Virtual Laboratories:
Collaborative Environments and Facilities-On-Line

C.E. Thomas, Jr., J.S. Cavallini, G.R. Seweryniak, R. J. Aiken,
T.A. Kitchens, D.A. Hitchcock, M.A. Scott, L.C. Welch, R.L. Stevens

Oak Ridge National Laboratory*
P.O. Box 2008
Oak Ridge, TN 37831-6006
(615)574-1155

IEEE 1995 Conference on
Real Time Computer Applications
in Nuclear and Plasma Sciences
East Lansing, Michigan
May 21-26, 1995

*The submitted manuscript has been authored by a contractor of the U.S. Government under contract No. DE-AC05-84OR21400. Accordingly, the U.S. Government retains a non-exclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

Managed by Martin Marietta Energy Systems, Inc., for the U.S. Department of Energy under contract DE-AC05-84OR21400

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.
Virtual Laboratories:  
Collaborative Environments and Facilities-On-Line

C.E. Thomas, Jr.,  
I&C Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831

J.S. Cavallini, G.R. Seweryniak, R.J. Aiken, T.A. Kitchens, D.A. Hitchcock,  
M.A. Scott, L.C. Welch  
Mathematical, Information, and Computational Sciences Division,  
USDOE, ER-31, 19901 Germantown Rd., Germantown, MD 20874

R.J. Aiken, NERSC, Lawrence Livermore National Laboratory, Livermore, CA  

R.L. Stevens  
Mathematics and Computer Sciences Division, Argonne National Laboratory

V. 2.0, May 23, 1995

ABSTRACT

The Department of Energy (DOE) has major research  
laboratories in a number of locations in the United States,  
typically co-located with large research instruments or  
research facilities valued at tens of millions to even billions  
of dollars. Present budget exigencies facing the entire nation  
are felt very deeply at DOE, just as elsewhere. Advances  
over the last few years in networking and computing  
technologies make virtual collaborative environments and  
conduct of experiments over the internetwork structure a  
possibility. We believe that development of these  
collaborative environments and facilities-on-line could lead  
to a "virtual laboratory" with tremendous potential for  
decreasing the costs of research and increasing the  
productivity of our capital investment in research facilities.  
The majority of these cost savings would be due to increased  
productivity of our research efforts, better utilization of  
resources and facilities, and avoiding duplication of  
expensive facilities. A vision of how this might all fit together  
and a discussion of the infrastructure necessary to enable  
these developments is presented.
I. What is a Virtual Laboratory??

There will be many definitions of “virtual laboratory” and perhaps hundreds of implementations as researchers evolve their ideas. Dick Kouzes, our colleague at Pacific Northwest Laboratories (PNL), has recently published a paper on “Creating the Cyberspace Laboratory” in which he uses the term “collaboratory,” a “‘center without walls’ in which the nation’s researchers can perform their research without regard to geographical location—interacting with colleagues, accessing instrumentation, sharing data and computational resources, and accessing information in digital libraries.” This term was coined by Professor William Wulf of the University of Virginia in 1989. So not only do we dream of Virtual Laboratories in 1995, but others have been dreaming of them since at least 1989, probably much earlier.

The Galvin Report on the Department of Energy’s National Laboratory System makes several mentions of the term Virtual Laboratory, including:

*The Task Force also recognizes that there is considerable potential in achieving greater coordination of R&D expertise across the lab complex. The national laboratories are equipped with the information technologies and the culture of communication via computer networks which could provide the basis for close programmatic integration. The Task Force believes that the goal of efficiency in utilization of the national laboratories mandates that these institutions be managed better as a system, and that complementary strengths be integrated to the extent possible through the establishment of “virtual laboratories” via computer networks and lead laboratory assignments. As the laboratories are given more discrete missions which result in enhanced R&D focus, the creation of “virtual laboratories” will be an important means for retaining complex, multi-disciplinary approaches to problem solving across the laboratory system.*

In the summer of 1994, the Lawrence Berkeley Laboratory of DOE, representing the then DOE Office of Energy Research, Office of Scientific Computing (DOE/ER/OSC; the new name is considerably longer, Office of Energy Research/Office of Technology and Computational Research/Mathematical, Information, and Computational Sciences Division, DOE/ER/CTR/MICS) issued a Request for Proposals (RFP) for research on “Distributed Collaboratory Experimental Environments.” The theme of this RFP was to produce tools useful in distributed collaborative environments for the DOE national laboratories. One of the winning proposals resulting from this RFP was “LabSpace, A National Electronic Infrastructure,” from Argonne National Laboratory in
collaboration with Northeastern University. This proposal outlines in some considerable detail proposed research to develop virtual electronic laboratory spaces to "...be a center of activity, creating an electronic community enabling researchers to congregate and interact." 4

In late March of 1995, fusion researchers from the Alcator C-Mod tokamak experiment at Massachusetts Institute of Technology (MIT), in collaboration with researchers at Lawrence Livermore National Laboratory (LLNL), successfully operated the MIT fusion experiment from 3,000 miles away:

Over several days, MIT staff, ..., took over full physics control of the C-Mod tokamak. The remote staff controlled all waveforms for plasma current, position, shape, and density as well as RF power waveforms and the operation of a fast scanning Langmuir Probe. Experimental data was available to the remote team using the same software tools that are used in the C-Mod control room. We used a video link, running over ESnet, (the DOE Energy Sciences network) to help coordinate activities between the remote and local control rooms. Livermore personnel participated in the experiment and provided computers and software support, ..., The purpose of the experiment was to demonstrate the feasibility of remote operations and to begin addressing the communications and human interaction issues that will be crucial for future distributed control rooms 5.

While many of the elements that we believe will be necessary for remote control and participation were missing from this demonstration, it was indeed a powerful and convincing example of what can be done over a well-run and reliable internet.

It is our belief that a "distributed collaborative environment" plus "facilities-on-line" will equal a "virtual laboratory," and we will discuss these ideas here.

Goal: To improve the efficiency and productivity of our national laboratories by making all forms of collaborative work and sharing of resources among our hard working researchers easy, attractive, available, and ubiquitous.

Vision: To define a "virtual laboratory," where the boundaries of space, time, distance, local facilities, and researchers on-site are irrelevant, it is only necessary to imagine a real modern laboratory and go as far beyond that as thought and creativity will take you.
• In a virtual laboratory researchers will be able to easily and
instantaneously locate one another. They will be able to leave
messages for one another in persistent electronic spaces designed
for experimental work, or for data visualization, or for theoretical
investigations, or collaborative writing of papers, or whatever work
their minds and hearts are intent on.

• Locating other researchers interested in a particular topic, or all of
the research papers available on that topic, will be a matter of a
few keystrokes, or a few mouse-clicks, or perhaps just speaking up
and asking your computer for "all the research in the last 10 years
on plasma turbulence in magnetic fusion, please."

• Having found another researcher in an area of interest, it will be
possible to send them e-mail, or if they are "in" (where they
indicate to the system if they want to be in or not), to invite them
to a video-conference or a suitable virtual collaborative electronic
space to find out the latest on their research and how it might be
applied to the present problem.

• It will be possible to share experimental facilities with researchers
10,000 miles away or with a university or high school class five
states away.

• Seminars by world experts on the latest work in passively safe
nuclear reactors, massively parallel computing with DNA, or
tunable X-ray light sources will be available over the network
during the seminar, the day after, or 5 years after the time the
seminar is actually delivered.

• Experimental control of instrumentation or facilities will be
accomplished as necessary across oceans or state and
international boundaries.

• Resources of all kinds, including massively parallel computers,
"CAVEs" (CAVE Automatic Virtual Environment, an immersive 3-D
environment with control wand) for three-dimensional immersion
in experimental simulations, or the latest in heat transfer analysis
software will be sharable and schedulable across the network.
Sites that have spare compute cycles will be able to make them
available to students or researchers in "compute-challenged"
locations.

• All of the research work that we do now, and work that we have yet
to imagine will be accomplishable over the network through our
"distributed collaborative environment" and "facilities-on-line"
implementation of "virtual laboratories." Researchers will be able
to pursue their ideas from their homes, from a conference in
II. A Distributed Collaborative Environment

A. Some of the Pieces

1. Universal Directory Services

Those of us who surf the Internet are well aware of how much driftwood and how many abandoned and neglected shells exist out there. While it might be claimed that “all of the information is useful to somebody,” it is certainly also true that “Not all of the information is useful to everybody.” To have a truly useful network for interactive collaborations, remote measurement and experimentation, and instantaneous communication, it is crucial that easy-to-use, fast, and accurate directory services be available to:

- Easily locate all researchers and collaborators by Name, Location, Research Interests, Electronic Presence or Electronic Space.
- Easily locate and review in depth all ongoing and past related research or other work. Strong universal links to research libraries and electronic reference materials are mandatory.
- Easily locate all facilities and resources available for sharing.

Without these services, our distributed collaborative environment would be an impenetrable saltmarsh swamp, not a wonderland.

2. Universal Security Services

Without reasonable security, collaboration will be impossible. Nobody wants their products (research papers, data, ideas, programs, hardware) subject to damage from mischievous or malevolent hackers. It is, of course, possible to carry security far beyond the sublime: the absolutely secure computer is locked away in a safe ruminating on its own bytes, never used by anyone (if they have been listening to DOE, certainly with the power turned off!). Security for open scientific research must leave the resources to be secured in a usable state (similarly concerning research for national security, but “usable” starts taking on a whole new meaning). The following security services should be seamlessly and transparently available in easy to use forms across all collaborative or shared resource sites:
• Authenticated logon (encrypted public key/private key password or smart card service) from any ESnet/InterNet site. No unencrypted passwords over the net. Must be interoperable across all collaborative sites. One logon receives all authorized services, whether at work, at home, at a collaborative site, while traveling to a conference, or on a mountain-top.

• Verification of the user and associated privileges at reasonable intervals during collaborative sessions.

• Encryption services for all sensitive or proprietary data or communications and e-mail. Interoperable across all collaborating sites. Encryption should be transparently easy to use and trivial to start or stop.

• Use of verified and certifiable digital signatures for e-mail and other important documents.

Security must be the base and the starting point on which all collaborative environments and shared resources are built. Without security, no collaborative environment or shared resource would be used or shared by any serious professional. Except for technology demonstrations, production quality collaborative tools and shared resources must have integrated security as their starting point.

3. **Shared Resources (CPUs, CAVEs, MPPs, ???)**

In order to reduce duplication and increase efficiency and productivity, seamless and transparent sharing of electronic resources across sites and across continents is highly desirable:

• Where both desired and authorized, all of our network, computing, simulation software, communication, ..., or other resources should be available to any authorized ESnet or other collaborative authorized user.

• We should allow, enable, and encourage researchers to share expensive, unique, or underused resources such as MPP’s (massively parallel processors), simulation machines or software (e.g., CAVEs) and allow efficient use of all our resources, as many hours a day as they can operate (preferably 24 x 7).

• Universal directory and security services are a required and enabling infrastructure to make this possible.

With seamless resource sharing and security in place, authorized graduate students doing energy research at Colorado State will be able to use excess cycles on the massively parallel Paragon (1024 nodes) processor at Oak Ridge National Laboratory (ORNL) just as easily as
researchers in the Center for Computational Science (home of the Paragon) at ORNL.

4. Shared Electronic Spaces

Both the ideas of electronic presence and electronic spaces are significant parts of our virtual laboratory. Electronic spaces are not a totally new thing. Electronic game players are very familiar with MUDs (multiuser dungeons or multiuser dimensions) and MOOs (multiuser object-oriented dungeons or dimensions). While we are not interested in playing games with research funding, games and entertainment are often at the very cutting edge of technology. This technology can be applied to more serious problems. A number of these types of spaces are already in existence. AstroVR, a collaborative environment at Stanford, uses the MOO server from Xerox PARC (Palo Alto Research Center) to create a virtual laboratory for astronomers. BioMOO brings together biologists from around the world, and InfoPark is a similar construction for systems administrators from several different sites. The CAVE is an immersive three-dimensional electronic space at Argonne National Laboratory. Many of these environments are primarily text-based, far from our vision of multidimensional graphical electronic spaces, yet they are already significantly promoting collaboration. The sociological aspects of these systems are currently being studied, but informal feedback from user communities is very positive. We propose the following qualities and uses:

- Electronic presences should be persistent and follow the owner (the user's electronic environment should travel where he is). Additionally, the electronic presence should persist whether or not she is logged on to the system. Electronic presences might include personal electronic spaces, personal World Wide Web (WWW) pages, message baskets, audio or video clips on recent work and research topics, graphical representations of recent research, 3-D simulations, personal research bulletin boards, and other objects to be imagined.

- Collaborative electronic spaces should be persistent and adaptable by the users to the job to be performed. The space should remain ready and waiting whether anybody is presently home or not (powered down but not off). Examples of shared electronic spaces include an electronic laboratory (measurement, calculation, viewing, simulation, and data analysis tools), an electronic writing collaboration room (shared writing tools), an electronic meeting room (shared communications) and a theoretical collaboration room (shared tools for symbolic and computational math, shared viewing and control of 3-D simulations, suitable electronic space.
for meetings with whiteboard and appropriately passed "talk" token, theoreticians being notoriously long-winded).

5. Collaborative Software and Tools

Collaborative software and tools will be the enabling technologies that really make virtual laboratories fly. In addition to the electronic spaces and presences mentioned above, collaborative tools are necessary to make all of the high-technology network and computing infrastructure actually pay off.

- Communications of all kinds are the heart of improving efficiency through collaboration.

- Communications (e-mail, voice, videoconferencing, shared documents, shared graphics, shared applications) should be easy, instantaneous, and ubiquitous. Electronic mail (e-mail) already exists on almost every LAN, and often works over the InterNet, but must be seamless, secure, and interoperable across all sites. Attachments of all kinds must be transparently delivered in a form usable at the receiving end with a mouse-click, keystroke, or voice command. Similarly for videoconferencing and sharing of documents and applications, one or two mouse-clicks or keystrokes should be all that is necessary. What is not easy to use, reasonably secure, and transparent to the worker need not apply.

- Tools for collaboration should include: (1) programs for shared analysis and viewing of data, both 2-D and 3-D; (2) shared design tools of all kinds (drawing, symbolic or computational mathematics, flow charting, presentation, graphing and graphical design tools, 3-D graphical immersion design environments for studying object behavior under simulated or real conditions); (3) collaborative software development tools of all kinds (collaborative compilers, editors, debuggers, tracing tools, tools for diagnosing MPPs; these are the tools that should come first to enable development of our collaborative environment); (4) scheduling tools for network scheduling of shared objects of all kinds, including computers, software, MPPs, video facilities, and human beings.

The above discussion is meant to be representative and not at all exhaustive. We hope that some of the most important tools for remote networked collaboration have been mentioned.
6. Building a Culture of Collaboration, Shared Resources, and Interoperability

Our present laboratory culture is one of islands of high technology, interconnected by low-bandwidth, low-speed canals, to allow limited sharing of documents or other low-bandwidth work. Changing this to a single-continent culture of collaborative researchers interconnected by high-speed (fat!) pipes for all forms of cyberspace communication and shared resources will not happen instantaneously. One of the first blocks is the lack of interoperability between the laboratories. In many cases each lab has rolled its own solution for problems not directly pertaining to network compatibility. What works at one lab is not available at the next or not interoperable with another.

An absolutely necessary piece of infrastructure for our collaborative environment is an established standard for a seamless, secure, transparent to the user, wide area file system (WAFS). A WAFS can provide much of the infrastructure for the following services:

- Logon should provide the same resources, views, and environment to the user whether at work, at home, at a collaborating institution, or logging-in from on the road. A single set of application files, environment descriptions, privilege files, and initialization files available and used for the user regardless of login location can simplify this tremendously.

- Researchers and other workers should be able to easily have access to a repository of common tools that they share in their joint work.

- One of the more obnoxious problems dogging collaborative enterprises is keeping software up to date and at the same version level at all sites. Synchronization of software versions should be easy and automatic if desired. If a single location on a WAFS maintains the current production-level version of a software tool, then it is relatively trivial for sites to use the software from that site or keep their local software automatically up to date.

- Program or tool location should be transparent to the user and to the network, whether running locally from San Diego or on a compute server located at Argonne.

All of this can be made possible with an interoperable, secure, WAFS and standardized interoperable tools. The WAFS will make disk drives and file resources transparently available to authorized users across the campus, across laboratory borders and state lines, across the country, and across international boundaries, regardless of the physical location of files or disk drive.
The need for a WAFS has been recognized and the AFS (Andrew File System) developed at Carnegie Mellon University and the newer DFS (Distributed File System) with integrated Generalized Security Services (GSS) API, developed by the OSF/DCE are initial attempts to provide a secure and transparent WAFS. Discovering what bugs exist, what needs still aren't met, and implementation across a critical mass of sites of AFS/DFS or some future WAFS remains to be accomplished.

Establishing our collaborative culture and building or finding tools that are interoperable across all of our research sites will take time. During the transition it may be necessary to have multiple tools operating at some or all sites in order to achieve interoperability. One rather important example of this is a suitable multimedia file format for seamless sharing of documents across all sites. No such open-standards interoperable format presently exists. While we are working on developing or finding a format for seamless multimedia document interchange (preserving all formatting and font information, along with the graphics and any audio or video material and hypertext links), it will probably be necessary to support multiple file/document formats and multiple applications to work with these formats. Eventually we will evolve this to a single multimedia format, interoperable among various commercial or freeware tools that compete not on their proprietary formats, but on the level of performance, features, and interoperability.

These ideas can be generally applied to all of our tools for collaboration and sharing of resources. We may initially have to support multiple noninteroperable tools, formats, and protocols at each site in order to have interoperability across all sites.

By constantly working together, temporarily supporting multiple solutions, and striving for compatible, interoperable, open-standards solutions for all of our tools and infrastructure, we can asymptotically approach the high-speed, secure, open-standards and interoperable "virtual laboratory" of our dreams.

III. Facilities-On-Line

**Vision:** Our vision of facilities-on-line is that unique or expensive research instruments, large research facilities, and even manufacturing sites can be made available over the network. Authorized researchers, engineers, and educators will be able to participate in or actively control measurement, experimental, and engineering and manufacturing demonstration or national defense facilities.

**Goal:** One of the strategic goals of the DOE/ER (Energy Research Strategic Plan, December 1994) is to "Sustain scientific leadership and
improve U.S. competitiveness by providing world class scientific facilities ...
A method of providing these facilities at reduced cost to more
researchers with increased utilization is to make them available over
ESnet, the InterNet, and the NII (national information infrastructure).
The objective of this effort would be to improve the infrastructure,
capabilities, and services (and use the already established
ESnet/InterNet infrastructure) to enable researchers to access these
expensive instruments/facilities in a fashion that was "as good or better
than being there."

It is our expectation that truly immersive implementations of facilities-
on-line will lead to improved productivity of research, increased
utilization and more efficient utilization of facilities, synergistic increases
in research productivity, decreased travel and facilities costs, and
increased effectiveness of our educational systems and enrichment of the
educational experience. "I read and I wonder, I see and I believe, I do
and I understand." Nothing is more exciting to a student or leads to
better understanding than seeing and participating in the subjects of her
studies.

Our concept of facilities-on-line can be split into at least two different
parts, one involving remote participation (for instance, for researchers
who are offering advice and examining data, but not actively controlling)
and the other involving not only participation, but actual remote
measurement, experimentation, and control.

A. **Remote Participation**

Remote participation in measurements or experiments could include at
least the following elements:

- Multiple streams of video and audio from the experimental site
  (control room, on-site collaborator, experiment or instrument itself,
  gauges, oscilloscopes, real-time data displays).

- Shared applications (whiteboard, data display, experimental
  waveform definition, ...).

- Shared data visualization tools.

- Shared data analysis tools.

Ultimately these tools could be melded into a truly immersive 3-D
environment, giving the sensation of total participation and involvement,
with displays on two, three, or all surfaces and true 3-D vision and
manipulation. This much interactivity is not always necessary or
desirable, and less immersively a researcher might participate with
streams of audio, video, data, and shared applications on two or three
large screen monitors controlled from a single console, so that the researcher could focus attention as desired on a particular element of the experiment. Experience to date (e.g., the author's personal knowledge of remote collaborations from ORNL to TORE SUPRA, TFTR, General Atomic in San Diego, and the previously referenced ALCATOR C-Mod remote operation) suggests that for remote participation, it is important to have good contact with the on-site researchers and views of the control room and experiment and particularly important to have shared applications between the on-site researchers and the remote participants. On the other hand, even more elaborately, remote participation might include a complete remote control room with participation of tens or even a hundred or more remote researchers (e.g., for the proposed Large Hadron Collider or the proposed International Thermonuclear Experimental Reactor) immersively involved in the experiment via the duplicate remote control room ("actual reality").

The alert reader will note that these tools for remote participation are perhaps different in quality, but not different in kind, from the tools described for our distributed collaborative environment. There is a large overlap and synergy between the tools required for remote collaboration and the tools required for remote participation in experiments, measurements, or engineering test facilities.

**B. Remote Experimentation, Instrument Control, or Data Acquisition**

While remote participation is perhaps not that different from our distributed collaborative environment, true remote experimentation, measurement, and control are entirely another story and quite a different target. While control and data acquisition over the LAN (local area network) are well understood and available, as evidenced by

- **TFTR**— The Toroidal Fusion Test Reactor, a magnetically confined fusion energy project [fusion of the deuterium and tritium isotopes of hydrogen to produce helium and about 17.6 MeV (million electron Volts) of energy per reaction], is located at Princeton Plasma Physics Laboratory (PPPL). This experiment is entirely run by remote control over a LAN, and all diagnostics and data acquisition have been done remotely over the LAN since about 1982. The researchers controlling and measuring the results of the experiment are located hundreds of feet away from the test cell, because the fusion reaction also produces high-energy neutrons and other forms of radiation, not to mention possible hazards from high voltages and strong electric and magnetic fields.
• Other examples include the large superconducting magnetic fusion experiment **TORE SUPRA** at Cadarache in France and numerous high-energy physics experiments, including the Advanced Light Source (**ALS**) at Lawrence Berkeley Laboratory and the Continuous Electron Beam Accelerator Facility (**CEBAF**) at Newport News Virginia. Many factories also have considerable control, data acquisition, and diagnosis over LANs.

Control and data acquisition over the WAN are undeveloped (despite the impressive **ALCATOR C-MOD** work mentioned previously). Before data acquisition and control over the WAN can be routinely used, things must be developed that do not exist:

• Generalized nonproprietary open-standards application programming interfaces (**APIs**) and applications for remote control of experiments and instruments and for data acquisition over the WAN.

• Protocols for fault tolerant (network faults, software faults) control and data acquisition.

No standards or interfaces exist yet for this problem, so this is not a job of months, but rather one of years. The interfaces, standards, and protocols must all be developed.

Measurement and control over the WAN are fundamentally different from local measurement and control:

• While LANs are not particularly famous for their reliability, the problems with them are well understood, and systems are designed to be fail-safe and fault-tolerant to the expected failure modes. The case for WANs (the InterNet) is just the opposite. They are notoriously famous for their unreliability and constantly shifting routes and changing delay. It is not particularly unusual for network connectivity to disappear to a site for a day or even for a week (often due to fiber cuts in remote areas). Even more common is for a route to a site to disappear and be replaced by a different route in the blink of an eye (a millisecond) or for the network traffic to undergo a momentary burst, causing the delay time for transmission of control data to a site to change by 30, 40, or even 100 milliseconds. This introduces large issues of timing, feedback speed, and even experiment viability.

• In case of an incident, accident, or problem, local systems normally have emergency shutdown procedures available to the operator or experimenter. If one loses connectivity over the WAN, there is no chance of hitting a big red SCRAM button on the way out of the
room, hardwired to absolutely shut down all power and equipment, and possibly calling emergency aid.

- While there are some ongoing experiments in control over the WAN, they are just that, experiments. None of these efforts incorporate security, an absolute requirement. It is highly probable that in addition to authentication and verification, encrypted control sessions and encrypted sessions for data acquisition and transmission will be required. If collaborative environments require reasonable security, then control of delicate, possibly dangerous equipment, and data acquisition for sensitive, possibly proprietary, data over the WAN requires double the protection from itinerant or malevolent hackers. On the LAN, one can often assume isolation from the rest of the world. Over the public WAN, that is fundamentally impossible.

Presently, every time somebody discusses controlling something over the WAN, it is always a fresh start. We cannot afford to keep reinventing the wheel, so it is important that we start developing some standardized parts, protocols, and APIs for WAN remote control and measurement. Some of the parts to be developed were mentioned above. More exhaustively, the pieces to be developed for WAN control and measurement might include:

- Generalized open-standards API.
- Generalized open-standards remote graphical user interface (GUI) for control and data acquisition (the next X-Windows?).
- Open standardized remote client program (driven by the GUI).
- Security interfaces (authentication, verification, data and control encryption).
- Database interfaces (local, remote, SQL, ODBC??, open-standards, secure!!).
- Open standardized local server program;
- Specialized driver interfaces between server and instruments/controls/data acquisition.
- Safety and control protocols (it is likely that all control will be of the set-point, or downloaded waveform, type, with a "return to safe home" protocol in case of loss of connectivity and provision for safety or security scram by either the local or remote personnel).
- ???? (pieces to be discovered when we actually do it on a production basis).
Again, this is a non-exclusive list representing the ideas of the authors.

IV. Future Infrastructure Development

As we develop our virtual laboratories, many improvements to our infrastructure are going to be needed.

A. Physical Infrastructure

While our present networks, typically with links between labs at 1.5 megabits per second (Mbs), or very recently at 45 Mbs, are adequate for the low-bandwidth collaborations presently going on between laboratories, as the infrastructure for our distributed collaborative environments and facilities-on-line is developed, the requirement for connectivity will grow at a very high rate. Eventually data rates of gigabits per second (Gbs), or even hundreds of Gbs between our research sites and internal to the sites themselves, will be required. While these data rates may sound extreme now, it should be considered that the 155-Mbs circuits which are now just coming into use for production WANs were absolutely out of the question 10 years ago. The available bandwidth is reasonably likely to keep pace with our requirements for connectivity, and the price will fall into line with what we can pay.

In addition to network bandwidth, the collaborative infrastructure that we have described will also require improvements to video and audio computer hardware, and development of new high-speed data links and standards between computers and their I/O (input/output) systems and peripherals. Increased amounts of fast high-volume storage to serve our virtual laboratories will also be a necessity.

B. Security Infrastructure

The elements that will be necessary for a security infrastructure have been discussed above in some detail. Nonproprietary, open-standards, interoperable implementations of these elements will be required rather rapidly as the virtual laboratories grow. This is one of the first and most important pieces of infrastructure to put in place in order to make virtual laboratories a reality.

C. Facilities-On-Line and Collaborative Infrastructure

Some of the collaborative tools that will be desirable or necessary for our virtual laboratories have been discussed above. Almost none of these exist in forms that are presently usable in a production environment where they are tools rather than being an experiment themselves. Some exist in experimental forms or as demonstrations; many do not exist at
all. Video conferencing over the WAN/LAN is perhaps the closest to production use of any of the collaborative tools and as it comes into widespread use over the WAN will be one of the drivers for improving the network physical infrastructure.

D. **Sociological Infrastructure**

The kind of strong and close collaborations over long distances involving groups of tens, hundreds, or even thousands of persons that have been described above are a new thing. This is not something that has been done before in the history of the world, and it is likely that new sociological paradigms will be required. How does one supervise an employee a thousand miles away with only (hopefully close) network contact? What are the rules for social interaction in collaborative electronic spaces?? Who will tidy up and clean up in these collaborative electronic environments (there are likely to be lots of electronic artifacts left lying about, maybe even BrontoFLOPs!)?

The social aspects of making these virtual laboratories work are perhaps just as important as the hardware and software aspects of implementing them, and even more unknown. As we venture into our new world of virtual laboratories, we will also be venturing down a new path of social interaction and development. New norms for behavior and interaction must be developed and implemented along with the virtual laboratories.

V. **Private Industry**

A. **Yes! Whenever and wherever possible (COTS)**

Where private industry produces open-standards, interoperable tools, we should absolutely use them. These tools should be sought out, purchased and used, and their developers encouraged to make any necessary changes and to develop other useful tools. It is extremely important to have strong interaction between government customers and industry vendors, and to have industry buy-in to the standards and solutions developed for our "virtual laboratories." There is no point in doing work twice or developing a government solution where a commercial solution exists or is being developed. Where capitalism works, it is tremendously more efficient than any government organized effort. Common off-the-shelf software (COTS), along with standardized interoperable hardware, is the fuel that has driven the explosion of PC use around the world and is now fueling development of the InterNet. It should be remembered, however, that the InterNet began as government-sponsored research decades ago.
B. Can Private Industry Do It All?

For the same reasons that private industry does not provide and pay for public schools, public roads and public libraries or organize and pay for the national defense, it is also unlikely that private industry will provide all of the infrastructure and tools that we need for our virtual laboratories. Many of the tools that we need represent a 5 to 10 year development project. This is far beyond the event horizon of commercial entities in the United States. In general, they do not have the capital for investment in such long-term, large, high-risk (and high-payoff) development. They must show a profit on their efforts in 2 to 3 years or risk disappearing.

Even more than this, many of the tools needed for our virtual laboratories are not generally needed by industry itself. Industrial work tends to be insular for competitive reasons. Commercial developers would rather dynamite a computer than reveal their R&D efforts to competitors. For our virtual laboratories, we intend to encourage just the opposite culture. It is one government, and we are all working together to improve our world. It is absolutely imperative that our national laboratories have complementary and collaborative research efforts. We can no longer afford to have insular, competitive research in our national laboratories.

The horizon for government research efforts should normally be at least five years or even much longer, and profits are only to be realized far into the future by private industry, not by government labs. Government must work in ways far different from private industry.

Looking further into the future, the tools that are developed now for virtual laboratories need not be limited to our national laboratory structure. As we build the tools and learn how to use them they can be applied to all of the problems of government and allow tightly connected, collaborative efforts across all levels of government.

Imagine a government where resources are easily located (universal directory services); where individuals can easily find and access all government information that pertains to them and can make changes when the changes are verified as legitimate; where individuals cannot access information describing their neighbors and have no access to information that is not appropriate for public viewing (universal security services); where interchange of information between local, state, and federal governments is easy and instantaneous (high-bandwidth, interoperable, collaborative environment); and where government agencies can easily develop plans, arrange virtual meetings without leaving their offices, and simulate the effects of government policy on the environment, social structures, and the economy (collaborative tools and
shared electronic spaces); a world where citizens can go to the Drivers' License Bureau or the Social Security Administration and access their family records or discuss their situation with public employees, all from their personal computer over the National VBNS (very broadband network system), without ever leaving their homes (distributed collaborative environment). Consider a world where kindergarten students can participate in veterinary procedures at the San Diego Zoo from Sumter, South Carolina, or graduate students can make nuclear bremsstrahlung cross-section measurements at the National Superconducting Cyclotron Laboratory (NSCL) from the nuclear physics graduate school at Georgia Tech in Atlanta, Georgia (facilities-on-line).

The work that is done now to develop virtual laboratories can be the foundation of distributed collaborative environments and facilities-on-line for our governments of the future.

C. Precompetitive-Collaboration Between Industry and Government

The business of government is to provide necessary services and infrastructure that its citizens, either individually or collectively, cannot provide for themselves. This clearly applies to the role of government research in developing virtual laboratories. It is the job of government to provide for research and develop technology that its citizens are expected to need in the future (for instance, in order to keep our industries competitive, our government efficient, and our economy viable 20 years from now) but which is beyond the event-horizon of private industry. It is the job of government to do research which is required but will produce no profits in the near future or may never produce profits in the foreseeable future. The Large Hadron Collider (LHC), fundamental research in micromachines, or DNA chemistry for ultra-massively parallel computations are excellent examples of this type of research.

Our job is research and development that is precompetitive and to either deliver the technology to private industry or to turn it into production level tools for our citizenry.

What is presently missing from our ability to get industry involved in our virtual laboratories, and more grandly in government research for the global information infrastructure, is the ability to form close collaborations with industrial entities and government research labs. Many extremely bright and knowledgeable people work in commercial firms. In order to ensure the best use of our development resources, we should have the ability to involve these people in guiding the course of government sponsored, precompetitive, or basic research. The Federal Advisory Committee Act (FACA) presently makes it very difficult to involve industry in government development efforts for precompetitive
technology. We need a better infrastructure for industry and government collaboration in order to develop many of the technologies that are required for our “virtual laboratories” and for the global information infrastructure (GII).

VI. Getting Started

Developing the vision of virtual laboratories into reality will be a challenging and long-term process. Five years from now, significant parts of this dream should be operational, but it may be even ten years before all of it exists. The time to start is now! Resources (funding and staff) are going to be invested in computing, networking, and information technology, regardless. The way that these resources are invested, and choices of where to invest any additional funds, will govern the development of our vision. Our choices to work together, to collaborate for the common good on secure, interoperable, high-bandwidth tools, will determine whether or not we have made significant and measurable leaps 5 and 10 years from today.

A. The DCCC and DICCE

The DOE Energy Sciences network (ESnet) was founded on the principle of collaboration. In the late 1980s, the DOE Office of Energy Research determined that there was not money to fund separate networks for all of its research offices. The researchers were forced to work out details of a single network that would serve the whole community. Out of this rather fierce early collaboration came the concept of supporting multiple protocols (using the same routers and physical links), but with the end in sight of merging into a single unified structure as time and opportunity allowed. This has now happened. All but about 0.3% of the traffic over ESnet is now TCP/IP (Transmission Control Protocol/Internet Protocol). The other protocols (e.g., DECNET) have faded away to nearly nothing, and are expected to disappear within 18 months.

Out of this original collaboration came the Energy Sciences Network Steering Committee (ESSC) and the implementation body of site network managers appointed by the chief officer for networks and computing at each site, the ESnet Site Coordinating Committee (ESCC). These two groups have labored over the years to promote shared resources and interoperability in networking.

Over the last 24 months, the idea of increased collaboration and much closer sharing of resources, even far beyond sharing network structures, has been growing within the ESnet research community. From this movement has come the Distributed Computing Coordinating Committee (DCCC). The DCCC has taken on as its own personal mission:
"...embarking on an innovative project to create a Distributed Informatics, Computing & Collaborative Environment (DICCE), ...Twenty-one sites, facilities and universities, have collaborated ... . The first goal of the DCCC DICCE project is to initiate DICCE in the ESnet community for three projects:

1. Bring Facilities-On-Line for networked collaborations,
2. Setup Networked Collaborative Environments, and
3. Create a Distributed Tera/PetaFLOP Computer (DTPC).

To accomplish this goal, ESnet sites are bringing into operation the following environment layered on ESnet:

1. Key Distribution system
2. Authentication Services
3. Open Software Foundations Distributed Computing Environment and Distributed File System (OSF's DCE/DFS)
4. Functional seamless E-mail enclosures
5. Secure E-mail

With the above minimum system in place, end users will be able to:

- access with a single login all of the systems in this environment on which they have accounts,
- "see" all of the files on all of the systems having global access
- exchange binary attachments via e-mail, and
- use secure mail systems and initiate digital signatures.

This environment will form the first step in bringing Facilities-On-Line, supporting networked collaborations, and building towards a Distributed Tera/PetaFLOP Computer. Additional activities supported by this environment include:

1. Distributed Systems Management
2. Smart Logins, i.e., the login process would automatically set up the end user to have the correct configuration for the system being used.
3. Additional applications and services such as;
• load sharing,
• distributed information servers and electronic spaces,
• distributed capacity computing, and
• distributed experimental controls. (Ref. 13)

[Please note in passing that the proposed authentication services and security services are Kerberos (out of MIT) and the OSF/DCE GSS (generalized security services) API, and secure e-mail services will be from MIME/MOSS compliant agents with PGP or PEM security using the present InterNet simple mail transfer protocol, SMTP.]

The success or failure of the DCCC and its DICCE (pronounced “dicey”) efforts will be determined over the next several years. It is clear, however, that their goals are in line with our vision of virtual laboratories and are a step down the path toward developing the required infrastructure. We wish grace and luck to the efforts of the DCCC and DICCE.

VII. The Payoff

It is our belief that the potential payoff from developing a true virtual laboratory system will be huge and that the whole formed from all of the individual parts will be far greater than their sum. This kind of enabling, collaborative, communications-oriented structure has the potential of considerable synergistic improvements in our national laboratory efficiency and productivity, with the almost certain extension of these gains to private industry and all levels of government.

A. Efficiency

The primary way to use technology to reduce costs in the virtualization process is to reduce the number of people needed to perform certain tasks and to reduce the cost of contracted activities by reducing the need for services. For example, when travel is reduced, not only is the direct cost of the travel decreased, but also the need for administrative positions to process travel requests and process the expense forms. We believe that a "virtual" organization can reduce the administrative structure needed to support actual research and development activity by large factors. Additional costs will be saved by reducing the difficulty in communicating and sharing information, but these cost savings will typically not be reflected in reduced research manpower but instead in more productive research. For the same reasons that a virtual laboratory will increase the productivity of researchers, it will also increase the
productivity of administration and decrease the need for layers of administration and overhead expense.

**B. Decreased Duplication**

Virtual laboratories will enable sharing of expensive and unique resources, including people, facilities, and data. By enabling remote access that is as good or better than "being there" for all authorized levels of government, education, and industry, the costs of providing a local resource that is already available elsewhere can be avoided. Making these resources easy to locate, secure from predation and easy to use (transparent, seamless, intuitive interfaces) with instant response (high-bandwidth connectivity) are the keys to this distributed collaborative environment which will enable us to share and fully utilize our resources, whether they are located across the campus or across the continent.

**C. LHC, ITER??**

If virtual laboratory technology is done correctly, it will make the United States much more attractive as a location for international collaborations such as ITER (the International Thermonuclear Experimental Reactor). If we can provide access to our international colleagues to facilities in the United States, with telepresence that is "better than being there," then the United States becomes much more attractive as a site for these collaborations and has a considerable influence on its international partners.

In the case where we are collaborating on facilities located outside the United States (perhaps the LHC?), the technology of virtual laboratories will make the collaboration more effective at a reduced cost. Our researchers will spend their time working rather than traveling, and the expenses of travel will also disappear.

**D. Synergy**

Enabling the remote use of facilities and the sharing of networked resources will allow universities and industry to have full access to facilities and network resources without significant additional cost. A major concern is to maintain high utilization of expensive and unique facilities and equipment. The High Flux Isotope Reactor, HFIR, at ORNL, useful for all kinds of neutron scattering studies to understand the response and fundamental structure of materials, is a good example. Major industries already travel thousands of miles and lose hundreds of person-hours for a day or two of time on this facility. Additionally, projects can be conceived that require access to multiple distributed facilities that would be impractical in the past. For example,
simultaneous x-ray scattering studies of materials at the ALS (Advanced Light Source, Lawrence Berkeley Laboratory) and simultaneous analysis of the data on the 1024-node massively parallel Paragon computer at the Center for Computational Science, ORNL, becomes possible. Another considerable advantage is that facilities which support remote users via the network also save hundreds of person-hours in reduced safety training, expose considerably fewer personnel to safety hazards of materials and radiation, and thereby again reduce support overhead costs.

Allowing remote participation by universities can enrich the educational experience for our students, helping them to understand and making them more productive when they begin careers as engineers or in science. Facilities can be made available over the network which are otherwise impossible to reproduce at more than one location. Virtualizing our laboratories makes them available everywhere that the network goes.

Collaborations will be supported (whether for research or administrative purposes) independent of location. The virtual laboratory provides support for collaborations including: video/audio environments; electronic meeting rooms; shared virtual spaces for interaction and communication concerning shared objects and data; electronic communication of all kinds; and shared software applications and tools. These resources will just as easily enable two or more scientists to share data and do three-dimensional manipulation of theoretical abstracts as to allow two budget administrators to negotiate a research contract and share a spreadsheet itemizing what costs have been incurred and will be allowed. The tools deployed will, of course, provide special support for remote use of supercomputers and scientific instruments. However, the general ability to support collaborations with the tools to share data and communicate just as if located in the same physical space will benefit all users.

Administrative costs and overhead personnel requirements will be greatly reduced by virtual laboratory technology. Both within laboratories and across the laboratory system, technology which is as good or better than "being there" will decrease duplication of effort and increase the bandwidth of communications between administrators. Functions that previously had to be performed locally can be centralized or concentrated where desired. Outsourcing of administrative and overhead services will become much more feasible with the technology of virtual laboratories. This technology will also allow "telecommuting" to become a more general reality where it is effective (often for theoreticians, design personnel, administrative personnel, or with facilities-on-line, even for experimentalists, for instance), thereby reducing the need for office space.
and the costs of running these tremendous enterprises, our national laboratories (and into the future, all layers and levels of government).

Our national laboratory system will be allowed and enabled to act as a single entity in support of national missions and customer needs. Today the national laboratory system functions as a loosely federated set of research institutions. Often, they are just as likely to be competing with one another (the word “cutthroat” is even occasionally heard in the halls) as cooperating. In fact, it is often easier technically and administratively to compete than to collaborate. The difficulty in forming multilab efforts seriously diminishes the effectiveness of multilab projects (e.g., USCAR—improved efficiency of autos, USABC, AMTEX—improved textiles with reduced resources, the SEMATECH collaboration—improved efficiency and competitiveness in semiconductors). Often industrial consortia wish to deal with a single lab entity, and yet the current DOE laboratory culture and technical capability prevents or hinders this collaboration model, even when it is in our best interests to provide a single logical laboratory view to customers. The virtual laboratory effort will allow the national labs to act and behave as a single entity. We believe that in the future it will be increasingly the case that the national laboratories will want to act together. The development and deployment of virtual laboratory technology is essential for this.

Virtual laboratories will enable “quickness” in response to new opportunities or threats to national security. The lag induced by the overhead and bureaucratic layers of billion dollar organizations between recognition of a threat or an opportunity and response to that threat or opportunity can be minimized or eliminated. With virtual laboratory technology in place it will be easy to find, contact, and gather the resources and information to address new opportunities or to counter new threats. Communication will be easy, high-bandwidth, and instantaneous.

It is our considered belief that the whole of a virtual laboratory will be tremendously larger than the sum of its parts.
REFERENCES


12. R. Evard, "Collaborative Networked Communication: MUDs as Systems Tools,"
13. R. Whitney (Chairman DCCC) and the DCCC/DICCE collaboration. private communication, May 1995.