SANDIA REPORT

SAND99–8749
Unlimited Release
Printed February 2000

The Diesel Combustion Collaboratory: Combustion Researchers Collaborating over the Internet

RECEIVED
JUL 24 2000
OSTI

C. M. Pancerella, L. A. Rahn, C. Yang

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: 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, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent 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, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Prices available from (703) 605-6000
Web site: http://www.ntis.gov/ordering.htm

Available to the public from
National Technical Information Service
U.S. Department of Commerce
5285 Port Royal Rd
Springfield, VA 22161

NTIS price codes
Printed copy: A03
Microfiche copy: A01

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

SAND99-8749
Unlimited Release
Printed February 2000

The Diesel Combustion Collaboratory:

Combustion Researchers Collaborating over the Internet

by
Carmen M. Pancerella, Larry A. Rahn, Christine Yang

Combustion Research Facility
Sandia National Laboratories
Livermore, CA 94551

ABSTRACT

The Diesel Combustion Collaboratory (DCC) is a pilot project to develop and deploy collaborative technologies to combustion researchers distributed throughout the DOE national laboratories, academia, and industry. The result is a problem-solving environment for combustion research. Researchers collaborate over the Internet using DCC tools, which include: a distributed execution management system for running combustion models on widely distributed computers, including supercomputers; web-accessible data archiving capabilities for sharing graphical experimental or modeling data; electronic notebooks and shared workspaces for facilitating collaboration; visualization of combustion data; and video-conferencing and data-conferencing among researchers at remote sites. Security is a key aspect of the collaborative tools. In many cases, we have integrated these tools to allow data, including large combustion data sets, to flow seamlessly, for example, from modeling tools to data archives. In this paper the authors describe the work of a larger collaborative effort to design, implement and deploy the DCC.

The Diesel Combustion Collaboratory: Combustion Researchers Collaborating over the Internet

Carmen M. Pancerella[†], Larry A. Rahn[†], Christine L. Yang[†]
{carmen|rahn|clyang}@sandia.gov
Sandia National Laboratories
P.O. Box 969, Mailstop 9012|9056|9012
Livermore, CA 94551-0969

1. Introduction

Bill Wulf coined the term collaboratories as "centers 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" [1]. Collaboratories enable researchers to work together across geographic and organizational boundaries to solve complex, interdisciplinary problems; to share data across distributed sites; and to have secure access to resources (both hardware and software) at remote sites. To date collaboratories have been developed in many areas, including telemedicine [2], materials science [3], natural sciences [4], space physics [5], and environmental molecular sciences [6].

The Diesel Combustion Collaboratory (DCC) is a research pilot project that includes collaborators distributed among Department of Energy (DOE) national laboratories (Sandia National Laboratories (SNL}, Lawrence Berkeley National Laboratory (LBNL), Los Alamos National Laboratory (LANL), and Lawrence Livermore National Laboratory (LLNL)), the University of Wisconsin, and industry partners (Caterpillar, Cummins Engine, and Detroit Diesel). It is funded by the Department of Energy's DOE 2000 program in the Office of Advanced Scientific Computing Research. The project also benefits from a strong coupling to other DOE 2000 projects funded to develop collaborative tools and services. The DCC itself is a problem-solving environment for combustion researchers. The goal of the DCC is to make doing the science, engineering, and information exchange for the combustion researchers more efficient. High-speed networking, computer security, distributed computing technology (i.e. CORBA-based frameworks for connecting data and resources), data visualization, and collaborative tools are part of the underlying infrastructure for the DCC. At this time remote instrumentation is not part of the DCC. Unique to the Diesel Combustion Collaboratory is a software architecture that allows collaborative tools to be integrated in a secure fashion and facilitates data exchange among tools.

This paper discusses the applications area, the DCC requirements, the DCC architecture and design, DCC security, and DCC deployment.

2. Application: Combustion Modeling

Due to their high efficiency and reliability, diesel engines are the dominant power source for heavy-duty trucks and busses in the U. S. Heavy-duty diesel thermal efficiency is about 45%, versus 30% for production gasoline engines. Even higher efficiencies are possible and the implementation of small-bore diesel engines promises to greatly improve the fuel efficiency of autos and the rapidly growing light truck

[†] This work was supported by Sandia Corporation under Contract No. DE-AC04-94-AL85000 with the U.S. Department of Energy.

market. However, NOx and soot emission regulations threaten to compromise future implementation of diesel engines. The development of chemically predictive model-based design tools is a critical part of the DOE's strategy to meet the efficiency and emission requirements over the next decade. Diesel combustion modeling presents unique challenges, however.

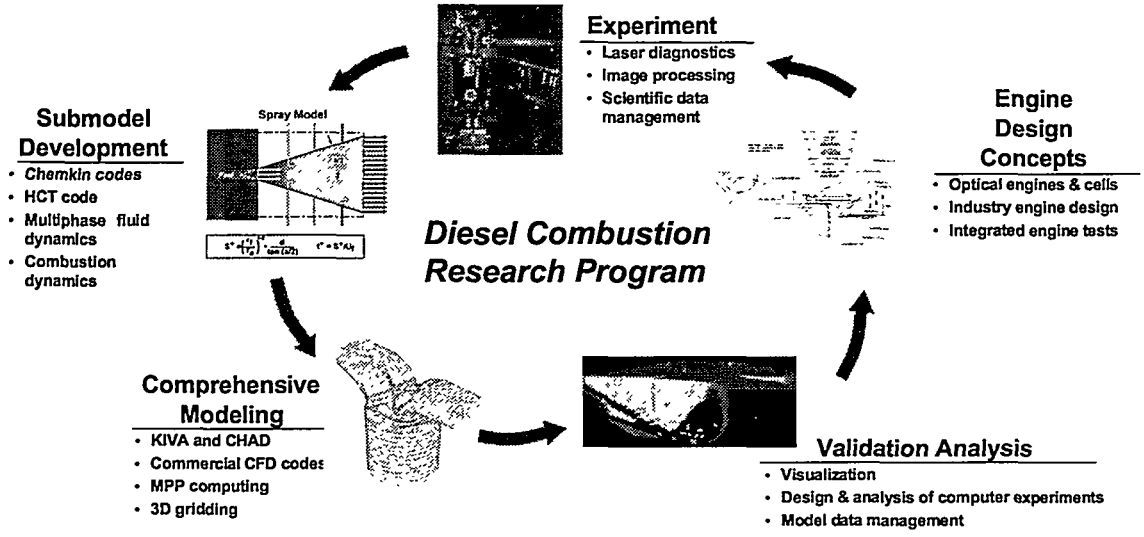

Figure 1. Combustion research process.

As seen in Figure 1, model development is a complex task that requires a close interplay among experimentation, development of physio-chemical submodels, development of numerical codes for engine simulation, model validation and formulation of new modeling concepts, the development of new engine concepts and hardware leading to new experiments. Ultimately, this process is envisioned to enable the simulation-based design that will revolutionize the production of new diesel engine technologies. Progress in this already difficult task is further complicated by the distributed and interdisciplinary nature of the required collaboration among the geographically distributed industry, national laboratory, and university partners. The initial focus of the DCC is on the DOE Heavy Duty Diesel Combustion Research CRADA (Cooperative Research and Development Agreement) involving three national laboratories (LANL, LLNL, SNL), the University of Wisconsin, three engine manufacturers (Cummins Engine Company, Caterpillar Inc., Detroit Diesel Corporation), and sponsored by the DOE Office of Transportation Technologies. Each institution offers a subset of the required interdependent capabilities that must be integrated for the overall success of the program.

The goal of the Diesel Combustion Collaboratory is to make doing the science, engineering, and information exchange for the Diesel Combustion Research CRADA partners more efficient. To improve on an already successful collaborative effort, our objective is to implement, evaluate, test, and use a set of collaborative tools. Researchers will be able to tackle new problems with their existing methods simply because these techniques are easier to use, and they will be able to access new methods through collaborations that were previously inaccessible because of geographic or expertise limitations.

3. Diesel Combustion Collaboratory Requirements

The DCC has adopted a unique set of requirements in order to accomplish these goals. The requirements are that researchers at the distributed sites be enabled to:

1. Share graphical data easily using desktop workstations.
2. Discuss modeling strategies and quickly exchange model descriptions between groups.
3. Archive collaborative information in a web-accessible electronic notebook.

4. Utilize a Distributed Execution Management System to run combustion models at widely separated locations.
5. Quickly analyze experimental data and modeling results in a web-accessible format.
6. Videoconference one-on-one collaborations and group meetings using desktop workstations.

Each of these six tasks must be done securely and efficiently across the Internet. The scale of these activities varies with the computational and data archive resources required by the combustion applications. Many of the zero- and one-dimensional chemistry simulations (such as Chemkin and HCT) may run for hours to days on high-end workstations and produce Kbyte to Mbyte data sets. For a particular problem, many of these data sets are typically produced (10-100) and each visualized using simultaneous 2-D plots of multiple variables. Two- and three-dimensional combustion simulations, on the other hand, are typically limited by the available computational hardware. Parallel versions of these codes are being developed for full engineering simulations of a diesel combustion chamber that require modern super computers and produce large (many Gbytes) data sets. These 3-D data sets are already a challenge to analyze by visualization. As the vision of the DOE Scientific Simulation Plan for combustion simulation [7] is fulfilled over the next 3-5 years, these simulations will fully utilize the anticipated 5-40 Teraflop supercomputers with 45-400 Tbytes of on-line disk storage [8]. This scale of activity presents significant challenges for distance collaboration technologies.

In addition to dealing with the scale of these activities, many of the collaborative tools (e.g., data archiving and modeling tools) must be integrated to allow data, including large chunks of experimental or simulation data, to flow seamlessly between tools. The DCC architecture must support rapid and easy integration of new modeling and visualization tools and both legacy software and newly developed models. Modeling and visualization tools include commercially developed software and software developed by some of the DCC partners (often written in Fortran). Collaboratory tools must be easy to deploy throughout distributed facilities. Users must be able to locate and combine resources and data. Users must be able to work together as teams across time and space. In addition to the tool integration requirement, Collaboratory users have a variety of hardware platforms and operating systems.

4. Design of the DCC Architecture

The DCC requires a framework to seamlessly provide distributed computing services between the various distributed tools being integrated in the Collaboratory. The underlying architecture is critical to how collaborative, modeling, and visualization tools are integrated. Without a planned integration and a robust and scalable architecture, an ad-hoc integration may crumble as the number of tools increase. An architecture makes the addition of new tools and the creation, modification, and maintenance of users interfaces more manageable for the DCC developers.

The DCC has selected the Product Realization Environment (PRE) [9], developed at Sandia National Laboratories, as a software architecture for tool integration and data exchange. PRE is a lightweight, horizontal framework for a broad variety of electronic resources (for example, databases, data translators, modeling codes, and visualization tools). PRE is built on top of CORBA [10] and available vendor-developed CORBA services. Applications are "wrapped" for use in the PRE framework as reusable components. PRE has a common API which all applications share; hence, in order to "wrap" an application into PRE, a developer simply implements the framework-compliant API in order to expose the functionality of the application.

The PRE architecture consists of several major pieces including uniform data objects and transport, a trading service, security, a conversion broker, integrated applications, and user interfaces. The data factory, distributed file manager, trader, conversion broker, and security are part of the PRE core services. Table 1 gives a quick definition of each component; the same architecture is shown in Figure 2.

Component	Description
Uniform data objects and data transport	PRE provides a persistent, uniform data container for structured information. Data/information is exchanged between PRE applications and services via data objects. PRE has a file manager to handle very large data files efficiently.
Integrated applications	PRE provides a standard application API for wrapped applications. An application's functionality is accessible to PRE clients and other PRE applications through this wrapper.
User Interfaces	User interfaces can be developed as stand-alone applications, applets, or web front ends to PRE core services and applications.
Trader Service	PRE provides a 'yellow pages' registry for the location and general information about the applications integrated into PRE.
Conversion broker	This PRE service employs a reasoning algorithm to determine the data translation steps necessary to convert data from one format to another, given a registry of PRE converters.
Security	PRE security supports a variety of security models, selected by plug-able library modules. Hence, PRE security is extensible.

Table 1. Major components of PRE.

As mentioned in Table 1, PRE has a file manager [9] to handle very large data files (multi-megabyte files) efficiently since CORBA is not the most efficient way to transmit many megabytes of file data. The PRE file manager is an optimized data transfer service built on top of the PRE framework. The property in a PRE data object can be one of two things: either the actual data or a reference describing the location of a file in which to find the actual data. The PRE file manager implements two very different protocols: one that uses CORBA, and one that uses a separate authenticated TCP connection. With the CORBA-based protocol, the data travels in-band with other PRE data, which may make implementation of firewalls easier. The other protocol is a raw-sockets based "blob streaming" protocol, which achieves FTP speed in typical use.

A further optimization in the PRE file manager is possible if the programs exchanging data are located on the same machine. In this case the file manager can allow end-user applications to get direct access to the files in the File Manager's store. Two communicating applications do not need to know when they are collocated since the decisions are made by the PRE framework; the applications can be written in a location-independent way, and get the benefit of optimization when it is possible.

The PRE development team is addressing many other high performance distributed computing needs, such as running jobs on teraflop computing platforms. PRE is an evolving framework. Most of the APIs remain the same, yet the underlying functionality and core services are becoming more advanced. The DCC will benefit from performance enhancements to PRE. Furthermore, as we explain in detail in Section 4.1, the Diesel Combustion Collaboratory has developed PRE servers that allow combustion modeling codes to be transparently run on supercomputers.

The resulting DCC architecture, built on top of PRE, can be seen in Figure 2. Combustion modeling codes, electronic notebooks, data converters, and plotting and data visualization tools are being wrapped in PRE and made available as servers distributed across remote sites.

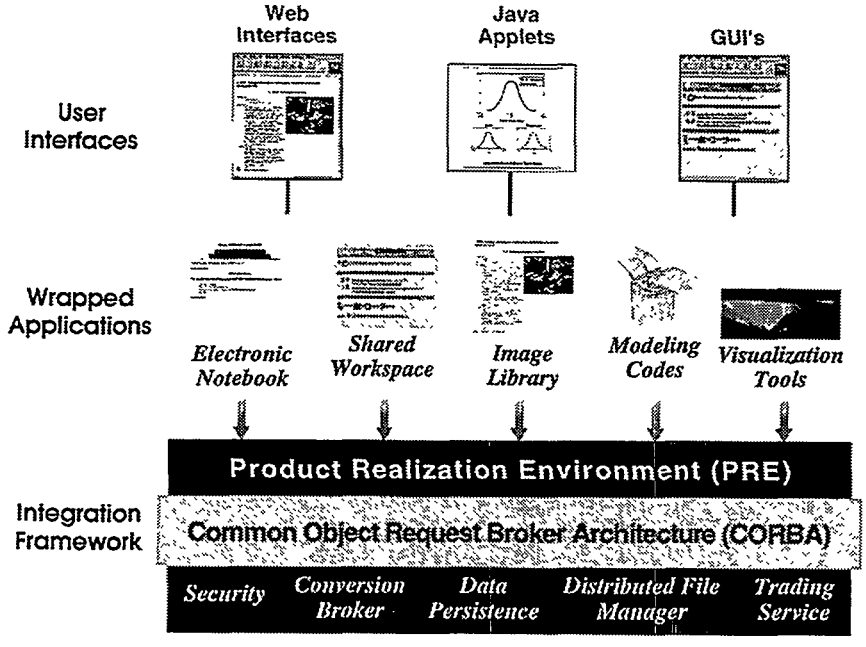

Figure 2. Diesel Combustion Collaboratory Architecture. Distributed Execution Management System

4.1 Distributed Execution Management System

The Distributed Execution Management System (DEMS) consists of modeling codes integrated into PRE, visualization and data archiving applications integrated into PRE, and customized modeling clients. DEMS takes advantage of PRE's distributed file manager to efficiently handle large modeling files.

Since PRE provides asynchronous execution of servers, a client can terminate while a large computation is executing and the results are not lost. A second monitoring client can access status data while the computation continues. We have used this feature in the DEMS modeling servers so that feedback can be provided to users as modeling codes are executing.

From a GUI or web browser on the user's desktop, Collaboratory tools appear at the user's fingertips even though they are distributed across the Internet. The tools themselves are integrated, allowing seamless data exchange and data conversion when necessary. For example, a combustion researcher can execute a modeling code, and the returned results will be converted to a different data format, and displayed as a 2D plot. Further, there is built-in security such that data can be encrypted, applications can be protected, and users can be authenticated prior to using the Collaboratory applications.

As seen in Figure 3a, we have created two templates, a modeling server template and a modeling client template, for integrating new modeling tools into the DCC. The standard DCC modeling server template is a PRE wrapper that makes the integration of a new modeling code trivial. The standard DCC modeling client template allows clients for new modeling codes to be customized and deployed rapidly. Employing client and server templates allows new capabilities to be added easily, provides easy maintenance of DCC clients and servers, maximizes code reuse, and minimizes bugs.

Modeling Server Template

The modeling server template shown in Figure 3a is a PRE wrapper that provides the functionality that is needed for any modeling server. A software developer needs to add some modeling code-specific code to the modeling server template to create a new modeling server. The code-specific code creates an execution script for a modeling code and specifies the location of executables and new MIME types for required files. The modeling server template has the following properties:

- Asynchronous execution of modeling codes.
- Ability to run several codes at a time, including pre- and post- processors.
- New server executes for each client.
- Checking of required input files and capability to receive optional input files.
- Returns only those output files specified by client.
- Return of arbitrary sized output file sequences.
- Logging of all server accesses for DCC metrics.
- Return of partial results upon failure.

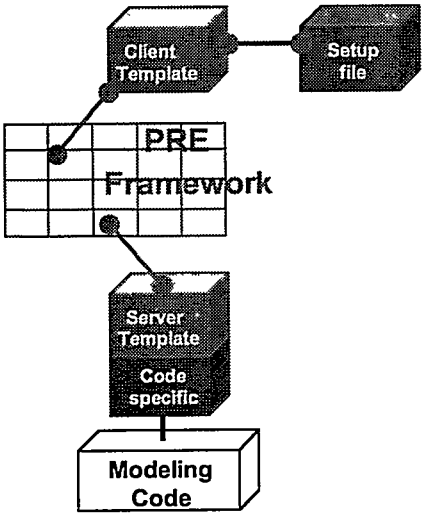

Figure 3a. Templates for modeling clients and servers in PRE.

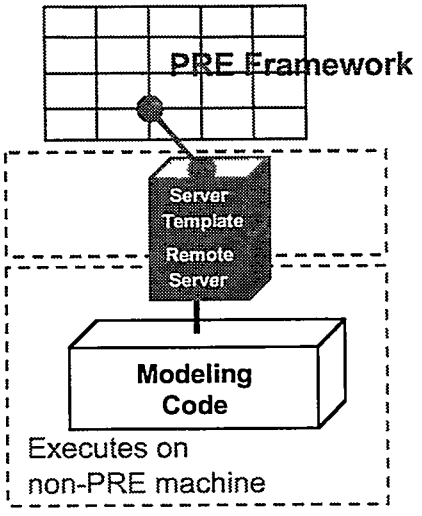

Figure 3b. Template for remote servers for the execution of modeling codes on non-PRE machines.

Modeling Client Template

The modeling client template is Java code that allows a new combustion modeling code to be integrated into the DCC. To create a new instance of a modeling client, a new setup file is created; the Java source code is not modified. Since there is a need for platform-independence in the Collaboratory, the use of Java and web technologies are central for creating portable and platform-independent user interfaces for all Collaboratory tools. All client development is done in Java to allow modeling codes to be deployed as applets, even though the client can execute as either an application or applet. The GUI for a modeling client is shown in Figure 4.

Properties of the modeling client template include these:

- Multi-threaded client.
- Monitoring of code status with periodic updates.
- Client can terminate and recover results at later time.

- Ability to run several codes including pre-processors and post-processors.
- Specification of required and optional input files.
- Specification of expected output files.
- Integration with 2D plotting.
- Applets deployed as signed JAR files to allow read/write permission to local file system.

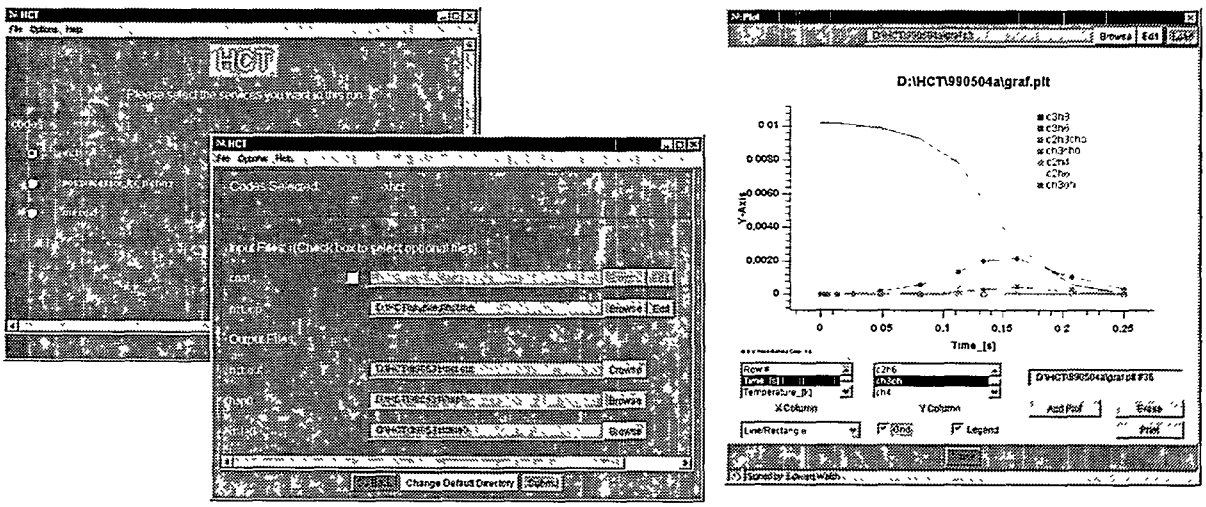

Figure 4. User Interface of Modeling Client Template.

Remote Modeling Server

The DCC plans to use supercomputers to execute many of the combustion analysis codes. Since many supercomputers do not have a CORBA implementation, we have developed a PRE server which executes on a Solaris server, and uses Java RMI [11] classes to send and receive input/output files to the supercomputer. Figure 3b shows a diagram of this remote modeling server. This remote server is actually a template for creating new modeling servers on a non-PRE machine. It has the same API and functionality as the modeling server template discussed above; hence, any modeling client can communicate with this server. This remote server then securely handles the transport of DCC data to and from the supercomputer via SSL. Hence, this gives the illusion that a supercomputer is available via PRE, and it makes wrapping modeling codes on a supercomputer possible.

DEMS Integration

To date, we have wrapped several combustion modeling codes and made them available to the DCC over the Internet: Chemkin, a chemical kinetics package for the analysis of gas-phase chemical and plasma kinetics; HCT, a one-dimensional, time-dependent code that models chemical reactions, thermal conduction, species diffusion, and hydrodynamics in the gas phase; Chemkin II/Soot, Chemkin with kinetic modeling of soot formation; and KIVA, a transient, three-dimensional, multiphase, multi-component code for the analysis of chemically reacting flows reacting flows with sprays. These codes are running at various distributed sites in the DCC, including LBNL, LLNL, SNL, and the University of Wisconsin. The DCC user does not need accounts on these machines in order run the codes. All security and authentication is done through a single login.

At this time, DEMS is not focusing on resource management and meta-computing issues such as those addressed by NCSA, NPACI, and the Grid Forum. As we move forward and integrate high performance combustion modeling codes into the DCC, we are looking at employing techniques such as Globus. Globus [12] is an infrastructure toolkit that provides resource location and allocation, and data access services, among other services.

4.2 Electronic Notebooks and Shared Workspaces

The DCC uses electronic notebooks [13], [14] and shared workspaces [15] as asynchronous data archives where researchers can place information that can be accessed by their collaborators using a web browser. Electronic notebooks can contain a rich variety of information - text and images, and links to information can be simultaneously viewed on the pages of a notebook. Pages can be edited, annotated and notarized. Like electronic notebooks, shared workspaces are accessible using a web browser. Within the DCC, shared workspaces are a very convenient way to quickly share files, such as presentations, with members of defined groups.

We are integrating the shared workspaces and electronic notebooks into PRE. By providing PRE interfaces to these archives, it is possible for a user to submit a modeling code and automatically document the run in an electronic notebook. This information could include input files, output files, a plot of the data and notes on the calculation just as a researcher would record in a traditional paper notebook. Since modeling codes, visualization tools and data archives are integrated into PRE, there is a seamless data flow from modeling to visualization and data archiving.

4.3 Visualization

We have delivered a 2-D plotting tool to all DCC users. A 2-D plotting package developed at Sandia, Java Xplot, has been integrated into the DCC modeling client template, such that users can interactively plot modeling results from several modeling runs. Users can then save plots as files or put plot files in an electronic notebook.

As we enter year 3 of the DCC, we plan to explore 3-D visualization needs of Collaboratory users. We expect new computational challenges as we explore the role of 3-D visualization in collaborative problem-solving environments.

4.4 Real-time Videoconferencing and Data Conferencing

The ability to be able to share information such as research data or an engineering drawing with a remote collaborator was named by our CRADA customers as one of the most immediately useful capabilities that the DCC could provide. Early in the development of the DCC, Mbone tools [16, 17] were used to provide audio and video conferencing. In addition, PostScript files could be shared using a whiteboard. These Internet multicast conferencing tools require access to the Mbone, a multicast backbone built on top of the Internet to support multi-party communication.

As the implementation of the DCC progressed, it was recognized that it was important to be able to support data and application sharing. Standards for multimedia teleconferencing have been established by the Telecommunication Standardization Sector (ITU-T) of the International Telecommunications Union (ITU). These standards support audio, video and data conferencing [18] and real-time, multipoint data communication [19]. Products which are compliant with these standards can interoperate, an important consideration for the DCC. Using products based on these standards, we are able to provide real-time, cross-platform conferencing and application sharing over both networks and ISDN.

5. Secure Collaboratories

The DCC customer is a DOE CRADA. Because information is proprietary to the CRADA, security is a critical requirement in the implementation of the DCC.

The DCC has developed security guidelines for setting up collaboratories between national laboratories, industry, and universities such that proprietary data can be shared when needed but protected at all other times during collaboration. These guidelines identify specific collaboratory capabilities to be protected and design guidelines which should be met in order to secure collaboratory computers and networks from potential threats. Specific security characteristics are outlined for electronic mail, sharing and execution of software tools, reading and writing of shared collaboratory files, conferencing, and reading and

execution of shared public files. Using this policy, we have developed a design for a secure Diesel Combustion Collaboratory.

Figure 5 shows our implementation of a secure collaboratory. A collaboratory partner that provides resources can set up secure PRE servers for modeling and data archiving, secure web servers for web-based DCC tools, and/or secure computers and supercomputers on the Internet. Collaboratory users, with either direct Internet connections or dial-up connections to the Internet, can securely access DCC data and resources from any secure server. There can be any number of resource providers and users.

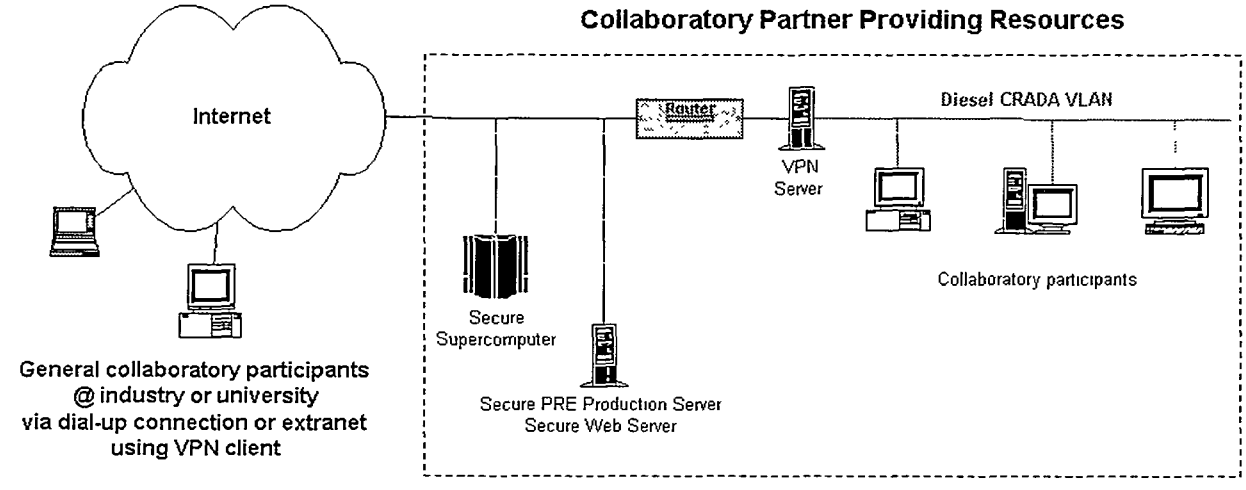

Figure 5. Secure Diesel Combustion Collaboratory.

Using the Akenti security model [20, 21] based on public key infrastructure (PKI), the DCC can provide access control to web-based resources and resources made available via the PRE architecture. Akenti is a security model and architecture that is intended to provide scalable security services in highly distributed network environments. Since the PRE security architecture is extensible and provides a way to integrate any security model, Akenti is being integrated into PRE.

Most video conferencing tools do not have built-in security. In order to provide secure real-time collaboration between two or more sites each with strong firewall protection, the DCC has implemented a virtual private network (VPN) using encrypted tunnels between the otherwise secure sites. While these tunnels are technically holes in the firewall, the fact that they are encrypted means that the data integrity is not compromised, and only members with appropriate keys are authorized through the tunnel.

6. Deploying Collaborative Tools in the DCC

At the beginning of the DCC project, a strong multi-institutional and multi-disciplinary collaboration on diesel combustion research had already been established, facilitated by continuous DOE funding over the past decade. The participants, however, found themselves adapted to almost daily encounters with the barriers introduced by the geographical distribution of participating institutions and/or by difficulties in sharing information among the varied disciplines and tools employed by their collaborators. Thus, the challenges involved in successful deployment of the DCC included development and implementation of an appropriate collaboratory infrastructure, integration of new and legacy disciplinary tools, and the deployment and adoption of the DCC among the research team. To accomplish this, we planned a phased approach to the development, integration, testing, and deployment of tools.

6.1 A Phased Approach

The first phase of deployment is of iterative development, testing and deployment of individual Collaboratory tools. Once Collaboratory team members have adequately tested each new Collaboratory tool, it is deployed to combustion researchers in the CRADA. Collaboratory integrators can then refine or replace tools, and deploy changes and new tools to the CRADA partners as needs are realized and new work approaches are adopted. The initial phase of deployment has been to directly involve the most interested combustion researchers in tool development and testing. This phase also involves deployment of infrastructure such as videoconferencing tools, multicast routing capabilities, and accessible web-based data archives to as many of the CRADA participants as possible.

The next phase is enabled by a full implementation of security in the Collaboratory. In this phase, the Collaboratory team focuses its efforts on concurrent CRADA collaborations, further implementing infrastructure, refining tools, and implementing new tools to facilitate these interactions. The primary focus here is on two-way synchronous and multi-way asynchronous collaborations. The Collaboratory has been found to have highest impact when geographically separated researchers are truly dependent on each other to achieve their goals. The focused effort of the Collaboratory team during this phase also provides the direct support to researchers as they confront some of the cultural changes inherent in electronic collaboration. These changes include not only the use of new technologies (such as secure networking) and the way work is accomplished by accessing new tools, but also the development of social behaviors in electronic communication and collaboration. Thus, this phase involves many of the issues discussed in the literature that must be overcome by each individual of the collaboration [1]. We have found that one challenge in particular is the balance of the increased access to each other's data, resources, and presence with the desire to invoke electronic security to protect the same.

The final phase of the Diesel Combustion Collaboratory project will be the broadening of the target research to multiple projects including a broader range of collaborators and the active dissemination of lessons learned, development processes, and new technology to other interested groups. This phase will be driven by the adoption of Collaboratory technology for the normal operation of DOE user facilities such as the Combustion Research Facility and by the needs of programs arising from the DOE Scientific Simulation Plan [7]. The broadening of the Collaboratory target research areas will provide a greater test of the generality of our approach and an opportunity to incorporate some advanced tools. Examples of the former are the applicability of our client/server template, usability of PKI security, and stability of our PRE infrastructure. Examples of the latter include multi-code sequences for simulations, advanced desktop visualization, and distributed group participation in technical discussions using collaborative visualization.

6.2 Illustrative Examples of DCC Deployment and Application

Descriptions of illustrative scenarios in the DCC serve to further describe its deployment and application in combustion research. Here we provide a few brief descriptions followed by a more detailed example:

A diesel combustion scientist can start a modeling run in collaboration with a distant colleague on Friday using video conferencing and application sharing to set up the run. The results can be stored in an electronic notebook for both to see the following Monday.

A diesel combustion scientist can select from a set of KIVA codes that execute on various computing platforms across the DCC, taking advantage of previously established collaborative relationships that provide access and the needed expertise.

A diesel combustion scientist can retrieve previous results generated by a collaborator from a shared workspace, interpret this data by visualizing with a collaborator, then rerun modeling code with new inputs, and put new results in the shared workspace that can be viewed by collaborator(s).

A diesel combustion scientist can retrieve visualization data from a personal electronic notebook and paste the data into a whiteboard allowing interactive markup and drawing during a data conference with a collaborator.

A diesel combustion scientist can retrieve previously generated modeling results from a personal electronic notebook or shared workspace and visualize this data, and create report or presentation material that is then deposited in a group shared workspace.

A team of diesel combustion scientists can securely share presentations and discuss results using multi-person audio-video conferencing and application sharing.

A specific example illustrating deployment and application of the DCC arises from research on homogeneous charge compression ignition (HCCI) engines. If successfully developed, this concept would allow for the high efficiency of compression-ignition engines with greatly reduced emissions. For this reason, HCCI engines are an area of interest in combustion research. Because of the importance of combustion chemistry in this research, Chemkin and HCT are actively used within the combustion research community for calculations of HCCI combustion. The codes were developed at SNL and LLNL, and the DCC has provided remote access to both codes via clients that are integrated into the Collaboratory environment.

In this example, the goal was to calculate the ignition delay for a mixture of natural gas at typical HCCI engine conditions using both Chemkin and HCT. The HCT code expert at LLNL collaborated with a colleague at SNL to develop the corresponding problem definition for Chemkin. Submodel development of chemical mechanisms for fuels is an important combustion research area. For these calculations, the Gas Research Institute (GRI) chemical kinetics mechanism was used in Chemkin while the LLNL chemical mechanism for natural gas was used in HCT. The combustion scientists submitted input files from local machines for remote execution of the code using the respective DCC clients. Output files were returned to the scientists' machines and the modeling results could be immediately analyzed using the DCC client's interactive 2D plotting package.

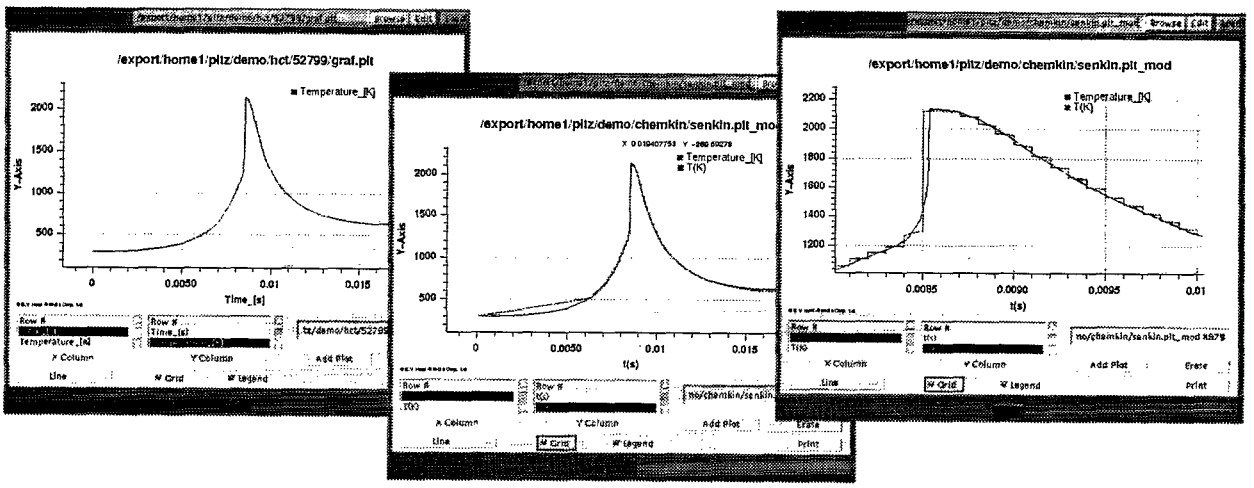

Figure 6. Comparison of temperature (Y-Axis) time-history for HCCI reaction volume, Chemkin (red) and HCT (blue). The plot on the far right is an expanded scale about the ignition point (sharp increase in temperature) achieved by using the interactive zoom capability.

During a collaborative session to discuss analysis of the modeling results, one researcher discovered that it was possible to plot data from two separate data files, originating from different modeling codes, on a single plot. He immediately recognized that he had the opportunity to directly compare the Chemkin results and the HCT results – something that had not been done before. Direct comparison of the output data (Figure 6) showed that calculations from the two codes using different chemical mechanisms produced almost identical ignition delays. The fact that the ignition delays matched so closely was a surprise to the HCT code expert. This was possible because the DCC provided 1) desktop access to both codes, and 2) a quick way to visualize modeling results from the desktop.

The calculation was documented in a web-based electronic notebook (Figure 7) that can be accessed by collaborators presenting a valid identity certificate. The combustion scientist used real-time collaboration tools to discuss these results with colleagues at Caterpillar and Sandia. In the course of this discussion, the Sandia scientist shared his private web-based data archive with his colleagues. The scientists identified modeling results for a different fuel of interest then downloaded and examined the data using the DCC client's built-in plotting package.

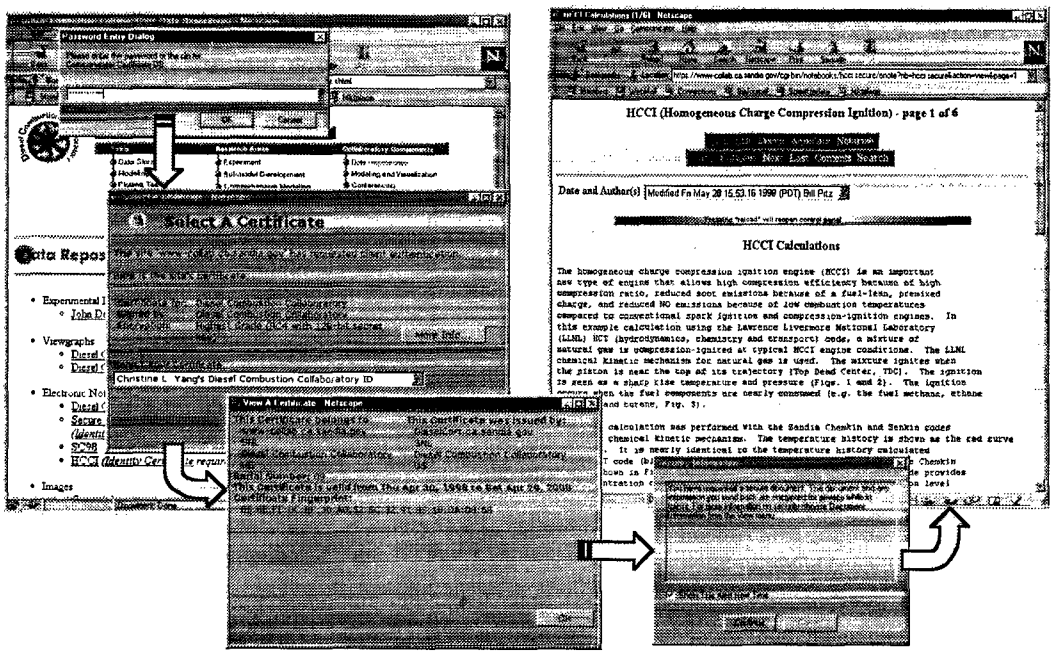

Figure 7. Process for secure access to the electronic HCCI notebook.

7. Status and Future Work

The DCC has successfully developed and deployed collaborative technologies to combustion researchers distributed throughout the DOE national laboratories, academia, and industry. Sandia's Product Realization Environment (PRE), the DCC client/server template, and modern browser technology have been successfully combined to make legacy combustion modeling codes available over the Internet. The environment also integrates security, notebooks, shared workspaces and other data archives, conversion utilities, and CORBA services. Current client/server technology delivers remote execution of models with local or remote input/output files and two-dimensional plotting. Security based on PKI is integrated into the environment allowing resource owners to manage access from their desktop. Commercial video conferencing and encryption solutions have been tested and adopted. Combining these tools, pairs of researchers can securely conference at their desktop and share applications. Using available software bridges, these capabilities are extended to group collaborations.

The current status of the Collaboratory project is early in Phase II of deployment, with many technological and social barriers to overcome to meet all of our objectives. User friendly and low maintenance security is a difficult challenge, and will require continuous development with user feedback for some time. We can offer much more real-time conferencing capability than we were able to at the beginning of this project (2 years ago). Never-the-less, desktop video and data conferencing among researchers is extremely dependent upon available Internet connection bandwidth, requires external security, and is not yet compatible across tool vendors and computer platforms. Group conferencing still requires expert assistance with software bridges.

Collaboratory technology is having an impact, however. A number of CRADA researchers now routinely share information through the web accessible workspace. We have found that one of the most significant impacts of the DCC is to extend the researcher's capabilities to use new tools by making it possible to use such tools in collaboration with an expert at another location. An example of this in combustion research is the situation where an experimentalist provides a problem definition to a computational expert to simulate in a modeling code such as Chemkin, and the experimentalist is also then able to run the code based on a notebook archive of the expert. The ability to leverage such knowledge resources is a key to interdisciplinary collaboration and an important characteristic of the DCC. Some are now planning future research around the new capabilities offered by the DCC.

Future work will build on the current successes in framework implementation, tool development, and deployment to researchers. We will work simultaneously for a greater range of capabilities and a higher level of user friendliness. This work will require close collaboration with developers of electronic notebooks to better integrate them into the Collaboratory environment. We need a directory service to locate data and a way to uniformly reference data so that data can be located and used anywhere on the network. Continued close interaction and testing with commercial suppliers of collaboration software and related tools is needed to speed development of capabilities that meet interoperability standards. New target research groups will involve international participation with new challenges for interoperability and bandwidth. We anticipate new client/server templates for complex multi-code simulations, and tools for code object development allowing quick integration of new submodels into large simulation codes. Finally, we anticipate new capabilities for collaborative visualization, computational steering, and group collaboration built on new tools, emerging standards for quality of service protocols, improved commercial tools, and broad availability of high bandwidth connections.

8. Acknowledgments

The scope of the work addressed in this project would not be possible without the efforts, support, and close cooperation of the group of talented and committed people associated with the DCC, the DOE Heavy Duty Diesel Combustion Research CRADA, the DOE Mathematical, Information, and Computational Sciences Division, within the Office of Advanced Scientific Computing Research, the Chemical Sciences Division within the Office of Basic Energy Sciences, and DOE Office of Transportation Technologies.

We specifically acknowledge these past and present contributors to the DCC:

SNL: Tuesday Armijo, Ann Gentile, Diane Gomes, Joe Grcar, Barry Hess, John Howard, Bruce Mah, Noel Nachtigal, Young Seiden, Tim Strayer, Edward Walsh, John Dec, Peter Kelly-Zion, Bob Carling, PRE development team
LANL: Justin Doak, Mike Springer, Dan Butler
LLNL: Bill Pitz, Larry Cloutman, Charlie Westbrook
LBNL: Bill Johnston, Mary Thompson, Keith Jackson, Ken Revzan, Nancy Brown
University of Wisconsin: Rolf Reitz, Jim Wiedenhoefer
Cummins Engine Co.: Roy Primus, Russ Durrett
Caterpillar Inc.: Bill Brown, Tom Briggs, Bill Hurt
Detroit Diesel Corporation: Nabil Hakim

9. References

[1] R. T. Kouzes, J. D. Myers, and W. A. Wulf, "Collaboratories: Doing Science On The Internet", *IEEE Computer*, Volume 29, Number 8, August 1996.

[2] D. G. Kilman and D. W. Forslund, "An International Collaboratory Based on Virtual Patient Records", *Communications of the ACM,* Volume 40, Number 8, August 1997.

[3] B. Parvin, J. Taylor, and G. Cong, "Deep View: A Collaborative Framework for Distributed Microscopy", *Proceedings of SC 98*, Orlando, Florida, November 1998.

[4] V. S. Sunderam, S. Y. Cheung, M. Hirsh, *et. al.*, "CCF: Collaboratory Computing Frameworks", *Proceedings of SC 98*, Orlando, Florida, November 1998.

[5] S. Subramanian, G. R. Malan, H. Sop Shim, *et. al.*, "Software Architecture for the UARC Web-Based Collaboratory", *IEEE Internet Computing*, Volume 3, Number 2, March/April 1999, pp. 46-54.

[6] D. A. Payne and J. D. Myers, "The EMSL Collaborative Research Environment (CORE) - Collaboration via the World Wide Web", *Proceedings of the Fifth Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE '96)*, Stanford, California, June 1996.

[7] "The DOE Strategic Simulation Initiative: Combustion Simulation and Modeling" (Draft Plan and Summary), http://www.ca.sandia.gov/csmi/plan/summary.pdf, July 1998.

[8] "Computer Science and Enabling Technologies for the DOE Scientific Simulation Plan" (Draft Plan and Summary), "http://www.mcs.anl.gov/cset/plan/, September 1998.

[9] R. A. Whiteside, E. J. Friedman-Hill, and R. J. Detry, "PRE: A framework for enterprise integration", *Proceedings of Design of Information Infrastructure Systems for Manufacturing 1998 (DIISM '98),* Fort Worth, Texas, May 1998.

[10] The Common Object Request Broker: Architecture and Specification, OMG Technical Document PTC/96-03-04, Object Management Group, Framingham, Massachusetts, July 1995.

[11] Java RMI reference "Java Remote Method Invocation – Distributed Computing for Java" (White Paper), http://java.sun.com/marketing/collateral/javarmi.html, May 1999.

[12] I. Foster and C. Kesselman,"Globus: A Metacomputing Infrastructure Toolkit", *International Journal of Supercomputer Applications*, Volume 11, Number 2, pp. 115-128, 1997.

[13] A. Geist and N. Nachtigal, ORNL Electronic Notebook Project, http://www.epm.ornl.gov/~geist/java/applets/enote/.

[14] J. D. Myers, C. Fox-Dobbs, J. Laird, *et al*, "Electronic Laboratory Notebooks for Collaborative Research", *Proceedings of the Fifth Workshop on Enabling Technologies: In5rastructure for Collaborative Enterprises (WET ICE '96)*, Stanford, California, June 1996.

[15] R. Bentley, W. Appelt, U. Busbach, E. Hinrichs, D. Kerr, S. Sikkel, J. Trevor and G. Woetzel, "Basic Support for Cooperative Work on the World Wide Web", *International Journal of Human-Computer Studies 46(6): Special issue on Innovative Applications of the World Wide Web*, June 1997, pp. 827-846.

[16] S. McCanne, "Scalable Multimedia Communication Using IP Multicast and Lightweight Sessions,"*IEEE Internet Computing,* Volume 3, Number 2, March/April 1999, pp. 33-45.

[17] V. Kumar, *MBone: Interactive Multimedia on the Internet*, New Riders Publishing, Indianapolis Indiana, 1996.

[18] "A Primer on the H.323 Series Standard", http://www.databeam.com/h323/h323primer.html.

[19] "A Primer on the T.120 Series Standard", http://www.databeam.com/ccts/t120primer.html.

[20] M. R. Thompson, W. Johnston, S. Mudumbai, *et. al.*, "Certificate-based Access Control for Widely distributed Resources", *Usenix Security Symposium '99*, Washington, D.C., August 1999.

[21] W. Johnston, S. Mudumbai, and M. Thompson, "Authorization and Attribute Certificates for Widely Distributed Access Control", *Proceedings of IEEE 7th International Workshops on Enabling Technologies: Infrastructures for Collaborataive Enterprises (WET ICE `98),* Stanford, California, June 1998.

Distribution:

Tim Strayer
BBN Technologies
10 Moulton Street
Cambridge, Mass 02138

Bill Brown
Caterpillar Inc.
PO Box 610, MOS 1
Mosvillle, IL 61552-0610

Bill Hurt
Caterpillar, Inc.
Technical Center, Bldg. F
PO Box 1875
Peoria, IL 61656-1875

Tom Briggs
Caterpillar, Inc.
Technical Center, Bldg. L
P.O. Box 1875
Peoria, IL 61656-1875

Rey Agama
Caterpillar, Inc.
Technical Center, Bldg. L
P.O. Box 1875
Peoria, IL 61656-1875

Roy Primus
Cummins Engine Co.
PO Box 3005 MS 50180
Columbus, IN 47202-3005

Russ Durett
Cummins Engine Co.
PO Box 3005, MS 50180
Columbus, IN 47202-3005

Nabil Hakim
Detroit Diesel Corporation
13400 Outer Drive
Detroit, MI 48239-4001

Peter Kelly-Zion
Trinity University
Dept. of Engineering Science
715 Stadium Drive
San Antonio, TX 78212-7200

Nancy Brown
Lawrence Berkeley National Lab
1 Cyclotron Road MS 70-108B
Berkeley, CA 94720

Keith Jackson
Lawrence Berkeley National Lab
1 Cyclotron Road MS 50B-2239
Berkeley, CA 94720

Bill Johnston
Lawrence Berkeley National Lab
1 Cyclotron Road MS 50B-2239
Berkeley, CA 94720

Stu Loken
Lawrence Berkeley National Lab
1 Cyclotron Road MS 50B-4230
Berkeley, CA 94720

Ken Revzan
Lawrence Berkeley National Lab
1 Cyclotron Road MS 90-3058
Berkeley, CA 94720

Mary Thompson
Lawrence Berkeley National Lab
1 Cyclotron Road MS 50B-2239
Berkeley, CA 94720

Larry Cloutman
Lawrence Livermore National Laboratory
P.O. Box 808, L-353
Livermore, CA 94551

Bill Pitz
Lawrence Livermore National Laboratory
P.O. Box 808, L-353
Livermore, CA 94551

Charlie Westbrook
Lawrence Livermore National Laboratory
P.O. Box 808, L-353
Livermore, CA 94551

Justin Doak
Los Alamos National Laboratory
P.O. Box 1663, MS B272
Los Alamos, NM 87545

Dan Butler
Los Alamos National Laboratory
P.O. Box 1663, MS B210
Los Alamos, NM 87545

Mike Springer
904 W. 16th Lane
Portales, NM 88130

Professor Gregory J. McRae
Department of Chemical Engineering
Massachusetts Institute of Technology
Room 66-372
77 Massachusetts Avenue
Cambridge, MA 02139-4307

Jim Wiedenhoefer
1011 Engineering Research Building
1500 Engineering Drive
Madison, WI 53706

Rolf Reitz
1018A Engineering Research Building
1500 Engineering Drive
Madison, WI 53706

William Kirchhoff
U.S. Department of Energy, Headquarters
19901 Germantown Road, MS 20
Germantown, MD 20874-1290

Mary Anne Scott
U.S. Department of Energy, Headquarters
19901 Germantown Road
Germantown, MD 20874-1290

Gurpreet Singh
U.S. Department of Energy,
Headquarters Forrestal
1000 Independence Avenue, S.W.
MS 6A-116
Washington, DC 205

8300
MS 9054 W. J. McLean, 8300
MS 9051 J. F. Grcar, 8345
MS 9051 L. A. Rahn, 8351 (10)
MS 9055 F. P. Tully, 8353
MS 9405 J. M. Hruby, 8355
MS 9671 Walter Bauer, 8358
MS 9052 D. R. Hardesty, 8361
MS 9053 R. W. Carling, 8362
MS 9053 J. E. Dec, 8362
MS 9051 R .J. Gallagher, 8366

8900
MS 9003 K. E. Washington, 8900
MS 9012 Carmen Pancerella, 8900
MS 9012 C. L. Yang, 8900
MS 9011 A. C. Gentile, 89102
MS 9011 B. V. Hess, 8910
MS 9011 J. D. Howard, 8910
MS 9011 B. A. Mah, 8910
MS 9011 P. E. Nielan, 8920, (5)
MS 9012 E. J. Walsh, 8920
MS 0916 R. L. Hartley, 8920
MS 0807 R. J. Detry, 8920
MS 9012 R. A. Whiteside, 8920
MS 0316 N. L. Leishman, 8920
MS 9012 E. J. Friedman-Hill, 8920
MS 9012 S. C. Gray, 8930
MS 9037 J. C. Berry, 8930-1
MS 9037 Diane Gomes, 8930-1
MS 9019 B. A. Maxwell, 8940
MS 9011 Juan Mesa, 8950
MS 9217 M. L. Koszykowski, 8950
MS 9019 Mike Rogers, 8960
MS 9019 J. A. Larson, 8970
MS 9051 Y. S. Seiden, 8970
MS 9012 T. A. Armijo, 8990
MS 9012 K. R. Hughes, 8990
MS 9214 N. M. Nachtigal, 8990

9900
MS 9003 D. L. Crawford, 9900
MS 9006 J. E. Costa, 9903
MS 9012 L. J. Prantil, 9903

MS 9018 Central Technical, Files, 8940-2
(3 copies)
MS 0899 Technical Library, 4916
MS 9021 Technical Communications Dept,
8528/Technical Library, MS 0899, 4916
MS 9021 Technical Communications Dept, 8528 For
DOE/OSTI

THIS PAGE LEFT INTENTIONALLY BLANK