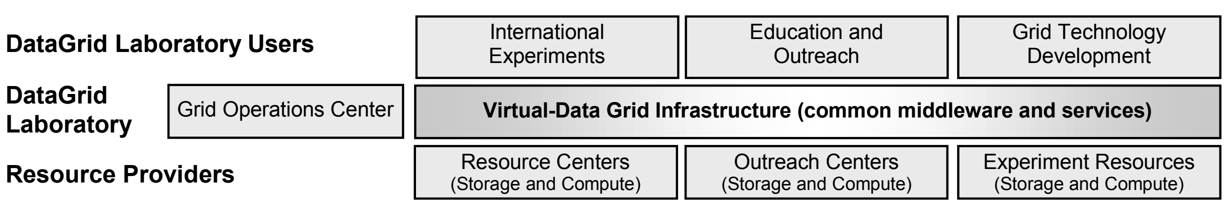
C.1. Introduction: The International Virtual-Data Grid Laboratory

We propose to establish and utilize an international Virtual-Data Grid Laboratory (iVDGL) of unprecedented scale and scope, comprising heterogeneous computing and storage resources in the U.S., Europe—and ultimately other regions—linked by high-speed networks, and operated as a single system for the purposes of interdisciplinary experimentation in Grid-enabled[1,2] data-intensive scientific computing[3,4].

Our goal in establishing this laboratory is to drive the development, and transition to every day 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, information technology, Data Grid software tools, and application sciences, as well as substantial investments in data-intensive science now underway in the U.S., Europe, and Asia. We expect experiments conducted in this unique international laboratory to influence the future of scientific investigation by bringing into practice new modes of transparent access to information in a wide range of disciplines, including high-energy and nuclear physics, gravitational wave research, astronomy, astrophysics, earth observations, and bioinformatics. iVDGL experiments will also provide computer scientists developing data grid technology with invaluable experience and insight, therefore influencing the future of data grids themselves. A significant additional benefit of this facility is that it will empower a set of universities who normally have little access to top tier facilities and state of the art software systems, 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 interdisciplinary efforts such as the NSF-funded GriPhyN[5] and European Union (EU) DataGrid projects[6] are engaged in the R&D of the basic technologies required to create working data grids. Missing are (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 the development of large-scale Data Grid applications and application design methodologies, thereby slowing the transition of data grid technology from proof of concept to full adoption by the scientific community. *Our proposed laboratory will enable us to overcome these obstacles to progress*.

The following figure illustrates the structure and scope of the proposed virtual laboratory. Laboratory users will include international scientific collaborations such as the Laser Interferometer Gravitational-wave Observatory (LIGO) [7,8,9], the ATLAS[10] and CMS[11] detectors at the Large Hadron Collider (LHC) at CERN, the Sloan Digital Sky Survey (SDSS)[12,13], and the proposed National Virtual Observatory (NVO)[14]; application groups affiliated with the NSF PACIs and EU projects; outreach activities; and Grid technology research efforts. The laboratory itself will be created by deploying a carefully crafted data grid technology base across an international set of sites, each of which provides substantial computing and storage capability accessible via iVDGL software. The 20+ sites, of varying sizes, will include U.S. sites put in place specifically for the laboratory; sites contributed by EU, Japanese, Australian, and potentially other international collaborators; existing facilities that are owned and managed by the scientific collaborations; and facilities placed at outreach institutions. These sites will be connected by national and transoceanic networks ranging in speed from hundreds of Megabits/s to tens of Gigabit/s. An international Grid Operations Center (iGOC) will provide the essential management and coordination elements required to ensure overall functionality and to reduce operational overhead on resource centers. The system represents an order-of-magnitude increase in size and sophistication relative to previous infrastructures of this kind[15,16].



Specific tasks to be undertaken in this project include the following. (1) Construct the international laboratory, including development of new techniques for low-overhead operation of a large, internationally distributed facility; (2) adapt current data grid applications and other large-scale production data analysis applications that can benefit from

Data Grid technology to exploit iVDGL features; (3) conduct ongoing and comprehensive evaluations of both data grid technologies and the Data Grid applications on iVDGL, using various (including agent-based[17,18,19]) software information gathering and dissemination systems to study 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[20], and at a later stage to guide the operation of iVDGL itself. The experience gained with information systems of this size and complexity, providing transparent managed access to massive distributed processing resources and data collections, will be applicable to large-scale data- and compute-intensive problems in a wide spectrum of scientific and engineering disciplines, and eventually in industry and commerce. Such systems will be needed in the coming decades as a central element of our information-based society.

We believe that the successful completion of this proposed R&D agenda will result in significant contributions to our *partner science applications* and to *information technologists*, via provision of, and sustained experimentation on, a laboratory facility of unprecedented scope and scale; to the *nation's scientific "cyberinfrastructure*," via the development and rigorous evaluation of new methods for supporting large-scale community-based, cyber-intensive scientific research; and to *learning and inclusion* via the integration of minority institutions into the IVDGL fabric, in particular by placing resource centers at those institutions to facilitate project participation. These significant contributions are possible because of the combined talents, experience, and leveraged resources of an exceptional team of leading application scientists and computer scientists. The strong interrelationships among these different topics demand an integrated project of this scale; the need to establish, scale, and evaluate the laboratory facility over an extended period demands a five-year duration.

C.2. IVDGL Motivation, Requirements, and Approach

Petabyte Virtual Data Grid (PVDG) concepts have been recognized as central to scientific progress in a wide range of disciplines. Simulation studies[21] have demonstrated the feasibility of the basic concept and projects such as GriPhyN are developing essential technologies and toolkits. However, the history of large networked systems such as the Internet makes it clear that experimentation at scale is required if we are to obtain the insights into the key factors controlling system behavior that will enable development of effective strategies for system operation that combine high resource utilization levels with acceptable response times. Thus, for PVDGs, the next critical step is to create facilities and deploy software systems to enable "at scale" experimentation, which means embracing issues of geographical distribution, ownership distribution, security across multiple administrative domains, size of user population, performance, partitioning and fragmentation of requests, processing and storage capacity, duration and heterogeneity of demands. Hence the need for the international, multi-institutional, multi-application laboratory being proposed.

For many middleware[22] and application components, iVDGL will represent the largest and most demanding operational configuration of facilities and networks ever tried, so we expect to learn many useful lessons from both iVDGL construction and experiments. Deployment across iVDGL should hence prove attractive to developers of other advanced software. This feedback will motivate substantial system evolution over the five-year period as limitations are corrected.

C.2.a. Requirements

These considerations lead us to propose a substantial, focused, and sustained investment of R&D effort to establish an international laboratory for data-intensive science. Our planning addresses the following requirements:

- *Realistic scale*: The laboratory needs to be realistic in terms of the number, diversity, and distribution of sites, so that we can perform experiments today in environments that are typical of the Data Grids of tomorrow. We believe that this demands 10s (initially) and 100s (ultimately) of sites, with considerable diversity in size, location, and network connectivity.
- *Delegated management and local autonomy*. The creation of a coherent and flexible experimental facility of this size will require careful, staged deployment, configuration, and management. Sites must be able to delegate management functions to central services, to permit coordinated and dynamic reconfiguration of resources to meet the needs of different disciplines and experiments—and to detect and diagnose faults. Individual sites and experiments will also require some autonomy, particularly when providing cost sharing on equipment.
- *Support large-scale experimentation*: Our goal is, above all, to enable experimentation. In order to gain useful

results, we must ensure that iVDGL is used for real "production" computing over an extended time period so that we can observe the behavior of these applications, our tools and middleware, and the physical infrastructure itself, in realistic settings. Hence, we must engage working scientists in the use of the infrastructure, which implies in turn that the infrastructure must be constructed so as to be highly useful to those scientists.

- *Robust operation*: In order to support production computation, our iVDGL design must operate robustly and support long running applications in the face of large scale, geographic diversity, institutional diversity, and high degree of complexity arising from the diverse range of tasks required for data analysis by worldwide scientific user-communities.
- *Instrumentation and Monitoring.* To be useful as an experimental facility, iVDGL must be capable of not only running applications but also instrumenting, monitoring, and recording their behavior—and the behavior of the infrastructure— at different granularity levels over long periods of time[23].
- *Integration with an (inter)national cyberinfrastructure*. iVDGL will be most useful if it is integrated with other substantial elements of what seems to be an emerging national (and international) cyberinfrastructure. In fact, iVDGL, if operated appropriately, can make a major contribution to the establishment of this new infrastructure both as a resource and as a source of insights into how to operate such facilities.
- *Extensibility*. iVDGL must be designed to support continual and substantial evolution over its lifetime, in terms of scale, services provided, applications supported, and experiments performed.

C.2.b. Approach

We propose to address the requirements listed above by creating, operating, and evaluating, over a sustained period of experimentation, an international research laboratory for data-intensive science. This unique experimental facility will be created by coupling a heterogeneous, geographically distributed, and (in the aggregate) extremely powerful set of iVDGL Sites (iSites). A core set of iSites controlled by iVDGL participants, and in many cases funded by this proposal, will be dedicated to iVDGL operations; others will participate on a part-time basis on terms defined by MOUs. In all cases, standard interfaces, services, and operational procedures, plus an iVDGL operations center, will ensure that users can treat iVDGL as a single, coherent laboratory facility.

The set of participating sites will be expanded in a phased fashion, expanding over the five years of this proposal from 6 to 15 core sites and from 0 to 45 or more partner sites. Details of the hardware purchases, local site commitments, and partnership agreements that we will use to achieve these goals are provided in Section F (Facilities). In brief, we expect that the laboratory will, by year 3, comprise 30 sites in four continents, and many more than that by year 5. These sites will all support a common VDG infrastructure, facilitating application experiments that run across significant fractions of these resources.

We approach the construction of iVDGL via focused and coordinated activities in four distinct areas, which we describe briefly here and expand upon in subsequent sections.

<u>Define a Scalable, Extensible, and Easily Reproducible Laboratory Architecture</u>: We will define the expected functionality of iVDGL resource sites along with an architecture for monitoring, instrumentation and support. We will establish computer science support teams charged with "productizing" and packaging the essential Data Grid technologies required for application use of iVDGL, developing additional tools and services required for iVDGL operation, and providing high-level support for iVDGL users.

<u>Create and Operate a Global-Scale Laboratory:</u> We will deploy hardware, software, and personnel to create, couple, and operate a diverse, geographically distributed collection of locally managed iSites. We will establish an international Grid Operations Center (iGOC) to provide a single point of contact for monitoring, support, and fault tracking. We will exploit international collaboration and coordination to extend iVDGL to sites in Europe and elsewhere, and establish formal coordination mechanisms to ensure effective global functioning of iVDGL.

<u>Evaluate and Improve the Laboratory via Sustained, Large-Scale Experimentation</u>: We will establish application teams that will work with major physics experiments to develop, apply, and evaluate substantial applications on iVDGL resources. We will work in partnership with other groups, notably the NSF PACIs[24], DOE PPDG[25] and ESG, and EU DataGrid project, to open up iVDGL resources to other applications. These studies will be performed in tandem with instrumentation and monitoring of middleware, tools, and infrastructure with the goal of guiding development and optimization of iVDGL operational software and strategies.

<u>Engage Underrepresented Groups in the Creation and Operation of the Laboratory</u>: We will fund iSites at institu-

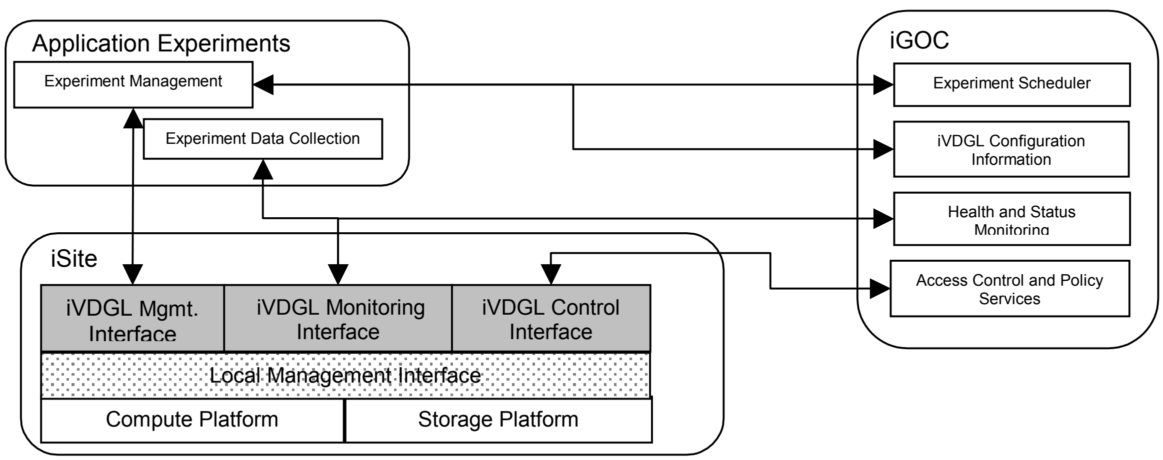
tions historically underrepresented in large research projects, exploiting the Grid's potential to utilize intellectual capital in diverse locations and extending research benefits to a much wider pool of researchers and students.

C.3. Define a Scalable, Extensible, and Easily Reproducible Laboratory Architecture

We have developed a detailed design for most iVDGL architecture elements, including site hardware; site software; global services and management software; and the grid operations center. This design builds on our extensive experience working with large-scale Grids, such as I-WAY[26], GUSTO[27], NASA Information Power Grid[28], NSF PACI's National Technology Grid[29], and DOE ASCI DISCOM Grid[30], to develop an iVDGL architecture that addresses the requirements above. We do not assert that this architecture addresses the requirements perfectly: it is only through concerted experimentation with systems such as iVDGL that we will learn how to build robust infrastructures of this sort. However, we do assert that we have a robust and extensible base on which to build.

As illustrated in Figure 1, our architecture distinguishes between the various locally managed iVDGL Sites (iSites), which provide computational and storage resources to iVDGL via a common set of protocols and services; an International Grid Operations Center (iGOC), which monitors iVDGL status, provides a single point of contact for support, and coordinates experiments; a set of global iVDGL services, operated by the iGOC and concerned with resource discovery, etc.; and the application experiments (discussed in Section C.5).

Figure 1. iVDGL architecture. Local sites support standard iVDGL services that enable remote resource manage-



ment, monitoring, and control. The iGOC monitors the state of the entire iVDGL, providing a single point of access for information about the testbed, as well as allowing the global apparatus to be configured for particular experimental activities. The application experiments interact with iVDGL first by arranging to conduct an experiment with the iGOC, and then by managing the needed resources directly via iVDGL services.

C.3.a. iSite Architecture: Clusters, Standard Protocols and Behaviors, Standard Software Loads

Each iSite is a locally managed entity that comprises a set of interconnected processor and storage elements. Our architecture aims to facilitate global experiments in the face of the inevitable heterogeneity that arises due to local control, system evolution, and funding profiles. However, current hardware cost-performance trends and the widespread acceptance of Linux suggest that most sites will be Linux clusters interconnected by high-performance switches. In some cases, these clusters may be divided into processing and storage units. The processing capacity of these clusters may be complemented by smaller workgroup clusters, using technologies such as Condor.

From our perspective, standardization of hardware is less important than standardization of software and services. Nevertheless, we will define a general space of recommended hardware configurations for iVDGL partners who request this information—and will deploy this platform at iVDGL sites directly supported by this proposal. This hardware configuration will consist of a Linux based cluster, with a high-speed network switch, running a standard set of cluster management environment, such as that defined by NCSA's Cluster in a Box project.

We transform the heterogeneous collection of iVDGL resources into an integrated laboratory by defining a common set of protocols and behaviors, supported by all iSites. These protocols and behaviors make it possible for iVDGL applications to discover, negotiate access to, access, manage computation on, and monitor arbitrary collections of iVDGL resources, subject of course to appropriate authorization. While the services provided will evolve over the course of this proposal, we start initially with the following set, building heavily on the proven and widely used protocol suite and code base offered by the Globus Toolkit[28]. All protocols use Grid Security Infrastructure mechanisms for authentication[31,32] and support local authorization.

Management services enable applications to allocate and manage computational and storage resources at the sites. We adopt the GRAM protocol for resource management (e.g., start and stop computations, transfer data) and computation control; and the GridFTP protocol for data movement[4]

Monitoring services support discovery of the existence, configuration, and state of iVDGL resources. We adopt MDS-2 registration and access protocols for discovery[33,34] and access to configuration and performance data.

Control services support global experiment management and testbed configuration. These services will include access control and policy enforcement via the Community Authorization Service (CAS) protocols currently under development at ANL, U.Chicago, and USC/ISI, along with remote configuration capabilities.

To facilitate the deployment of operational iSites, we will develop standard software loads implementing the required protocols and behaviors, leveraging "Grid in a Box" software produced by NCSA and its partners.

C.3.b. iVDGL Operations Center: Global Services and Centralized Operations

The effective operation of a distributed system such as iVDGL also requires certain global services and centralized monitoring, management, and support functions. These functions will be coordinated by the iVDGL Grid Operations Center (iGOC), with technical effort provided by iGOC staff, iSite staff, and the CS support teams. The iGOC will operate iVDGL as a NOC manages a network, providing a single, dedicated point of contact for iVDGL status, configuration, and management, and addressing overall robustness issues. Building on the experience and structure of the Global Research Network Operations Center (GNOC) at Indiana University, as well as experience gained with research Grids such as GUSTO, we will investigate, design, develop, and evaluate the techniques required to create an operational iVDGL. The following will be priority areas for early investigation.

Health and status monitoring. The iGOC will actively monitor the health and status of all iVDGL resources and generate alarms to resource owners and iGOC personal when exceptional conditions are discovered. In addition to monitoring the status of iVDGL hardware, this service will actively monitor iSite services to ensure that they are compliant with iVDGL architecture specifications.

Configuration and information services. The status and configuration of iVDGL resources will be published through an iVDGL information service. This service will organize iSites into one or more (usually multiple) "virtual organizations" corresponding to the various confederations of common interest that apply among iVDGL participants. This service will leverage the virtual organization support found in MDS-2[34].

Experiment scheduling. The large-scale application experiments planned for iVDGL will require explicit scheduling of scarce resources. To this end, the iGOC will operate a simple online experiment scheduler, based on the Globus slot manager library.

Access control and policy. The iGOC will operate an iVDGL-wide access control service. Based on the Globus CAS, this service will define top-level policy for laboratory usage, including the application experiments that are allowed to use the laboratory.

Trouble ticket system. The iGOC will operate a centralized trouble ticket system to provide a single point of contact for all technical difficulties associated with iVDGL operation. Tickets that cannot be resolved by iGOC staff will be forwarded to the support teams of the specific software tool(s).

Strong cost sharing from Indiana allows us to support iGOC development at a level of two full-time staff by FY2003. Nevertheless, sustained operation of iVDGL will require a substantially larger operation. To this end, we will establish partnerships with other groups operating Grid infrastructures, in particular the DTF, European Data Grid, and Japanese groups. We will also seek additional support. In addition, we are hopeful that some degree of 24x7 support can be provided by the Indiana GNOC, however further discussion is required to determine details.

C.4. Create and Operate a Global-Scale Laboratory

As we indicated above, and describe in more detail in Section F (Facilities), iVDGL will ramp over time to approximately 20 "core" and 20-30 "part-time" sites, each having computing, storage, and network resource. These sites span a range of sizes and serve a range of communities, ranging from national-level centers, typically operated by a national laboratory or research center on behalf of a large national or international project, to smaller systems operated by individual universities or even research groups. Our deployment strategy combines incremental rollout, functional evolution, iGOC support structure with software troubleshooting teams (discussed in Section C.4.e) to provide technical, managerial, and political solutions raised by a rich and diverse range of resources, software infrastructure, management styles, and user requirements.

We address in the following the responsibilities of an individual iVDGL site; our plans for iVDGL networking; and our plans for iVDGL operation. See also Section H (Facilities) for a list of partner sites, Section E (International Collaboration) for a discussion of international partnerships, and Section C.11 for a discussion of management structure.

C.4.a. Defining and Codifying iVDGL Site Responsibilities

Effective iVDGL operation will require substantial cooperation from participating sites. The responsibilities of an iVDGL site will be codified in an MOU that will commit a site to: (1) implement standard iVDGL behaviors, as defined by the iVDGL Coordination Committee, e.g. by running standard iVDGL software loads; (2) make available a specified fraction (100% for core sites, a lesser fraction for other sites) of its resources for iVDGL purposes, in exchange for access to iVDGL resources; and (3) provide personnel for operating the site resources, participating in tests, and for responding to problem reports.

C.4.b. iVDGL Networking

Future scientific infrastructures will benefit from dramatic changes in the cost-performance of wide area network links, as a result of recent revolutionary developments in optical technologies, in particular DWDM. We will exploit these developments, taking advantage of existing and proposed deployments of high-speed networks both within the U.S. and internationally. This focus on high-speed networks allows us to address demanding application requirements and also makes iVDGL extremely interesting from a research perspective.

Within the U.S., we will exploit moderate speed production networks (Abilene, ESnet, CALREN, MREN) and high-speed research and production networks (e.g., IWIRE within Illinois, NTON, and the proposed DTF TeraGrid network; both involve multiple 10 Gb/s connections). The operators of these networks strongly support iVDGL goals.

Internationally, we plan to leverage numerous planned high-speed optical connections, many of which will peer at the STAR TAP and StarLight facilities in Chicago headed by Prof. Tom DeFanti. STAR TAP provides up to 622 Mb/s connectivity to a variety of international sites, including CERN at 155 to 622 Mb/s. StarLight will provide high-speed international optical connections; the first of these, a 2.5 Gb/s research link from SURFnet in the Netherlands, will be in place by July 2001 and others are planned, including a 2.5 Gb/s research link to CERN. Within Europe, the GEANT network will provide a 10 Gb/s backbone; in the U.K., Super JANET 3 provides comparable connectivity. The operators of all these networks have agreed in principle to support iVDGL operations.

C.4.c. Integration with PACI Resources and Distributed Terascale Facility

NCSA and SDSC have submitted a joint proposal to the NSF to establish a Distributed Terascale Facility (DTF) linking a small number of large compute-storage sites with a high-speed network, all configured for Grid computing. We propose to integrate this DTF and other PACI resources with iVDGL via transparent gateways that allow PACI applications to call upon iVDGL resources for large-scale experimentation and data distribution purposes, and iVDGL partners to call upon DTF resources for ultra-large scale computing. The combined DTF resources we will call upon (for limited periods for tests and "peak" production) are expected to be comparable in compute power and complexity to a future production national-scale center for the LHC, and hence to be within one order of magnitude of the initial global capacity of a major LHC experiment when it begins operation in 2006. The enclosed MOU lays out the rationale for, and essential elements of, this proposed collaboration, which we believe will be strongly mutually beneficial both for the participating scientists and for the research community as a whole as we explore what it means to create a national and international cyberinfrastructure.

C.4.d. Creating an Global Laboratory via International Collaboration

International participation is vital to our iVDGL goals, for three reasons: (1) it allows us to increase the scale of iVDGL significantly, with respect to number of sites and geographic extent; (2) the challenging science projects with which we wish to work all have international extent; and (3) international participation enables us to address the challenging operations and management issues that arise in this context.

The strong collaborative links that we have established with international partners, and the compelling nature of iVDGL vision, have enabled us to obtain firm commitments from a number of international partners that translate into (1) committed iVDGL sites in the EU, Japan, and Australia; (2) resource commitments, in the form of six visiting researchers, from the UK; (3) access to international networks; and (4) strong additional interest from other sites. More details are provided in the supplementary information, which also discusses management issues.

C.4.e. Supporting iVDGL Software

iVDGL software is, for the most part, being developed under other funding by iVDGL participants and others. In particular, building on the Globus and Condor toolkits and incorporating pioneering new "virtual data" concepts, they are constructing the Virtual Data Toolkit (VDT), a software system for experimentation with virtual data concepts in large-scale Data Grid applications. However, that work (primarily GriPhyN) is not funded to produce production quality software or to support large-scale use by the substantial scientific communities that will use iVDGL. In addition, the unique nature of iVDGL will require the development of additional software components for the management of large-scale testbeds and for data collection during application experiments. Both software hardening and development will be addressed within iVDGL. Both activities have the potential to be tremendously demanding and the resources available here are unlikely to be sufficient to tackle them in their entirety. However, by providing staff dedicated to these tasks within iVDGL, by leveraging the participation of VDT developers and by integrating the support staff with the developers, we can provide VDT code with sufficient stability, scalability and support for everyday production use by application scientists.

Software development activities will be split across two teams: a VDT Robustification and Troubleshooting team (VRT) and a Large Scale Operations (LSO) team. To leverage GriPhyN funding and to maximize effectiveness and efficiency, the VRT will be led by GriPhyN VDT lead Miron Livny, and will include participation by USC and UWisc. The VRT will not develop new software but instead will work to enhance the releases of the GriPhyN VDT software, as follows: (1) develop and maintain a VDT test suite; (2) define and implement a unified and fully integrated error reporting framework across all VDT components; (3) equip the VDT with dynamically configurable event logging capabilities; (4) extend the VDT with new components required for specific application purposes; (5) help users to maintain and trouble shoot the VDT software, (6) provide documentation, and (7) create procedures and tools for reporting, tracking, and fixing bugs.

The LSO will be lead by Ian Foster and will have participants from U. Chicago and USC. The LSO will focus on the design and development of tools unique to iVDGL, relating particularly to testbed management, scalability, and usability issues that arise in iVDGL operation. These tools will ultimately transition to the iGOC and the LSO team will work in collaboration with the iGOC in software design and deployment. Issues to be addressed will include (1) development of automatic distribution and update services, (2) tools for testing iSite compliance with iVDGL configuration guidelines, (3) agent based monitoring and diagnostic tools for tracking overall iVDGL workflow, and (4) instrumentation and logging services for collecting and archiving long-term iVDGL performance metrics.

C.5. Evaluate and Improve the Laboratory via Sustained, Large-Scale Experimentation

iVDGL will be used to conduct a range of large-scale experiments designed to improve our understanding of three vitally important topics: (1) the design, functioning, and operation of large-scale Data Grid applications; (2) the performance and scalability of Grid middleware infrastructure; and (3) the construction, maintenance, and operation of an international-scale cyberinfrastructure. These three topics will be studied together and separately, via a combination of "production" application use and more targeted experiments. The results of this experimentation will provide feedback and guidance to partner application projects, partner middleware technology projects, and iVDGL itself, as well as contributing to the effective development of national and international Grid infrastructures.

We describe here the goals and approach for large-scale experimentation. We address first the foundation of our experimental strategy, namely the direct involvement of four international physics projects who are already committed to using Data Grid concepts. Next, we indicate how we will expand on this set of applications via other partnerships. Finally, we describe our strategy and goals for using iVDGL for investigations of middleware and operations.

C.5.a. Production Use by Four Frontier Physics Projects

The centerpiece of our experimental strategy is the intimate involvement of four major experimental physics projects whose effective operation depends on the creation and large-scale operation of international Data Grid infrastructures. By involving these experiments, we ensure that: (1) individual iVDGL nodes are maintained as working systems and (2) iVDGL is constantly stressed by realistic application use. In addition, NSF investment in iVDGL serves double duty, as it provides a substantial contribution to the production goals of these four projects.

Table 1 shows the characteristics of these four frontier projects in physics and astronomy: the CMS and ATLAS experiments at the LHC at CERN, Geneva which will probe the TeV frontier of particle energies to search for new states and fundamental interactions of matter, LIGO which will detect and analyze, nature's most energetic events sending gravitational waves across the cosmos and SDSS, the first of several planned surveys that will systematically scan the sky to provide the most comprehensive catalog of astronomical data ever recorded, as well as the planned NVO, which will integrate multiple sky surveys into a single virtual observatory. The NSF has made heavy investments in LHC, LIGO, and SDSS.

Table 1: Characteristics of the four primary physics applications targeted by iVDGL

Application	First Data	Data Rate MB/s	Data Volume (TB/yr)	User Community	Data Access Pattern	Compute Rate (Gflops)	Type of data
SDSS	1999	8	10	100s	Object access and streaming	1 to 50	Catalogs, image files
NVO	2004	100	100-500	1000s	Object access, distributed joins	3,000	Distributed catalogs, images, spectra
LIGO	2002	10	250	100s	Random, 100 MB streaming	50 to 10,000	Multiple channel time series, Fourier transformations
ATLAS /CMS	2006	100-400	5000 each	1000s	Streaming, 1 MB object access	120,000	Events, 100 GB/sec simultaneous access

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 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. The intrinsic complexity of the problem requires tighter integration than is possible by scaling up present-day solutions using Moore's-law technological improvements.

The projects that we partner with here are already planning their futures in terms of Data Grid structures, defining national and international structures, in some cases hierarchical due to the existence of a single central data source (e.g., for CMS and ATLAS) and in other cases more oriented towards the federation of multiple data sources. A common requirement, which forms the focus of the GriPhyN project, is to be able to access and manage *virtual* data, which may not physically exist except as a specification for how it is to be calculated or fetched.

C.5.a.1. ATLAS and CMS

The ATLAS and CMS Collaborations, each including 2000 physicists and engineers from 150 institutes in more than 30 countries will explore a new realm of physics up to the TeV energy scale, including the Higgs particles through to be responsible for mass, supersymmetry and evidence for extra dimensions of space-time. Within the first year of LHC operation, from 2006, each project will store, access, process and analyze 10 PB of data, using of order 200 Teraflops of fully utilized compute resources situated at the Tier-N centers in their Grid hierarchies, situated throughout the world. The LHC data volume is expected to subsequently increase rapidly, reaching 1 Exabyte (1 million Terabytes) and several Petaflops of compute power consumed full time by approximately 2015.

LHC physicists need to seamlessly access their experimental data and results, independent of location and storage medium, in order to focus on the exploration for the new physics signals rather than the complexities of worldwide data management. Each project is in the process of implementing object-oriented software frameworks, database systems, and middleware components to support the seamless access to results at a single site. Both CMS and ATLAS will rely on the GriPhyN VDT, Grid security and information infrastructures, and the strategies to be developed in the course of this project, to provide global views of the data and worldwide rapid access to results, as the foundations of their scientific data exploration. For these experiments, iVDGL provides a realistic, wide-area distributed environment in which their Grid-based software can be prototyped, developed, refined, and evaluated, and optimized using simulated data production and analysis on increasing unprecedented scales from 2002 onwards. The field trials on increasing scales coincide with the major milestones of the physics projects during the construction phase, which is now underway. These milestones serve to verify the capabilities of the detector and online event filters used to select events in real time, to set final design details, and to provide a testing ground for the development of each project's overall data reconstruction and analysis systems. Major CMS data production and analysis milestones (ATLAS milestones are similar) include a 5% complexity data challenge (Dec 2002), 20% of the 2007 CPU and 100% complexity (Dec 2005), and start of LHC operations (2006).

ATLAS is developing an analysis framework and an object database management system. A prototype of this framework, ATHENA[35,36,37], released last May, does not support virtual data access directly, but is extensible and cleanly separates the persistency service, which must be grid-aware, from the transient data service, through which user algorithms access data. Virtual-data support is now being worked on by a number of ATLAS-related grid projects; this proposal will help provide the interfaces and develop protocols needed to integrate grid middleware into the ATLAS framework. In particular, we plan to integrate each yearly release of the VDT into ATHENA.

The main iVDGL experiments needed by ATLAS are the large scale exercising of the data-intensive aspects of the analysis framework culminating in the data challenge milestones mentioned above. These experiments will involve large distributed data samples of actual data (from test beams) and large simulated data samples produced by distributed, CPU-intensive Monte Carlo programs. The main goals of these experiments will be to insure that the most advanced implementations of grid middleware (virtual data, replica management, metadata catalogs, etc.) all work seamlessly with the ATLAS framework software at the largest possible scale.

CMS is at an advanced stage of developing several major object-oriented subsystems that are now entering a third round of development, notably the Object Reconstruction for CMS Analysis (ORCA) program, the Interactive Graphical User Analysis (IGUANA) code system, and the CMS Analysis and Reconstruction Framework (CARF). Persistent objects are handled by CARF using an ODBMS. CARF supports virtual data access. Monitoring tools are planned that will provide helpful feedback to users formulating requests for data, setting appropriate expectations as a function of data volume and data storage location; ultimately these functions may be automated. Efforts are underway to integrate Grid tools into the CMS core software, to help produce the Grid tools, and to monitor, measure and simulate Grid systems so as to identify good strategies for efficient data handling and workflow management.

Grid-enabled production is a near term CMS objective that will entail development of Globus-based services that tie resources and facilities at CERN, FNAL and existing Tier2 Centers, and that will enable CMS production managers to schedule the required large scale data processing tasks amongst the facilities in a generic, convenient, and robust fashion. The use of Globus as an integral component for this and more general user software facilitates global authentication schemes via unique Grid identities for all users and groups, leading to a secure system for global production and analysis sometime in 2002. The first prototype for this system will make use of a metadata catalogue that maps physics object collections to the database file(s) in which they are contained. The file view will then be used to locate suitable instances of the data in the Grid. The system will also include a suite of tools that allow task creation, deletion and redirection, in some cases automatically handled by the system itself using real time measurements of the prevailing load and availability of the computing resources in the Grid. After experience has been gained with the components of the prototype system, decisions will be made in 2004 on candidates for inclusion in the production Grid system for CMS. This will be followed by construction of the system itself, in time for startup in 2006, but with limited deployment in the intervening year.

CMS and ATLAS are committed to developing their software according to the iVDGL concepts presented in this proposal. Additionally, a strong emphasis is being placed on coordinating and retaining full consistency with other Grid-based efforts such as the EU Datagrid and PPDG, as this will lead to a unified set of Grid software and a common set of Grid operations-support tools, as well as policies that will allow for resource sharing in an equitable way.

C.5.a.2. LIGO/LSC

LIGO will become fully operational as a scientific facility during the period covered by this proposal. The Tier 2 centers that will be built at LIGO Scientific Collaboration (LSC) sites in the US will be designed as production systems from the beginning because the LSC cannot afford to disrupt operations to rebuild its centers in a second phase. With this in mind, the LIGO component of iVDGL will begin to provide useful scientific service to the collaboration within the first few years of operations. The LSC will rely on these resources for a variety of needs not now being met. These include (1) the ability to replicate and cache large volumes of LIGO data outside the Tier 1 data center of LIGO Laboratory; (2) the ability to stage computationally intense analyses[38,39], such as large-sky-area searches for continuous wave gravitational sources that cannot be accommodated at the LIGO Tier 1 center; (3) extending the range of minimum mass for inspiraling binary coalescences and using multiple interferometer data streams for coherent signal processing of a vector of signals generated by an array of detectors. Each of these needs can be fulfilled by successively larger scale use of iVDGL.

Data mirroring and caching will be performed using the Globus tools for generating virtual data replica catalogs and transmitting data over the Grid. Grid-based searches for continuous wave sources and inspiraling binary coalescences will be implemented using Condor-G[40], with executables and manageable data sets distributed to run independently on many nodes. The network-based analysis of multiple interferometer streams brings into play the major European interferometric gravitational wave detector projects. These projects have begun to explore how to use the UK/EU Grids for their respective programs. LIGO is collaborating with both the GEO Project with a 600m interferometer in Hanover, Germany, and the Virgo Project, with a 3000m interferometer in Cascina, Italy, to jointly analyze the datastreams from all interferometers. Use of iVDGL for such analyses will enable transportation and replication of the data volumes generated remotely by each observatory. In addition, the availability of large-scale computational resources on the grid will enable coherent processing of data on a scale that is beyond the capabilities of the individual data centers of each of the three projects separately.

The LIGO baseline data analysis software design does not incorporate inter-process communications between geographically isolated resources. A primary goal for the proposed LIGO/LSC effort will be to Grid-enable this software suite, using iVDGL to extend the currently restricted functionality in four major ways: (1) integration of LIGO core software API components with the GriPhyN Virtual Data Toolkit to enable Grid-based access to the LIGO databases; (2) adapting the LIGO search algorithm software to use iVDGL distributed computing capabilities; (3) replication of LIGO data across the iVDGL using VDT capabilities; (4) joint work with LIGO's international partners in Europe (GEO600[41] in UK/Germany, Virgo[42] in Italy/France) to establish a network-based analysis capability based on the grid that will include sharing of data across iVDGL.

C.5.a.3. SDSS/NVO

The NVO effort will initially use the early data releases from the SDSS project, integrated with other data sets as an early testbed of the iVDGL concept. This testbed for the NVO will have two initial, smaller-scale sites for iVDGL, with different functionalities, equivalent to a single Tier2 node. The Fermilab node will create large amounts of Virtual Data through reprocessing the 2.5 Terapixels of SDSS imaging data. It 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. We will modify the existing pipelines to be compliant with iVDGL middleware. This will be an extremely useful learning experience in the design of the NVO services. The JHU node will consist of the parallel catalog database and perform Virtual Data computations consisting of advanced statistical tools, measuring spatial clustering and their dependence on galaxy parameters[43,44,45]. These analyses will lead to new insights into the galaxy formation process: are galaxy types determined by "nature or nurture", and will measure the fundamental parameters of the Universe, like the cosmological constant. 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) interface the database to the iVDGL environment, (b) create a Grid-enabled version of the advanced statistical tools (c) support analyses within the iVDGL environment. Once the two test sites are functional, we will establish connections to other sites, in particular to Caltech, where NVO and iVDGL activities are both present as well. After the testing phase we will implement a virtual data service based on the SDSS experience 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 are seen as major customers of our VDG technology. Early access for the astronomy community to iVDGL resources will accelerate the development of Virtual Observatories over the whole world.

C.5.b. Experimental Use by Other Major Science Applications

Physics experiments are far not alone in their requirements for the management and analysis of large data sets: we find similar problems in earth sciences (e.g., climate modeling, earthquake modeling), biology (structural genomics, neuroscience), engineering (NEES), and many other disciplines. We will exploit the partnerships that we have established with other major consortia—in particular, the NSF-funded PACIs and NEESgrid, the EU-funded European Data Grid, and the DOE-funded Particle Physics Data Grid and Earth System Grid projects—to open up iVDGL to other application projects, with the twin goals of increasing the user community for this substantial NSF investment and broadening the range of usage scenarios investigated during studies of middleware and operations.

We give four examples to illustrate the range of application groups that we expect to work with in this way; we expect this list to grow substantially over the 5-year duration of the project.

Earthquake Engineering: The NSF's Network for Earthquake Engineering Simulation (NEES)[46] will connect a large user community with the nation's earthquake engineering facilities, community data archives, and simulation systems. NEESgrid principals wish to use iVDGL storage resources for mirroring common large datasets for collaborative analysis and iVDGL compute resources for computationally demanding analyses.

Biology: The NSF PACIs are building data grids to support distributed data collections[47] (NLM Digital Embryo[48], NIH Biology Imaging Research Network), federation of digital libraries[49] (NIH Human Brain project[50]), and distributed data ingestion[51] (NIH Joint Center for Structural Genomics[52]), with distributed collection management provided by the SDSC Storage Resource Broker[63] (SRB). SRB and VDT will be interfaced to extend these grids to iVDGL.

Climate: The DOE-funded Earth System Grid (ESG) project, involving DOE laboratories, NCAR, and universities, seeks to construct a national-scale infrastructure for access to and remote analysis of large climate model datasets. (In contrast to previous systems[53], computationally demanding analyses are supported.) iVDGL resources will allow ESG personnel to experiment with these issues at a dramatically enhanced scale.

Astrophysics: The Astrophysics Simulation Collaboratory is an international collaboration involving scientists at Washington University, the Max Planck Institute for Gravitational Physics, and others[54]. This group has already created a "Science Portal" that supports transparent remote access to computers, and is experimenting with automatic resource selection techniques; iVDGL will allow them to expand the scale of these experiments dramatically.

C.5.c. Experimental Studies of Middleware Infrastructure and Laboratory Operations

As discussed above, iVDGL is constructed from a significant software base that includes substantial middleware services and a range of data grid tools. This infrastructure is constructed from a combination of existing elements such as the Globus Toolkit, Condor, GDMP, etc., and other elements that are under development or planned for the future. Because of both its scale and the demanding nature of the proposed application experiments, iVDGL will push many elements of the infrastructure into previously unexplored operational regimes where one may expect to see new and subtle issues arrive. For example, there were problems with the ARPANET routing protocols that were not discovered until the network grew to a significant size. While we do not expect to see anything as dramatic as a "system meltdown" we do believe that iVDGL will expose issues of scalability, reliability, maintainability, and performance that would be difficult to otherwise observe. iVDGL provides a unique opportunity to study these questions *before* they become obstacles to scientific progress.

We will maximize this opportunity by incorporating instrumentation and logging at all levels within the iVDGL software infrastructure (a VTR team task) along with measurement of the performance of all of the underlying iVDGL resources and iVDGL workflows. We will record this information in persistent measurement archives (an LSO responsibility). Analysis of archived performance data can help us to detect and diagnose a variety of performance and scalability issues. This data will provide us realistic workloads and execution traces for use by groups studying, for example, Data Grid replication strategies, scheduling, fault tolerance strategies or computational markets[55,56].

An added benefit of the archived historical data is in addition to indicating how well the infrastructure is working, it can help us determine how well iVDGL is being used as a whole. The past 30 years have provided a huge amount of experience with the creation, maintenance, and operation of large-scale networks, and there are well-understood social and technical protocols for managing such large shared infrastructures. In contrast, we have little experience with operating large shared distributed processing and data storage infrastructures, and what experience we do have suggests that the effective operation of such systems is complicated by complex issues of control that arise when a resource is simultaneously local managed and owned, and shared with larger virtual organizations. The achieved usage data can help answer these questions. We plan to engage sociologists to investigate social issues relating to the sharing and collaborative work issues that arise in these settings. We will seek funding from other sources to this end: for example, senior personnel Livny has submitted a proposal with Daniel Kleinman of Wisconsin.

We plan to supplement measurements with active diagnostic tools. Many network problems have been diagnosed with tools such as `ping` and `traceroute`; analogous tools can help identify and diagnose middleware problems.

Finally, we will develop middleware benchmarks that can be used to study the behavior of iVDGL middleware.

These benchmarks consist of specialized application experiments that are designed to stress specific aspects of the middleware. These benchmarks can be used in combination with measurements and diagnostic tools to better understand the behavior exhibited by the benchmark. We note that there is currently little understanding of how to construct meaningful Grid benchmarks, making this activity especially useful.

C.6. Engage Underrepresented Groups in the Creation and Operation of the Laboratory

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. Our education and outreach program has several thrusts, which we describe in some detail below.

Integration of MSI sites and personnel into the laboratory: We will place iSites at three minority serving institutions (MSIs) historically underrepresented in scientific research projects to allow them to participate in our multi-level iVDGL. Placing iSites at these institutions will allow computer science and physics students to contribute in an extremely concrete, hands-on way to the creation, operation, and scientific exploitation of the iVDGL strategy that we believe will get students excited and informed about the Grid, advanced technology, and science.

Each of the three partner MSI institutions has some existing research, computing, and networking infrastructure to build on, and has ties to our partner physics applications and prior involvement with education and outreach activities. We will utilize the 96-node Linux cluster that is currently being built at UT Brownsville, a Hispanic Serving Institution (HSI) in Brownsville, Texas that has close ties with the LIGO project and the education and outreach component for GriPhyN. This cluster is supported by funds obtained from a previous NSF grant [PHY-9981795], but it can also be used as a test-bed for Grid software. We will also construct small (~32 node) clusters at Hampton University in Hampton, Virginia and Salish Kootenai College in Pablo, Montana, a Historically Black College and University (HBCU) and a member of the American Indian Higher Education Consortium (AIHEC), respectively. Hampton University has two faculty members doing research for the ATLAS experiment. Salish Kootenai College has a faculty member trained in gravitational physics who is interested in doing LIGO-related research at the LIGO Hanford Observatory. They are also currently working on a separate proposal for an Internet2 connection.

In support of this overall strategy, we will organize training courses on iSite installation and operation, summer student internships at other iVDGL partner institutions, and collaborative projects between MSI and other personnel aimed at both iVDGL measurement and analysis, and scientific studies. We are also discussing with the NSF PACIs ways in which this program could be expanded to other institutions if it proves as successful as we expect. One possibility is a combination of equipment funds from industrial sponsors, REU funds from NSF, and training by the PACI-EOT[57] and iVDGL partners, as a means of cloning this program elsewhere.

Other activities: In addition to constructing clusters at MSIs, we will be actively involved in other aspects of education and outreach. These include: (1) course development at both the undergraduate and graduate levels, with an emphasis on new advances and innovations in Grid computing; (2) commitment of all iVDGL participants to give seminars and/or colloquia at other institutions; (3) submission of a proposal for an REU supplement starting Summer 2002, giving undergraduate students the opportunity to participate in grid-related research at other participating iVDGL institutions; (5) leverage of Manuela Campanelli of UT Brownsville, who coordinates E/O activities for GriPhyN, to coordinate E/O for this proposal; (6) taking advantage of QuarkNet[58], a professional development program primarily for high school teachers funded by NSF and DOE and based at Fermilab, to provide additional resources and a grid-related component to existing QuarkNet activities at iVDGL universities; (7) utilization of the Johns Hopkins – Microsoft collaboration to develop a website, modeled after the Terraserver[59] to provide color images of the sky using SDSS data, tied to an object catalog and utilizing a special interface to support educational activities, e.g., by producing on-the-fly custom datasets representing images of the sky as viewed in different frequency bands; (8) an innovative partnership with ThinkQuest[60] to develop special challenge projects for student competitions. We would provide resources in the form of interesting datasets (e.g., SDSS images or LIGO data) and/or "sandbox" CPUs that students could use to create innovative Web-based educational tools based on that data.

C.7. Relationship to Other Projects

Strong connections with other projects will increase our chances of major success and magnify the impact of the proposal on scientific, industrial, and international practice. We speak elsewhere of our close linkages with application and international science and technology projects; here we address other connections.

Information Technology Research Community. The ambitious goals proposed for iVDGL are possible only because we are able to leverage and build upon a large body of prior work in data-intensive computing, distributed computing, networking, resource management, scheduling, agent technologies, and data grids. 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. The iVDGL principals are deeply involved in these various projects and so are well placed to coordinate closely with them.

iVDGL PIs and senior personnel lead the Condor[61], Globus[62], and Storage Resource Broker[63] projects, which are developing the technologies required for security, resource management, data access, and so forth in Grid environments. These projects are already partnering with iVDGL 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. These iVDGL personal are also participants in the NSF ITR GriPhyN project, which is conducting CS research on virtual data, request planning, request execution, and performance estimation technologies; integrating "best of breed" technologies from this work and elsewhere into a Virtual Data Toolkit. GriPhyN project CS participants are already collaborating with many of the iVDGL application experiments. Finally, we note that Foster and Kesselman are both PIs in the GraDS project, which is developing software development environments for Grid applications. Not only will the tools be developed by GraDS be of great use to developing application experiments, but the iVDGL can provide an essential evaluation environment for GraDS technology.

The NSF PACI program. The NSF sponsored PACI programs have been developing and deploying elements of Grid infrastructure across the NPACI and NCSA partnerships to form a *National Technology Grid*.. This has included common security infrastructure, information and resource discovery, resource management and collection management. iVDGL senior personal have played a pivotal role in the development and deployment of this Grid infrastructure. Both PACI centers have agreed to work with the iVDGL; given the reliance and adoption of common Grid technology, we can consider the resulting system to be a first step towards a national (and ultimately international) "cyberinfrastructure." We expect this cyberinfrastructure to evolve to link large, medium, and small compute and storage resources distributed nationally and internationally, so that individual communities can contribute to and/or negotiate access to elements to meet their own data sharing and analysis needs.

CAL-IT2 Initiative The California Institute for Telecommunications and Information Technology [Cal-IT2] was recently funded by the State of California. Centered at UCSD, its mission is to extend the reach of the current information infrastructure. Cal-IT2 has a particular focus on data gathered from sensing devices monitoring various aspects of the environment (e.g., seismicity, water pollution, snowpack level, health of the civil infrastructure) via wireless extensions of the Internet. Cal-IT2 had agreed to both participate as an iSite and contribute application experiments. This interaction will be coordinated through CAL-IT2 member USC/ISI.

Industrial Partners Industrial partners will play a strong role in the iVDGL, not only as technology providers (CPU, disk and tape storage, switches, networks, etc.), but also as partners in solving problems in low-cost cluster management, cluster performance, scaling to very large clusters (thousands of boxes), scheduling, resource discovery, and other areas of Grid technology. We have initiated discussions of collaboration with a number of different computer vendors to undertake collaborative activities related to iSite hardware configurations.

C.8. 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* that already exists among the physics and astronomy experiments in this proposal 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 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.

C.9. Schedules and Milestones

We describe here the specific tasks to be undertaken during each of the five years of the project. We note that due to the unique nature and global appeal of the iVDGL, we anticipate many additional benefits accruing that are not identified here; the following is hence a *subset* of the overall outcomes. The scale and diversity of these outcomes emphasizes why we request five year's funding: the ambitious goals of the iVDGL and of iVDGL partner scientific projects demand a sustained effort over this period. Our work during this period will be focused roughly as follows:

Year 1: Establish the Laboratory

Architecture: Initial design. Document the iSite and iGOC software designs, indicating protocols and behaviors. In collaboration with PACIs, Globus, GriPhyN projects, produce software loads for standard hardware configurations.

Deployment: Create U.S.-based multi-site prototype. Instantiate iSite software at the first ten sites in the U.S, including prototype URC facilities at LHC affiliated universities. Negotiate MOUs among participating sites and with four frontier applications. Develop initial iGOC services including monitoring for failure detection and reporting, certificate authority support, and trouble ticket system. Establish contact with analogs at DTF, CERN, Japan.

Experiments: Test infrastructure and conduct first application experiments. *ATLAS*: Initial Grid-enabled software development. 1% scale data challenge simulating group-based production and individual-physicist analysis chains on iVDGL. *CMS*: Deploy first interfaces between the CMS CARF framework (and ODBMS) and Globus APIs. Complete a CMS production cycle between multiple iVDGL sites with partially Grid enabled software. *LIGO*: Develop Globus-based APIs to mirror data among URCs; develop virtual data replication and extraction methods in order to deliver reduced data subsets. *SDSS/NVO*: Grid-enable galaxy cluster finding code correlation function and power spectrum code. Test replication of SDSS databases using existing grid infrastructure. Demonstrate galaxy searches on 6-site iVDGL prototype. *Middleware/Tools*: Verify correct operation on scale of ten sites. Benchmark studies to document performance of data movement, replication, and other basic functions. *Operations*: Verify ability to detect resource and network failures in small-scale system.

C.9.a. Year 2: Demonstrate Value in Application Experiments

Architecture: Refine design; interface to PACIs and DTF. Refine software design and software load definition based on Year 1 experience and parallel developments. Define interfaces to PACI and DTF resources and negotiate MOUs with them and with EU participants.

Deployment: Expand iVDGL to 18 sites, including international. Extend to 13 sites in the U.S. and 5 in Europe, all running second-generation software including new URCs. Establish high-speed network connectivity within U.S. and across the Atlantic using the SURFNET and DATATAG research links (OC-48).. Establish iGOC trouble ticket system. Develop iVDGL monitoring systems which handle collection and presentation of grid telemetry data. Regression tests of all components. Start logging of iVDGL trouble tickets for human factors studies. Evaluate reliability and scalability of iVDGL monitoring. Deploy a set of infrastructure support services (bandwidth management services) being developed by other organizations such as Internet2 e.g. QoS, DiffServ, and Multicast.

Experiments/Applications: First large-scale application runs across 15 sites and Gb/s networks. *ATLAS*: Continued integration of grid services with the Athena framework. Data Challenge 2: 10% complexity scale involving 5-10 iVDGL sites. Performance and functionality tests will be used in ATLAS computing technical design report. Validate the LHC computing model[64]. *CMS*: 5% complexity data challenge using 10 iVDGL sites and approximately 50 users. Use DTF to explore use of NRC class facilities. Completion of a CMS production cycle where half the efforts are completed using Grid tools, including: first pre-production set tools for task monitoring, optimal task assignment to sites, in addition to the tools used in the previous year. *LIGO*: Port LIGO scientific algorithm code to the iVDGL for pulsar searches over large portions of the sky; port code to perform lower-mass inspiraling binary coalescence searches; work with EU partners to replicate this capability on the EU grid; *SDSS/NVO*: Tests of code migration between iVDGL sites. Grid-enable gravitational lensing application code. Integrate first SDSS data release. Run prototype power spectrum calculations in production mode, on medium scale. *Other Apps*: Work with PACIs and others to define experiments. *Middleware/Tools*: Deploy instrumentation archives. Start collecting and analysis usage patterns on iVDGL services, scalability studies.

C.9.b. Year 3: Couple with Other National and International Infrastructures

Architecture: Increased iGOC function, Improved virtual data support. Expand iGOC design to support coordination with DTF, Europe, Asia GOCs. Development and deployment of measurement tools designed to troubleshoot a specific grid "path" to isolate grid bottlenecks, including preliminary designs agent-based monitoring services.

Deployment: Expand to 30 sites and Asia. 15 in the U.S. and 15 in Europe. Establish coupling with DTF and demonstrate resource sharing on a large scale. Operate over 10 Gb/s networks within the U.S. and over multi-Gb/s networks internationally.

Experiments: Large-scale international experiments. *ATLAS*: Establish full chain tests of the grid-enabled Athena control framework with increasing size and complexity. Execute Athena-Grid production for the ATLAS "Physics Readiness Report" (January – June 2004). *CMS*: Completion of a CMS production cycle between multiple sites where by default all of the CMS production efforts are completed using Grid tools. System evaluation with diverse studies of ~10^8 fully simulated and reconstructed events. *LIGO*: Port LIGO scientific algorithm code to the iVDGL for multiple interferometer coherent processing; stage these analyses in both EU and US to perform complementary analyses. *SDSS/NVO*: Implement grid-enabled tools for statistical calculations on large scale structure. Begin full scale tests with additional iVDGL sites. Integrate second SDSS data release. *Other Apps*: CAL-IT(2). *Middleware/Tools*: develop benchmarking applications, performance tuning, develop workload models.

C.9.c. Year 4: Production Operation on an International Scale

Architecture: Incorporation of final GriPhyN VDT results. Active iSite compliance checking tools

Deployment: Add sites in Japan and Australia. Deployment of security tools and centralized monitoring of security services deployed in the iVDGL.

Experiments: Large-scale production runs on an international scale. iVDGL resources will be used for large-scale, long-duration computing runs designed to stress test multiple aspects of the infrastructure. *ATLAS*: 20% scale full production capability realized involving 10's of iVDGL sites in the U.S., Europe and Japan. *CMS*: first year of development of the Production Grid System for CMS Physics.. Large scale data productions, including runs with a level of complexity (number of processors, disks, tape slots) which are 50% of the LHC production levels.. *LIGO*: Begin transition to full-scale operations. Link to Australian ARIGA project. *SDSS/NVO*: Begin integration of National Virtual Observatory infrastructure with iVDGL technology. Large scale production runs on core science using half SDSS dataset. Update to second generation hardware. Integrate third SDSS data release. *Other Apps*: Climate modeling. *Middleware/Tools*: Grid diagnostic tools. Performance evaluation of impact of Asia/Europe long-hall links. Data replication reliability study.

C.9.d. Year 5: Expand iVDGL to Other Disciplines and Resources

Deployment: Cloning of iVDGL and iGOC functionality to other disciplines and sites, allowing expansion to potentially 100s of sites.

Experiments: Continued production runs; large-scale experiments involving other scientific disciplines. *ATLAS*: LHC startup during 2006. Support full-scale production of Monte Carlo and detector calibration analysis activities using the iVDGL . *CMS*: Deployment of the unified collaboration-wide CMS Grid system, to be used during LHC operations. Final testing and development stages, with continual scaling up and progressive integration of all CMS sites between Tier0 and Tier2 (30+ sites) into the production Grid system. Test interfaces to an increasing number of institute servers (up to 150) and to ~2000 desktops in CMS, by the time of the first LHC physics run, in Summer 2006. *LIGO*: Complete transition to full-scale operations. *SDSS/NVO*: Large scale production runs on core science using full SDSS dataset. Conduct initial joint SDSS/NVO analyses. Integrate final SDSS data release. *Middleware/Tools*: Analysis of usage and sharing patterns. Continued scalability analysis.

C.10. Broader Impact of Proposed Research

Promoting Scientific Discovery: Scientific discovery in many areas of research depends increasingly on collaborations of nationally and internationally distributed teams of researchers who must manage and access enormous data collections. The iVDGL provides teams with an information technology infrastructure that promotes efficient interaction and sharing of results, leading inevitably to higher scientific productivity.

Developing University Partnerships: Universities will make up approximately half of the total resources in the iVDGL, giving them a powerful voice in their large-scale scientific experiments. The universities in this program will also lead the effort to integrate grid tools with their scientific applications, increasing the scientific productivity of the experiments. Finally, iVDGL middleware will lower the barrier faced by researchers accessing data from distant sites and empowers smaller institutions with limited budgets to contribute intellectually to the analysis.

Building Global Partnerships: International partnerships are interwoven throughout the iVDGL, as seen by the geographical spread of the iVDGL sites, the collaborations with other international Grid projects, and the global nature of the scientific collaborations that will participate in the iVDGL.

Promoting Teaching and Learning: We promote learning and inclusion both via the integration of other under-represented institutions into the iVDGL fabric (by placing resource centers at those institutions) and via the integration of a diverse set of U.S. universities into the scientific program of application and computer science experiments.

C.11. iVDGL Management Plan

C.11.a. Overview

The U.S. iVDGL project will be a complex and challenging enterprise. In addition to carrying out the ambitious plan of work presented in this proposal, effective coordination needs to be maintained with several other organizations. Coordination with the GriPhyN project is essential since the iVDGL is designed to put tools developed in the GriPhyN research project to practical use by ensuring integration with software applications of the four participating physics projects. The iVDGL will also build prototype virtual data grids in which the physics collaborations and other application groups can begin to use GriPhyN and other Data Grid tools for their own needs and in the process put them to stringent tests in the real world of scientific applications. Second only to GriPhyN, strong coordination must be maintained with the four physics collaborations participating in the project, so that they can use the resulting iVDGL effectively and so that iVDGL developments mesh with the deliverables and milestones established in each experiment. Finally, close coordination is required with other Data Grid projects, especially the DOE-supported Particle Physics Data Grid and the EU DataGrid and related national projects, and with other partners, in particular the NSF PACIs and the proposed Distributed Terascale Facility.

In this section we describe a management organization to meet these goals. It is closely tied to the management organization of the GriPhyN project, which is modified slightly to mesh with that of iVDGL. While we believe that the managements of GriPhyN and iVDGL must be closely connected, the proposed organization is robust enough to provide strong management attention to each project independently, while ensuring that their goals and work will mesh where needed. Management of GriPhyN and iVDGL are combined at the top level of the Project Directors and the advisory Collaboration Board and External Advisory Panel. Below this level, the management and other organization of the two projects are separate.

We note that while we have taken pains here to define responsibilities clearly, the center will emphasize participatory decision making and, as far as possible, consensus-based management.

A pictorial layout of the management of iVDGL and GriPhyN together is included at the end of this section.

C.11.b. Project Leadership

Directors: Project PIs Paul Avery and Ian Foster will act as Co-Directors, with Avery serving as the primary point of contact for NSF and external relations and providing overall scientific leadership, and Foster providing technical leadership and also being responsible for coordinating activities with other relevant CS activities. The Co-Directors will make resource allocation decisions jointly and are ultimately responsible for the success of the Project.

Project Coordinators: iVDGL and GriPhyN separately will have full-time Project Coordinators who manage the day-to-day activities of each project under the overall management of the Directors. The principal duties are guiding and overseeing progress in the various technical areas, tracking costs and technical progress toward deliverables and milestones, calling meetings and providing minutes, maintaining the Project web page and identifying areas that require attention. The Project Coordinators will be appointed by the Project Directors on the advice of the Collaboration Board.

Project Coordination Group: The iVDGL Project Coordination Group (PCG) will be formed from the Directors, the Project Coordinator, the Systems Integration Coordinator, the Education Outreach Coordinator, the Technical Area Coordinators, a representative of each physics collaboration designated by the lead computing manager of that collaboration, a representative from each URC funded by this project, and additional members nominated to represent the NSF PACIs and potentially other interested application groups. This membership will result in some overlap with the already in place for GriPhyN, but it will be a goal to minimize overlap so that each project gets dedicated attention and guidance. The PCG will advise the Project Coordinator and will meet usually weekly and usually via teleconference or videoconference to set goals, plan integration milestones, and discuss progress.

Advisory Committees: The Directors and Project Coordinators will be assisted in their duties by two committees that are common to the two projects: (1) the Collaboration Board, which makes policy and strategic decisions for the Projects and (2) the External Advisory Board, which is an external group of experts that provides strategic guidance. These are described further below. In addition the NSF will provide committees that review the project either together or separately as best meets the needs of the agency.

Collaboration Board: The Collaboration Board (CB) will serve an internal advisory role. The CB will consist of the local project PI from each collaborating site and will be charged with advising the PIs on long-term strategy relating to resource allocations, relationships with other projects, and other matters of strategic importance. The CB will meet 2 or 3 times per year, mostly by videoconference. Most matters will be discussed electronically through the appropriate mailing lists and voting will normally be conducted electronically, except when possible at meetings.

External Advisory Board: The External Advisory Board (EAB), appointed by the Project Directors in consultation with NSF, will serve as a source of strategic advice on project directions. We expect this board to meet annually face-to-face and to participate in teleconference reviews on two separate occasions during the year. The EAB will consist of senior IT and physics researchers and managers and will meet once a year face-to-face at a project all-hands meeting and on additional occasions via video or teleconference.

NSF Review Committee: NSF will convene a Review Committee charged with assessing the progress of the project. This committee will meet approximately once a year.

C.11.c. Technical Areas

The day-to-day work of the Project is organized as four main areas: (1) iVDGL Design and Deployment, (2) University Research Centers and University Research Groups, (3) the International Grid Operations Center, and (4) Integration with Applications. Each area is led by an *Area Coordinator*, responsible for monitoring work within that area and for negotiating requirements and handoffs with other subprojects. The iVDGL Technical Areas are :

iVDGL Design and Deployment: This technical area will carry out the architectural definition of Section **Error! Reference source not found.** of the main proposal and the iteration and improvement of the architecture described in Section C.5, as well as the operation and experiments to be carried out with the iVDGL.

University Research Centers and University Research Groups This technical area will coordinate the design, implementation, and operation of the regional centers. It will be closely coordinated with the facilities organizations of the participating physics collaborations.

International Grid Operations Center: An International Grid Operations Center (iGOC) will be established to monitor the operation of the iVDGL and to facilitate rectification of difficulties through use of sophisticated monitoring and diagnostic tools, with tracking implemented through a trouble ticket system.

Integration with Applications This technical area will bring computer scientists active in datagrid software development together with those developing science applications to enable the scientific communities to make full use of the datagrid tools and capabilities

C.11.d. Liaison with Application Projects and Management of Facilities

The partner application projects are represented in the Project in several ways. First, each project has a representative in the Planning Coordination Group and therefore has frequent interactions with the research efforts and is able make its concerns heard by the Directors. Each project is also represented in the Application Projects research area. Finally, the Directors maintain a formal liaison with the computing heads of the applications or their designees. These mechanisms are meant to provide clear communication channels between the Project and the applications and to ensure that the desires of the applications are met.

Each of the developmental 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 physics collaborations. Each center will be associated with one of the collaborations. The local operators of each center will be part of the facilities organization of the associated collaboration, which will manage it for the iVDGL project. These operators will have the responsibility of operating the center to meet the needs of both iVDGL and of the particular collaboration. In order to permit the necessary integration of data grid capabilities with each collaboration'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 collaboration, with operations assigned by the collaboration. 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.

The dual responsibilities of each regional center will be described and governed by Memoranda of Understanding (MOUs) between iVDGL and the computing managements of the participating physics collaborations. The MOUs will describe the capabilities of each regional center and the commitment of that center to participate in the iVDGL. Each MOU will also describe how the regional center is managed, what responsibilities are assigned within the center, and how any conflicts between the needs of the physics collaborations and those of iVDGL will be resolved.

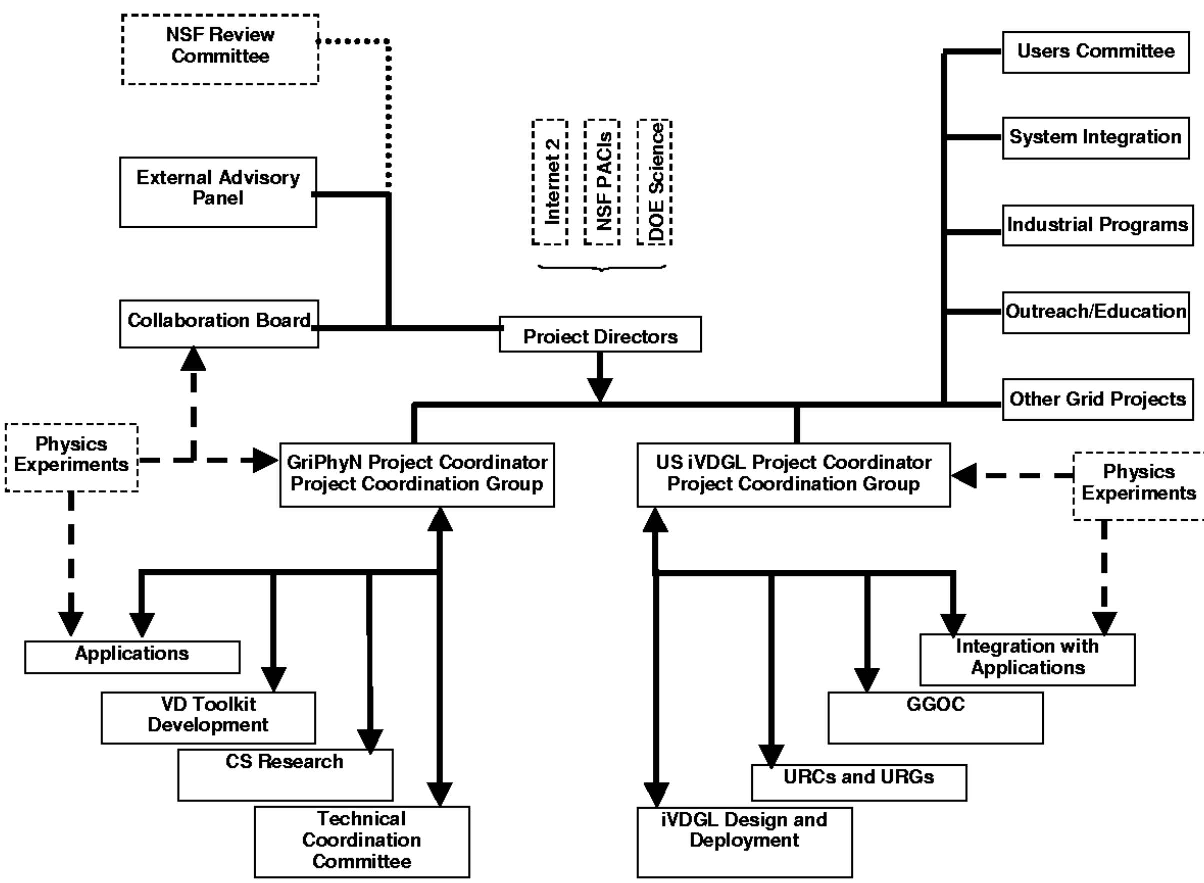
C.11.e. Additional Coordination Tasks

Special coordinators handle additional coordination tasks. These include (1) *Systems Integration*, responsible for ensuring that the various results from CS research and the Virtual Data Toolkit groups are consistent and complete enough to form a working software release; (2) *Industrial Programs*, responsible for developing relationships with vendors to further the Project's aims; (3) *Education and Outreach*, exposing faculty and students at other institutions to iVDGL research; and (4) Other Grid Projects, providing liaison and coordination with related projects including PPDG, EU DataGrid, the PACIs, and Globus. Finally, a users committee will be established on an international basis to drive and carry out the experiments and measurements that will be made on the iVDGL worldwide and on its various component grid systems.

C.11.f. Coordination with International Efforts

The four experimental collaborations 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 proposal, 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. One major meeting has already been held to organize structures to assure coordination and cooperation among the datagrid projects and a follow-up meeting is planned. The major projects worldwide are committed to developing compatible architectures and to conducting extended joint testbeds to ensure and demonstrate interoperability.

Organization of iVDGL (with connections to GriPhyN organization)



D. Results from Prior NSF Support

Paul Avery: (PHY-9318151: Nile National Challenge project) developed a scalable framework for distributing data-intensive applications, including a Java-2 based control system that uses CORBA. Project Director for GriPhyn Project (NSF ITR-0086044), funded in FY2001. Responsible for overall mission of GriPhyN and for application of virtual data in CMS experiment, where virtual data concepts are to be used.

Ian Foster: (CCR-8899615: Center for Research in Parallel Computation). Funding provided under this NSF funded science and technology center was used to conduct research in parallel programming languages and the Nexus runtime system. This work let to the development of the development of the Globus toolkit which provides mechanisms for communication, resource management, security, data access, and information in high-performance distributed environments. Globus is used extensively within the National Technology Grid being established by the two NSF PACIs as well as by many NSF-supported research projects. (NSF ASC 96-1920 National Computational Science Alliance) Funding from this PACI center was used to deploy Globus in production environments, and develop advanced Grid applications. Foster is also part of the CGrADS project (NSF-EIA 9975020), which is investigating the construction of software development environments for Grid applications. Work in GrADS has focused on the development of so called virtual organization tools for structuring Grid-based execution environments for program execution. Finally, as Co-PI on the GriPhyn Project (NSF ITR-0086044), funded in FY2001, Foster is leading research in the use of "virtual data" as a basic usage paradigm for data-intensive science. To date, this work has produced documents detailing a virtual data grid reference architecture and catalog structures for representing virtual data. In addition, graduate students supported under GriPhyN have started simulation studies of data replication strategies and high-performance data transport protocols.

Rob Gardner: GriPhyn Project (NSF ITR-0086044), funded in FY2001. Research and development of virtual data grid technologies for high-energy physics data analysis. Communicating requirements, integration of GriPhyN virtual data tools for the ATLAS experiment at CERN.

Harvey Newman: KDI proposal, GriPhyN (Ongoing KDI- and GriPhyN-related developments to be applied in the iVDGL. Work at Caltech in collaboration with Bunn, Messina, Samar, Litvin, Holtman Wilkinson, et al., as well as the PPDG and DataGrid projects): development of ODBMS-based scalable reconstruction and analysis prototypes working seamlessly over WANs; Grid Data Management Pilot distributed file service used by CMS in production (together with EU DataGrid); Grid-optimized client-server data analysis prototype development (Steenberg et al.), MONARC[65] simulation systems and application to optimized inter-site load balancing using Self Organizing Neural Nets (Legrand et al.); development of a scalable execution service (Hickey et al.); modeling CMS Grid workloads (Holtman, Bunn et al.); optimized bit-sliced TAGs for rapid object access (Stockinger et al.); development of a DTF prototype for seamless data production between Caltech, Wisconsin and NCSA (Litvin et al.; with Livny at Wisconsin and Koranda at NCSA).

Alex Szalay: (KDI-9980044: Large Distributed Archives in Phyics and Astronomy) in a collaborative effort between JHU (Szalay, Vishniac, Pevsner in Physics and Goodrich, CS) with Caltech (**Newman**, **Bunn** and Martin), Fermilab (Nash, Kasemann, Pordes) and Microsoft (Jim Gray) seeks to create a distributed database environment with intelligent query agents. The project is exploring different data organization to speed up certain types of distributed queries[66,67,68], comparing object-oriented and relational databases[69,70], and building intelligent query agents[71]. GriPhyn Project (NSF ITR-0086044), funded in FY2001.

E. International Collaborations

The proposed iVDGL will benefit from what is, for an information technology project, an unprecedented degree of international collaboration that will permit us to establish a truly international virtual laboratory of a scale significantly larger than could be established within the U.S. alone. Furthermore, the creation of this facility and its global scope will itself make major contributions to the practice and infrastructure of international science.

Note: *This section is supported by letters in the Supplementary Documents section of the proposal.*

E.1. International Collaborators: Programs

Our international collaborators comprise resource providers, network operators, and scientific collaborations. Section F (Facilities) provides details on the resources committed by these collaborators; here, we provide an overview of their significance.

European Data Grid (Fabrizio Gagliardi, Project Director, CERN): This 10M Euro flagship European Union project is charged with establishing a European analog to GriPhyN, a Data Grid focused on the analysis of data from LHC experiments, European Space Agency satellites, and biomedical programs. As stated in the enclosed letter of support, the EDG project director has committed to working with us to establish an iVDGL in which EDG sites (of which some 40 are envisioned by Fall 2001) participate as equal partners. This commitment is feasible because the EDG has already committed to the use of the same Globus infrastructure in use by U.S. projects, and because the EDG project includes a significant "testbed" program. The partnership with EDG is important for three reasons: it allows us to increase the scale, and hence the interest from a research perspective, of iVDGL; it increases significantly the interest of the iVDGL to our partners in high energy physics, due to the connection to CERN; and it connects us with strong environmental and biomedical communities.

U.K. eScience (Tony Hey, Director Core Programme, Neil Geddes, Director, PPARC eScience): The 100M pound U.K. eScience program has been established to support advanced Grid-based technologies across a wide range of U.K. science and engineering. As stated in the enclosed letter of support, the director of the Core Programme of the eScience project has committed to working with us to establish an iVDGL in which U.K. sites participate as major partners. This support includes 6 staff who will work *in the U.S.* on iVDGL activities, a Globus support center in the U.K., and support for U.K. iVDGL nodes. The partnership with the U.K. program is important not only because it adds significantly to the size of the program, but because it provides us with another connection with strong biomedical and environmental groups.

INFN Grid (Mirco Mazzucato, Project Directory, INFN): This large Italian project is developing Grid capabilities across some 20 centers and universities in Italy in support of the EU DataGrid project. They have been aggressive and early adopters of Condor and Globus technologies. As stated in the enclosed letter of support, the project director has committed to five iVDGL sites *and* 5-6 supporting staff positions within Italy. This commitment represents a major contribution to the scope and operation of iVDGL.

Japan (Satoshi Sekiguchi, Tsukuba Advanced Computing Center; Satoshi Matsuoka, Tokyo Institute of Technology): Japan does not as yet have a formal government Grid research program, but several institutions are active in Grid research and international scientific collaborations that require Grid capabilities. Japanese scientists have committed to establishing two iVDGL nodes within Japan and to participating in iVDGL experiments. This partnership is important as it both doubles the geographic reach of the iVDGL and supports a range of international data-intensive science projects that involve Japanese collaborations, including LHC experiments.

Australia (John O'Callaghan, Director, Australian Partnership for Advanced Computing): Australia has a significant investment in high-performance computing and networking and plays a significant role in certain international science projects, including astronomy (Sloan Sky Survey) and gravity wave research (ACIGA).

STAR-TAP and STARLIGHT (Tom DeFanti, UIC/EVL): We discuss this U.S.-based program here because of its importance as an enabler of international science. STAR TAP (www.startap.net) currently provides connectivity and transit services to 19 countries in Europe, Asia, and the Americas. StarLight, operational from July 1, 2001, will provide high-speed all-optical connectivity on an international scale. Dr. Tom DeFanti, STAR TAP and StarLight Director, has committed to providing advisory support on international networking issues, access to his facilities, and engineering support as required and when available. StarLight will provide an essential element of our iVDGL, allowing for data movement to international sites at speeds approaching this supported within the U.S.

E.2. International Collaborations: Application Projects

IVDGL also integrates, and provides support to, major international science projects and international collaborations. We list just a few of the more significant of these here.

Large Hadron Collider. The ATLAS and CMS detectors at the LHC are both extremely large international collaborations, comprising thousands of scientists in tens of countries. iVDGL facilities will make significant contributions to the success of these projects, as attested to by the enclosed letters of support.

Gravitational Wave Observatories. We have explained how iVDGL facilities will be used to enable collaboration between LIGO, VIRGO, and GEO, and to facilitate international access to data produced by these various experiments—for example, from Australia, via connections to the ACIGA project.

Computer Science. Substantial international collaborations in computer science are being planned in support of anticipated iVDGL activities. These include experimental investigations of high-speed wide area protocols on transoceanic links; development of Data Grid simulation tools; investigations of agent technologies for Data Grid monitoring and operation; developments of Data Grid management tools; etc.

E.3. International Collaborations: Management and Oversight

A project as complex, broad-ranging, and multi-institutional as the iVDGL will require extremely careful management if it is to be successful. We have already taken a first significant step towards this goal via the establishment of an *International Data Grid Coordination Board*. This body met for the first time in Amsterdam in March 2001, under the chairmanship of Larry Price of Argonne National Laboratory, at the time of the Global Grid Forum meeting, and discussed mechanisms for coordinated testbed development and experimentation. 25 participants from Europe, Japan, and the U.S. participated. The next meeting is scheduled for June 23rd, in Rome, immediately following EuroGlobus; the meeting after that will occur at the Global Grid Forum in Washington, D.C., in July, 2001.

In addition, we note that strong interpersonal and inter-project relationships have been established involving many of the participants. For example, Foster and Kesselman are on the project management board of the EU DataGrid project and the Technical Advisory Board of the U.K. eScience project; Gagliardi is on the External Advisory Committee for GriPhyN. Numerous staff exchanges have occurred within the past year. Participating application projects have similarly strong linkages.

The Global Grid Forum, founded by Co-PI Foster and involving many iVDGL participants, provides another body that will contribute to coordination of iVDGL activities.

E.4. International Synergies and Benefits

We conclude be explaining the three reasons why we believe that the iVDGL is of critical importance for international science.

A motivator for, and enabler of, international collaboration. International collaboration on Grid technologies is of vital importance due to the complexity of the problems involved, the importance of the international science projects that depend on those technologies, and the high costs of lack of cooperation in the form of inconsistent standards. However, international collaboration on information technology typically does not just happen, due to the significant cultural, funding, and technical barriers involved. The creation of iVDGL will provide both an extremely attractive experimental system that will engage the most talented scientists, and a shared task that will create the personal bonds needed for long-term success. The result will be a significant strengthening of international cooperation in both discipline sciences and information technology.

A unique experimental facility. iVDGL will represent an experimental testbed of unprecedented scale and scope, and as such will enable IT research investigations that would just not be possible in its absence. As is the case with the LHC and other contemporary multinational experimental facilities, this new capability can be created only via international cooperation.

A prerequisite for international discipline science. Last but certainly not least, iVDGL will provide the infrastructure needed for large-scale international collaboration in a number of large application science projects, in such areas as astronomy, physics, earth sciences, and bioinformatics. iVDGL facilities will accelerate scientific progress in a range of scientific disciplines.

F. IVDGL Facilities

The iVDGL computing facilities are of course a central part of the project, forming the core of the proposed international virtual laboratory. These facilities will comprise computer and storage systems located at university sites and national and international laboratories in four continents. We describe here the types of facilities included in iVDGL, list the locations and funding status of these facilities, and describe the deployment plan by which these facilities will be integrated into iVDGL.

F.1. iVDGL Facilities Overview

We distinguish in our description of facilities between large, national-level National Resource Centers (NRC), typically hosted by national laboratories or research centers on behalf of large national or international efforts; University Resource Centers (URC) that serve a user community of several research groups who are part of the same collaboration or belong to the same institution and small, often dynamic Group Resource Centers (GRC) operated by individual research groups. This proposal will either fully or partially fund the establishment a number of URCs and GRCs and will partner with other groups to obtain access to NRCs as well as to additional URCs and GRCs. Together, these three classes of sites span a rich and diverse range of resources, software infrastructure, management styles, and user requirements; the aggregate computing power and storage capacity of the systems that we integrate to form iVDGL are unprecedented.

Some iVDGL resources already exist at U.S. institutions, but are inadequate for anticipated future demands; others will be created under funding requested in this proposal, or will be contributed by international collaborators. The majority of these systems are, or will be, cost-effective compute clusters, constructed of commodity PC-type components, using the Linux operating system, and publicly available cluster computing tools such as MPICH, Condor, and Globus. Experience has shown that these clusters offer the best value for the “embarrassingly parallel” computations needed for data analysis and experimental simulation. In some cases, driven by already large data and simulation volumes, these commodity clusters are connected to specialized large-scale data storage systems.

iVDGL sites will, ultimately, include:

- URCs and GRCs funded by this proposal: 8 URCs at US universities, 3 GRCs at small US colleges and universities;
- URCs and GRCs at another 4 participating university sites, funded from other sources
- NRCs and URCs at laboratory sites operated by US agencies (Fermilab, Brookhaven, Caltech)
- NRCs and URCs at European Data Grid testbed sites (at least 15, at CERN and in UK, France, Germany, Italy, perhaps elsewhere; ultimately up to 40 or more)
- NRCs and URCs in Australia and Japan (4 initially)

Additional existing GRCs actively engaged in the development of application software frameworks and Grid systems will be invited to join iVDGL on a case-by-case basis. 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, as described in Section C.4.b.

These sites, all at universities and laboratories in the US, Europe, and Asia, and funded by their respective national agencies, will comprise the iVDGL foundation, organizing activities for the different application experiments and occasionally arranging large-scale exercises utilizing a large fraction of the total sites, from national laboratories to small clusters.

F.2. iVDGL Facilities Details

Existing URCs and GRCs are summarized in the Table 1, which lists their primary characteristics, including number of nodes, processor type and speed, aggregate memory, disk, and tape storage, and internal and external networking bandwidth.

A typical Tier 2 center will contain many of the following elements:

- A hundred or more compute nodes. These will typically be Intel Pentium, AMD Athlon, or Compaq Alpha-based workstations or PCs, running the linux operating system.
- A private 100 Mbps or Gbps network with a fully meshed 100 Gbps intersection bandwidth dedicated switch.
- A tape robot with a capacity of tens to hundreds of Tbytes.
- Specialized RAID servers that provide reliable access to very large data sets.
- Front-end nodes that control the computational cluster and serve as world-visible portals or entry points.
- External network connections that scale up from OC3 and OC12 (2001) over time.

URCs and GRCs vary in their physical characteristics according to the needs of their primary applications. For example, analysis of high-energy experimental data has fairly low bandwidth requirements for the local private network, and compute throughput is largely limited by the speed of integer operations. Thusthe centers for these purposes will probably use 100 Mb/s Ethernet and dual-processor nodes. On the other hand, the analysis of gravitational wave data using frequency-domain techniques uses larger data sets, and is largely limited by the speed of floating-point arithmetic and access to memory. For these centers, single processor systems and Gb/s networking may be more suitable.

Because commodity computer hardware is evolving so rapidly, it is neither possible nor desirable to give a detailed design for centers that will be built more than a year in the future. Optimal choices of CPU type and speed, networking technology, and disk and tape architectures vary with a time scale of a year. Because the needs of the physics experiments drive the deployment of Tier 2 hardware resources, this means that iVDGL will be composed of diverse and heterogeneous systems.

F.3. iVDGL Deployment Schedule

Initial iVDGL deployment will target six existing URCs and GRCs, selected based on capabilities available local expertise available. (Caltech, FNAL, U.Chicago, UT Brownsville, UW Madison, and UW Milwaukee are current targets.) These centers will be provide resources for Grid computation and testing.

In the second year, we will include the outreach centers and roughly five of the European centers that have pledged support as detailed in the supplementary documents part of this proposal. Each European center is expected to have equipment similar to what is proposed here, and candidate centers will be integrated on a timeline based again on capabilities and local expertise. If DTF is funded, then these resources will also be made available for periods of peak or burst operation.

In the third year, the number of US and international sites within the iVDGL will increase sharply, to of order fifteen US sites and fifteen European sites, including NSF PACI sites.

In the fourth year, Japan and Australia will be added, as will additional sites in the US.

By the fifth year, we anticipate that there will be of order 100 Tier 2 centers in iVDGL, in addition to the ones funded here.

The proposed initial hardware purchases for the iVDGL compute facilities are summarized in Table 2. The type and scope of later purchases will be determined at a later time.

Table 1: Existing Facilities

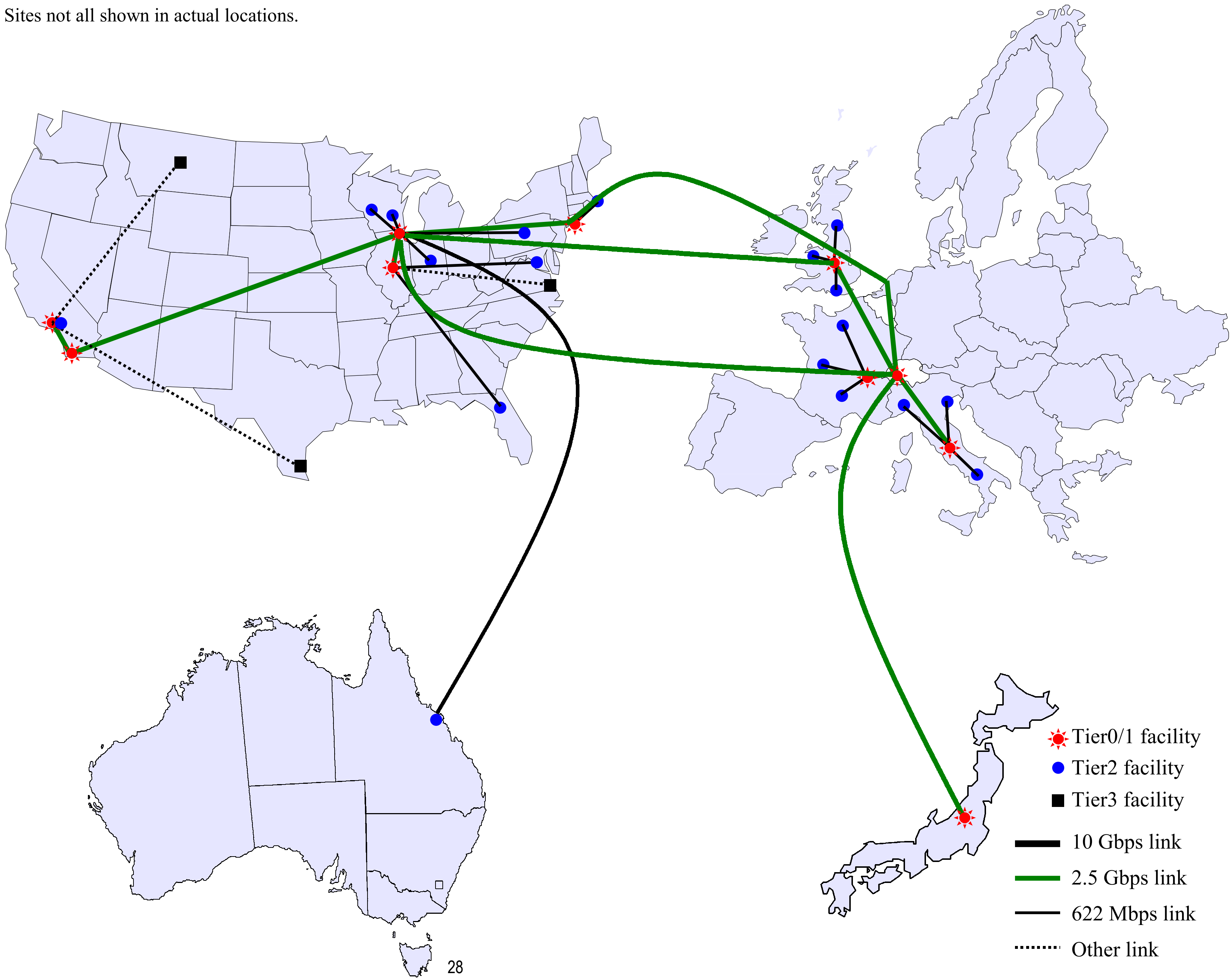
Institution	Exp	Processors	Peak	Total Memory	Total Local Disk	Storage	Network
BNL	ATLAS	124 Intel Pentium III @ 850 MHz	3K Si95	60 GB	3 TB	1 TB HPSS	OC3 to ESnet
BU	ATLAS	192 MIPS R10000 @ 195 MHz	75 Gflops	24 GB	250 GB	8 TB tape robot	OC3 to Abilene and vBNS
BU	ATLAS	64 IBM RS6000 @ 375 MHz	95 Gflops	8 GB	512 GB		
BU	ATLAS	10 Intel Pentium III @ 850 MHz	241 Si95	3 GB	300 GB		
IU	ATLAS	32 Intel Pentium II @ 400 MHz	12 Gflops	8 GB	300 GB	150 TB tape robot	OC12 to Abilene/Internet2
FNAL	CMS	8 Sun E4500 @ 400 MHz	3 Gflops	4 GB	500 GB	2 TB RAID, 10 TB HSM tape robot	OC3 to ESnet
FNAL	CMS	12 Intel Pentium III @ 550 MHz	6 Gflops	6 GB	500 GB		
FNAL	CMS	80 Intel Pentium III @ 750 MHz	60 Gflops	40 GB	3 TB		
UCSD	CS	1162 Power3+ @ 375 MHz	1700 Gflops	512 GB	10 TB	500 TB HPSS	OC48, OC12-48 (CalREN-2), ATM (OC12), Esnet
UFL	CMS/LIGO	144 Intel Pentium III @ 1 GHz	144 Gflops	36 GB	5.3 TB	0.7 TB RAID	OC12 to Internet2
UWM	LIGO	48 Alpha 21164 @ 300 MHz	29 Gflops	9 GB	70 GB	1.5 TB tape robot	OC3 link to MREN
UWM	LIGO	250 Intel Pentium III @ 1 GHz	250 Gflops	125 GB	21 TB		
UTB	LIGO	96 Intel Pentium III @ 800 MHz	76.8 Gflops	48 GB	3.8 TB	1.5 TB tape robot	4 T1
CIT	LIGO Lab	400 Intel Pentium IV @ 1.5 GHz	600 Gflops	400 GB	8 TB (IDE)	38 TB RAID, 600 TB HPSS	ATM (OC-12) - Caltech
JHU	SDSS/NVO	32 Intel Pentium II @ 330 MHz	10 Gflops	16 GB	1.5TB	600 GB SCSI	vBNS
JHU	SDSS/NVO	24 Intel XEON @ 1 GHz	24 Gflops	12 GB	1.5TB		
FNAL	SDSS	10 Pentium II @ 600 MHz	6 Gflops	7 GB	600 GB	3 TB RAID, 40 TB tape robot	OC3 to ESnet
ANL	CS	20 Intel Pentium III @ 850 MHz	17 Gflops	10 GB	600 GB	5 TB	MREN (OC3), Esnet (OC12+), I-WIRE (OC48+)
UC	CS	20 Intel Pentium III @ 850 MHz	17 Gflops	10 GB	600 GB	5 TB	
CACR	CS	128 PA8500 @ 440 MHz	1.5K Si95	64 GB	0.4 TB	3 TB RAID	NTON (OC48), CalREN-2 (OC12-48), ATM (OC12), Esnet
CACR	CS	256 PA8000 @ 180 MHz	3.2K Si95	64 GB	1 TB	> 600 TB HPSS	
CACR	CS	42 Intel Pentium III @ 850 MHz	1.5K Si95	25 GB	5 TB		
WISC	CS	1350 of various (Condor pool)	N/A	N/A	N/A	250 GB	vBNS, ESNet

Table 2: iVDGL sites to be funded under this proposal (initial purchases)

Institution	Exp	Processors	Peak	Total Memory	Total Local Disk	Storage	Network
IU	ATLAS	32 Intel Pentium IV @ 1.5 GHz	48 Gflops	16 GB	1.0 TB	-	-
UCSD	CMS	40 Intel Pentium III @ 800 MHz	3K Si95	40 GB	2.5 TB	-	-
FNAL	CMS	-	-	-	-	6 TB RAID	-
UFL	CMS/LIGO	100 Intel Pentium IV @ 1.5 GHz	150 Gflops	25 GB	3.7 TB	5 TB RAID	OC48
UWM	LIGO	96 nodes	192 Gflops	96 GB	8 TB	1 TB tape robot	OC12 to MREN
JHU	SDSS/NVO	64 Intel Pentium IV @ 1.5 GHz	96 Gflops	32 GB	1.0 TB	-	Abilene
FNAL	SDSS	64 Intel Pentium IV @ 1.5 GHz	96 Gflops	32 GB	6 TB	-	-
PSU	LIGO	96 nodes	192 Gflops	96 GB	8 TB	1 TB tape robot	OC12 to MREN
HU	ATLAS Outreach	32 Intel Pentium IV @ 1.5 GHz	48 Gflops	16 GB	1 TB	-	-
SKC	LIGO Outreach	32 Intel Pentium IV @ 1.5 GHz	48 Gflops	16 GB	1 TB	-	-

F.4. iVDGL Map (Circa 2002-2003)

Sites not all shown in actual locations.



G. List of Institutions and Personnel

Co-PIs

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

Collaborators

Only a fraction of these people would be funded by this proposal. No laboratory personnel will be funded.

Miron Livny	University of Wisconsin
Jenny Schopf	Northwestern University
Valerie Taylor	Northwestern University
Carl Kesselman	University of Southern California
Reagan Moore	San Diego Supercomputing Center
Ruth Pordes	Fermi National Accelerator Laboratory
Richard Mount	Stanford Linear Accelerator Center
Lothar Bauerdick	Fermi National Accelerator Laboratory
Vivian O'Dell	Fermi National Accelerator Laboratory
James Branson	University of California, San Diego
Ian Fisk	University of California, San Diego
Julian Bunn	California Institute of Technology
Alan George	University of Florida
Torre Wenaus	Brookhaven National Lab
Rich Baker	Brookhaven National Lab
John Huth	Harvard University
Jim Shank	Boston University
Lawrence Price	Argonne National Lab
Stephen Kent	Fermi National Accelerator Laboratory
Albert Lazzarini	California Institute of Technology
Roy Williams	California Institute of Technology
L. Samuel Finn	The Pennsylvania State University
Bruce Allen	University of Wisconsin, Milwaukee
Scott Koranda	University of Wisconsin, Milwaukee and NCSA
Manuela Campanelli	University of Texas, Brownsville
Tim Olson	Salish Kootenai University
Keith Baker	Hampton University

H. References

[1] Foster, I. and Kesselman, C. (eds.). *The Grid: Blueprint for a New Computing Infrastructure*. Morgan Kaufmann, 1999.

[2] Foster, I. Internet Computing and the Emerging Grid. *Nature Web Matters,* 2000. http://www.nature.com/nature/webmatters/grid/grid.html.

[3] Moore, R., Baru, C., Marciano, R., Rajasekar, A. and Wan, M. Data-Intensive Computing. In Foster, I. and Kesselman, C. eds. *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann, 1999, 105-129.

[4] Chervenak, A., Foster, I., Kesselman, C., Salisbury, C. and Tuecke, S. The Data Grid: Towards an Architecture for the Distributed Management and Analysis of Large Scientific Data Sets. *J. Network and Computer Applications,* 2001.

[5] GriPhyN Project home page, http://www.griphyn.org/.

[6] European Union DataGrid Project home page, http://grid.web.cern.ch/grid/.

[7] Abramovici A, Althouse WE, Drever RWP, Gursel Y, Kawamura S, Raab FJ, Shoemaker D, Sievers L, Spero RE, Thorne KS, Vogt RE, Weiss R, Whitcomb SE, Zucker ME, LIGO - The Laser-Interferometer-Gravitational-Wave-Observatory, Science 256: (5055) 325 (1992).

[8] B.C. Barish and R. Weiss, LIGO and the Detection of Gravitational Waves, Phys. Today, 52: (10), 44 (1999).

[9] The LIGO Experiment, http://www.ligo.caltech.edu/.

[10] P. Jenni, Spokesperson, The ATLAS Experiment, A Toroidal LHC ApparatuS, atlasinfo.cern.ch/Atlas/Welcome.html.

[11] M. Della Negra, Spokesperson, The CMS Experiment, A Compact Muon Solenoid, cmsinfo.cern.ch/Welcome.html.

[12] The Sloan Digital Sky Survey : http://www.sdss.org/.

[13] Szalay,A.S.: The Sloan Digital Sky Survey, CiSE, Vol 1, No 2, 54 (1999).

[14] National Virtual Observatory: http://www.srl.caltech.edu/nvo/

[15] A. Ricciardi, M. Ogg and K. Marzullo. The Nile System Architecture. *Proceedings of the Eleventh International Conference on Systems Engineering*, 9-11 July 1996, Las Vegas, NV, USA, pp. 414-419.

[16] E.g., National Scalable Cluster Project, http://nscp.upenn.edu/data_mining.htm

[17] Nelson Minar, Kwindla Hultman Kramer, and Pattie Maes, Cooperating Mobile Agents for Mapping Networks, *Proceedings of the First Hungarian National Conference on Agent Based Computation*, 1999.

[18] Y Yang, O F Rana, D W Walker, C Georgousopoulos, R Williams, Mobile Agents on the SARA Digital Library, *Proc. IEEE Advances in Digital Libraries 2000*, Washington DC, May 2000.

[19] Brian Tierney, William Johnston, Brian Crowley, Gary Hoo, Chris Brooks, Dan Gunter, The NetLogger Methodology for High Performance Distributed Systems Performance Analysis, *HPDC'99*, 1999.

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

[21] Simulation tools. http://monarc.web.cern.ch/MONARC/sim_tool/Publish/CMS/, http://monarc.web.cern.ch/MONARC/sim_tool/Publish/SONN

[22] R. Aiken, J. Strassner, Brian Carpenter, I. Foster, C. Lynch, J. Mambretti, R. Moore, B. Teitelbaum, Terminology for describing middleware for network policy and services, Internet RFC, 1999.

[23] R. Wolski, Forecasting Network Performance to Support Dynamic Scheduling Using the Network Weather Service, Proc. 5th IEEE Symposium on High-Performance Distributed Computing, IEEE Press, 1996.

[24] Partnership for Advanced Computational Infrastructure, http://www.ncsa.uiuc.edu/, http://www.npaci.edu/

[25] The Particle Physics Data Grid, http://www.cacr.caltech.edu/ppdg/

[26] I. Foster, J. Geisler, W. Nickless, W. Smith, S. Tuecke, Software Infrastructure for the I-WAY Metacomputing Experiment, *Concurrency: Practice and Experience*, 10(7):567-581, 1998.

[27] I. Foster and C. Kesselman, The Globus Project: A Status Report, *Proceedings of the Heterogeneous Computing Workshop*, IEEE Press, 4-18, 1998; also see www.globus.org

[28] W. Johnston, D. Gannon, B. Nitzberg, Grids as Production Computing Environments: The Engineering Aspects of NASA's Information Power Grid, *Proc. 8th High Performance Distributed Computing Symposium*, IEEE Press, 1999; see also www.nas.nasa.gov/IPG

[29] Stevens, R., Woodward, P., DeFanti, T. and Catlett, C. From the I-WAY to the National Technology Grid. *Communications of the ACM, 40*(11):50-61. 1997.

[30] Beiriger, J., Johnson, W., Bivens, H., Humphreys, S. and Rhea, R., Constructing the ASCI Grid. In *Proc. 9th IEEE Symposium on High Performance Distributed Computing*, 2000, IEEE Press.

[31] Foster, I., Kesselman, C., Tsudik, G. and Tuecke, S. A Security Architecture for Computational Grids. In *ACM Conference on Computers and Security*, 1998, 83-91.

[32] Butler, R., Engert, D., Foster, I., Kesselman, C., Tuecke, S., Volmer, J. and Welch, V. Design and Deployment of a National-Scale Authentication Infrastructure. *IEEE Computer, 33*(12):60-66. 2000.

[33] S. Fitzgerald, I. Foster, C. Kesselman, G. von Laszewski, W. Smith, and S. Tuecke, A Directory Service for Configuring High-performance Distributed Computations, *Proc. 6th IEEE Symposium on High-Performance Distributed Computing*, 365-375, IEEE Press, 1997.

[34] Czajkowski, K., Fitzgerald, S., Foster, I. and Kesselman, C. Grid Information Services for Distributed Resource Sharing, 2001.

[35] ATHENA Architecture and Framework, ATLAS Collaboration, http://atlas.web.cern.ch/Atlas/GROUPS/SOFTWARE/OO/architecture/

[36] Gaudi Architecture and Framework, LHCb Collaboration, http://lhcb-comp.web.cern.ch/lhcb-comp/Components/html/GaudiMain.html

[37] LHCb Collaboration, http://lhcb.web.cern.ch/lhcb/

[38] P.R. Brady and T. Creighton, *Phys.Rev.* D61 (2000) 082001

[39] P.R. Brady, T. Creighton, C. Cutler, B.F. Schutz, *Phys.Rev*. D57 (1998) 2101-2116

[40] Frey, J., Foster, I., Livny, M., Tannenbaum, T. and Tuecke, S. Condor-G: A Computation Management Agent for Multi-Institutional Grids, University of Wisconsin Madison, 2001.

[41] GEO600 gravitational wave experiment home page, http://www.geo600.uni-hannover.de/.

[42] VIRGO gravitational wave experiment home page, http://www.virgo.infn.it/.

[43] Vogeley,M.S. and Szalay,A.S.: Eigenmode Analysis of Galaxy Redshift Surveys, Theory and Methods, *Astrophys. J.*, 465, 34 (1996).

[44] Tegmark,M., Hamilton,A.J.S., Strauss,M.A., Vogeley,M.S. and Szalay,A.S.: Measuring the Galaxy Power Spectrum with Future Redshift Surveys, *Astrophys. J.*, 499, 555 (1998).

[45] Matsubara,T., Szalay,A.S. and Landy,S.D.: Cosmological Parameters from the Eigenmode Analysis of the Las Campanas Redshift Survey, *Astrophys. J.*, in press (2000).

[46] Network for Earthquake Engineering Simulation (NEESgrid): www.neesgrid.org

[47] C. Baru, Managing Very Large Scientific Data Collections, poster session, *5th International Conference on High Performance Computing HiPC'98*, Dec. 17-20, 1998, Chennai, India.

[48] Digital Embryo Library and Collaboratory Tools, http://www.archimuse.com/mw2001/abstracts/prg_95000313.html

[49] C. Baru, A. Gupta, V. Chu, B. Ludäscher, R. Marciano, Y. Papakonstantinou, P. Velikhov, XML-Based Information Mediation for Digital Libraries, demonstration session, *Digital Libraries '99*, http://www.npaci.edu/DICE/Pubs/

[50]Ellisman, M. Martone, M. et. al, NPACI neuroscience thrust area, Federating Brain Data.

[51] Rajasekar, A., R. Marciano, R. Moore, Collection Based Persistent Archives, *Proceedings of the 16th IEEE Symposium on Mass Storage Systems*, March 1999.

[52] Joint Center for Structural Genomics is a collaboration between Stanford Synchrotron Radiation Laboratory, University of California, San Diego, and The Scripps Research Institute, http://www.jcsg.org/.

[53]Davis, F., W. Farrell, J. Gray, C. R. Mechoso, R. W. Moore, S. Sides, M. Stonebraker, EOSDIS Alternative Architecture, Final Report submitted to HAIS, Sept 6, 1994.

[54] Benger, W., Foster, I., Novotny, J., Seidel, E., Shalf, J., Smith, W. and Walker, P., Numerical Relativity in a Distributed Environment. In *Proc. 9th SIAM Conference on Parallel Processing for Scientific Computing*, 1999.

[55] T. Malone, J. Yates, R. Benjamin, Electronic Markets and Electronic Hierarchies, *Communications of the ACM*, 30(6):484-497, 1987.

[56] Rich Wolski, James S. Plank and John Brevik. g-Commerce -- Building Computational Marketplaces for the Computational Grid, Technical Report UT-CS-00-439, University of Tennessee, April, 2000

[57] NPACI-EOT home page, http://www.eot.org/.

[58] QuarkNet home page, http://quarknet.fnal.gov/.

[59] TerraServer home page, http://terraserver.homeadvisor.msn.com/default.asp.

[60] ThinkQuest home page, http://www.advanced.org/thinkquest.html. Support letter in Supplementary Documents.

[61] The Condor Project, High Throughput Computing, www.cs.wisc.edu/condor/.

[62] I. Foster and C. Kesselman, The Globus Project: A Status Report, *Proceedings of the Heterogeneous Computing Workshop*, IEEE Press, 4-18, 1998; also see http://www.globus.org/.

[63] Baru, C., R, Moore, A. Rajasekar, M. Wan,The SDSC Storage Resource Broker, Proc. CASCON'98 Conference, Nov.30-Dec.3, 1998, Toronto, Canada; also http://www.npaci.edu/DICE/SRB/.

[64] Report of the Steering Group of the LHC Computing Review (Hoffman Panel), S. Bethe (Chair), CERN/RRB-D 2001-3, February 22, 2001.

[65] L. Barone et al. (The MONARC Architecture Group), *Regional Centers for LHC Computing*, MONARC-DOCS 1999-03, monarc.web.cern.ch/MONARC/docs/monarc_docs/1999-03.html (1999)

[66] Chaudhary,A., Duncan,C., Goodrich,M. and Szalay,A.S.: Generalized Balanced Aspect Ratio Trees in R^2, *4th CGC Workshop on Computational Geometry*, (1999).

[67] Szalay, A.S., Kunszt, P.Z., Thakar, A., Gray, J., and Slutz, D.: The Sloan Digital Sky Survey and its Archive, *Proc. ADASS 99*, in press (1999).

[68] Kunszt, P.Z., Szalay, A.S., Csabai, I., and Thakar, A.: The Indexing of the SDSS Science Archive, *Proc. ADASS 99*, in press (1999).

[69] Szalay, A.S., Kunszt, P. Thakar, A., and Gray, J.: Designing and Mining Multi-Terabyte Astronomy Archives: The Sloan Digital Sky Survey, Microsoft Technical Report, MS-TR-99-30, (1999).

[70] Szalay, A.S., Kunszt, P. Thakar, A., Gray, J., Slutz, D., and Brunner, R.J.: Designing and Mining Multi-Terabyte Astronomy Archives: The Sloan Digital Sky Survey, *SIGMOD 2000 Conference*, (2000).

[71] Thakar, A., Kunszt, P.Z. and Szalay, A.S.: Multi-threaded Query Agent and Engine for a Very Large Astronomical Database, *Proc. ADASS 99*, in press (1999).