

Appendix C: Excerpt from “Workshop on New Visions for Large-Scale Networks”¹

2.6 High-Energy Physics

High-energy physics (HEP) has pushed against the limits of networking and computing technologies for decades. Twenty years ago, the largest HEP experiment involved 100 physicists from many nations and acquired tens of thousands of magnetic tapes of data per year; graduate students spent months reading those tapes to perform data queries. Life is not so different for today’s physicists. The new BaBar detector at the Stanford Linear Accelerator (SLAC) was designed by a large international collaboration of physicists at institutions. The BaBar collaboration enables hundreds of physicists worldwide to query its 300-terabyte and rapidly growing database in hours or days rather than months. In the next 10 years, the Large Hadron Collider (LHC) experiments at CERN, the European Physics Laboratory, where some 600 U.S. physicists form the largest national group, will face the challenge of distributed analysis of hundreds of petabytes of data.

2.6.1 High-Energy Physics Scenario

The physics community greatly values being able to distribute digitized data electronically at the rate at which it is produced from the site of an experiment to collaborators worldwide who can analyze them. The HEP community has the goal of using affordable network and computational resources to provide physicists with transparent access to a distributed data analysis system that uses all available resources as efficiently as possible. By 2005 to 2010, HEP computing will involve queries on databases containing exabytes (10^{18} bytes) of data structured as up to 10^{16} individually addressable objects. These massive amounts of data will require the distribution of terabits per second of real-time data to major HEP data analysis centers.

Challenging networking and other information technology research needed to enable distribution of data, analysis, and collaboration includes:

- Multicast service delivered to multiple remote centers with diverse firewall filters
- Network error rate and robustness control *without* impacting the experiment’s data acquisition system
- Massive applications software – e.g., 3 million lines of BaBar C++ code
- Commercial object database management software
- Interfaces of the database with the network and storage
- Technology improvements including:
 - Computing technologies
 - Computer science
 - Networking

¹ “Workshop on New Visions for Large-Scale Networks: Research and Applications”, Large Scale Networking (LSN) Coordinating Group Of the Interagency Working Group (IWG) for Information Technology Research and Development (IT R&D), March 12-14, 2001, Vienna, Virginia

- Computing system-to-network interfaces
- Fiber technologies
- Data storage

Improvements in HEP applications must be accomplished at minimal incremental costs. To help contain costs, network engineering labor, required to configure, optimize, and maintain networks, should be minimized by developing automated network engineering and management.

HEP collaborations are increasingly international in composition. It is difficult to adopt standards across the resulting international boundaries, so that the implementation of uniform, collaboration-wide middleware, security, or hardware technologies is almost always unrealistic. The best that can be achieved is the adoption of a set of protocols and interfaces to link components that will almost certainly be implemented in different ways.

The international HEP research community is increasingly using Grid technologies, an integrated suite of services developed with Federal IT R&D funding. The Grid is a set of middleware tools and capabilities that enable seamless end user access to applications, data storage, and compute resources to support high-end modeling. The Globus project (<http://www.globus.org/>) is one state-of-the-art example of Grid development. Grid middleware faces many hard computer science problems. Vertical integration of existing components to provide Grid services to demanding, well-defined communities is essential to progress on Grid architecture and technologies.

2.6.2 High-Energy Physics Networking and Networking Research Needs

Networking underlies many of the services and applications being developed to support HEP. Progress in networking is expected to be evolutionary over the next five years, with revolutionary capabilities being developed over the longer term. The following table presents the current state of the art in various networking areas supporting HEP, what could evolve by around 2006, and the requirements to approach meeting the HEP goals. The current HEP capabilities are what is affordable, not what could be obtained with unlimited funding.

Current HEP Capabilities	Evolution to 2006	HEP Goals
<u>Links Between Major Centers</u> • 1 or 2 x 1.55 Mbps	• 10 Gbps	• 1 Tbps
<u>Bulk Transfer Protocol</u> • TCP/IP + fixes	• TCP/IP + more fixes	• New, <i>widely adopted</i> , transport protocol
<u>Differentiated Services (CoS, QoS, Mixture of Packet and Circuit Switching, etc.)</u> • Provide ~1.1 differentiated services (best effort + some Voice over IP (VOIP))	• Provide ~2 differentiated services	• Provide ~6 differentiated services that are application-negotiated, on-demand, and responsive to cost and policy
<u>Network Measurement, Analysis, Interpretation, and Action and Network Modeling</u> • Limited measurement, analysis, and modeling	• More/better measurement and analysis, and some interpretation • Models begin to predict non-obvious failure modes	• Automated measurement, analysis, and interpretation • Automated action based on measured and modeled information
<u>Support for Collaboration</u> • Some proof-of-concept (PoC) prototypes • Some commercial tools	• New PoC prototypes • Some mature components • Still incomplete	• Collaborations form via the Internet • Real sense of working together
<u>Data-Grid: Authentication and Authorization</u> • Local and manual	• Cross-authentication via proxies	• New approaches to regulating access to resources
<u>Data-Grid: Information Infrastructure (Replica Catalog, Resource Catalog, Software Catalog, Operation/Task Catalog, etc.)</u> • Manual and local • Limited <i>ad hoc</i> automation	• Evolution of Globus by 2+ generations	• Efficient distributed information management for more than 10^{16} virtual objects using millions of operations each using millions of lines of code (MLoC)
<u>Data-Grid: Data Payload Infrastructure (Exabyte Databases, Reliable Replication, Storage Management, etc.)</u> • Few x 100-Tbyte databases • PoC replication prototypes • PoC storage management	• Bleeding-edge 10^{19} Byte databases • Grid replica management • Grid storage management	• Industry-standard exabyte databases, replication, and storage management
<u>Data-Grid: Resource Discovery</u> • Telephone, e-mail	• Telephone, e-mail, partial automation	• Automated discovery • Standardized information models
<u>Data-Grid: Distributed Resource Management, Distributed Job (Task, Operation) Management</u> • Local batch systems • Prototype systems	• Grid job management • Early distributed resource management	• New approaches to regulating access to resources
<u>Data-Grid: Virtual Data</u> • Conceptual phase	• Starting to work for cutting-edge HEP experiments	• A generally accepted and implemented paradigm
<u>The Grid an Integrated "Network" Service</u> • Manually integrated services have been in use for more than 10 years	• Vertical integration of fabric and data payload services • Incomplete information services • Incomplete resource management services	• Easy creation of vertically integrated, worldwide information management and processing systems from standard industry components

Notes:

1. HEP technologies that work well locally but do not become widely adopted and supported may inhibit collaboration and prove costly. Qualifiers like "widely adopted, industry-standard," and "generally accepted" are vitally important.
2. Elegant approaches to authentication and authorization appear to be available for organizations that are part of a single administrative structure. Worldwide collaborations seem unlikely ever to fit this model. Discussion identified that a totally new approach to regulating access to resources might foster more open scientific research.

The Role of Industry in HEP Networking R&D

Wherever possible, high-end science takes advantage of capabilities that are developed and commercialized by industry. For example, the HEP community has benefited from

cost reductions and reliability increases provided by industrial commercialization of individual middleware components, such as databases and well-defined information systems. Also, the HEP community has benefited from the availability of commercial high-end computing systems, high-bandwidth networks, and extensive middleware. It is likely that higher bandwidth will be more affordable in the future due to economies of scale, greater supply, and competition among providers. Carriers are beginning to make individual wavelengths available to major customers. Affordable links between major HEP computer centers should exceed 10 Gbps within five years and may approach 1 Tbps in less than a decade. However, it is likely to be difficult to exploit the available bandwidth using industry-standard transport protocols. TCP/IP requires fixes such as multiple streams to use today's affordable bandwidth. Additional fixes will be needed to accommodate the expected increases in numbers of users, number of nodes, and network traffic. It is possible to develop a new protocol or to extend TCP to work over dedicated links, but the extensive investment of industry and users in the current protocols would likely hinder acceptance of alternatives.

Workshop participants identified a need for a vertically integrated HEP solution for managing and processing the massive amounts of data expected from HEP experiments. Networking research, development of faster computing systems and more capable computational algorithms, and commercial development and marketing (productization) together deliver components that provide part of this vertically integrated HEP solution. New component technologies emerging from networking research and computer science are funded normally only to the proof-of-concept stage and fall short of the level of product hardening and support needed to provide technologies that can be reliably integrated into a complex operational system. Collaboration by network researchers, computer scientists, and application scientists required to provide vertical integration of the component capabilities are also research and development and, in the view of the workshop participants, should be funded by the Federal IT R&D funding agencies.

The HEP community is rapidly taking advantage of the Grid infrastructure to enable transparent, distributed, and international collaborations, resulting in improvements in the ability to cooperatively carry out science and to analyze increasingly large volumes of HEP data. However, the Grid primarily has been developed in universities and industry is currently largely decoupled from development of an integrated Grid capability. Thus, Grid software and infrastructure have not benefited from the standardization, cost reductions, and increased reliability often provided by commercial productization. This productization will take place only if industry perceives the potential for profitably marketing the technologies. Federal funding could help bridge the gap between the proof-of-concept prototype and the point at which successful vertical integration has demonstrated commercial viability.