

## **FutureGrid**

### **Quarterly Report – Q2 2010 – 1 January to 31 March 2010**

---

Geoffrey Fox, Indiana University (PI)

Kate Keahey, University of Chicago (co-PI)

Jose Fortes, University of Florida (co-PI)

Andrew Grimshaw, University of Virginia (co-PI)

Warren Smith, University of Texas (co-PI)

## **Introduction**

This quarterly report provides a summarization of the following bi-weekly reports submitted to NSF each Monday:

January 4, 2010

January 18, 2010

February 2, 2010

February 15, 2010

March 1, 2010

March 15, 2010

March 29, 2010

The organization of this report follows our biweekly reports:

1. Science Highlights
2. Systems Administration and Network Management Committee
3. Software Committee
4. User Requirements Committee and User Advisory Board
5. Performance Analysis Committee
6. Training, Education, and Outreach Committee
  - a. Events
  - b. Publications
7. User Support Committee
8. Operations and Change Management Committee

## Science Highlights

### **SPEC MPI2007**

A collection of performance benchmarks were executed on the IBM iDataPlex cluster (the UCSD/SDSC FutureGrid machine) using two different operating systems. Windows HPC Server 2008 (WinHPC) and Red Hat Enterprise Linux v5.4 (RHEL5) were compared using SPEC MPI2007 v1.1, the High Performance Computing Challenge (HPCC) and National Science Foundation (NSF) acceptance test benchmark suites. Overall, we found the performance of WinHPC and RHEL5 to be equivalent but significant performance differences exist when analyzing specific applications. The focus was on presenting the results from the application benchmarks and including the results of the HPCC microbenchmark for completeness.

The paper was submitted to the Standard Performance Evaluation Corporation (SPEC) Workshop to be held in Paderborn Germany in October 2010.

### **Cloud Computing Paradigms for Pleasingly Parallel Biomedical Applications**

A team from the Indiana University SALSA group compared MapReduce on FutureGrid with an alternative implementation using the “master-worker” approach offered on the commercial cloud infrastructure service based virtual machine utility computing models of Amazon AWS and Microsoft Windows Azure. On FutureGrid they used the MapReduce based computing frameworks Apache Hadoop (deployed on raw hardware as well as on virtual machines) and Microsoft DryadLINQ. This involved running Windows Linux and Xen on the same FutureGrid system -- the UCSD/SDSC IBM iDataPlex hardware. They compared performance showing strong variations in cost between different EC2 machine choices and comparable performance between the utility computing (spawn off a set of jobs) and managed parallelism (MapReduce). The MapReduce approach offered the most user friendly approach.

This paper was published as Thilina Gunarathne, Tak-Lon Wu, Judy Qiu, and Geoffrey Fox, [Cloud Computing Paradigms for Pleasingly Parallel Biomedical Applications](#) March 21 2010. Proceedings of Emerging Computational Methods for the Life Sciences [Workshop](#) of ACM [HPDC](#) 2010 conference, Chicago, Illinois, June 20-25, 2010.

## **Systems Administration & Network Management Committee**

### **Compute and Storage Systems**

#### **IU Cray XT5m**

- Delivered and installed
- All acceptance and stability tests passed
- PAPI functionality testing completed
- Verified and replicated Cray acceptance benchmarks
- MPI and Resource Management functionality complete
- Notice of acceptance released to Cray
- Initiated contact with NERSC regarding Hadoop execution using Cray libraries
- Lustre DC-WAN configuration is needed for compute nodes to mount the file system. This will be limited by a 1Gb link, and additional hardware may be desired if Lustre WAN use is requested

#### **IU iDataPlex**

- Delivered and installed
- OS and cluster management installation complete
- Acceptance testing on-going. Status as of March 29 biweekly report:

#### **Status of IU iDataPlex Benchmarking and Reliability**

The applications for the acceptance tests are HOMME, MILC, WRF, PARATEC and the HPCC benchmark. The acceptance test consists of two parts, proving for different data sets of the above mentioned applications and different core counts that they are meeting the target numbers and running a 14 day stability test with selected data sets and core counts of those applications. IBM is working on Infiniband and PARATEC problems.

The following paragraphs summarize the tests with the individual applications. For all applications, acceptance test numbers for 256, 672 and 1024 core jobs are defined.

HOMME has been run with the standard data set at all three core counts. It meets the acceptance test numbers.

MILC has been run at 256 and 1024 cores using the medium and the large dataset, and it meets the acceptance test numbers. MILC cannot be run at 672 cores because the communication pattern in the MILC code requires each MPI process to have a block of lattice points of the same size and shape. The closest number of cores that can be used is 512. The target number for 672 cores was used for a 512 core run, and MILC meets that number.

WRF has been run with the standard data set on 256 and 672 cores and meets the acceptance test numbers. The same is true for the large dataset and 672 cores. We have not yet been able to run WRF at 1024 cores, but we are confident that once we get it to run, we can meet the acceptance test numbers. We are seeing a combination of WRF specific and Infiniband related error messages when running WRF at 1024 cores.

PARATEC has been run at 256, 672 and 1024 cores using the medium and the large data set. It does not meet the acceptance test numbers in any of the tests.

HPCC has been run for 8, 16, 32, 64, 128, 256, 512, 672 and 1024 cores. Acceptance test criteria exist for the 256, 672 and 1024 core runs. For the 256 and 672 core run, HPCC meets all the acceptance test numbers. In the 1024 core run, the only component that fails to meet the criteria is the MPI Random Access component. We are confident this can be fixed.

We have been doing a 2 day dry run of the 14 day stability test but so far have not been able to achieve the job completion rate required by the acceptance test criteria. We had no problem with the uptime criteria.

### **UC iDataPlex**

- Delivered
- Final cluster networking in place
- Central management servers and storage servers built
- Node installations completed

### **UF iDataPlex**

- Delivered, installed, and ready for performance testing when personnel resources are available

### **SDSC iDataPlex (from IU)**

- Delivered and installed. The cluster is now referred to as “Sierra,” following the FutureGrid NATO phonetic alphabet naming scheme
- With UCSD, Sun storage integration nearly complete; additional network cards installed and some failed components replaced
- A resource manager (Rocks) is running and performance verification tests have begun
- System re-installed with XCAT on February 22<sup>nd</sup> and 23<sup>rd</sup>
- In March, UCSD moved the Sun storage to the iDataplex rack and configured the management network (managed via Rocks). Infiniband cards were purchased for the storage so it could be connected to Sierra’s cluster nodes over the Infiniband network

### **LDAP Server**

- Installed for account provisioning using SSH keys for authentication
- Internal accounts can be provisioned on IU Cray, IU iDataPlex, and SDSC iDataPlex

### **TACC Dell**

- Purchase order sent to Dell on March 24 (TACC subaward is now in place)

### **UCSD DDN Storage**

- UCSD finished its acceptance of the two 48 TB Sun x4540 storage boxes in January

## **Networks**

- IU Global NOC monitoring of central networking equipment is active
- FutureGrid connection between Chicago (Starlight) and IU is active
- NLR agreement is fully executed and circuit provisioning has begun
- FutureGrid IP space secured and allocated for each site where desired
- Central Router and Network Impairment Device installed at Starlight in Chicago
- UCSD worked with IU to order Juniper switches for the external network for Sierra, which were installed and configured in mid February.

## **Software Committee**

### **Staffing**

J. Voekler (ISI) joined the team and replaced R. Bolze. UC: JA. Younge (IU) joined FG. UV: J. Otey was named a project manager at UV and was introduced to the IU team. IU lost one of its staff members due to personal reasons. This staff member was replaced by Fugang Wang (IU).

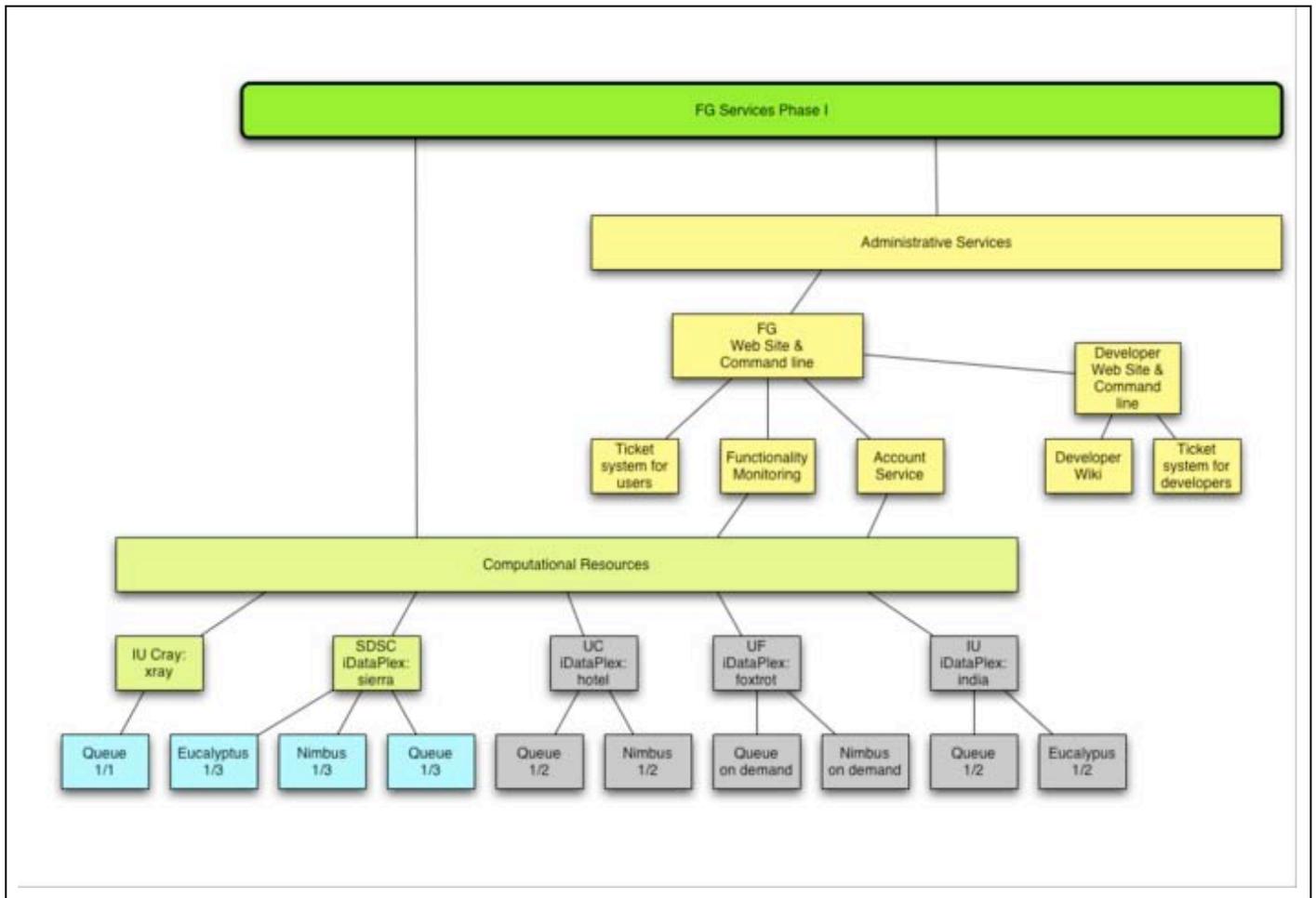
### **Software Task Management**

A security issue rendered the FutureGrid web site server (written in Drupal) unusable for development activities. As a result, the Software Committee created standalone Drupal and wiki servers to facilitate and continue the collaborative work. The issue tracking software and a development wiki has been set up. The wiki has been augmented with a number of very useful plug-ins, including a UML drawing tool and a customized bridge to the jira systems so that issues can be included through search queries into wiki pages. All tasks have been revisited and team members have been asked to organize all software related activities through jira. We have also set up SVN archives. We have provided templates for architecture documents, command line commands, and use cases. In order to facilitate better documentation we integrated much of the documentation into the software wiki allowing administrators, developers, and management to instantaneously have access to it

### **Architecture**

We completed the Phase I document and handed it over to the deployment team at IU. Phase I contains the creation of a Eucalyptus, and Nimbus cloud and HPC software on all available hardware. Hardware acceptance testing has delayed the execution of the architecture plan during Q2. During this period no hardware was available for a significant amount of time.

## FutureGrid Architecture for Phase I:



### Architecture Subsystems

**Nimbus:** Several Nimbus software release were conducted. The Nimbus installation process was much improved.

**Experiment Management:** We have started and revised the experiment harness document

**Authentication:** We have designed a role based accounting system using LDAP with the plan to integrate it with InCommon integration via CiLogon. As an interim step, a simple copy system based on key authentication has been deployed.

**Image Management:** We completed a short survey of configuration management tools. We have started an Image management architecture documenting how more secure images can be created semi-automatically.

**Dynamic Provisioning:** The Phase II document has been initiated for the RAIN service. We have developed a first testbed and a virtual cluster setup via XCAT and VirtualBox. The cluster is based on 4 new Dell Optiplexes. The software group felt it was necessary to have a development platform during the time the hardware is not available. We have now a template available how to create such an environment, which can also be replicated on FutureGrid.

**Software tasks related to performance activities:** The following activities took place as part of the performance committee that have an impact on software:

- Setting up automated benchmarking of HPCC in order to detect performance problems.
- Developing an Inca reporter to execute the HPCC benchmarks in continuous mode.
- A Vampir workshop was held at IU with about 15 participants. This included an overview of the Vampir components and a hands-on tutorial session. Further details are at <http://iu-pti.org/hpa/vampir-workshop-2010>.
- LDAP, FG machines, SDSC NOC historical data
- Inca was set up at <http://inca.futuregrid.org>.

## **User Requirements Committee and User Advisory Board Summary**

The candidate membership of the UAB was finalized and invitations sent out. Thus far, all but two of the candidate members responded in the affirmative. Thus, the board consists of the following members:

- |                         |                     |
|-------------------------|---------------------|
| 1. Ann Chervenak        | ISI                 |
| 2. Andrew Grimshaw      | UVA -- Chair of UAB |
| 3. Shantenu Jha         | LSU                 |
| 4. Steven Newhouse      | EGI                 |
| 5. Ruth Pordes          | OSG                 |
| 6. Morris Riedel        | Unicore 6/FZJ       |
| 7. John Towns           | NCSA/TIS            |
| 8. Jon Weisman          | UMN                 |
| 9. Rich Wolski          | Eucalyptus/UCSB     |
| 10. Nancy Wilkens-Diehr | SDSC                |
| 11. Frederic Desprez    | Inria               |

A teleconference with early users was held during Q2, with the objective of generating the discussion of user requirements for FutureGrid software and use. Early adopters to date who have registered to use FutureGrid resources include the following people:

- |                     |                                    |
|---------------------|------------------------------------|
| 1. Daniel LaPine    | NCSA                               |
| 2. Ian Stokes-Rees  | Harvard                            |
| 3. Shantenu Jha     | LSU                                |
| 4. Jon Weissman     | University of Minnesota            |
| 5. Katerina Stamou  | LSU (grad student of Shantenu Jha) |
| 6. Nick Edmonds     | Indiana University                 |
| 7. Andrea Matsunaga | University of Florida              |
| 8. John Karpovich   | University of Virginia             |
| 9. JP Navarro       | Argonne National Lab               |
| 10. Ian Gable       | University of Victoria, Canada     |
| 11. Steven Newhouse | EGI                                |
| 12. Amy Apon        | University of Arkansas             |
| 13. Martin Swany    | University of Delaware             |

## Performance Analysis Committee

### Summary

The goals of the Performance Analysis committee are to 1) evaluate the FutureGrid infrastructure as it is developed and deployed and 2) help users to leverage the FutureGrid infrastructure for their experiments. In January, the committee finished our architecture document, led by UCSD with help from its members from IU, TU-D, and ISI. The architecture document was published on the FutureGrid's Drupal website and in March, the document was migrated to FutureGrid's new MediaWiki website. During January, we also developed a set of detailed milestones for year 1, which were uploaded into the project's JIRA repository for tracking. The three main activities defined for the group are to select:

- 1) A set of benchmarks that can be used to establish and analyze the baseline performance of FutureGrid and deploy a framework that can automate periodic execution of the benchmarks in order to detect performance problems. We will start with traditional benchmarks like HPCC and the SPEC suite and then try application benchmarks in both the bare metal and VM environments.
- 2) An instrumentation tool to collect performance measurements from FutureGrid software in order to validate its usability and detect performance problems. This will include providing documentation and deploying a server to store the data.
- 3) A set of performance tools for the FutureGrid runtime environment to help users analyze the performance of their software and/or applications. We will start with tools we can fully support such as Vampir and PAPI and then deploy tools we can support at best effort such as IPM and Tau.

In February, the team acquired a VM instance for our software activities and deployed an Inca installation for FutureGrid. The Inca web status pages can be found at <http://inca.futuregrid.org>. In March, we developed an Inca "reporter" to execute HPCC and report the results. It was tested on TACC's Ranger machine and then deployed to the IU Cray machine, xray, when it became available. During March, we also deployed functionality monitoring to FutureGrid. This included a small set of three external tests that verify connectivity (ping) to the machine, verify login capability (SSH), and verify the machine's DNS alias name. Four additional tests were also deployed on machines to verify prerequisites needed to execute HPCC since the acceptance tests were completed. An additional FutureGrid service test was also added to test the availability of the LDAP server used to store user login information.

Members of the group also ran and continue to run the acceptance tests for the Cray and iDataplex machines as they became available and prepared for a Vampir workshop at IU given on April 21<sup>st</sup>. We also are collaborating with the LBNL team working on performance activities on the DOE Magellan project to share our progress.

## **Training, Education, and Outreach Committee**

### **Summary**

The UF team has packaged and tested a new version of the virtual grid appliance, which streamlines deployment in cloud-enabled resources, in preparation for integration with FutureGrid infrastructure. The new appliance has been successfully tested on virtual clusters transparently spanning private resources and Amazon EC2 cloud resources.

Improved versions of educational virtual 'grid' appliances have been deployed, with tutorials guiding users on how to create their own Condor, MPI and Hadoop virtual clusters. The improved version uses deb packages and allows simple installation of grid appliances from baseline Ubuntu 9.10 images; UF is maintaining the repository of grid appliance packages. The MPI virtual cluster tutorial is pre-loaded with files to support activities described in the CI-Tutor TeraGrid MPI online tutorial; a video has been created demonstrating how to create an MPI cluster and is available on YouTube.

An “early adopter” survey that seeks information from potential users of FutureGrid with respect to their education and training needs has been completed. The survey is linked from the FutureGrid TEOS page - <http://www.futuregrid.org/outreach>.

A video on the process of customizing the Grid appliance for redistribution has been created and uploaded to YouTube.

Grid appliances are being used in a high-performance computing class at University of Arkansas.

Grid appliance team is finishing packaging and documentation of a front-end Web appliance with open-source Web/CMS software that supports management of users and groups and generation of appliance configuration files (this front-end is currently serving <http://www.grid-appliance.org>) based on interest expressed from early adopters who would like to create their own resource pools.

## Summary of EOT Events: Q1 and Q2

Type of Event		Title	Location	Presenter	Type of Audience	Month
Workshops and Tutorials	Q1	Introduction to the Grid Appliance	FutureGrid web site	University of Florida	On-line tutorial	Nov-Dec 2009
		Creating Grid Appliance clusters	FutureGrid web site	University of Florida	On-line tutorial	Nov-Dec 2009
		Building a Ubuntu-based Grid Appliance on cloud or local resources	FutureGrid web site	University of Florida	On-line tutorial	Nov-Dec 2009
		Deploying Grid Appliances using Nimbus	FutureGrid web site	University of Florida	On-line tutorial	Nov-Dec 2009
		Virtual MPI clusters with the Grid Appliance and MPICH2	FutureGrid web site	University of Florida	On-line tutorial	Nov-Dec 2009
		Introduction to Hadoop using the Grid Appliance	FutureGrid web site	University of Florida	On-line tutorial	Nov-Dec 2009
	Q2	Performance Analysis Using the Vampir Toolchain	Innovation Center, Indiana University	Robert Henschel, Indiana University; Thomas William, ZIH, Dresden	Technical	April 2010
Presentations and Talks	Q1	FutureGrid	Open Grid Forum 27, Banff, Canada	Andrew Grimshaw, University of Virginia	Advanced technical	Oct 2009
		FutureGrid Overview	CCA-09 Cloud Computing and its Applications Workshop, Chicago, IL	Geoffrey Fox, Indiana University	Technical	Oct-2009
		FutureGrid and Green Aware Computing	Keynote at 3 <sup>rd</sup> International Conference on Networks & System Security, Gold Coast, Australia	Gregor von Laszewski, Indiana University	Technical	Oct 2009
		FutureGrid and Green Aware Computing	Northern Illinois University, DeKalb, IL	Gregor von Laszewski, Indiana University	Technical	Oct 2009
		FutureGrid: An Experimental High-Performance Grid Testbed	NCSA, Champaign-Urbana, IL	Craig Stewart, Indiana University	Technical	Oct 2009
		FutureGrid Overview	SC09 Conference, Indiana University booth, Portland, OR	Geoffrey Fox, Indiana University	Mixed technical	Nov 2009
		FutureGrid Overview	SC09 Conference, AIST booth, Portland, OR	Geoffrey Fox, Indiana University	Mixed technical	Nov 2009
		FutureGrid Overview	SC09 Conference, Indiana University booth, Portland, OR	Geoffrey Fox, Gregor von Laszewski, Marlon Pierce, Judy Qiu, Indiana University	Mixed technical	Nov 2009
		FutureGrid Cloud Technologies and Bioinformatics Applications	Keynote at 1 <sup>st</sup> International Conference, CloudCom 2009, Jiaotong University, Beijing, China	Geoffrey Fox, Indiana University	Mixed technical	Dec 2009
		Cloud Technologies and GeoScience Applications, including FutureGrid	International Symposium on Geo-Computation and Analysis (ISGA) 2009, Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, China	Geoffrey Fox, Indiana University	Technical	Dec 2009

## FutureGrid Quarterly Reports: Q2 2010 – 1 January to 31 March 2010

Type of Event		Title	Location	Presenter	Type of Audience	Month
Presentations and Talks		FutureGrid and Applications	Innovation Center, Indiana University	Geoffrey Fox, Indiana University	Mixed technical	Dec 2009
	Q2	Clouds and FutureGrid	All Hands Meeting, Minority Serving Institutions – Cyberinfrastructure Empowerment Coalition (MSI-CIEC), SDSC, San Diego, CA	Geoffrey Fox, Indiana University	Mixed technical	Jan 2010
		Building Effective CyberGIS: FutureGrid	National Science Foundation TeraGrid Workshop on Cyber-GIS, Washington, DC	Marlon Pierce and Geoffrey Fox, Indiana University	Technical	Feb 2010
		FutureGrid: An Experimental High-Performance Grid Testbed	TeraGrid Quarterly Meeting, Tampa, FL	Craig Stewart, Indiana University	Technical	Mar 2010
		FutureGrid Introduction	All Hands Meeting, Open Science Grid, Fermilab, Batavia, IL	Gregor von Laszewski, Indiana University	Technical	Mar 2010

## Publications

Andre Luckow, Lukasz Lacinski, Shantenu Jha, [SAGA BigJob: An Extensible and Interoperable Pilot-Job Abstraction for Distributed Applications and Systems](#), Center for Computation & Technology, Louisiana State University; Department of Computer Science, Louisiana State University; e-Science Institute, Edinburgh, UK

Thilina Gunarathne, Tak-Lon Wu, Judy Qiu, and Geoffrey Fox, [Cloud Computing Paradigms for Pleasingly Parallel Biomedical Applications](#) March 21 2010. Proceedings of Emerging Computational Methods for the Life Sciences [Workshop](#) of ACM [HPDC](#) 2010 conference, Chicago, Illinois, June 20-25, 2010.

## **User Support Committee**

### **Summary**

The original tool for inserting individual FGKB documents into the FutureGrid web site was rewritten. The format is <http://futuregrid.org/kb/ayzv> the last four letters being a unique FGKB doc id. The display page includes a search of the FGKB. Future features will include a feedback/comment option that can submit to the Knowledge Management team for action or posted as part of the public presentation of the documents.

The first version of the FGKB widget for the portal was created.

A set of web pages was created, providing a template to be used to collect information about specific pieces of software when available. The template includes the institution supporting the software, the homepage, support contact, links to existing faq pages, manuals, training, and presentations. Population of the pages with information has begun. Software owner will be responsible for completing them.

The list of contacts for FutureGrid systems has been provided to the Global NOC and they have created an alert/announcement list

### **FutureGrid Knowledge Base**

- Active content in the FutureGridKB now includes:
  - How do I contact FutureGrid?
  - On the FutureGrid site, how do I use the mailing lists?
  - Mailing list for FutureGrid network notifications
  - For FutureGrid support, what standard reports are available in IU Research Technologies' Request Tracker (RT)?
  - Network monitoring on the FutureGrid
  - On the FutureGrid site, how do I add or change personal information?
  - How can I get a FutureGrid account, or propose a project?
  - How do I add a document to the FutureGrid Knowledge Base?
  - Are there any RSS feeds available for FutureGrid news?
  - How do I create a ticket for FutureGrid support?
  - What hardware is part of FutureGrid, and when will it be available?
  - On the FutureGrid site, how do I add an event to the calendar?
  - On the FutureGrid site, how do I add myself to a committee?
  - What is FutureGrid?
  - How will FutureGrid help me do my research?
  - Workflow plan for the FutureGrid - Global Research NOC service desk

FutureGrid Quarterly Reports: Q2 2010 – 1 January to 31 March 2010

- FutureGrid: Support information for VirtualAppliance
- For the FutureGrid, how can I upload tasks into Jira from the command line?
- Adding FutureGridKB documents to Drupal
- Information about IU research computing systems for use in grant application

## Operations & Change Management Committee Summary

All subaward documents have been reviewed, updated where necessary, and approved by NSF to execute. All subaward documents distributed to partner institutions for requisite signatures. Final subaward purchase orders set for execution in Q3.

TACC to provide updated estimates for procurement, installation, and implementation of the Dell cluster. The purchase order was sent on March 24, but no delivery date has been set.

Transfer of ownership documents for Chicago and Florida iDataPlex machines in progress.

ZIH (Zentrum für Informationsdienste und Hochleistungsrechnen) desires to accelerate their involvement in FutureGrid (i.e. VAMPIR) in Y1 of the project. A Statement of Work has been developed and will be submitted to the Operations Committee and the FutureGrid PI for approval.

Early adopter allocation process has been started. We have requests now for the Cray computer at Indiana.

## Expenditures Report Actuals thru March 31, 2010

Expenditures Report	Q1		Q2		Q3		TOTAL	
	NSF \$	Match \$	NSF \$	Match \$	NSF \$	Match \$	NSF \$	Match \$
Total Salaries and Wages	15,150	722	118,007	69,817			133,157	70,539
Total Fringe Benefits	3,489	51	37,081	27,531			40,570	27,582
Total Salaries, Wages and Fringe Benefits	18,639	773	155,088	97,348			173,727	98,121
Total Equipment	0	266,981	0	528,593			0	795,574
Total Travel (Domestic)	5,745	56	4,464	3,554			10,209	3,610
Total Participant Support	10,804	0	867	0			11,671	0
Total Other Direct Costs	0	0	0				0	0
Subawards	0	0	0	0			0	0
Other (Networking, etc.)	0	32,288	0	48,047			0	80,335
<b>Total Direct Costs</b>	<b>35,188</b>	<b>300,098</b>	<b>160,419</b>	<b>677,542</b>			<b>195,607</b>	<b>977,640</b>
Total Indirect Costs	13,167	0	86,158	0			99,325	0
<b>Total Direct and Indirect</b>	<b>48,355</b>	<b>300,098</b>	<b>246,577</b>	<b>677,542</b>			<b>294,932</b>	<b>977,640</b>