

# Particle Physics Data Grid Collaboratory Pilot Tier 1/A Grid Site Authentication, Authorization and Accounting Proposal

## Principal Investigators:

Bob Cowles (Stanford Linear Accelerator Center)

Dane Skow (Fermi National Accelerator Laboratory)

## Contents

1 Introduction.....	1
2 Site work plans .....	2
2.1 Brookhaven National Laboratory.....	2
2.2 Fermi National Accelerator Laboratory.....	4
2.3 Lawrence Berkeley National Laboratory-NERSC .....	5
2.4 Stanford Linear Accelerator Center.....	7
2.5 Thomas Jefferson National Accelerator Facility.....	8
3 Participants.....	11
4 Relationship to other projects .....	11
5 Funding Requests: .....	11
6 Deliverables:.....	12
7 Management .....	12

## 1 Introduction

Research efforts in the Grid Community have developed an impressive model for globally distributed computing. The Particle Physics Data Grid (PPDG) Collaboratory Pilot project strives to take these efforts and to progress to the next level of early user production in a community that has not traditionally been tightly coupled with the Computer Science research community: High Energy and Nuclear Physics (HENP). Successful deployment will require significant investments in integrating both policies and mechanisms with existing infrastructure to accommodate this new approach. Of particular concern is the authentication, authorization and accounting infrastructure where the urgent demands of the hostile Internet environment and the presently increasing levels of threat and attacks has meant that computer security specialists at already short-staffed organizations have not yet had opportunity to evaluate the impact of the Grid Security Infrastructure (GSI) on their business processes and operations.

This proposal requests funds for an extension of the PPDG project to include efforts by computer security experts at HENP sites across the US to examine the impact on their

particular site of implementing GSI-based services. This effort will include evaluation of the architectural designs and their implications, as well as integration of some pilot grid-enabled services. The Globus team, as collaborators in PPDG and participating in this activity, will receive the results of these efforts at this early stage of development. The proposal should provide the opportunity to reduce the costs of future grid service deployment by providing documented lessons learned, and by suggesting improvements to developers and architects.

Since the nature of this effort is largely integration it is expected that much of the work will be done locally within the context of each site. However, we already note, beyond the common need for such effort, several common themes: Kerberos local infrastructure, a focus on mass storage resources as pilot services, etc. These common areas of interest will warrant inter-laboratory communication and technical sharing. We have created an archived mailing list<sup>1</sup> within the PPDG framework and have begun to use this to facilitate that communication, most notably for the creation of this proposal. This will also serve as a common channel for discussion with the Globus developers and other US HENP grid projects with which PPDG collaborates (GriPhyN and iVDGL), as well as our European, South American, and Asian colleagues.

The work plans for each of the participating sites are included below. In total, the request to DOE is for \$500K in funds to be equally divided among the 5 sites. The vast majority of this funding will go toward personnel costs to accomplish this work. As in many cases desirable individuals are already identified who may not be available after long delay, we request that approval to begin work be given by 15 April and that funds arrive by mid-May. The work and funding requests are based on this assumption.

## 2 Site work plans

### 2.1 Brookhaven National Laboratory

**Project responsibility:** Tom Throwe, Rich Baker

#### 2.1.1 Overview:

Our group at Brookhaven National Laboratory operates both the RHIC Computing Facility (RCF) and the US ATLAS Tier 1 Computing Facility (ACF). Both of these facilities operate a variety of authentication protocols including NIS, Kerberos (for AFS) and DCE (for HPSS). This project will focus on mapping a Grid X.509 certificate onto these local authentication protocols so that the Grid request is executed with the appropriate tokens and privileges while logging sufficient information to allow a meaningful security audit.

---

<sup>1</sup> Ppdg-siteaaa mailing list  
[Ppdg-siteaaa@ppdg.net](mailto:Ppdg-siteaaa@ppdg.net)  
<http://www.ppdg.net/mailman/listinfo/ppdg-siteaaa>

### 2.1.2 Tasks:

1. Authorization policy: Develop authorization policies for local resources. This will require using the X.509 certificate to determine the level of access that should be granted to a Grid request. Some requests will come from users without local accounts. These requests may be executed based on a trust relationship with the Certificate Authority that granted the certificate.
2. Unix account mapping: Map each Grid request to a local user account. For Grid requests that come from preexisting local users, these requests can be mapped onto the user's preexisting local account and this is already done using current Globus tools. The more interesting case is how to handle Grid requests from users who are not known in advance. This may require using a pool of accounts that can be mapped dynamically and recycled.
3. Kerberos: Ensure that Grid processes obtain required Kerberos tokens to allow appropriate AFS access.
4. HPSS: Ensure that Grid process can access the HPSS system for read/write as appropriate.
5. Security Audit: Develop necessary logging tools to ensure that actions caused by Grid requests can be traced back to a specific user.

### 2.1.3 Deliverables:

1. BNL authorization policy for local resources
2. Feedback to Certificate Authorities and Grid developers on what information is required in a certificate
3. Implement account mapping for preexisting site users
4. Prototype design of dynamic account mapping
5. Prototype design of mechanism to grant Kerberos tickets based on X.509 certificate
6. Prototype design of mechanism to grant HPSS (DCE) authorization based on X.509 certificate
7. Prototype logging tools for security audit purposes
8. Contribute to PPDG requirements documents

### 2.1.4 Resources:

- Request:
 

Funding is requested for one FTE for the duration of this project plus travel expenses to allow participation in relevant meetings.
- Contributed resources
 

The BNL computing facilities (RCF/ACF) will provide the computing environment and test platform including all standard hardware and software. In addition, this project will draw heavily on the experience and expertise of the existing staff.

## 2.2 Fermi National Accelerator Laboratory

**Project responsibility:** Dane Skow, Matt Crawford

### 2.2.1 Overview:

Fermilab has recently implemented a site authentication/authorization infrastructure based on Kerberos 5. We have developed a policy model of varying authentication requirements depending on the level of access being made available. When investigating use of GRID technologies we are concerned about the ability of individual sites to make appropriate authorization decisions for GRID services and what information is required. We propose to evaluate the levels of access needed for useful collaboration, the authorization information needed and the Globus Security Infrastructure (GSI) for implementing these. We will deploy two pilot services as part of this evaluation.

### 2.2.2 Tasks:

1. Evaluate impact of GRID Reference Architecture Requirements on Grid Resource providers in context of D0 and CMS current and documented planned computing systems.
2. Deploy necessary infrastructure and integration in support of a Grid Mass Storage Resource based on the FNAL Enstore mass storage system. Evaluate the software stack to determine whether existing hooks are sufficient to accommodate site policy restrictions.
3. Deploy a production class infrastructure for the appropriate translation of FNAL Kerberos tickets into X.509 certificates acceptable to the DOE Science Grid. Identify the issues needing resolution for a full, bi-directional mutual acceptance.

### 2.2.3 Deliverables:

1. The deliverable from this task will be a documented issues analysis and work estimate to each of the major development groups: Globus, D0, and CMS. This work will require close collaborative discussions with all 4 major parties (FNAL Site, D0, CMS, Globus) The current goal is to have the initial draft of this analysis complete in time for discussion at the July 2002 GGF/PPDG meetings.
2. Deliverable will be the presence on the GRID of a fully compliant Mass Storage Resource for use in further development. This will be a joint deliverable with work under the SRM project run by Don Petravick. The specific deliverable for this project is the necessary infrastructure support.
3. Deliverable will be an operational translation service for FNAL Kerberos ticket to certificates acceptable to the DOE Science Grid. Initial tests of operation of this service will be: short lived certificates for use in Web authentication, and generation of acceptable proxy certificates from Kerberos authenticated clients.

## 2.2.4 Resources:

- Request:

Funds sufficient to cover 1 FTE contract programmer for the period from 15 April until 30 September 2002. A suitable candidate already familiar with FNAL site has been identified and is coming to the end of current contract. An alternate plan is to utilize personnel on staff and backfill effort with contract support.

Travel funds sufficient for the required collaboration efforts cited above and the presentation of the results of this work at an appropriate open forum (e.g. GGF5).

- Contributed resources

All hardware necessary to implement above tasks.

All admin, monitoring and site services needed for their operation.

Expert consultants on details of site policy and current infrastructure operation.

Supplemental effort as required beyond the 1 FTE programmer to accomplish the goals above.

## 2.3 Lawrence Berkeley National Laboratory-NERSC

**Project responsibility:** Doug Olson

### 2.3.1 Overview:

The test-bed for this project at LBNL-NERSC will be authenticating and authorizing grid users to PDSF and HPSS. PDSF uses NIS for authentication and HPSS uses DCE/Kerberos internally but user access is from outside the DCE domain. Accounts for these and other NERSC systems are managed with a centralized user management system called NIM.

NERSC has a project, lead by Steve Chan, beginning in FY2002 to evaluate and phase deployment of grid services to the entire NERSC community over the next few years. The scope of work described in this proposal is identified as those tasks which go beyond the scope of general NERSC grid effort and specifically address the needs of the PPDG (HENP) community. STAR and ATLAS are the two experiments participating in PPDG which are also users of NERSC resources.

Issues that we need to address:

- The current model of user access is for all users to be registered in NIM and have individual accounts. The broad scope of the HENP collaborations (VO's) appear to require outside services defining these VO's, such as the EU DataGrid ldap VO server from INFN or the Globus Community Authorization Service currently under development. We need to understand the interface and impact of an external (to NERSC) VO server on NERSC policies and procedures. Will we still need separate accounts for individuals accessing resources from the grid or is there an acceptable

mechanism to share or lease local accounts? What requirements are there on traceability, accountability and usage accounting? Is there a better model?

- If a user belongs to several VO's, how does he communicate under which VO his task should run?
- What is required from the application – for example is it acceptable for a leased user account to write a file to a disk pool? To Tape? What further is required of the application or the basic infrastructure to permit determination of to whom a file belongs, or who is using the quota for a group?
- What modifications would be necessary to the NIM accounting system if shared or leased accounts are used for grid access?
- Do we need a shared acceptable use policy for grid users?

### 2.3.2 Tasks:

The tasks listed below have some interactions but can all proceed in parallel.

1. Evaluate external VO server, initially the INFN Idap server and then the Globus CAS when available or appropriate. This will be done in cooperation with Conrad Steenberg who is also testing the INFN Idap server at Caltech, and with the Globus people working on CAS.
2. Set up and run a local Idap server for local authentication/authorization. Investigate procedures for synchronizing or applying updates to local Idap server from external VO server. Investigate how a myproxy server integrates with this structure.
3. Compile, evaluate, understand and document requirements. Contribute to STAR, ATLAS, PPDG requirements as applied to NERSC. Document technical requirements of NERSC systems and procedures to accommodate PPDG. Recommend any necessary changes to NERSC Acceptable Use Policy or other policies.
4. Test GSI enabled ssh.

### 2.3.3 Deliverables:

1. NERSC requirements document
2. Contribute to the PPDG requirements document that defines the PPDG model for authorization, which could be based on CAS, could be based on the INFN VO LDAP service, or some other mechanism – this definition is part of a deliverable
3. Demonstrate a working GSI-based basic infrastructure that includes authentication/authorization to local services via a local Idap server.:
  - a. Access LSF batch services on PDSF from globus job submission.
  - b. Access HPSS from grid access to PDSF
4. Analysis document, or deployment of GSI enabled ssh, depends on test results.

### 2.3.4 Resources:

- Request: 1 FTE + travel
- Contributed resources
  - PDSF development and hardware
  - Experts on NERSC systems, PDSF, HPSS, NIM for consultation
  - myproxy server, ldap server
  - Doug Olson (requirements, coordination)

## 2.4 Stanford Linear Accelerator Center

**Project responsibility:** Bob Cowles

### 2.4.1 Overview:

The test-bed for this project at SLAC will be authenticating and authorizing Grid users for use of high performance file transfer to move simulation and experimental data associated with the BaBar experiment.

Issues that we need to address:

- What policies are appropriate for authentication in HENP collaborations where the full weight of X.509 certificates may not be required? Work will include an examination of the minimal assumptions HENP labs might make in performing authentication and what information controls are necessary to balance various levels of risk.
- What policy changes need to be implemented to provide various levels of computing resources to appropriately authorized members of a grid enabled collaboration? Work in this area will require review of DoE policies, SLAC policies, development of models for useful levels of computing resources and evaluating the risk based on types of security controls.
- What policies and mechanisms need to be in place to reduce the risk of inter-lab trust relationships to an acceptable level?
- What is the mechanism for local users to obtain access to even locally enabled grid applications and resources?

### 2.4.2 Tasks:

1. Review relevant policies and recommend changes.
2. Review authentication architecture for HENP and provide recommendations.
3. Work with the BaBar Collaboration to arrive at a suitable set of levels for computing resource availability based on adequate mitigation of risk factors.
4. Install suitable site security infrastructure (e. g. Kerberos 5, GSI gateways systems) to provide for authentication of grid users and use of local grid resources by local users.

5. Test the use of grid-enabled security mechanisms in conjunction with high performance file transfer software for movement of simulation and experimental data.

### 2.4.3 Deliverables:

1. Revised policies for inclusion SLAC business processes dealing with levels of computer resource usage, and account activation and termination.
2. SLAC requirements document for authentication in HENP collaborations.
3. Operational procedures for BaBar collaborators to be enrolled in and authorized for using resources associated with SLAC and the BaBar Tier A centers.
4. Initial definition of appropriate set of levels for computer resource access.
5. Contribute to the PPDG requirements document for GSI authentication and authorization services.

### 2.4.4 Resources:

- Request:
  - Approximately 1 FTE level of effort between 15 April 2002 and 30 September 2002.
  - Travel funds for coordination meetings with other labs and discussion of proposals for authentication at GGF5.
- Contributed resources
  - All required hardware, software and network resources.
  - Experienced/knowledgeable staff, expert in: project management, requirements, grid architecture, authentication and authorization, PKI, Kerberos, GSI, and high performance file transfer software; including Chuck Boeheim, Gary Buhrmaster, Bob Cowles, Andy Hanushevsky, Adil Hasan and Doug Smith.

## 2.5 Thomas Jefferson National Accelerator Facility

**Project responsibility:** Ian Bird, Robert Lukens

### 2.5.1 Overview:

The initial test-bed for this project at Jefferson Lab (JLAB) will be authentication and authorization of grid users to enable the use of the SRM interface to the JLAB mass storage system – JASMine. A secondary test would be high-speed file transfer, either initiated by the SRM system or in a stand-alone mode. JLAB is somewhat unique in that it does not use Kerberos at all (including AFS), but uses NIS for authentication, making extensive use of Unix group membership and NIS netgroups for authorization. Users may belong to many groups and netgroups. We intend to investigate the relative merits of the CAS and INFN LDAP models for user authentication mapping.

Issues that we need to address:

- What is the appropriate authorization mechanism – do we (JLAB) require grid users to be mapped to real local users, or is (for example) a leased pool of VO users acceptable, and if so what requirements are there on traceability and accountability? Is there a better model?
- If a user belongs to several VO's, how does he communicate under which VO his task should run?
- What is required from the application – for example is it acceptable for a leased user account to write a file to a disk pool? To Tape? What further is required of the application or the basic infrastructure to permit determination of to whom a file belongs, or who is using the quota for a group?
- At JLAB all of the infrastructure will be provided by the Computer Center, including setting up the VO for the experiments. For this reason we require a straightforward mapping between Unix group membership and the authorization mechanism (but this could equally be a simple registration service).

### **2.5.2 Tasks:**

1. Review relevant security and usage policies and recommend changes,
2. Review authentication models based on the Globus Community Authorization Service and the INFN VO ldap service in the context of the JLAB environment
3. Deploy infrastructure to support a grid-enabled mass storage resource based on the JLAB mass storage system, JASMine, with an SRM interface
4. Demonstrate a working GSI-based basic infrastructure that includes the end-to-end process of user registration, certificate issuance and use that is integrated into the JLAB user validation, auditing and account granting mechanism.

### **2.5.3 Deliverables:**

1. JLAB requirements document for grid user authentication and authorization,
2. Revised policies and procedures (CSPP) to include grid users and associated issues,
3. Contribute to the PPDG requirements document that defines the PPDG model for authorization, authentication, and accounting,
4. A grid-enabled pilot service of an SRM interface to JASMine that:
  - Accepts GSI certificates as authentication and integrates that with the standard NIS user database
  - Grants authorization to an authenticated user based on the agreed PPDG model defined above
  - Accounts for usage of the resource by user, group, etc as necessary, and provides sufficient usage audit trail to satisfy the requirements of application and site security.

#### **2.5.4 Resources:**

- Request:

Provide contract programmer effort during the period April – September 2002 to work directly on this project or to back-fill diversion of existing effort, and to cover travel expenses associated with the project.
- Contributed resources

Ian Bird (project management, requirements), Bryan Hess (implementation over SRM interface), Robert Lukens (integration with NIS, requirements and policies). All hardware and resources to test and deploy the grid-enabled mass storage system, including systems and software support.

### 3 Participants

Listed below are the people who are expected to participate at varying levels in this project. It is a characteristic of this effort that there are many people at the sites with expertise in the various aspects of authentication, authorization and accounting who will contribute to the requirements and policy analysis. It is clear that, for the funding requested, there will be fewer people who spend a significant amount of time and the people listed below will not all be funded by this proposal.

<b>ANL:</b>	Von Welch
<b>BNL:</b>	Rich Baker, Tom Throwe, Jason Smith, Dantong Yu, Razvan Popescu, Shigeki Misawa
<b>Caltech:</b>	Conrad Steenberg
<b>FNAL:</b>	Dane Skow, Matt Crawford, Rich Wellner, Igor Mandrichenko, Don Petravick, Gabrielle Garzoglio, Sinisa Vesili, Igor Terekhov, Lothar Bauerdick, Ruth Pordes, Lee Lueking
<b>JLAB:</b>	Ian Bird, Robert Lukens, Bryan Hess, Andy Kowalski
<b>LBNL:</b>	Doug Olson, Shane Canon, Steve Chan, Iwona Sakrejda, Mary Thompson
<b>SLAC:</b>	Bob Cowles, Chuck Boeheim, Gary Buhrmaster, Andy Hanushevsky, Adil Hasan, Douglas Smith

### 4 Relationship to other projects

- The participants in this activity include representatives of the following SciDAC projects: DOE Science Grid
- Security and Policy for Group Collaboration
- Distributed Security Architectures

Since most of the HENP experiments participating in PPDG are collaborations extending beyond the U.S., particularly to Europe, it is important to interact with groups working on the same authorization and accounting issues in Europe for these same experiments. The EU funded DataTAG Project ([www.datatag.org](http://www.datatag.org)) Work Package 4 group headed by Roberto Cecchini from INFN is our primary contact for this interaction.

### 5 Funding Requests:

This table lists the funding request for each laboratory.

<b>Site</b>	<b>Request</b>
BNL	\$100K for contract labor and travel
FNAL	\$100K for contract labor and travel
LBNL	\$100K for contract labor and travel
SLAC	\$100K for contract labor and travel
TJNAF	\$100K for contract labor and travel
<b>Total</b>	<b>\$500K</b>

## 6 Deliverables:

Funding of this project allows the sites to produce the following deliverables:

- Grid resources available at each site and operating in pilot or production mode – these are the test-bed projects used to validate the requirements at each site.
- Site requirements list – each site will list the requirements they fed back to the developers in the course of performing the work on this project.
- Summary of site policies changed – each site will summarize the local policies that needed to be created or updated.
- Early draft of PPDG requirements document – the combined experience of the labs will provide a view of common problems that need solutions to be provided in the GSI core.
- Progress report at GGF5 – early experience with the test-bed projects provides valuable feedback to developers and “lessons learned” for future grid implementations.
- Wrap-up meeting in September – review progress and future directions

## 7 Management

The work proposed here will be included as an extension to the PPDG scope of work and carried out as one more of several existing project activities<sup>2</sup> within PPDG. Bob Cowles and Dane Skow will be the PPDG team leaders for this work and will be responsible for reporting on this work for the PPDG quarterly reports as well as organizing meetings and teleconferences for this activity. Doug Olson will assist Bob and Dane in compiling documentation of the deliverables.

The PPDG executive team (Ruth Pordes, Doug Olson, Miron Livny) will track the progress of this activity and the PPDG Steering Committee assess the deliverables and help plan the future directions.

---

<sup>2</sup> <http://www.ppdg.net/pa/ppdg-pa/projects.htm>