UltraLight Technical Report
28 February 2005
http://UltraLight.caltech.edu/portal/html/

UltraLight Annual Report for 2004 – 2005

The UltraLight Collaboration

NSF Grant 0427110

Director:	Harvey Newman, California Institute of Technology

Table of Contents

1 Overview

UltraLight is a collaboration of experimental physicists and network engineers whose purpose is to provide the network advances required to enable petabyte-scale analysis of globally distributed data. Current Grid-based infrastructures provide massive computing and storage resources, but are currently limited by their treatment of the network as an external, passive, and largely unmanaged resource. The goals of UltraLight are to:

- Develop and deploy prototype global services which broaden existing Grid computing systems by promoting the network as an actively managed component.
- Integrate and test UltraLight in Grid-based physics production and analysis systems currently under development in ATLAS and CMS.
- Engineer and operate a trans- and intercontinental optical network testbed, including high-speed data caches and computing clusters, with U.S. nodes in California, Illinois, Florida, Michigan and Massachusetts, and overseas nodes in Europe, Asia and South America.

This report is being written at the start of UltraLight's fifth month of funding. During these first five months, we worked to set a realistic scope for the project, started an aggressive network testbed deployment, began to integrate our activities within High Energy Physics CMS collaboration, and established close ties to external groups and industrial vendors. We demonstrated several important milestones at the November 2004 Supercomputing Conference, held in Pittsburgh, where data transfers of over 100 Gb/s were achieved and prototype services for distributed analysis of CMS simulated data were tested. In December 2004, we held a collaboration-wide meeting, with broad participation from engineers, graduate students, physicists and computer scientists. We also conducted a smaller, more focused week-long workshop-style meeting between grid researchers and physicists in January 2005 to develop an UltraLight Analysis Architecture and to establish an initial grid-enabled data analysis testbed for CMS. In January 2005 we participated in an NSF visit with positive feedback. Finally, we have begun to integrate our efforts within the Open Science Grid and to partner with newly proposed synergistic projects, like the Data Intensive Science University Network (DISUN) and the Global Information Systems and Network-Efficient Toolsets (GISNET, [1]).

The overall project is directed by Harvey Newman of the California Institute of Technology. UltraLight is managed by a core team, coordinated by Rick Cavanaugh of the University of Florida. Shawn McKee of the University of Michigan leads the Network Engineering Group, Frank van Lingen of the California Institute of Technology leads the Applications Services Group, Laird Kramer of Florida International University leads the Education and Outreach activities, Dimitri Bourilkov of the University of Florida leads the Physics Analysis User Community, and Steven Low of the California Institute of Technology functions as a liaison with Wan-In-Lab. Support for the project and for the development of web communication resources, including a permanent videoconferencing VRVS (Virtual Rooms VideoConferencing System) [2] room for enhanced intra-project collaboration, is supplied by staff at the California Institute of Technology.

This document is structured as follows: first, we describe the details of the UltraLight Network as it is being planned and deployed, second we discuss activities and goals for the Applications Services which interface high energy physics applications with the UltraLight Network, third we present the status and plans for Education & Outreach activities, and finally we describe how a group of early adopter physicists are preparing to exercise UltraLight for physics analyses.

2 Network Engineering

The core network of UltraLight is not a standard core network with static links and fixed bandwidth connecting nodes but will dynamically evolve as a function of available resources on other backbones such as NLR (National LambdaRail) [3], HOPI (Hybrid Optical and Packet Infrastructure) [4], Abilene [5] or ESnet (Energy Sciences Network) [6] . Appropriate mechanisms will be deployed to dynamically re-configure and manage the backbone when new capacities become available. Figure 1 shows some the network resources that UltraLight plans to utilize.

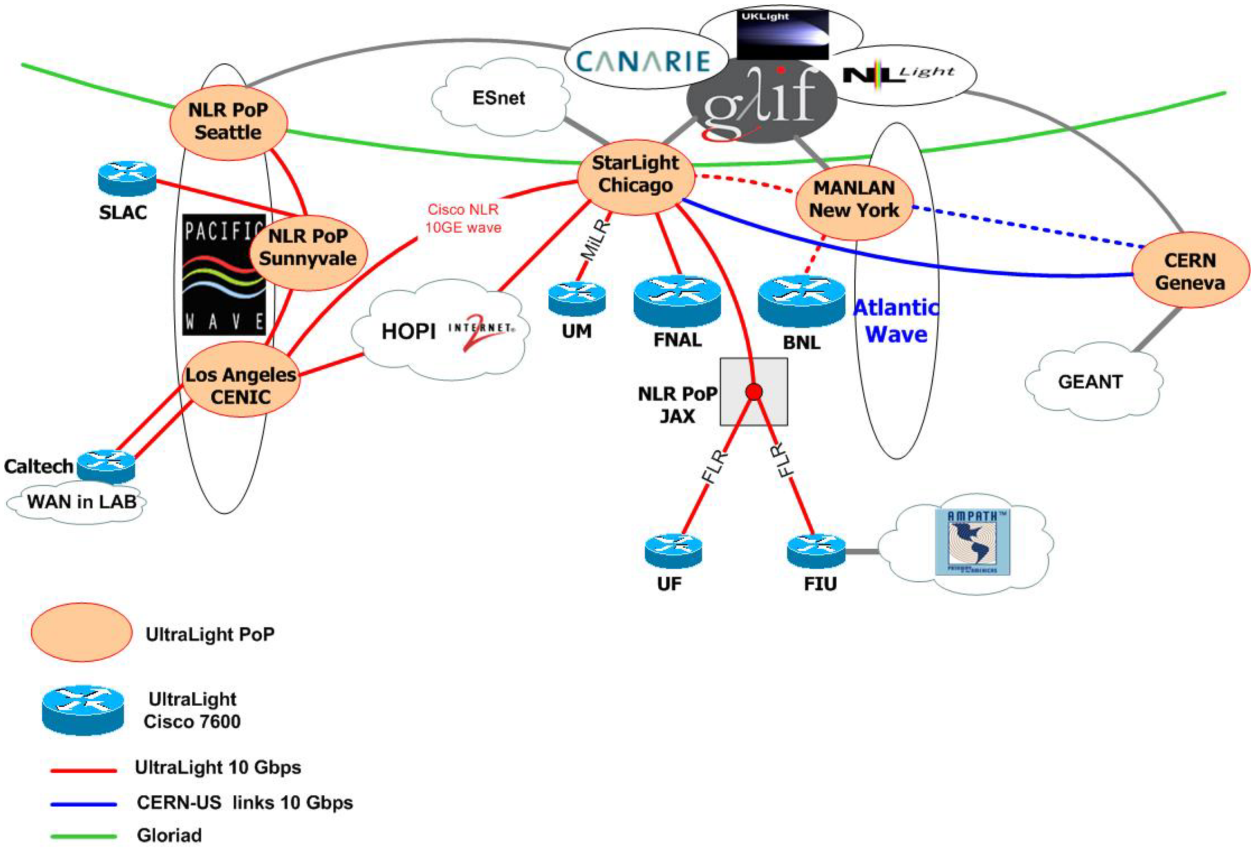


Figure 1 Connectivity Diagram for UltraLight showing International and National Partners

We will use the optical hybrid global network represented by the UltraLight core, its international partner network projects, and the production networks with it peers, to serve the needs of data-intensive science community by efficiently partitioning data flows by size and requirements among traditionally routed Layer 3, Ethernet switched Layer 2, and optically switched Layer 1 paths. In particular, the largest or most time-critical flows, typically instrumented with 10GbE (Gigabit Ethernet) interfaces in the end systems, that need reliable, quantifiable high performance will be switched over the UltraLight core paths provisioned via a reservation systems directly interfaced with the data transfer application (one or a very few flows per 10G "wave"). Medium to large-sized flows matched to end-systems will typically be relegated to MPLS (Multi-Protocol Label Switching) LSPs (Label-Switched Paths) [7] with QoS

(Quality of Service) attributes. Numerous small flows, either within each LSP or in the general traffic mix on our partner production networks (e..g. Abilene) will be managed through the use of adaptive, advanced fair-sharing protocols.

The NLR-CISCO wave between Chicago and Los Angeles will be dedicated to UltraLight for the first 6 months of its operation, until the Summer of 2005, and UltraLight will then share its use. HOPI and UltraScience Net waves connecting Chicago to respectively Los-Angeles and Sunnyvale will be provisioned on demand. UltraLight nodes in Los Angeles (LA) and Chicago will be designed and configured to allow such a provisioning.

In order to transparently cross conventional IP (Internet Protocol) networks such as Abilene or ESnet we hope to build layer 2 VPNs (Virtual Private Networks) based on MPLS between UltraLight nodes. Those point-to-point connections could be used as a backup when no dedicated waves are available. They also offer additional capacity.

The "core" resources for UltraLight are:

- LHCnet (IP, L2VPN, CCC (layer 2 emulation from Juniper))
- Abilene (IP, L2VPN)
- ESnet (IP, L2VPN)
- Cisco NLR wave (Ethernet)
- HOPI NLR waves (Ethernet; provisioned on demand)
- UltraLight nodes: Caltech, SLAC, FNAL, UF, UM, StarLight, CENIC PoP at LA, CERN

2.1 UltraLight Site Details

2.1.1 Brookhaven National Laboratory (BNL)

BNL's redundant WAN routers and high-availability Core and Distribution Layers are comprised of Cisco Catalyst 6513's, providing scalable 10 GbE transport to the RHIC/USATLAS Computing Facility. This 10 GbE transport will be integrated into BNL's infrastructure with the addition of Cisco Supervisor 720 Engines and 10 GbE board/Xenpack modules. The current campus architecture will easily scale to meet the ever growing bandwidth/service needs of our scientific community for both production and experimental work, for years to come.

2.1.2 California Institute of Technology (Caltech)

The Caltech Ultralight local loop connection is done with 2 10 GbE waves from Caltech campus to CENIC PoP (Point of Presence). One of the waves can be dedicated for production (Abilene, NLR) and other wave can be used for test traffic. Both waves are converged at downtown Los Angeles on a Cisco 7606 and can later be connected to a Calient Optical switch. It is also planned to deploy 2 OC-48 (Optical Carrier-48: 2.4 Gbps) local waves for the WAN-In-Lab connection extending the WAN-In-Lab reach to Sunnyvale and Seattle.

2.1.3 CERN

Ultralight has already deployed a few powerful end-systems at CERN directly connected at Layer-2 to the Ultralight testbed. Connectivity to the OpenLab testbed [8] will be guaranteed via conventional Layer-3 routing but direct Layer-2 channels could also be deployed on demand. OpenLab consists of 100 HP dual processors machines equipped with Intel's Itanium Processor, Enterasys's 10-Gbps switches and a high-capacity storage system based on IBM's Storage Tank system.

2.1.4 Florida International University (FIU)

Florida International University plans a multi-phased implementation of its connection to UltraLight: first a Layer-3 peering through Abilene, and subsequently, as the Florida Light Rail (FLR) optical network comes online, FIU will provision a 10GbE LAN-PHY wavelength through FLR to the OXC (Optical eXChange) in Jacksonville, where FLR and NLR will meet. From there, FIU as well as UF will use a 10 GbE LAN-PHY wavelength across NLR to connect to UltraLight. Ultimately, in a third and final stage, a shared 10GbE LAN-PHY wavelength between FIU and the University of Florida will traverse NLR, ultimately connecting to the UltraLight optical core in Chicago.

2.1.5 Fermi National Accelerator Laboratory (FNAL)

FNAL's internal network architecture is based on work group LANs. Major experiments at the facility (CDF, D0, US-CMS Tier-1 Center) have dedicated computing resources consolidated within their own LAN. The network infrastructure of each work group is made up of high performance switching fabric, consisting of Cisco Catalyst 6500s interconnected with 10 Gb/s links. By default, production network traffic for each experiment is sent over ESnet. Research and high impact data movement traffic, such as UltraLight would be expected to carry, is routed over the Laboratory's StarLight infrastructure by the core network device within a specific work group. The US-CMS work group has a 10 Gb/s path to StarLight in place and in use; the CDF and D0 work groups are scheduled to have 10 Gb/s paths to StarLight in place by March 1. FNAL is leading a research project, called LambdaStation [9], that will dynamically reroute select traffic over high capacity alternate wide-area network paths, such as UltraLight will provide. LambdaStation will facilitate early use of UltraLight for access to production-use storage facilities and other resources belonging to FNAL experiments.

2.1.6 Internet2

Internet2 plans to implement a test facility that will support the functionality of a HOPI node. Included will be a fiber cross connect, an Ethernet switch, and a collection of support PCs. The support PCs will provide a development platform for experimentation, a measurement platform for testing flows and latency, and a support platform for use by projects such as the Ultralight project. Implementations of software and control plane functionality can be tested using the support platform.

2.1.7 MIT Haystack Observatory

Haystack Observatory houses the Mark 4 VLBI Correlator system, capable of simultaneously processing 1 Gb/s/station from up to 16 stations simultaneously (120 baselines). The correlator system is at the heart of the e-VLBI (Very-Long Baseline Interferometry) capabilities currently being developed by Haystack Observatory and connected via an OC-48 connection to the external world. Additionally, the 20 m diameter Westford radio telescope, about 1.5 km from Haystack Observatory, is connected via a 10 Gb/s link to Haystack Observatory.

2.1.8 SLAC

SLAC intends to procure 1 rack worth of space and install the Cisco router switch plus 2-4 servers at the co-location area. One of the servers will be used for monitoring on a regular basis. Two others will be used for performance measurements on 10 GbE links (e.g. comparing the effects of different TCP (Transmission Control Protocol) stacks, UDT [10], QoS etc.). In addition we will be deploying monitoring (using IEPM-BW) of the production links between the UltraLight sites.

2.1.9 University of Florida (UF)

The University of Florida is in the process of developing and deploying a 20 Gb/s research network under a grant from the National Science Foundation. This network will initially link 6 facilities at 4 major sites on campus using a 20 Gb/s Ethernet backbone and provide 10 Gb/s and 1 Gb/s ports at the edge for tributary sites, storage, and cluster/grid computing research. This campus network will be used to deliver, among other research traffic, Ultralight services to the campus edge at 10 Gb/s and will be made up of equipment of a similar capability to the Ultralight network. The network is designed to be rapidly reconfigured to meet the needs of the researchers, and may be modified to meet the needs of specific UltraLight experiments. This includes MPLS, QoS, and other control and forwarding plane work.

2.1.10 University of Michigan (UM)

The University of Michigan will connect to UltraLight via a wavelength on MiLR [11] (Michigan Light-Rail). The UltraLight chassis is located at the Physics Research Laboratory on the main campus. Single mode fiber is already in place connecting this lab with the MiLR PoP at the University of Michigan. The UltraLight chassis is in place in the lab and has its redundant power supplies both connected to conditioned power mains. The scheduled date to initiate the UltraLight connection is March 11, 2005. After the UltraLight connection is brought up, Michigan plans to be active in a number of different areas including MPLS/QoS configuration and testing, disk-to-disk transfers, network monitoring and deployment and testing of UltraLight middle- and upper-ware and especially its interface with ATLAS software.

2.2 International Connections and Partners

An important aspect of UltraLight is the extensive international collaboration we have brought together. In Table 1 we summarize these details for quick reference.

Partner	Peering	OCP	URL (prefix with http://)
AARNet	Seattle/LA	No	www.aarnet.edu.au/
Brazil/HEPGrid	AMPATH	No	www.hepgridbrazil.uerj.br/
CA*net4	StarLight	UCLP[1]	www.canarie.ca/canet4/
GLORIAD	StarLight	No	www.gloriad.org/
IEEAF	MANLAN	-NA-	www.ieeaf.org/
Korea	StarLight	UCLP	www.kreonet.re.kr/
NetherLight	StarLight	UCLP	www.surfnet.nl/info/innovatie/netherlight/home.jsp
UKLight/ESLEA	StarLight		www.uklight.ac.uk - www.mb-ng.net/eslea/

Table 1 International Partner Summary

---

[1] User Controlled Light-Paths

2.3 UltraLight Research and Development Goals

2.3.1 Basic Network Services

One of the goals of UltraLight is to augment existing grid computing infrastructures, currently focused on CPU/storage, to include the network as an integral Grid component that offers guaranteed services and can be reserved.

UltraLight will provide on demand dedicated bi-directional data paths between UltraLight nodes. Data paths can be either dedicated Layer-2 channels (with guaranteed bandwidth, delay…) or paths shared with other traffics. In both cases, the only constraint will be the Ethernet framing and connections will be point-to-point. UltraLight will try to be as transparent as possible to end-users. Users should be able to run whatever protocols over Ethernet.

UltraLight will focus on point-to-point connections provisioned on demand but could also offer single broadcast domains to users who want to deploy their own Layer-2 private network connecting several sites. Underlying technology used will be VPLS or tagged VLAN (Virtual LAN).

UltraLight will dedicate a few Layer-2 channels to connect each site and offers IP (soon to include IPv6) services. UltraLight has its own address space (192.84.86.0/24) and autonomous system (AS, 32361). This will help to interconnect the UltraLight testbed to conventional IP networks and facilitate access to the testbed from sites not connected to UltraLight. UltraLight will peer with other backbones at Chicago, Los Angeles, New York and Seattle.

2.3.2 Data transport protocols

The protocols used to control the information flow across the network are one of the important areas UltraLight plans to explore. The most widely used protocol, especially for reliable data transport, is TCP. TCP, its variants, limitations and extensions will be examined by UltraLight in conjunction with the FAST team [12].

The UltraLight testbed is the ideal place to evaluate and test new TCP stacks at 10 Gb/s speed. Efficiency, the requirements and effect on end-hosts, the ability to coexist stably with other TCP implementations and the ability to share the bandwidth fairly will be evaluated. HSTCP [13], TCP Westwood+ [14], HTCP [15], and FAST TCP [16] are some of the new implementations we are going to test.

For the last three years, our team has been working closely with the FAST TCP team. The UltraLight testbed is an excellent opportunity to re-enforce the collaboration between the FAST team implementing new algorithms, and our experience on real Layer-2 and Layer-3 long-distance networks.

Another approach to overcome TCP's limitations is to use UDP-based data transport protocols. The best known protocol is UDT proposed by B. Grossman. Collaboration with the SABUL/UDT team is under discussion. Some servers dedicated to UDT tests have already been installed at CERN. Other servers may also be installed at Los Angeles and directly attached to the UltraLight backbone.

2.3.3 MPLS/QoS Services and Planning

UltraLight plans to explore the full range of end-to-end connections across the network, from best-effort, packet-switched through dedicated end-to-end light-paths. This is because the scientific applications supported by UltraLight have a wide variety of transfers that must be supported, ranging from the highly predictable (movement of large-scale simulated data between a few national centers) to the highly dynamic (analysis tasks initiated by rapidly changing teams of scientists at dozens of institutions).

Current network engineering knowledge is insufficient to predict what combination of “best-effort” packet switching, QoS-enabled packet switching, MPLS and dedicated circuits will be most effective in supporting these applications. We intend to engineer the most performant, reliable, and cost-effective combination of networking technologies, test them in a unique integrated environment, and lay the groundwork for deploying the resulting mix to meet the networking needs of the LHC community by first-collisions in 2007.

For UltraLight we plan to enable a combination of QoS on the LAN and MPLS “pipes”(network paths) across the network to support such intermediate flows. Using QoS and MPLS allows us to dynamically construct these pipes sized appropriately for the underlying application flow. We will work closely with the network control plane efforts within UltraLight to integrate QoS/MPLS configuration capabilities into our system. In addition we will be working closely with DoE funded efforts (like those of the TeraPaths project [17]) to find common extensible solutions to deploy and managing such virtual pipes across UltraLight.

2.3.4 Optical Path Management Plans

Emerging “light path” technologies are becoming more and more popular in the Grid community because they can extend and augment existing grid computing infrastructures, currently focused on CPU/storage, to include the network as an integral Grid component. Those technologies seem to be the most effective way to offer network resource provisioning on-demand between end-systems.

A major function we wish to develop in UltraLight nodes is the ability to switch optical paths across the node, bypassing electronic equipment if possible. For example, in a node that simply patches two 10 GbE paths together, there is usually no need to have the path go through two expensive ports on an Ethernet switch. Rather, the fiber cross connect provides the ability to bypass the electronics if possible. The ability to switch dynamically provides additional functionality and also models the more abstract case where switching is done between colors (grid lambdas).

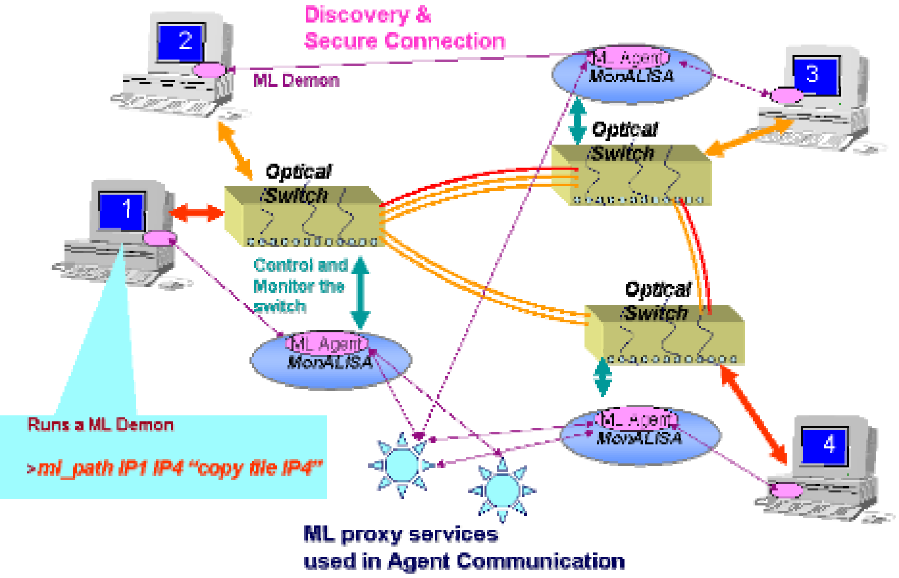
2.3.5 Optical Testbed

The California Institute of Technology and CERN have each deployed a photonic switch in their infrastructure and formed an optical testbed.

Since the number of transatlantic and transcontinental waves is limited, connections between the two sites are not deterministic, and bandwidth has to be shared with production and other experimental traffics. To overcome this limitation, the concept of a “virtual fiber” has been introduced, to emulate point-to-point connections between the two switches. A virtual fiber is a layer 2 channel with Ethernet framing. From the photonic switch, the virtual fiber appears like a dedicated link but the bandwidth, the path, the delay and the jitter are not guaranteed and the path is not framing-agnostic.

The goal of the testbed is to develop and test a control plane to manage the optical backplane of future networks with multiple waves between each node. The control plane being developed is based on the MonALISA (http://monalisa.caltech.edu), making it easy to interface with other environment such as HOPI or UltraNet. Within the MonALISA framework we developed dedicated modules and agents to monitor and control Optical Switches. These modules are used now for the CALIENT switch at CALTECH and the GLIMMERGLASS switch at CERN. The monitoring modules use the TL1 language to communicate with the switch and they are used to collect specific monitoring information. The state of each link and any change in the system is reported to dedicated MonALISA agents which are dynamically loadable modules running inside MonALISA services.

The distributed set of MonALISA agents is used to control the system. The agents use a discovery mechanism to find each other and they communicate with each other using proxy services. Each proxy service can handle ~1000 messages/sec and the architecture uses more than one such service, achieving very reliable communication between agents. The agent system is used to create a global path, or tree, as it knows the state of each link, inter-site connections, and the cross connections. The routing algorithm provides global optimization and can be extended to handle priorities and pre-reservations.



```
>ml_path IP1 IP4 "copy file IP4"
```

Figure 2

The system is integrated in a reliable and secure way with the end user applications and provides simple shell-like commands to map global connections and to create an optical path / tree on demand for any data transfer application. A schematic view of how the entire system works is shown in Figure 2.

2.3.6 Optical Exchange Point

Caltech and CERN have added a new dimension to their fiber cross connect points in Los-Angeles and Geneva. Each of these points of presence not only provides Layer 2 and Layer 3 connectivity, but now on-demand optical connections at Layer1 as well. This new architecture is a first step toward a hybrid circuit- and packet- switched network.

2.3.7 Network Monitoring

Network monitoring is essential for the UltraLight project. We need to understand our network infrastructure and its performance both historically and in real-time to enable utilization of the network as a managed robust component in our infrastructure. There are two ongoing efforts we intend to leverage to help provide us with the monitoring information required: IEPM and MonALISA.

As part of the UltraLight project we plan to install the Internet End-to-end Performance Monitoring (IEPM see http://www-iepm.slac.stanford.edu/bw/) toolkit at major UltraNet sites. This will provide a realistic expectation for network performance on the production networks between UltraLight sites, plus a powerful trouble shooting and planning tool.

The MonALISA framework will allow us to collect a complete set of network measurements and to correlate these measurements from different sites to present a global picture. We developed a real time network topology monitoring agent in the MonALISA system. It provides complete picture of the connectivity graphs and delay on each segment for routers, networks and AS.

2.3.8 Network Management and AAA

Since our project has many interactions with the HOPI project, we propose to follow the same implementation plan

- In phase-one, the control plan will be manually configured by the UltraLight engineering team for each request. Users' requests will be addressed via phone or email. Reconfiguration will be done by remotely logging into the network equipment. Bandwidth provisioning on other advanced backbones (such as HOPI) and interconnection configuration will be manually done by following procedures and mechanisms defined by each of the Network Operations Centers (NOCs). For example, during the first phase of the HOPI deployment, the UltraLight engineering team will address service requests to HOPI via emails or phone.
- In the phase-two, the network resources provisioning process is expected to be more sophisticated, automated and distributed. Provisioning software, appropriate protocols and routing/switching architecture will be deployed to locate suitable paths, schedule the resources in an environment of competing priorities, detect failures, etc. UltraLight will evaluate emerging light path technologies such as GMPLS (Generalized MPLS, targeted toward optical networks) [18], UCLP (User Controlled LightPath) [19] and deploy them as appropriate. Interfaces between the UltraLight environment and other environments will be developed.
- Phase-three will attempt to address the issues associated with building end-to-end light paths on a dynamic, global basis and doing so in an operationally sustainable fashion. Our intent is to provide dynamic light-path construction "on demand" as individual flows warrant their construction. Authentication, Authorization and Accounting (AAA) will play a crucial role in this phase.

UltraLight equipment should be accessible out-of-band via a conventional IP network such as LHCnet, Abilene or CENIC. Where an out-of-band access is not possible, UltraLight will try to provide an in-band access.

2.3.9 Disk-to-disk transfers: Breaking the 1 GByte/s barrier

One of the goals of UltraLight is to enable high performance disk-to-disk data transfers across the UltraLight network. This is a critical capability for data intensive science and an area we think we can make significant contributions in.

2.4 Milestones and Timeline

The initial UltraLight network was operational on February 1, 2005 with all sites to be connected by June 2005. There are a significant number of milestones for the Phase 1 network listed here:

1. Protocols:

    a. Integration and of FAST TCP (V.1) into testbed (July 2005)
    b. New MPLS and optical path-based provisioning methods (August 2005)
    c. New TCP implementations
        i. Working closely with the FAST TCP team; test the new TCP stack and give feedback for improvements (August 2005)
        ii. Testing new implementations like HSTCP, TCP Westwood+ or HTCP (September 2005)

2. Optical Switching:
    a. Install and commission optical switch at the Los Angeles CENIC/NLR (along with an optical switch at CERN) (May 2005)
    b. Develop dynamic connections of servers at the ends of a path to support Terabyte transactions with several 1G or 1-2 10G waves (September 2005)
3. Storage and Application Services:
    a. Evaluate and optimize drivers and parameter settings for I/O filesystems (April 2005)
    b. Evaluate and optimize drivers and parameter settings for and 10 GbE server NICs (June 2005)
    c. Selecting appropriate hardware
        i. Testing 10 GE network adapter (S2io, Chelsio, Intel)
        ii. PCI-X 266 & 533 GHz, PCIe
        iii. TCP offload engine
        iv. Raid controllers, disks
    d. Closing the gap between memory-to-memory and disk-to-disk transfers:
        i. Tuning end-systems
        ii. Fixing bugs in network drivers and file systems software
    e. Breaking the 1 GByte/s barrier:
        i. Experimenting with 802.3ad (filling a 10 Gbps pipe with a single pair of end-hosts)
        ii. Testing new PCI-express network adapters
    f. Compare 10 GbE NIC performance and CPU load (Intel, Neterion, Chelsio and others), with and without Transport Offload Engines (TOE) (August 2005)
4. Monitoring and Simulation:
    a. Deployment of end-to-end performance monitoring framework. (August 2005)
    b. Integration of tools & models to build simulation testbed for network fabric. (December 2005)
5. Agents:
    a. Start development of Agents for resource scheduling (June 2005)
    b. Match scheduling allocations to usage policies (September 2005)
6. Wan-In-Lab:
    a. Connect Caltech Wan-In-Lab to testbed (June 2005)
    b. Develop the procedure to move new protocol stacks developed in an instrumented network laboratory into field trials on the UltraLight testbed (June 2005)

The following table details the physical connection milestones over the next year for the UltraLight Network:

Date	Milestone (Bold indicates completion)
January 2005	**NLR Cisco wave connecting LA to CHI**
	BGP peering with Abilene & ESnet at CHI.
	MPLS upgrade at CHI.
	Two virtual fibers connecting CERN to Caltech
	End-systems at Chicago in UltraLight domain.
	Layer 2 connection to Abilene at Los-Angeles
	Extension of UltraLight Network to CERN
February 2005	**Connection to HOPI at 10 GE**
	Move Caltech switch to CENIC PoP
March 2005	10 GE link to UM via MiLR
	Connection to FLR at 10 GE
	Connection to BNL at OC48
May 2005	Connection to MIT at OC48
April 2005	“Manual” provisioning across the UltraLight backbone
August 2005	“A degree of automation” in the provisioning process.
September 2005	Connection to SLAC at 10 GE
	Switch NLR Cisco wave from exclusive to scheduled use (1 shift per day)

Table 2 Physical connection milestones for UltraLight

2.5 Year 3 and 4 outlook

The initial effort for the network focuses on creating the underlying monitored infrastructure which is the basis of UltraLight. Once a layer 1 and layer 2 network is operational among the core UltraLight sites, we intend to focus first on testing, integrating and hardening our core network services and capabilities and then on moving toward production with UltraLight.

By the middle of the third year we plan to have refined a beta version of Hybrid Network Provisioning (HNP) services which emphasize dynamically constructed optical light paths. The following list our goals for specific areas on this timescale:

1. ***Protocols***: Evaluate & optimize FAST TCP(Version 2), GridDT [20] and other TCP variants, integrated with MPLS and GMPLS and optical light path construction techniques.
2. ***Optical Switching***: Develop methods for wide optical switching (Lambda Grids) with Translight; particularly UIC, CANARIE, CERN, Netherlight and UKLight, using multiple optical switches
3. ***Storage and Application Services:*** Evaluate then-current generation 10GbE NICs, drivers and TOEs. Acquire and test servers with PCI Express buses.
4. ***Monitoring and Simulation:*** Refine end-to-end performance monitoring framework and integration with Global Services. Adapt system to more wavelengths and greater emphasis on optical paths.
5. ***Agents:*** Refine Agents for bandwidth and other resource scheduling, and matching allocation profiles to usage policies. Refine, deploy agents for global system optimization and policy-matching.

During the final 18 months of UltraLight we will focus on creating the first production HNP services to manage a combination of shared packet-switch and many dynamically constructed light-paths across the U.S., Atlantic and Pacific. Our goals for this time period are:

1. ***Protocols:*** Integration and deployment of production-ready "ultrascale TCP" stack with MPLS/GMPLS and optical lightpaths.
2. ***Optical Switching:*** Develop and deploy extensive methods for optical lightpath construction on demand across global wide-area network paths, together with partners. Develop full-scale production software matched to many wide area 10G waves (or several 40G waves, if available).
3. ***Storage and Application Services:*** Evaluate next generation 10GbE (or 40 Gbps) NICs, drivers and TOEs. Release new protocol stacks resident in onboard NIC processors.
4. ***Monitoring and Simulation:*** Production version of the full-scale end-to-end performance monitoring framework.
5. ***Agents:*** Production-ready Agent architecture for bandwidth and other resource scheduling, and global system optimization**.**

3 High Energy Physics Application Services

Besides the proposed work on network infrastructure and on new ways of provisioning the network, UltraLight provides the foundation for a coherent end-to-end environment for LHC data analysis, which is the primary focus of the Applications Technical Group. Within the complex and resource constrained environment, this group aims to enable an easy and coherent access to the wide range of data which comprises LHC physics, from high-level physics analysis to low-level detector studies.

Due to the limited resources available within the UltraLight project, the workgroup will focus primarily on the High Energy Physics CMS experiment to enable end-to-end analysis. Despite this restriction the group aims at creating generic interfaces that are broader applicable and can potentially be used within software stacks of other experiments.

The work within the Applications Technical Group is based on the CAIGEE[2] [21] project that resulted in the Grid Analysis Environment (GAE) architecture [22]. The GAE describes a high level architecture to support end-to-end physics analysis. Ultralight is extending the GAE to the Ultralight Analysis Environment (UAE) to make the network an integrated managed resource through end-to-end monitoring.

3.1 First 6 Months

The first phase of UltraLight was focused on implementation of the essential services and functionality:

- The core Clarens [23] Grid Service framework has been extended with a Shell service. The Shell provides a secure way for authorized clients to execute shell commands on the server. The command is executed by a designated local system user. The shell service allows so called "power users" to get low level access to resources. The shell service also enables the exposure of "local" applications through authorized, access controlled Grid Services. Besides utilization within the UltraLight project, the shell service is being used within the Monte Carlo Processing Service (MCPS) [24] and the Hot-Grid [25] project.

- The Sphinx scheduler [26] has been tested on scalability and a stable release has been integrated within the Java version of Clarens [27]. Scheduling decisions are based on information from MonALISA [28] montoring components.

---

[2] CAIGEE: CMS Analysis: an Interactive Grid-Enabled Environment

- A File catalog service has been created based on the interface developed within the LHC POOL project [29]. This catalog stores metadata, physical file names and logical file names associated with data sets.
- In collaboration with INFN[3], the BOSS[4] job submission tool [30] has been extended to operate in a distributed service environment. BOSS also provides detailed job monitoring information that is stored in a database. Work has started to integrate client analysis applications such as PhySH[5] and CRAB[6] into the UAE
- A first version of the steering service has been created to enable job interaction with the user or autonomous components during the execution of a job.
- A first version of the discovery service has been implemented within Clarens. Real time web service information is published within MonALISA [28]. The discovery service enables location independent interaction between services and is also part of the spring release of the open science grid (OSG) [31]
- Instant Message (IM) functionality has been integrated within Clarens to support interactive analysis: Applications and users can send asynchronous messages to their jobs. IM functionality will become useful within for example future versions of the steering service.

Although the services described above were not fully integrated, several of these services interact with the monitoring framework MonALISA and thus provide the first phase in establishing end-to-end monitoring as described in the UltraLight proposal. Most services described above, have been packaged and can be deployed using the Clarens service installer.

During supercomputing 2004 in Pittsburgh, a demonstration was given involving two farms that both contained the job submission service. Jobs where continuously submitted and progress was monitored via MonALISA and the job monitoring provided by the job submission application.

From January 10th until January 14th an application workgroup workshop was held at Caltech, to start addressing integration of the components discussed above and identified within the GAE architecture. A developers testbed was set up and the first components where deployed. At the end of the workshop Sphinx (scheduler) had been integrated with BOSS (job submission) and an analysis client was able to submit a dataset name and analysis code to the scheduler which would then execute the job at the location where the data resides.

The CODESH (Collaborative Development shell [32]) client has been integrated with a dedicated Clarens persistent back-end CVS server, and the analysis client could log working sessions on the server, to be accessed remotely by and shared with the members of a collaborating group.

3.2 Second 6 Months (February-July)

A first prototype end-to-end analysis system has been created, however the workshop showed several limitations that will be addressed in the upcoming months:

Secure high performance data transfer is important in order to support hundreds to thousands of users doing physics analysis. Not only the “raw” data transfer is important (e.g. number of Gigabytes) but also the administration of it through catalogs within a distributed service environment. Examples of administration include: updating and modifying catalogs, monitoring failed and

---

[3] INFN: Istituto Nazionale di Fisica Nucleare

[4] BOSS: Batch Object Submission System

[5] PhySH: Physics Shell

[6] CRAB: Cms Remote Analysis Builder

successful data transfers. PhEDEx[7] [33] is a data transfer and administration system developed within CMS for this purpose. Within these 6 months PhEDEx and its associated catalogs PubDB[8] and RefDB[9] [34] will be deployed at University of Florida (UFL), UCSD and Caltech to enable transfer of CMS data between regional centers. During this period, PhEDEx, PubDB and RefDB will be wrapped into a web service to enable authorized, access controlled access in a distributed service environment.

Integration between SPHINX and BOSS will be completed and a first prototype policy service will be created. This prototype will enable the specification of constraints on resource usage for individual users or Virtual Organizations (VOs).

Once users utilize services for analysis, it is possible that errors occur due to pathologies in the code a user submitted, the service itself, or other causes. It is important that users get feedback of these anomalies to detect errors in the distributed service environment created while executing jobs and tasks on the users' behalf. A first version of a "logger" service will be developed to enable users to store logging information about their sessions and use this information to "debug" the distributed services or job.

Packages developed in the previous 6 months and refined in these 6 months will be made available through the Virtual Data Toolkit. The (VDT) [35] is an ensemble of grid middleware that can be easily installed and configured. The goal of the VDT is to make it as easy as possible for users to deploy, maintain and use grid middleware. The VDT package is used by both CMS and ATLAS on many sites for deployment of software and an ideal vehicle for distribution of UltraLight components.

Integration between CODESH and Clarens will be extended by providing the full set of services available to a user connected to a local persistent back-end to users using Clarens web services to log and share their work sessions. The users will be able to move and copy session logs between private local stores and group stores based on Clarens. The CAVES project [36] provides functionality similar to CODESH for users performing data analysis with the ROOT analysis toolkit. The CAVES/ROOT client will be integrated with Clarens to provide a group log-book of persistent sessions complementary to local or cvs based persistent stores. In addition, the fast access of large data volumes from a ROOT/Clarens/CAVES (RC3) client to Clarens data servers over fast networks will be investigated.

3.3 Third 6 Months (August-January)

At the start of this period we will have the data movement[10] capability in place together with a scheduling/job submission prototype for submitting analysis jobs. Within these 6 months the focus will be at enabling users to perform analysis on the large datasets, and strategically moving data around for that purpose.

- Tools and services will be developed to enable users to select (partial) datasets and submit their analysis code and dataset selection to a grid scheduler.

---

[7] PhEDEx: Physics Experiment Data Export

[8] PubDB: Publication Database (for analysis data)

[9] RefDB: Reference Database (for Monte Carlo production)

[10] Data movement does not refer to using protocols like http or GridFTP to move arbitrary data, but refers to the capability of moving physics datasets (which consists of a collection of files) around while keeping track of their replicas and associated meta data

- (Authorized) users will be able to move datasets (a collection of files) from one site to another and be able to monitor the progress of the jobs they submit.
- Several large scale data movements (using real datasets) between sites will be performed and the performance of these data movements will be monitored (e.g. throughput, failure, etc). Based on these measurements first algorithms/heuristics can be developed to enable applications to use the network as a managed resource

3.4 Year 3 and 4 outlook

The necessary infrastructure is in place (network, schedulers, data movement software, etc) and several tests have been conducted in monitoring the movement of large datasets several other issues can be addressed:

- End to end error trapping and diagnosis: cause and effect. Give useful feedback to users when something goes wrong with their task in a distributed environment, but do not overwhelm them with information if things go right.
- Strategic Workflow re-planning:
- Adaptive steering and optimization algorithms for scheduling of jobs and (network) resources to enable efficient usage of these resources.

3.5 Synergistic Activities

The Clarens and MonALISA frameworks and its services are not only used within the UltraLight project but have a broader user and developer base. These synergistic activities enable the developers to get early feedback on services being developed and deployed, but also to reuse existing services and software.

MonALISA has been deployed on more than 40 farms world wide and is part of the VRVS monitoring system [2]. Recent developments include a distributed intrusion detection system and WAN topology layout.

Clarens is being used within several projects such as MCPS, Lambda Station, and Hotgrid .

The Monte Carlo Processing Service (MCPS) project at Fermilab addresses the difficulties associated with producing simulation data and processing results of simulation data outside of the official CMS production system. MCPS will be geared toward user-initiated production and analysis of simulated data, and provides a user friendly service front end to a collection of services involved in Monte Carlo production. MCPS is based on the workflow management tool RunJob and the Monte Carlo production tool MOP (a Monte Carlo Production database). The services that have been identified within the MCPS will be constructed using the Clarens web service framework.

The Lambda Station [37] project at Fermilab will allow clients to gain awareness of potential alternate network paths to a file transfer peer. Clarens based web services will provide clients with characteristics of such an alternate path that will help them decide whether or not to request use of the alternate path. The service will gather such information from network monitoring packages such as MonALISA. If the alternate path is requested and granted, Lambda Station will configure the necessary network equipment such that specified data flows are routed to the alternate path.

The Clarens based web service will have a client interface for configuring the allocation (or scheduling parameters) of the alternate path which will be based on, for example, Virtual Organization affiliation. Lambda Station will make use of the authentication and security utilities provided via the Clarens framework.

For many scientists, the Grid is perceived as difficult to use. In return for this difficulty is the promise of access to great computational power, but this is only available for a small number of dedicated people who understand both Grid and Science. Conscious of these difficulties, the HotGrid [25] project makes domain-specific easy of use the driving factor in developing science gateways. The approach to science gateways will provide a graduated and documented path for scientists, from anonymous user, to the weakly authenticated HotGrid user, through to the power user already comfortable with the Grid. The Clarens framework provides a basis for HotGrid science gateways, in developing web services and user interfaces on top of these services. Two science gateway prototypes have been developed within the HotGrid project for the astronomy community to support both image processing and data processing.

4 Education and Outreach Status

The education and outreach component of UltraLight has been designed and implementation plans are underway. The goal is to train aspiring undergraduate and graduate computer and discipline science students in state of the art network and distributed system science and technologies using the UltraLight interactive toolkit. To accomplish this, we will hold yearly one-week tutorial workshops followed by immersive collaborative research projects based on the toolkit. This framework has the added advantage of providing rigorous testing of UltraLight tools by a set of first adopters.

The workshop will be held annually at the beginning of summer. Topics will include network engineering, network research and monitoring, and applications, *ie*, those underlying the UltraLight project. Speakers and tutorial leaders will be composed of the proposed GISNET PIs and Senior Researchers as well as several invited guests. Support for 15 students is included in the project with additional self-supported participants welcome as well.

Research projects will follow the workshop to further immerse students in the UltraLight experience. Students from all UltraLight institutions will prepare collaborative projects utilizing the UltraLight toolkit. Their participation will be structured as collaborators in the project, where projects will be defined in terms of outcome with students working as part of the research community. The goal is to emulate the professional researcher role, thus providing deep insights into the nature of network research, and its key role and impact on leading-edge international projects in the physics and astronomy communities.

Project teams will be organized and assignments detailed during the workshop. The research projects will begin upon students returning to their home institution. Groups will maintain communication among themselves and with the collaboration through regular VRVS conferencing and the persistent VRVS UltraLight meeting space. At the end of the summer, results will be presented to the collaboration and archived.

Support for the summer research projects will be provided through several identified existing REU-like programs as well as a dedicated REU proposal to be submitted. Limited support is also included in year one of the UltraLight budget for a prototype program.

At the present time, planning is underway for the summer workshop. Date selection will be complete by the end of February and the agenda is being developed. A research project web page is being developed where collaborators can identify projects suitable for undergraduates. Time estimates, level of complexity, and priority will be collected through the web form.

5 Physics Analysis User Community

The physics analysis user community in UltraLight is completing the initial phase of planning, to be followed by implementation plans. Our main goal is to establish in the next six months a community of early adopters and users. They will come first from within UltraLight, and can be considered as expert users; as often is the case, they will have roles overlapping to some extent with the HEP application services group. Later, with the maturing of the UltraLight software, the community will grow to include outside users. This community will use the system being developed e.g. will start actual physics analysis efforts exploiting the test-bed, and provide a user perspective on the problems being solved.

The user community group will organize the early adoption of the system. It will play an active role by identifying the most valuable features of the system from the users' perspective, to be released early at production quality level, or at useful level of functionality. This is "where the rubber will meet the road", and will provide rapid user feedback to the development team.

A key component of this group's work will be an evolving dialog with the HEP application services group. As a result if this collaboration the group will have it's input on the planning and scope of software releases, guided by what is expected to be most valuable for physics analysis and aligned with the milestones of the experiments, and helping to set priorities for implementing features.

We envisage the development, in collaboration with the applications group, of an expanding suite of functional tests. In contrast to unit tests, they provide a user view on the system, and can be very useful for measuring the progress of the project, as well as for educating new users and making it easier for them to pass the threshold for adopting the system. Users should be encouraged to provide new tests for important new features under implementation.

The physics analysis user community will study in depth the software framework of HEP applications (e.g. ORCA/COBRA for CMS, ATHENA for ATLAS), the data and metadata models of the experiments, stressing commonality where feasible and/or practical, and the steps to best integrate the UltraLight services in the experimental software systems. As the frameworks of the experiments are evolving and changing rapidly right now, this will be a continuing activity, which can provide an important input to the experiments about the exciting new possibilities which will be made possible by UltraLight.

In order to arrive to optimal integration of the UltraLight activities in the software systems of the experiments we will maintain close contacts with the people in charge of the software development in the experiments and respond to their requirements and needs.

The activities of the physics analysis user community will come fully to fruition by contributing to the ATLAS/CMS Physics preparation milestones, utilizing the services developed by UltraLight. UltraLight members are already active in LHC physics studies and are leading several analyses, officially recognized in CMS for the Physics Technical Design Report, which will be completed by the end of 2005. CMS will conduct a Data Challenge in 2006 to prepare for first beams in 2007. The activities of ATLAS users and milestones, such as Data Challenge 3 and the Physics Preparedness Review, will also be aligned with UltraLight work.

6 Relationships between UltraLight and External Projects

6.1 WAN-in-LAB Liaison Group

WAN-in-LAB is a high-speed long-haul optical network testbed built to aid protocol research, such as FAST TCP, by providing multi-Gb/s bandwidth, real propagation delay and active real-time monitoring. WAN-in-LAB includes an array of 4 Cisco 7609 core routers, 13 ONS 15454 chassis, and hundreds of line cards interconnected via OC-48/192 and 1&10GbE.

Over the last 6 months, WAN-in-LAB is becoming a reality. We finalized our network design and topology such that it matched our budget. We have obtained strong support from Cisco Systems to provide the bulk of the networking equipment, and from Corning to provide over 2400 km of high performance LEAF [38] fiber. In addition, we worked with various engineers and contractors on the physical build out of the lab including power and cooling requirements.

The physical construction of lab officially started on October 6, 2004 and was completed December 13, 2004. Since then the actual networking components have been installed and are currently being tested. The testing phase should last until the end of March 2005, when we will move into the production phase. We expect to connect WAN in Lab to the UltraLight infrastructure by Fall 2005, making it an integral part of the global research and education testbed. WAN in Lab will serve as a development platform for network debugging where programs that are too disruptive to run on production networks can be debugged in WAN in Lab before they are tested and demonstrated on real testbeds. The topology, routing, delay in WAN in Lab are also configurable, providing a flexible environment for debugging and testing.

6.2 Synergies with the Proposed DISUN Shared Cyber-infrastructure

DISUN (Data Intensive Science University Network), is a proposed integrated tier of university-based distributed regional computing centers (“Tier-2”) for the CMS experiment and other science communities that have massive data storage and data movement requirements. One of the central DISUN concepts is that of a data cache, distributed amongst four DISUN Tier-2s (Caltech, U. California-San Diego, U. Florida, and U. Wisconsin-Madison), enabling data analysis and simulations to be performed which require much larger resources and throughput than is available locally. UltraLight provides the critical network advances which are required to link the four DISUN sites together, forming the coherent distributed data cache amongst the Tier-2 regional computing centers. Further, due to an already existing close working relationship, UltraLight will naturally work intimately with DISUN in the ground-braking work required to provide the early Application Services for High Energy Physics, enabling the very first physics analyses at LHC turn-on and subsequent low-luminosity physics runs. In return, DISUN will expand the Application Services significantly beyond the UltraLight scope, developing and deploying commonly useful tools of increasing sophistication for Grid-based analysis, and eventually delivering a full scale Grid Analysis Environment on a timescale compatible with the evolution to full LHC luminosity (and hence full capability) of CMS data analysis.

6.3 Connections to Open Science Grid

The Open Science Grid (OSG) aims to create a large-scale, multi-disciplinary U.S. Grid computing infrastructure and to provide an open and pragmatic governance structure to be attractive to the broadest possible audience. UltraLight is partnering with the OSG to become an integrated infrastructural component of OSG. By partnering with OSG, UltraLight will have a much broader impact across a wide range of scientific and educational disciplines. To that end, UltraLight is planning to take on leading roles within the OSG development in the testing of new networking concepts and services. In particular, UltraLight researchers already actively participate in several OSG technical groups such as the Monitoring and Information Services Group and we are currently in the early stages of forming and leading a new OSG Networking Technical Group.

7 Justification for the Early Release of Year 2 Funding

In response to partial funding of the original UltraLight proposal, we crafted the budget profile to maximize the project's impact. We kept virtually all of the UltraLight chasses foreseen, to take full advantage of the donation of equipment provided by Cisco Systems, and phased in the manpower (at a reduced level). This plan enabled us to get a rapid start and in many cases we have exceeded our goals for our first "year" of funding, as presented at NSF January 26 in detail, but left us with only a few months of FY04 funding after the September 2004 start date.

The push to achieve our planned milestones and goals as rapidly as possible is driven by the requirement to have the UltraLight infrastructure "in place" to meet the tight schedule for completion and turn-on of the Large Hadron Collider in a little over two years from now. The UltraLight network is already beginning operation, our management structure is in place and we have well defined work plans for each of our primary areas within UltraLight. To maintain our ramp-up and continue to synchronize with the LHC schedule, we need to continue the project tempo that we have already set during this Year 1. Advancing the funding schedule and release of the funds within the next two months will allow us to effectively capitalize on the efforts we have made to date, proceed with the project as foreseen, and to continue to meet our project milestones on schedule.

8 Summary

During these first five months, we have started a rapid network testbed deployment, begun to integrate our activities within High Energy Physics CMS collaboration, and established close ties to external groups and industrial vendors. The UltraLight network is already beginning operation, our management structure is in place and we have well defined work plans for each of our primary areas within UltraLight. Finally, UltraLight physicists are already active in LHC physics studies and are leading several analyses, officially recognized in CMS for an upcoming Physics Technical Design Report. In order to efficiently profit from this and stay in phase with the LHC schedule, we are asking for an early release of the Year 2 funding for the UltraLight Project.

In addition to the stated work plans for Year 2, several important activities are planned or already underway in 2005, including: forming a Network Working Group in Open Science Grid, preparing a major presence at Super Computing 2005 (to be held at Seattle) in cooperation with Caltech CACR and CERN IT, and participating in iGRID2005. In addition, we anticipate a strong participation in the GridNets 2005 Conference.

Appendix: Publications Related to UltraLight Research

[1] H. Stockinger, Flavia Donno, Giulio Eulisse, Mirco Mazzucato, Conrad Steenberg, "*Matchmaking, Datasets and Physics Analysis*", Submitted to Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA), June 14-17, 2005

[2] F. van Lingen, J. Bunn, I. Legrand, H. Newman, C. Steenberg, M. Thomas, A. Anjum, T. Azim, *"The Clarens Web Service Framework for Distributed Scientific Analysis in Grid Projects"* , Submitted to Workshop on Web and Grid Services for Scientific Data Analysis (WAGSSDA), June 14-17, 2005

[3] "*Meeting the Challenges of High-Energy Physics, (How the UltraLight Consortium is Finding Answers to the Universe's Oldest Questions)*" CENIC Interact Winter 2005, Partnership award extract

[4] Dimitri Bourilkov, Julian Bunn, Rick Cavanaugh, Iosif Legrand, Frank van Lingen, Harvey Newman, Conrad Steenberg, Michael Thomas " *Grid Enabled Analysis for CERN's CMS Experiment at the Large Hadron Collider*", GlobusWorld 2005, Boston

[5] M. Thomas, C. Steenberg, F. van Lingen, H. Newman, J. Bunn, A. Ali, A. Anjum, T. Azim, W. Rehman, F. Khan, J. In, "*JClarens: A Java Framework for Developing and Deploying Web Services for Grid Computing"*, Submitted to the International Conference on Web Services, 2005, Orlando.

References

---

[1] GISNET proposal at http://ultralight.caltech.edu/gaeweb/portal/proposals/2005/01GISNET/GISNET.doc

[2] VRVS, http://www.vrvs.org

[3] National Lamba-Rail (NLR) homepage http://www.nlr.net

[4] The Hybrid Optical and Packet-switched Infrastructure (HOPI) project homepage http://networks.internet2.edu/hopi

[5] The Abilene network homepage http://abilene.internet2.edu

[6] The Energy Sciences Network homepage http://www.es.net

[7] MPLS Resource Center homepage http://www.mplsrc.com/

[8] CERN OpenLab project homepage http://proj-openlab-datagrid-public.web.cern.ch/proj-openlab-datagrid-public

[9] LambdaStation project homepage http://www.lambdastation.org

[10] See http://www.ncdm.uic.edu/sabul.html

[11] Michigan Light Rail (MiLR) See http://www.lib.umich.edu/aael/news.php?newsID=58

[12] The FAST homepage http://netlab.caltech.edu/fast

[13] S. Floyd "*HighSpeed TCP for Large Congestion Windows*" RFC 3649, Experimental, December 2003

[14] L. A. Grieco and S. Mascolo,"*A Mathematical Model of Westwood + TCP Congestion Control Algorithm*",18th International Teletraffic Congress 2003 (ITC 2003)

[15] R.N. Shorten, D.J. Leith, J. Foy, R. Kilduff, "Analysis and design of congestion control in synchronized communication networks" Proc. 12th Yale Workshop on Adaptive & Learning Systems, May 2003

[16] S. H. Low, F. Paganini, J. Wang and J. C. Doyle, "Linear Stability of TCP/RED and a Scalable Control," Com-puter Networks Journal, 43(5):633-647, December 2003

[17] See http://www.atlasgrid.bnl.gov/terapaths/

[18] GMPLS Resource Center web page http://www.polarisnetworks.com/gmpls/

[19] UCLP Software web page http://www.canarie.ca/canet4/uclp/uclp_software.html

[20] GridD(ata)T(ransport): For more information and downloads see http://sravot.home.cern.ch/sravot/GridDT/GridDT.htm

[21] CAIGEE, http://pcbunn.cacr.caltech.edu/GAE/CAIGEE/default.htm

[22] GAE Architecture, http://UltraLight.caltech.edu/gaeweb/portal/proposals/2004/01GAE/GAEArchitecture.pdf

[23] C. Steenberg, J. Bunn, I. Legrand, H. Newman, M. Thomas, F. van Lingen, A. Anjum, T. Azim "*The Clarens Grid-enabled Web Services Framework: Services and Implementation*" CHEP 2004 Interlaken

[24] MCPS, http://www.uscms.org/SoftwareComputing/Grid/MCPS/

---

[25] R. Williams, C. Steenberg, J. Bunn, " *HotGrid: Graduated Access to Grid-based Science Gateways*",Proceedings of IEEE Supercomputing Conference, Pittsburgh, 2004

[26] J. In., P. Avery, R. Cavanaugh, M. Kulkarni, S. Ranka, ,"*SPHINX: A Scheduling Middleware For Data Intensive Application on a Grid*," In the proceedings of Computing in High Energy Physics, Interlaken, Switzerland, 2004

[27] A. Ali, A. Anjum, R. Haider, T. Azim, W. ur Rehman, J. Bunn, H. Newman, M. Thomas, C. Steenberg "*JClarens: A Java Based Interactive Physics Analysis Environment for Data Intensive Applications*." In the Proceedings of ICWS, the International Conference of Web Services, San Diego, USA, 2004

[28] Legrand, I., Newman, H., Galvez, P., Voicu, E., Cirstoiu, C., "*MonaLISA: A Distributed Monitoring Service Architecture*", Computing for High Energy Physics, La Jolla, California, 2003

[29] Duellman, D. "*POOL Project Overview*", Computing for High Energy Physics, La Jolla, California, 2003

[30] Grandi, C., Renzi, A., "*Object Based System for Batch Job Submission and Monitoring (BOSS)"*, CMS note 2003/005

[31] OSG, http://www.opensciencegrid.org

[32] D. Bourilkov, ``*THE CAVES Project - Collaborative Analysis Versioning Environment System; THE CODESH Project - Collaborative Development Shell*'', arXiv:physics/0410226 ; To be published in Int. J. Modern Physics

[33] T. Barras, A. Afaq, W. Jank, O. Maroney, S. Metson, D. Newbold, K. Rabbertz, J. Rehn, L. Tuura, T. Wildish, Y. Wu, C. Grandi, D. Bonacorsi, C. Charlot, M. Ernst, A. Fanfani, I. Fisk, "*Software agents in data and workflow management"*, In Proceedings of CHEP, Interlaken Switzerland 2005

[34] V. Lefebure, J. Andreeva, "*RefDB: The Reference Database for CMS Monte Carlo Production*" In Proceedings of CHEP La Jolla, California, 2003

[35] Virtual Data Toolkit, http://www.cs.wisc.edu/vdt//index.html

[36] D. Bourilkov, "*The CAVES project: Exploring virtual data concepts for data analysis*", arXiv:physics/0401007 ; presented at ROOT 2004 Users Workshop, SLAC, Stanford, USA, 2004

[37] Lambda station, http://www.lambdastation.org/

[38] LEAF fiber: http://www.corning.com/opticalfiber/products__applications/products/leaf.aspx