



# CEDPS: Center for Enabling Distributed Petascale Science



## A DOE SciDAC Center for Enabling Technology

Annual Progress Report for the Period  
October 1, 2006, through October 1, 2007

### Executive Summary

This report summarizes ongoing collaborative work, overall progress, and presentations carried out by CEDPS during the period October 1, 2006, through October 1, 2007. Significant progress has been made in each of the three CEDPS subareas: Data, Services, and Troubleshooting, with a particular focus on development of code prototypes and initial collaborations with the Open Science Grid (OSG) and Earth System Grid (ESG). In our first year we also published five refereed papers, three posters, and five design documents and gave over 20 presentations of CEDPS material. In the next year, we plan significant work to expand the number of users of CEDPS technology.

The Data area provided key technologies for petascale science in DOE in terms of policy-driven, end-to-end placement of large-scale data. One focus of the Data area was the Managed Object Placement Service (MOPS), which provides reliable data movement with resource management. MOPS underpins the higher-level Data Placement Service (DPS), which will offer an end-to-end service to orchestrate the placement of large-scale data, a particularly acute problem for end-users with large data sets, such as the high energy physics and climate application communities, the Spallation Neutron Source at ORNL, and the Advanced Photon Source at Argonne. DPS will improve the performance of DOE applications by reducing the time to stage data in and out of computational resources. Prototypes of both MOPS and DPS are available for alpha use.

The Services area extended work in two directions. Front-end work involved extensions of pyGridware and CLWrapper, two Python tools that allow legacy applications to run in distributed environments. Back-end work, in close collaboration with the high energy physics applications STAR and Alice, resulted in the deployment of a proof-of-concept infrastructure to adapt virtualization technologies as part of the Workspace service.

The Troubleshooting area completed a logging “best practices” document, which was incorporated into many Globus software tools. The Troubleshooting group also designed and began deployment of a centralized logging service based on syslog-ng in collaboration with the Open Science Grid, and the group designed the Trigger service functionality in support of the Earth System Grid.

More details about the CEDPS project are available on the website at <http://www.cedps.net>.



## Contents

1	Overview.....	4
2	Data Area .....	4
2.1	Year 1 Accomplishments.....	4
2.2	Year 2 and Future.....	5
3	Services Area .....	6
3.1	Year 1 Accomplishments.....	6
3.2	Year 2 and Future.....	7
4	Troubleshooting Area .....	8
4.1	Year 1 Accomplishments.....	8
4.2	Year 2 and Future.....	9
5	Application and Center Outreach.....	9
5.1	ESG.....	9
5.2	OSG.....	10
5.3	Centers .....	10
5.4	Additional Collaboration in Year 1.....	11
5.5	Outreach in Year 2.....	11
6	Metrics for Success for Year 2.....	12
7	Management and Staffing.....	13
8	Publications and Presentations.....	13
8.1	Refereed Papers .....	14
8.2	Presentations .....	14
8.3	Posters .....	16
8.4	CEDPS Design Documents.....	16
9	Software Releases .....	16
	Appendix 1: Milestones Year 1 .....	18
	Appendix 2: Milestones Year 2 .....	24
	Data Area Year 2 Milestones.....	24
	Data Area Changes from Original Plan .....	25
	Service Area Year 2 Milestones .....	25
	Services Area Changes from Original Plan .....	26

Troubleshooting Area Year 2 Milestones .....	26
Troubleshooting Area Changes from Original Plan .....	27
Appendix 3: Partner Listing.....	28

# 1 Overview

The Center for Enabling Distributed Petascale Science (CEDPS) is producing technical innovations for (1) rapid and dependable *data management* within a distributed high-performance environment, (2) convenient construction of *scalable science services* that enable reliable, high-performance processing of computation and data analysis requests from many remote clients, and (3) *troubleshooting* of these and other related ultra-high-performance distributed activities, from the perspectives of both performance and functionality. CEDPS activities have been defined in consultation with leading DOE application groups and will be deployed, applied, and evaluated in collaboration with major DOE projects in climate, high energy and nuclear physics, combustion, astrophysics, fusion, biology, and other sciences.

The project team brings together researchers and technologists from three DOE laboratories (Argonne National Laboratory, Fermi National Accelerator Laboratory, and Lawrence Berkeley National Laboratory) and two universities (the University of Southern California and the University of Wisconsin).

In the first year we focused on putting in place the management structure for this five-site center; understanding current technology and interactions; working with existing collaborations to understand user requirements, establishing initial outreach to new communities, and extending prototypes to releases for use by additional alpha testers. In the second year we will focus on working with additional application communities and SciDAC centers and on developing and deploying production code.

## 2 Data Area

The Data area is working to provide key technologies for petascale science in DOE for policy-driven, end-to-end placement of large-scale data sets. One focus of the Data area is the Managed Object Placement Service (MOPS), which will provide reliable data movement with resource management. MOPS underpins the higher-level Data Placement Service (DPS), which will offer a policy-driven, end-to-end service to orchestrate the placement of large-scale data and improve the performance of DOE applications by reducing the time to stage data in and out of computational resources.

### 2.1 Year 1 Accomplishments

In the first year, GridFTP data transfer functionality was extended and integrated with the NeST storage system management to provide efficient resource management inside the object mover. The GridFTP team at Argonne and the NeST group at Wisconsin worked together to create a baseline architecture and agree on a set of protocols for a MOPS service. The GridFTP server was modified by the team at Argonne to support local resource management—in particular, to better manage network connections and memory usage. The NeST team at Wisconsin modified the existing NeST implementation to integrate it into Condor and to make it suitable for integration into the MOPS implementation. Argonne and Wisconsin combined the two components into an integrated MOPS implementation, with a release scheduled for October 1, 2007, for use by alpha users.

The Data Placement Service offers a policy-driven end-to-end service to orchestrate the placement of large-scale data sets and decide where to stage and replicate data sets to improve the performance of applications and workflows. The ISI team prepared a design document [33] for team comment; this document formed the basis of the ongoing development work. The ISI and Wisconsin teams collaborated on a study of placement service options that evaluated the performance of a data placement service in combination with the Pegasus workflow management system. The results of this study were published in a Grid 2007 conference paper [5]. In addition, the ISI team modified an existing Grid simulator to experiment with different data placement algorithms in conjunction with a workflow manager. A Data Placement Service 1.0 release is scheduled for October 1, 2007.

Additional work in support of DOE applications, especially CMS and APS users, was completed, primarily to improve the performance of GridFTP for large-scale file transfers. The Argonne GridFTP group implemented an optimization to improve the performance of transfers of many small files. A mechanism for streamlining multiple small files into an open channel all but eliminates the file transfer channel startup overhead and thereby enables bundles of small files (> ~ 50 KB) to transfer as fast as single large files. The Argonne team also developed a prototype of a GridFTP service that allows the use of SSH tunneling for the control channel instead of the standard GSI approach, thereby easing the burden for some sites in using GridFTP. Both mechanisms are available in the Globus Toolkit 4.1.2 development release. The Fermi team implemented check summing in its DCache GridFTP server and contributed a Java client that supports check summing for file transfers. This will be available in the GridFTP GT 4.1.3 release.

The Data area continues to work closely with various applications groups. Extensive discussions (through site visits as well as at collaborative meetings) were held with the STAR project, OSG, ESG, and others to understand their requirements and make sure that the MOPS and Placement service work will contribute to improving the performance of these applications.

## **2.2 Year 2 and Future**

Our work in year 2 will extend the functionality of the MOPS and data placement services begun in year 1. A significant focus work will be outreach to application communities, helping them to deploy and evaluate CEDPS data services and getting feedback on features and changes needed by applications.

For the Managed Object Placemen Service, we plan to extend the functionality available in release 1.0 to provide more complete integration of the GridFTP and NeST code bases, including improving the functionality for storage and connection management. We also will work closely with application communities to deploy and evaluate MOPS version 1.0. In addition, we plan to deliver a MOPS 2.0 release that includes additional optimizations such as LOSF optimizations in GridFTP, dynamic resource registration, non-striped connection management, and basic troubleshooting instrumentation. This release will also include an enhanced interface for exposing resources for use by higher level services.

For the Data Placement Service, we will extend the work of year 1 to provide additional functionality. DPS accepts client requests, places outstanding data transfer operations on priority queues, and issues data transfer operations efficiently. We will implement additional logic to accept hints about data placement requirements from clients such as workflow managers. We will also add functionality to manage competing placement requests, for example, those related to multiple workflows performing data staging operations in parallel as well as data distribution operations required by virtual organization. We will evaluate a variety of algorithms for data placement using simulation and experimentation in cooperation with application communities to identify which algorithms are most efficient. Release of version 2.0 of the placement service is expected in October 2008, with one or more intermediate versions likely earlier in the year.

We will also investigate the practicality of incorporating mechanisms for quality of service into storage management and data placement services. The goal is to determine whether using contracts or agreements could improve storage performance and reliability.

### **3 Services Area**

The Services area work comprises two major components. The front-end, developed at LBNL, generates Web service interfaces turning existing applications into network services. The back-end, developed at Argonne, provides services for flexibly provisioning resources to ensure on-demand service deployment.

#### **3.1 Year 1 Accomplishments**

In the first year the front-end work focused on providing a set of Python tools, called Command Line Wrapper (or CLWrapper) tools, that automatically generate Web service interfaces for applications. CLWrapper works by generating (1) stubs for the server side that call out to the actual application and allow the application developer to make the code available as Web service and (2) remote client code that invokes remote services. Services support was enabled for basic GridFTP actions, allowing easy data transfers from a Web service environment and improved data management for Web-based services. The basic support needed to add a pyGlobus-style GRAM interface was also added, as well as delegation for credential management.

In addition to work on the core functionality, front-end usability work was done to provide a portal generation capability (portals being the interface of choice for many application groups), allowing existing Grid services to be easily adapted into a Web user interface. Clients were developed for both Java and Python.

The back-end work focused on developing services for flexible provisioning of resources to ensure on-demand service deployment. Workspaces, execution environments that can be dynamically deployed in the Grid, can be deployed and managed by the Workspace Service (<http://workspace.globus.org>), which implements workspaces as virtual machines (VMs). Work in this area began by prototyping basic provisioning for the STAR application on the Teraport cluster at the University of Chicago, as demonstrated at SC'07 [27]. As the work progressed, we identified missing capabilities, engaged in design and development efforts to provide them, produced workspace releases

and assisted application groups in their deployment and use, as well as expanding our application end-users to include Alice. For STAR we developed multiple capabilities, including partition-based image management and methods for deployment and management of multiple virtual machines as ensembles (virtual clusters).

To give scientific communities easy access to relevant VM images, we initiated the development of the Workspace Service Marketplace (<http://workspace.globus.org/library-01.html>), which contains a library of workspace images for scientific applications, with STAR, Alice, ATLAS, and example codes currently available.

We also developed a workspace back-end capable of leasing resources from the Amazon's Elastic Compute Cloud (EC2) Service to enable VM-based runs on large numbers of resources enabling production runs for STAR. Multiple trial runs took place in August, and in early September 2007 our efforts resulted in the first 100-node VM-based STAR production run.

Science enabled by work in the Services area has been presented by both ATLAS and STAR at CHEP 2007: "Deploying HEP Applications Using Xen and Globus Virtual Workspaces," by Agrawal et al.; and "STAR and Virtual machines," by Lauret and Hajdu.

### **3.2 Year 2 and Future**

Because of changes in personnel and management, CEDPS management decided to refocus the front-end technology work away from the Python-based tools pyGlobus and pyGridware and instead extend the Java-based RAVi (Remote Application Virtualization Infrastructure) tool. These tools have similar current functionality, but RAVi offers several benefits over the pyGlobus/pyGridware tools:

- Authorization framework: PyGlobus does not have the extensive GT4 authorization framework, while that RAVi gets for free from its use of the Introduce tool and GT4 core.
- Strong typing: PyGlobus used only primitive types where as RAVi is integrated with Global model exchange to create strongly typed services.
- Hosting container maturity: Python as a language is still catching up with Java in terms of hosting containers and their stability. In general, Java as a platform has far richer tooling than does Python.

RAVi is currently in use by several end-users in the bioinformatics caGRID project, and we felt this use might encourage some of the SciDAC applications in that space as well.

The focus of the front-end efforts will be first and foremost to work with SciDAC users, especially those with legacy codes for which this work is well suited. RAVi will need extensions to work with the Delegation service (and user credentials) because it is currently using community certificates; some work with data staging also may be needed.

The focus of the back-end efforts will be on refining the capabilities developed in year 1, providing support and requirement-driven feature development within the existing framework, and improving usability of the provided capabilities. We plan to continue our work with the STAR community, focusing on policy and deployment mechanisms for

virtual clusters (and access to Amazon EC2) in collaboration with OSG and STAR scientists. We will also develop and make available more VM images as well as assist application groups in developing their own images. This work, combined with LBNL work on providing easy-to-use portal interfaces, will greatly simplify VM-based computing for DOE communities.

## **4 Troubleshooting Area**

The CEDPS Troubleshooting area is developing tools to assist in the automated failure detection for Grid services, the end-to-end failure detection and analysis of user jobs (including log file discovery and management), and automated performance degradation detection for both Grid services and user jobs. Our approach has two parts: (1) defining better logging practices, ways to collect logs for ease of use, and tools to do better troubleshooting with access to the logs, and (2) providing alerts on errors and performance problems to notify of failures before they are discovered by users.

### **4.1 Year 1 Accomplishments**

The Troubleshooting effort for year 1 of the CEDPS project resulted in several key accomplishments. Of particular note was the definition and implementation of a Logging Best Practices document [31] for distributed middleware and application logging. This common log format and methodology document facilitates troubleshooting and makes it possible to write higher-level anomaly detection tools. Preparation of the document involved meetings and discussion with many members of the community, including key people from the Open Science Grid, EGEE, Open Grid Forum, STAR project, and Globus project. All Globus components are being modified to follow this logging practice, which will be included in the upcoming GT 4.1.3 development release, as are other middleware components such as CLWrapper.

Another major accomplishment was the design of a central log collection system based on syslog-ng. We worked closely with members of the OSG to design and package a syslog-ng configuration for VDT and are currently deploying this on the OSG. Having OSG logs in a single location will make it much easier to troubleshoot Grid jobs on OSG and will be useful for security analysis.

We designed and wrote prototype tools that parse existing logs and load them into a SQL database. These tools make it possible to query the database for interesting log information.

We also designed and implemented a general-purpose log summarization library. With this library a developer can instrument code to perform detailed logging but to output only periodic summaries, thereby minimizing perturbation of the program (see [4] for details). We are currently instrumenting the GridFTP client and server codes with this summarization library, which will allow us to tell the user or system administrator if a transfer bottleneck is the read disk, write disk, or network.

In addition to the logging work, we continued work with ESG to extend the current deployment of the Globus MDS4 Trigger service, which provides alerts on errors. In the first year of CEDPS we extended the number of information sources available to the

Trigger service and reworked the underpinnings to provide additional functionality in terms of how many triggers are set and their notification methods.

## **4.2 Year 2 and Future**

Our main goal for FY2008 is to deploy a useful central log database for OSG and other Grid projects. This will include extensions for OSG for security and a database interface for additional higher-level services for anomaly detection and understanding performance baselines. We will also package and document the parsing tools. In addition, we are working with NERSC to design and deploy a centralized log collection and alert framework based on the work with OSG.

Other work will include additional implementations of the logging best practices, including Condor, and extensions for generation and propagation of globally unique IDs

Trigger service work with ESG will include a broader deployment, user-centric tools, and development of action scripts to take actions based on results from the anomaly detection tools. We also plan to deploy this service on at least one additional application platform.

## **5 Application and Center Outreach**

In the first year of CEDPS, application and outreach work focused on strengthening previous ties, especially with the SciDAC CET Earth System Grid project and the SciDAC Sustaining and Extending the Open Science Grid: Science Innovation on a Petascale Nationwide Facility. We also began contacting additional partners, with STAR, ORNL, and NERSC being the largest of these.

OSG, ESG, and CEDPS colleagues are working together to define areas of interaction and complementary work at a higher level in order to strengthen overall distributed science work within SciDAC. Face-to-face meetings with the three groups took place in February, March, and June; and informal conversations between pairs occurred more frequently. The meetings have provided the basis for ensuring that user requirements of these communities are properly aligned with upcoming CEDPS deliverables.

We have also delivered several tutorials in order to present CEDPS work more broadly. The largest CEDPS-specific of these was at the June SciDAC meeting in Boston, where a general overview was given, and both the Data and Troubleshooting groups detailed their current and future work. CEDPS material is also part of the current general Globus outreach tutorials.

### **5.1 ESG**

To address the critical issues of ensuring reliable delivery of climate data to the climate community, we worked with ESG to identify opportunities for improved data delivery and problem detection in the ESG services.

The Data area presented CEDPS plans for the Managed Object Placement Service and higher-level placement services.

The Troubleshooting area continued ongoing work with ESG in the extension and deployment of the Trigger service, which performs warnings on errors. This service was improved to address scalability concerns for ESG, generalized from specific ESG services to be more broadly applicable, and extended to allow ESG users to receive alerts as well as system administrators.

## **5.2 OSG**

Data work with OSG includes contacts with both end-user applications and the infrastructure team. In terms of end-user applications, the Data area continued working with several HEP applications (especially CMS and ATLAS) that use Globus version 2 data tools to evaluate the use of current and future tools for data management. Work with these partners was instrumental in motivating GridFTP improvements.

The Services area worked with the HEP STAR and Alice applications on developing workspaces and services allowing authorized clients to provision a platform ready to execute the application. The workspace-based approach is attractive to these communities because the complexity of installation and maintenance of their applications currently prevents the researchers from leveraging additional resources. We are continuing our interaction with the STAR experiment to define conditions that must be met to bring the workspace-based provisioning model into production.

A two-day meeting was arranged with the STAR application, primarily Jerome Laurent and David Alexander from Tech-X, who have an SBIR grant together. In a meeting with the Troubleshooting group, STAR showed strong interest in the best practices approach to logging and suggested a follow-on discussion in six months to discuss next steps. In a full-day meeting with the Services area, we identified the next steps that would allow STAR to use the Workspace Services later in the year: definition of an image, virtual cluster organization, user interface concerns, and ways to further work with OSG.

The Troubleshooting area has been coordinating with the OSG infrastructure group to deploy a centralized logging system first on the OSG Validation TestBed and then on the Integration TestBed, with plans for a full rollout to follow. Troubleshooting has also been working with the NERSC center to discuss a central log file collector for NERSC.

## **5.3 Centers**

Significant effort was spent with ORNL to better understand their user requirements, especially with respect to the newly operational SNS. Jennifer Schopf and Dan Fraser spent two days visiting ORNL in February, presenting CEDPS work and meeting with applications groups. Follow-on contacts led to additional discussion of possible collaboration but have been delayed into the second year because of ORNL staffing issues.

We are working with NERSC to help design and possibly deploy a central log collection facility based on syslog-ng. We are also discussing the possibility of deploying our log database and anomaly detection tools and using NERSC logs to help test our algorithms.

We are also working with the Advanced Photon Source at Argonne. In particular on tomographic experiments, data is collected in groups of small files typically on the order of 2000 8 GB files per shot, every 15-20 minutes. The LOSF mechanism for GridFTP has proven extremely useful for transferring this data in near real time to enable distributed processing capabilities.

#### **5.4 Additional Collaboration in Year 1**

Several smaller collaborations also were begun, in many cases through discussions initiated at the SciDAC PI meeting in Atlanta. The Data area began exploring a collaboration with the Petascale Data Storage Institute to see whether CEDPS tools can provide mechanisms for file systems to convey data using GridFTP servers for better performance, especially with how the data is mapped on disk, a feature of MOPS. The Data area has also begun discussions with the Performance Engineering Research Institute to help identify possible application users with performance problems in the data space.

The Services group met with the fusion software development community at Tech-X, who are exploring the use of CEDPS tools to expose the new fusion code as a Grid service, including making their stand-alone Fortran code into a network-accessible service using standard protocols. Minor variation in compiler/library versions can cause significant differences in the results produced by the applications.

#### **5.5 Outreach in Year 2**

In year 2 we will focus on attracting additional end users. This work will involve increased interactions with the current centers (OSG and NERSC) in order to discuss our tools with their end users. We also envision six site visits to SciDAC centers, with representation from each of the three CEDPS areas to better understand and encourage software uptake. This work will be in addition to the following area-specific outreach efforts .

The Data area plans to explore a possible collaboration with the combustion application group based at LBNL. In addition, the Data area scientists will work with current users of the Globus Data Replication Service to see whether these users have additional requirements for higher-level data placement services.

The Services area will meet with ESG to explore exposing several of their major climate codes as Grid services using pyGridWare. Since the climate community is moving more and more to Python as their standard scripting language, this strategy will allow them access to all of their major codes from their preferred environment. In addition, Workspace service work will be extended for ALICE and CCSM.

The Troubleshooting area plans to start working with STAR and other OSG users, as well as the Scientific Data Management Center for Enabling Technologies in the next year.

## 6 Metrics for Success for Year 2

In addition to the milestones listed for each area (in Appendix 1), the overall project will be evaluated by the following metrics of success.

1. *How well is the outreach to application groups and compute facilities such as OSG and NERSC functioning as a mechanism for two-way exchange of information—both in letting users know about the new functionality offered in CEDPS software and in providing a set of requirements drivers for development?*

This will be evaluated by requesting feedback from each of the named applications groups for the midterm and annual reports and by evaluating contacts made to new user groups. Any tutorials or outreach meetings will also be listed in the reports.

2. *How well are CEDPS software and tools enabling SciDAC users to overcome the barriers of petascale science?*

The most common way to measure this work is the traditional count of papers produced by the science end-users after deployment of the software. We will also report directly from users on new functionality or approaches enabled by the CEDPS tools.

3. *How well are the new data tools (MOPS, DPS) increasing the performance and usability of the SciDAC applications communities?*

This question will be answered by taking performance measurements throughout the projects work with applications groups and evaluating improvement in terms of number of files managed, transfer times, number of replicas used, number of users on the system, and user feedback in terms of response time and usability.

4. *How many applications are successfully using the Services tools to adapt their existing applications into network services?*

This will be measured through a count of deployments and through testimonials from user communities.

5. *How much has the CEDPS troubleshooting infrastructure improved job throughput and file transfer reliability and performance for a set of targeted applications communities (starting with OSG and ESG)?*

The key to this metric will be evaluating failure behavior both before and after CEPDS software is used. To date, we have been able to reliably establish metrics only after deployment has occurred, so this value may show improvement with updates, as opposed to that from a pre-CEDPS baseline.

6. *How well is CEDPS enabling the scalability of applications on the DOE leadership-class facilities and other resources?*

This will be measured in part by the number of deployments of CEDPS tools across SciDAC applications communities, reported yearly and in part by the number of interactions and deployments of software on selected environments.

## 7 Management and Staffing

Management of CEDPS is the overall responsibility of the principal investigator, Ian Foster, and the project director, Jennifer Schopf, both at Argonne. The project director is responsible for general project organization, tracking of cross-over issues, and moderating of priority discussions between application requests and technologists. Jointly with the PI, the project director coordinates general project outreach and direction.

Each of the three focus areas has an area lead. In year 1 these were Ann Chervenak, ISI, for Data; Kate Keahey, Argonne, for Services; and Brian Tierney, LBNL, for Troubleshooting. The area lead for Services is changing in year 2. Each area lead is responsible for milestones, vision, and coordination within the area.

Full project meetings took place at Argonne on October 25–26, 2006, and March 12–13, 2007, at USC. In addition, each area had at least one face-to-face meeting during the summer of 2007. An annual report, initially the management plan, was submitted on October 1, 2007, and an interim report was submitted to DOE in March 2007.

In year 2, the project staff will meet at Argonne on October 25–26, 2007, with a follow-on meeting to take place in the spring on the West coast.

The CEDPS yearly budget is \$2.4M. Table 1 shows approximate FTE levels and areas of concentration for staffing in the first year of the project.

**Table 1: CEDPS funding and staffing levels by site and work area.**

Site	Total Budget	FTE's (approx)	Total Personnel	Data Area	Svcs Area	TS Area	Mgt
ANL	\$800,000	3.0	10	1.65	0.55	0.30	0.5
LBNL	\$720,000	2.5	4		1.3	1	0.2
FNAL	\$120,000	0.8	4	0.75			0.05
ISI	\$480,000	1.75	5	1.05		0.5	0.2
Wisc	\$280,000	1.0	2 (+)	0.95			0.05

## 8 Publications and Presentations

CEDPS produced five refereed papers, 20 presentations, three posters, and five design documents during the past year.

## 8.1 Refereed Papers

- [1] Andrew Baranovski, Shishir Bharathi, John Bresnahan, Ann Chervenak, Ian Foster, Dan Fraser, Tim Freeman, Dan Gunter, Keith Jackson, Kate Keahey, Carl Kesselman, David E. Konerding, Nick Leroy, Mike Link, Miron Livny, Neill Miller, Robert Miller, Gene Oleynik, Laura Pearlman, Jennifer M. Schopf, Robert Schuler, Brian Tierney, *Enabling Distributed Petascale Science*, Proceedings of SciDAC 2007, Boston, MA, 2007.
- [2] Kate Keahey, Tim Freeman, Jerome Laurent, Doug Olson, *Virtual Workspaces for Scientific Applications*, poster paper, Proceedings of SciDAC 2007, Boston, MA, 2007. [http://workspace.globus.org/papers/SciDAC\\_STAR\\_POC.pdf](http://workspace.globus.org/papers/SciDAC_STAR_POC.pdf)
- [3] Brian L. Tierney, Dan Gunter, Jennifer M. Schopf, *The CEDPS Troubleshooting Architecture and Deployment on the Open Science Grid*, poster paper, Proceedings of SciDAC 2007, Boston, MA, 2007. <http://www.cedps.net/wiki/images/6/62/Scidac07-tierney-final.pdf>
- [4] J. Bresnahan, A. Brown, D. Gunter, J. M. Schopf, M. Swany, B. L. Tierney, *Log Summarization and Anomaly Detection for Troubleshooting Distributed Systems*, Proceedings of the 8th IEEE/ACM International Conference on Grid Computing (Grid 2007), Austin, TX, 2007. <http://www.cedps.net/wiki/images/e/ec/Grid2007.pdf>
- [5] Ann Chervenak, Ewa Deelman, Miron Livny, Mei-Hui Su, Rob Schuler, Shishir Bharathi, Gaurang Mehta, Karan Vahi, *Data Placement for Scientific Applications in Distributed Environments*, Proceedings of 8th IEEE/ACM International Conference on Grid Computing (Grid 2007), Austin, TX, 2007. <http://www.cedps.net/wiki/images/7/7f/Grid2007Placement.pdf>

## 8.2 Presentations

- [6] Brian Tierney, "Centralized Logging for Grid Troubleshooting," NERSC Invited Seminar, Sept 6, 2007, [https://www.nersc.gov/news/presentations/osf\\_lunch/cal/show.php?CLm=09&CLd=6&CLy=2007&c\\_num=OSF\\_Lunch](https://www.nersc.gov/news/presentations/osf_lunch/cal/show.php?CLm=09&CLd=6&CLy=2007&c_num=OSF_Lunch)
- [7] Ian Foster, "CEDPS Introduction," R. Lindsay visit to ANL, July 30, 2007, [http://www.cedps.net/wiki/images/0/0c/Cedps\\_overview.pdf](http://www.cedps.net/wiki/images/0/0c/Cedps_overview.pdf)
- [8] Dan Fraser, "CEDPS Data," R. Lindsay visit to ANL, July 30, 2007, [http://www.cedps.net/wiki/images/5/5d/CEDPS\\_Data\\_Overview.pdf](http://www.cedps.net/wiki/images/5/5d/CEDPS_Data_Overview.pdf)
- [9] Kate Keahey, "CEDPS Scalable services," R. Lindsay visit to ANL, July 30, 2007, [http://www.cedps.net/wiki/images/6/6a/CEDPS\\_Services\\_Overview.pdf](http://www.cedps.net/wiki/images/6/6a/CEDPS_Services_Overview.pdf)
- [10] Jennifer Schopf, "CEDPS Troubleshooting," R. Lindsay visit to ANL, July 30, 2007, [http://www.cedps.net/wiki/images/4/45/CEDPS\\_TS\\_Overview.pdf](http://www.cedps.net/wiki/images/4/45/CEDPS_TS_Overview.pdf)
- [11] Ian Foster, "Future Opportunities," R. Lindsay visit to ANL, July 30, 2007, [http://www.cedps.net/wiki/images/d/d3/Future\\_opportunities.pdf](http://www.cedps.net/wiki/images/d/d3/Future_opportunities.pdf)
- [12] Ian Foster, "Enabling Distributed Petascale Science," Annual SciDAC Meeting 2007, Boston, MA, June 28, 2007, [http://www.cedps.net/wiki/images/b/b9/070628\\_SciDAC\\_Foster.pdf](http://www.cedps.net/wiki/images/b/b9/070628_SciDAC_Foster.pdf)

- [13] Dan Gunter, "Troubleshooting Data Movement," CLADE: Challenges of Large Applications on Distributed Environments 2007, Monterey, CA, June 25, 2007, [http://www.cedps.net/wiki/images/e/e3/Clade2007\\_dkg.ppt](http://www.cedps.net/wiki/images/e/e3/Clade2007_dkg.ppt)
- [14] Dan Fraser, "A Managed Object Placement Service (MOPS) Using NEST and GridFTP," Condor Week, Madison, Wisconsin, May 1, 2007. [http://www.cs.wisc.edu/condor/CondorWeek2007/tuesday\\_condor.html](http://www.cs.wisc.edu/condor/CondorWeek2007/tuesday_condor.html)
- [15] Jennifer M. Schopf, "Logging Best Practices," Open Grid Forum 20, Manchester, UK, March 9 2007, <http://cedps.net/wiki/images/4/45/CEDPS-LoggingBestPracticesOGF20.pdf>
- [16] Ian Foster, "Service-Oriented Science: Scaling eScience Impact," Distinguished Lecture, Louisiana State University, November 27, 2006. Also presented as Keynote, Web Intelligence Conference, Hong Kong, December 18, 2006; and at the University of Capetown, March 29, 2007.
- [17] Ian Foster, "Grid," IFIP Summer School on Software Engineering and Computer Science, Gordon's Bay, South Africa, March 26, 2007.
- [18] Ian Foster, "System-Level Science: Scientific Exploration & IT Implications," Keynote, Workshop on Wireless Networking, Automated Information Processing, and Web & Grid Services, Puerto Rico, February 4, 2007.
- [19] Brian Tierney, "Logging Recommendations for Effective Troubleshooting," OSG Consortium All Hands Meeting, March 6, 2007. <https://indico.fnal.gov/contributionDisplay.py?contribId=120&sessionId=13&confId=468>
- [20] Ann Chervenak, "Data Services in the SciDAC CEDPS Project," OSG Consortium All Hands Meeting, March 6, 2007. <https://indico.fnal.gov/contributionDisplay.py?contribId=90&sessionId=53&confId=468>
- [21] Jennifer M. Schopf, "SciDAC Center for Enabling Distributed Petascale Science," Oak ridge National Laboratory, Oak Ridge, TN, February 8, 2007. <http://www-unix.mcs.anl.gov/~schopf/Talks/cedps-ornl-feb07.ppt>
- [22] Jennifer M. Schopf, "CEDPS and CDIGS: Two Globus Projects," Middleware And Grid Infrastructure Coordination (MAGIC), National Science Foundation, Arlington, VA, February 7, 2007. <http://www-unix.mcs.anl.gov/~schopf/Talks/cedps-nsf-feb07.ppt>
- [23] Dan Fraser and Ann Chervenak, "Data Services: Future Directions," Earth System Grid meeting in Boulder, CO, January 17-19, 2007.
- [24] Ann Chervenak and Dan Fraser, "Data Services: Future Directions," TeraGrid Data Meeting in San Diego, January 9-11, 2007.
- [25] Kate Keahey, "On-Demand Virtual Workspaces: Quality of Life in the Grid," 5th Meeting of Spanish Initiative in Grid Middleware, November 2006 [http://workspace.globus.org/papers/on\\_demand\\_workspaces\\_granada.ppt](http://workspace.globus.org/papers/on_demand_workspaces_granada.ppt)
- [26] Ann Chervenak, "Next Generation Data Services," Fermi National Laboratory, October 27, 2006
- [27] Tim Freeman and Kate Keahey, "Virtual Workspace Appliances," SC06 Booth Presentation, November 2006, [http://workspace.globus.org/papers/workspace\\_appliances\\_sc06\\_booth.pdf](http://workspace.globus.org/papers/workspace_appliances_sc06_booth.pdf)

## 8.3 Posters

- [28] Troubleshooting on the Open Science Grid, SciDAC Annual Meeting 2007, June 27, 2007, Boston, MA
- [29] Virtual Workspaces for Scientific Applications, SciDAC Annual Meeting 2007, Boston, MA, June 28, 2007
- [30] Dynamic Resource Provisioning in Grid Environments, Ioan Raicu, Catalin Dumitrescu, and Ian Foster, TeraGrid Conference 07, Madison, Wisconsin, June 2007.

## 8.4 CEDPS Design Documents

- [31] Best Practices Document, Brian Tierney and Dan Gunter, CEDPS Design Document, July 2007 <http://www.cedps.net/wiki/images/2/26/CEDPS-troubleshooting-bestPractices-17.doc>
- [32] Data Placement Service Design Draft, R. Shuler and A. Chervenak, July 2007, <http://www.cedps.net/wiki/images/e/ec/DPS4Staging.pdf>
- [33] Dynamic Resource Provisioning, Ioan Raicu and Ian Foster, CEDPS Design Document, October 2007 [http://people.cs.uchicago.edu/~iraicu/research/DRP/DRP\\_overview\\_v2.pdf](http://people.cs.uchicago.edu/~iraicu/research/DRP/DRP_overview_v2.pdf)
- [34] MOPS Architecture Overview, J. Bresnahan, R. Kettimut, D. Fraser, N. Leroy, July 2007, [http://www.cedps.net/wiki/images/e/e3/Mops\\_Architecture-1.pdf](http://www.cedps.net/wiki/images/e/e3/Mops_Architecture-1.pdf)
- [35] Virtual Workspace Roadmap, Sept 2007, [http://www.cedps.net/wiki/images/0/05/Virtual\\_Workspaces\\_Roadmap\\_Sept\\_2007.pdf](http://www.cedps.net/wiki/images/0/05/Virtual_Workspaces_Roadmap_Sept_2007.pdf)

## 9 Software Releases

### MOPS

- Latest release: MOPS v 1.0, October 1, 2007

### Data Placement Service

- Latest release: DPS 1.0, October 1, 2007

### Workspace Service

- Latest production release: Workspace Service 1.2.3, April 2007 <http://workspace.globus.org/vm/TP1.2.3/index.html>
  - Previous release 1.2.2 in January 2007 <http://workspace.globus.org/vm/TP1.2.2/index.html>
  - Previous release: 1.2.1 in October 2006 <http://workspace.globus.org/vm/TP1.2.1/index.html>

### GridFTP

- Latest development release GridFTP 4.1.2, June 2007 <http://www-unix.globus.org/ftppub/gt4/4.1.2/installers/src/gt4.1.2-all-source-installer.tar.gz>
- Latest stable release GridFTP 4.0.5, June 2007 <http://www.globus.org/toolkit/downloads/4.0.5/>

**NeST**

- Latest release NeST 0.9.8, Nov. 13, 2006 <http://www.cs.wisc.edu/condor/nest> .

**LotMan**

- Latest release LotMan 0.3.1, Aug. 28 2007, <http://www.cs.wisc.edu/~nleroy/pub/>

**dCache**

- Latest production release dCache 1.7.0.44, Sept. 19, 2007, <http://www.dcache.org/downloads.shtml>

**NetLogger**

- Development release <http://dev.globus.org/wiki/Incubator/NetLogger>
- Stable release <http://dsd.lbl.gov/NetLogger/>

**MDS4**

- Latest development release MDS4 4.1.2, June 2007 <http://www-unix.globus.org/ftppub/gt4/4.1.2/installers/src/gt4.1.2-all-source-installer.tar.gz>
- Latest stable release MDS4 4.0.5, June 2007 <http://www.globus.org/toolkit/downloads/4.0.5/>

**pyGlobus**

- Latest release: <http://dev.globus.org/wiki/Python/pyGlobus>

**pyGridware**

- Latest release: <http://dev.globus.org/wiki/PyGridWare>

## Appendix 1: Milestones Year 1

- **MS1:** *Work with application communities, including OSG, ESG, LIGO, and others, to understand their requirements and use these to drive service development.*

**Ongoing:** We have worked with all these application communities to understand their end-to-end data placement requirements. With OSG, we have worked to understand the problems that have occurred with lack of resource management in GridFTP servers, and we will address this work with the MOPS server that provides resource management. ESG and LIGO are both interested in the improved functionality that MOPS will offer over existing GridFTP, and both of these applications are discussing their long-term data placement requirements with us to help drive the design of our placement services. We have also presented work to TeraGrid data applications and at the SciDAC PI meeting.

- **MS2:** *Enable local resource management within GridFTP server: Prototype and implement basic capabilities within the GridFTP server for protecting the underlying system from resource exhaustion, in particular, to better manage network connections and memory usage.*

**Completed:** The GridFTP team at Argonne and the NeST teams at Wisconsin have created a baseline architecture and prototype for resource management within GridFTP. The GridFTP server has been enabled to respect memory limits, stream limits, and block size limits. The limits can be managed by configuration or by a resource plug-in, and allow basic memory and file system management in MOPS.

- **MS3:** *Prototype and implement NeST capabilities for managing the underlying storage component of GridFTP transfers. An important study will be to minimize the storage management impact on the GridFTP transfer rate.*

**Completed:** The Wisconsin team has integrated NeST into Condor as a precursor to integrating with GridFTP. This is available in the latest NeST release.

- **MS4:** *Combine enhanced GridFTP and NeST services into a MOPS release with basic functionality. This release will be made available for users to start utilizing a basic MOPS functionality as soon as possible. We will work with OSG on evaluating, testing, and incorporating MOPS into the VDT.*

**Completed:** The GridFTP and NeST teams prototyped development but did not formally release a 0.5 version, preferring to have a more complete version 1.0 only, as shown in MS5.

- **MS5:** *MOPS 1.0 release by October 2007 that builds on the MOPS 0.5 implementation and adds capabilities for exposing information on internally managed resources. We note that exposing resource information is only a first cut as some resource information may not be available via this mechanism. More sophisticated logic may be required to handle this problem at a higher level. MOPS 1.0 bundles will also be targeted for release in an upcoming OSG VDT release.*

**Completed.** The GridFTP and NeST teams have an initial working implementation of MOPS 1.0. It includes capabilities for exposing information on internally managed

resources. This resource information can then be used by higher-level services to more efficiently manage transfers in services such as RFT and to perform resource matching and scheduling in tools such as Stork. In addition, this release includes callouts made by GridFTP to ensure that file space is available when needed. Additional feedback mechanisms between NeST and GridFTP are being investigated to make sure the capability meets the anticipated user requirements for effectively reporting accurate utilization data, thereby allowing the system to effectively monitor resource utilization. Incorporation of this release into VDT is now a year 2 milestone.

- **MS6:** *Research optimization for transfer of Lots of Small Files (LOSF).*

**Completed.** Enhanced performance for small files is now available in the GT 4.1.2 release as part of GridFTP. Multiple small files are streamed into an open channel which all but eliminates the file transfer channel startup overhead. Previously, performance could drop by an order of magnitude as soon as files less than 100 MB were encountered. This drop now occurs when files are smaller than 100 KB.

- **MS7:** *Investigate the integration of SRM and dCache with GridFTP service into the Managed Object Placement Service. Develop a plan for this work.*

**Ongoing.** Argonne and Fermi have met to discuss the relationship of dCache and GridFTP, but a complete plan and implementation need to continue in year 2. Progress at Fermi includes work on checksum support in the Fermi GridFTP implementation for dCache. Additional integration with GridFTP/MOPS is still needed, as is work with SRM.

- **MS8:** *Design and prototype implementation of a higher-level, policy-driven Data Placement Service that interfaces with the reliable distribution layer.*

**Completed:** The ISI team distributed the draft: “A Data Placement Service for Data Staging: Scenarios, Interfaces and Design” on July 17, 2007. This design document with interface description for the data placement service is available at <http://cedps.net/wiki/index.php/Data>. We are implementing a data placement service based on this design document. An initial release of this placement service will be available on October 1, 2007, at <http://www.cedps.net/wiki/index.php/Software>.

- **MS9:** *Design interface between higher-level data placement services and the reliable distribution layer and implement a prototype.*

**Ongoing:** As part of the design process for the Data Placement service, we have been reconsidering the design of the reliable distribution layer. We are evaluating whether there is a pressing need for a separate reliable distribution layer or whether the use of simple GridFTP/MOPS data transfer operations will suffice, along with a mechanism to detect and retry failed transfers. We will be documenting these issues after the release of the Data Placement Service.

- **MS 10:** *Develop an architecture document for dynamic resource provisioning using GRAM.*

**Ongoing.** We discussed the overall architecture throughout the year. A CEDPS Design Document summarizing discussions at the fall meeting is available [33], and Raicu followed up this work with a poster presentation [30]. Additional discussion of

related technologies that would fulfill the roles of front-end and back-end components has continued, but this document needs updating and completion.

- **MS 11:** *Dynamic provisioning for HEP applications.*

**Ongoing:** See milestones 12, 13

- **MS 12:** *Dynamic provisioning for STAR*

**Completed:** Initial prototype work with STAR (demonstrated at SC 2006) has been made into a production release. In August 2007 the STAR scientists were able to stage multiple trial VM-based runs. In September 2007 the first large production run (over 100 virtual nodes) took place. This is currently deployed for both the University of Chicago Teraport and Amazon EC2 and uses the 1.2.3 release with some additional unreleased alpha-work for STAR.

- **MS 13:** *Dynamic provisioning for a second HEP application*

**Ongoing:**

1) Alice: Work with the Alice experiment has consisted of mostly trial-and-error work to configure their images. This work has stalled on the incompatibility between needed kernel versions between an Alice extension and the workspace services work (32-bit vs 64-bit), but work will continue to try to resolve this in year 2.

2) ATLAS. A group of ATLAS users is using the workspace service for ATLAS applications, with a small deployment on their own resources. We are supporting them through bug fixes and explaining the code base. A follow-on meeting is planned for October 2007 to discuss next steps.

- **MS 14:** *Develop an architecture and roadmap for dynamic provisioning using virtual machine technology.*

**Completed:** A roadmap is available at [http://www.cedps.net/wiki/images/0/05/Virtual\\_Workspaces\\_Roadmap\\_Sept\\_2007.pdf](http://www.cedps.net/wiki/images/0/05/Virtual_Workspaces_Roadmap_Sept_2007.pdf)

- **MS 15:** *Library of Workspace service images.*

**Completed.** Initial images are available at the Workspace Service Marketplace <http://workspace.globus.org/marketplace.html>. The marketplace is modeled on similar efforts in industry by VMware and Amazon's EC2. It is intended to help scientists find and download images needed to deploy their work on the Grid. The marketplace currently contains the images developed for application communities we actively engage with (STAR, Alice), along with other images useful for Grid computing, including the images of OSG edge services.

- **MS 16:** *Workspace service release.*

**Completed.** Three releases for the Workspace Service have taken place: 1.2.1 in October 2006, 1.2.2 in January 2007, and 1.2.3 in April 2007. Detailed change logs are available at the following websites: <http://workspace.globus.org/vm/TP1.2.1/index.html> and <http://workspace.globus.org/vm/TP1.2.2/index.html>.

- **MS 17:** *Network service requirements for pyGridware wrappers.*

**Completed.** We developed a set of requirements for what services an automatically generated science service will require. We developed tools to wrap several molecular modelling codes and used those codes as part of a complex biological workflow. This is available from SVN at <https://bosshog.lbl.gov/repos>. We also discussed requirements with members of both the ESG project and the Tech-X-led fusion SciDAC2 project.

- *MS18: Develop pyGridWare-based portal generation services.*

**Completed.** PythonCLServiceTool was renamed to CLWrapper to emphasize the cross-language feature support. The new features increase both the applicability of CLWrapper to multiple target languages and its usability, by adding portal generation capability allowing existing Grid services to be easily adapted into a Web user interface. Work included the following:

1. Java client binding generation. A Java-based command line client that behaves identically to the Python command line client was implemented on top of the client bindings. This enables the use of the client side of CLWrapper in pure-Java environments.
2. A JSR154-compliant servlet based on the Java client. This allows CLWrapper client functionality to be included in Java Web application hosting environments such as Tomcat.
3. A JSR168-compliant portlet based on the Java servlet. This allows CLWrapper client functionality to be embedded in portlet frameworks such as GridSphere.
4. A proof-of-concept portal based on CLWrapper, pyGridWare, and TurboGears, a Python Web application framework.

- *MS19, 20: CLWrapper and pyGridWare instrumented with the NetLogger toolkit.*

**Completed.** CLWrapper and pyGridWare were instrumented according to the “Best Practices Document” provided by the Troubleshooting area. This instrumentation logs all of the important calls in CLWrapper and pyGridWare including start and end of service operations, serialization and deserialization, and higher-level constructs such as service initialization, container operation, and notifications.

- *MS21: Adaptor classes for pyGlobus interfaces to legacy pre-Web Service tools.*

**Complete:** Support for Globus RFT / GridFTP and GRAM2 (pre-ws GRAM) was added to pyGlobus. This provides a mechanism to allow easy integration of GT2 components with a Web service environment. This is part of the current pyGlobus release.

- *MS22: pyGridWare support for additional high-level services.*

**Complete:** PyGridWare was enhanced to include a Delegation Factory Service and Delegation Service so that clients can use the pyGridWare container to delegate their credentials. Delegated credentials are required for the container to invoke services on the user's behalf. Client-side support for the Reliable File Transfer (RFT) service and the WS-GRAM was also implemented. Support for a GT4-compatible Reliable File Transfer Service is also complete. Support for these services will allow a pyGridWare container to be used interchangeably with a Java-based container.

- **MS23:** *Deploy new service-area services on OSG and ESG.*

Our virtual workspace work with OSG community has now included a deployment on the core OSG infrastructure. Both STAR and ATLAS are using it on local resources.

Support for the front-end pyGlobus system for the OSG LIGO project continues through fixing bugs and implementing features at their request. This is also deployed on the Open Science Grid Integration Testbed cluster to help test both OSG and CEDPS middleware.

- **MS24:** *Customer/User Application Interfacing.*

**Ongoing:** We continued support of the LIGO project's use of pyGlobus by fixing bugs and implementing features at their request. Deployed an Open Science Grid "Integration Testbed" cluster to help test both OSG and CEDPS middleware.

We are currently working with the Star project to develop a portal to the virtual workspaces service to enable dynamic provisioning of resources for HENP applications.

- **MS25:** *Log File Collection Service: This service will be used to gather middleware and application log files from multiple locations to a single host.*

**Complete:** We explored a number of open source tools, and determined that syslog-ng (<http://www.balabit.com/downloads/syslog-ng/>) fills all our requirements. We tested it thoroughly and put together a set of recommendations for Grid deployment at (<http://www.cedps.net/wiki/index.php/Syslog-ng>). We worked closely with OSG to come up with a deployment configuration for OSG and have deployed it on the OSG "Validation Testbed." This work went well, and we are currently are in discussions with OSG for initial deployments on the rest of OSG.

- **MS 26:** *Write an Instrumentation "best practices guide": This is advice to middleware and application developers on what to log for maximal troubleshooting abilities. This includes what data should be included in the log files, and what format the log file should be.*

**Complete.** A document is available from <http://www.cedps.net/wiki/index.php/LoggingBestPractices>. This document has been presented in a number of forums for comment, including the OSG All Hands meeting, the CEDPS all-hands meeting, the OGF-NN, Scidac tutorial, and a NERSC seminar.

- **MS 27:** *Develop an Event Log database. This database will be used to store log data and perform queries on log events to determine what was happening just before an error occurred.*

**Complete:** A prototype version of this is complete and can be seen here <http://netlogger.lbl.gov/phpsyslogng/>

- **MS28:** *Instrument Globus and Condor based on the Logging Best Practices guide.*

**Ongoing.** GT4 components prototyped the Best Practices Logging format and will be available in the 4.1.3 development release for C Core, Java Core, MDS (except the new Trigger service), GridFTP, and GRAM. Other components will be added in year 2, as will converting Condor logs to the new format.

- **MS29:** *Develop additional MDS4 Trigger Service Information Providers.*

**Complete.** Work on incorporating ESG-specific information providers has been completed and included in the GT 4.1.1 release.

- **MS30:** *Develop additional MDS4 Trigger Service action scripts (currently only send email).*

**Postponed:** Current Trigger work is focusing on usability, ease of deployment, and easier setup of Triggers, per the request of ESG directly. Additional actions will be part of CEDPS work in Year 2.

- **MS31:** *Deploy new services on OSG and ESG.*

**Ongoing:** We have been working extensively with OSG to deploy the log collection service, and are deployed an Open Science Grid “Integration Testbed” cluster to help test both OSG and CEDPS middleware. We have had initial discussions with them about deploying the Trigger service as well, but they are waiting on the new version of this code (available in GT 4.2).

ESG has currently deployed the Trigger service and is alpha testing the new functionality that we have developed for them. We have also had discussions with ESG about deploying the log collection work and expect this to continue into year 2.

- **MS32:** *Outreach to Grid application developers to instrument their applications.*

**Ongoing:** We have begun discussing additional logging with a number of groups, including STAR, Tech-X Corporation, XROOTD, and SRM/DRM developers.

- **MS34:** *Log file anomaly detection service prototype.*

**Complete:** A prototype is completed and was demonstrated at the CEDPS midterm meeting in March 2007, and is detailed in [4]. This is part of the planned deployment with OSG.

## **Appendix 2: Milestones Year 2**

### ***Data Area Year 2 Milestones***

- MS2.1.** Continue working with application communities (OSG, ESG) to understand requirements and deploy and test CEDPS services. (Continuation of MS1.)
- MS2.2.** Work with OSG & VDT teams to incorporate feedback and target MOPS 1.0 for inclusion in an upcoming VDT release (Continuation of MS4 and MS5)
- MS2.3.** Evaluate the need for separate reliable distribution layer in a CEDPS Design Document; discern whether simple GridFTP/MOPS data transfer operations will suffice, along with a mechanism to detect and retry failed transfers.(Continuation of MS 9)
- MS2.4.** Based on outcome of MS 2.3 technology evaluation, determine whether to implement a separate Reliable Distribution Service. If needed, we will produce the interface design and prototype implementation for this service.
- MS2.5.** Prototype a non-striped connection management capability using NeST.
- MS2.6.** Increase outreach efforts to identify and work with new and existing targeted opportunities such as the APS.
- MS2.7.** Develop a managed storage capability for non-striped MOPS.
- MS2.8.** Document use cases and performance capabilities of different ways to manage transfers on storage in a feature-rich provisioning environment. Given a use case, compare performance of MOPS 1.0 and dCache directly, RFT, SRM, etc.
- MS2.9.** Develop a CEDPS Design Document and prototype implementation of a common interface to underlying storage mechanisms (i.e. GridFTP, dCache, etc.).
- MS2.10.** Prototype methods of incorporating troubleshooting into MOPS.
- MS2.11.** Produce a CEDPS Design Document describing the practicality of incorporating mechanisms for quality of service in storage mechanisms and data placement services. The goal is to investigate whether using contracts or agreements could improve performance and reliability.
- MS2.12.** Deliver a MOPS 2.0 release that includes additional optimizations including LOSF in GridFTP, dynamic resource registration, nonstriped connection management, and basic troubleshooting instrumentation. This release will also have the capability of exposing resources for use by higher level services.
- MS2.13.** Release version 2.0 of the DPS with additional functionality. The placement service is responsible for accepting client requests, placing outstanding data transfer operations on priority queues, and issuing data transfer operations efficiently. We will implement additional logic in the placement service to manage competing placement requests, for example, those related to multiple workflows performing data staging operations in parallel as well as data placement operations required by virtual organization for data distribution. We will evaluate a variety of algorithms for data placement using simulation and experimentation in cooperation with application communities and identify which algorithms are most efficient.

**MS2.14.** Work with troubleshooting area to include additional data in logs for anomaly detection in placement services.

### ***Data Area Changes from Original Plan***

The changes in milestones with respect to the original plan are evolutionary and are largely based on lessons learned from our year 1 accomplishments. For example, based on our improved understanding of data placement services after implementing DPS1.0, we are evaluating whether we need a separate reliable distribution layer or whether this functionality should be contained within the DPS service itself. Additional functionality planned for MOPS extends and enriches functionality available in version 1.0. Moreover, we will investigate the practicality and utility of incorporating QoS mechanisms in our MOPS and placement services to improve data transfer performance and reliability.

### ***Service Area Year 2 Milestones***

**MS2.15.** Develop an architecture document (as a CEDPS Design Document) for dynamic resource provisioning using GRAM, extending the current design document [33]. This is a continuation of MS 10.

**MS2.16.** Develop an architecture document (as a CEDPS Design Document) integrating the Web service application infrastructure with provisioning backends.

**MS2.17.** Continued support for STAR application and workspace services, including host certificate and OSG integration issues. Continuation of MS2.12

**MS2.18.** Continued support of Alice and ATLAS with workspace service image development and deployment issues. Continuation of MS2.13.

**MS2.19.** (New) Define 1 possibly 2 new application groups for the Workspace Service group to work with within SciDAC Community (by February 2008).

**MS2.20.** (New) Define images and help deployment of application(s) identified in MS 24.

**MS2.21.** Prototype schema for Workspace bookshelf description of execution environments

**MS2.22.** Release additional 1.2.x versions of Workspace Service.

**MS2.23.** (New) Incorporate file staging into RAVi prototype.

**MS2.24.** (New) Work with GT security personnel to extend RAVi to work with individual certificates and the Delegation Service instead of community credentials

**MS2.25.** (New) Define 1 or possibly 2 new application groups for the RAVi group to work with within SciDAC Community (by February 2008).

**MS2.26.** (New) Support deployment of application(s) identified in MS2.30.

**MS2.27.** (New) Portal development and release for the Workspace Service for better user interfaces in support of STAR.

## ***Services Area Changes from Original Plan***

The CEDPS management team decided to refocus the effort in the Services area for FY2008 because of changes in personnel and management. Specifically, instead of the Python-based tools pyGlobus and pyGridware, work will focus on extending the Java-based RAVi (Remote Application Virtualization Infrastructure) tool.

Both the front-end and back-end teams will focus more strongly on work with end user applications. The front-end team will work with SciDAC end user applications, especially those with legacy codes for which this work is well suited. In addition, RAVi will need extensions to work with the Delegation service (and user credentials), as it is currently using community certificates, and some work with data staging may be needed.

The focus of the back-end work will be on refining the capabilities developed in year 1 and hardening that work for additional communities, instead of the previous focus on a large amount of additional functionality for the same user (STAR).

## ***Troubleshooting Area Year 2 Milestones***

**MS2.28.** Continue adaptation of Globus software to use Best Practices Logging format. Continuation of MS 28.

**MS2.29.** Work with Condor team to incorporate Best Practices Logging format. Continuation of MS 29.

**MS2.30.** Continue working with OSG to deploy centralized log file collection service based on syslog-ng. This includes support for removing or encrypting sensitive data.

**MS2.31.** Continue refinement of the logging “best practices guide” as we get more experience implement and deploying applications and middleware using this logging. This also includes continued work on how to best generate and propagate “global identifiers,” or GUIDs.

**MS2.32.** Version 1.0 of the log database service: This service is necessary to be able to track baseline performance, monitor hardware and software configuration changes, and perform anomaly detection.

**MS2.33.** Complete packaging and documenting log database service and log parsers tools for legacy (pre-“best practices”) logs, and work with OSG to deploy these tools

**MS2.34.** Continue work on anomaly detection tools, including methods to use the log database service to establish performance baselines, and trigger events if performance deviates too much from the baseline.

**MS2.35.** Develop MDS4 Trigger service action scripts to take actions based on results from the anomaly detection tools.

**MS2.36.** Continued support of ESG for use of new MDS Trigger service.

**MS2.37.** Deploy MDS4 Trigger service on a second infrastructure (possibly OSG or NERSC).

**MS2.38.** Continue to deploy new services on OSG, ESG, and others.

**MS2.39.** Continue outreach to Grid application developers to help instrument their applications

***Troubleshooting Area Changes from Original Plan***

No significant changes have been made to the Troubleshooting milestones except to clarify the language. Some year 1 work has been extended into year 2 (full development of logging work in all Globus and Condor components, some Trigger service work), but this will not affect the overall scope of the plan for this area.

## Appendix 3: Partner Listing

### Mathematics and Computer Science Division, Argonne National Laboratory

- Ian Foster, PI
- Jennifer M. Schopf , Project Director, Argonne Site lead
- Kate Keahey, Service lead 2006–2007
- John Bresnahan
- Dan Fraser
- Tim Freeman
- Raj Kettimut
- Mike Link
- Ravi Madduri, 2007–
- Neill Miller

### Lawrence Berkeley National Laboratory

- Keith Jackson, LBNL Site lead 2006–2007
- Brian Tierney, Troubleshooting Area lead, LBNL Site lead 2007-
- Joshua Boverhof, 2007-
- Dan Gunter
- David Koenerding, 2006-2007

### Information Sciences Institute, University of Southern California

- Ann Chervenak, ISI Site lead
- Carl Kesselman
- Laura Pearlman
- Robert Schuler
- Mei-Hui Su

### Dept of Computer Science, University of Wisconsin Madison

- Nick LeRoy, UW Site lead
- Miron Livny

### Computing Division, Fermi National Accelerator Laboratory

- Andrew Baranovski, Fermi Site lead 2007-
- Matt Crawford
- Gene Oleynik
- Ruth Pordes