

NISTIR 5232

Report of the NSF/NIST Workshop on NSFNET/NREN Security July 6 - 7, 1992

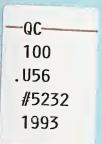
NIST PUBLICATIONS

Arthur E. Oldehoeft

Chairman Computer Science Department Iowa State University

for the

U.S. DEPARTMENT OF COMMERCE Technology Administration National Institute of Standards and Technology Computer Systems Laboratory Computer Security Division Gaithersburg, MD 20899





NISTIR 5232

Report of the NSF/NIST Workshop on NSFNET/NREN Security July 6 - 7, 1992

Arthur E. Oldehoeft

Chairman Computer Science Department Iowa State University

for the

U.S. DEPARTMENT OF COMMERCE Technology Administration National Institute of Standards and Technology Computer Systems Laboratory Computer Security Division Gaithersburg, MD 20899

May 1993



U.S. DEPARTMENT OF COMMERCE Ronald H. Brown, Secretary

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY Arati Prabhakar, Director



Report of the NSF/NIST Workshop

on

NSFNET/NREN Security

July 6-7, 1992

by

Arthur E. Oldehoeft Department of Computer Science Iowa State University

Edited by

Dennis K. Branstad Computer Systems Laboratory National Institute of Standards and Technology



FOREWORD

The Workshop on NSFNET/NREN Security was hosted by NIST and sponsored by NSF to address the need for improving the security of national computer networks. Emphasis was on identifying offthe-shelf security technology that could be implemented in the National Science Foundation Network, especially to control access to the super computers on the network. Steve Wolff initiated the workshop; Bob Aiken and Dennis Branstad organized it; and Arthur Oldehoeft reported on it.

The report sections are organized like the workshop sessions were organized: an introduction section, four technical sections and a recommendations section. Each session had a leader who led the discussion on the topic selected for the session. The viewgraphs used by the presenters on the topic are contained in the appendices.

Each participant had an opportunity to comment on the report. Suggested changes have been included. However, several people who were quoted or paraphrased by name did not comment. Therefore, the editor cannot attest to the complete accuracy of the statements as attributed to them in the report.

Work is expected to continue on developing and fielding the technology described in the report. A second workshop and supporting projects are anticipated as follow-ons to the recommendations of the workshop participants.

Please note that certain commercial equipment, products, instruments or materials are identified in this report in order to describe the discussion adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose. Tradenames are used for a specific product only when necessary to convey the context of the discussion.

Dennis K. Branstad Workshop Host and Report Editor



Contents

1	Executive Summary 1.1 Purpose of the Workshop 1.2 Structure of the Workshop 1.3 Workshop Recommendations	1 1 1 1	
2	Workshop Participants		
3	Introductory Session		
4	Session on Authentication in Distributed Networks 4.1 NIST Smart Card Authentication/Access Control - Jim Dray, National Institute of	7	
	Standards and Technology	7 11	
5	Session on Access Control in Distributed Networks 5.1 Kerberos Authentication/Access Control System - Jeff Schiller, Massachusetts Insti-	16	
	tute of Technology	16	
	 5.2 Supercomputer Center Access Control Requirements, Dan Nessett, Lawrence Liver- more National Laboratory 5.3 General Discussion on Access Control 	$\frac{20}{22}$	
6	Session on Application Security in Distributed Networks 6.1 NSF Super Computer Center Security - Gerard Newman, San Diego Supercomputer	27	
	 6.2 An Overview of Internet Privacy Enhanced Mail - Steve Kent, Bolt, Beranek, and 	27	
	 Newman	30 35	
7	Session on Security Management in Distributed Networks 7.1 SDNS Security Management - Wayne Jansen, National Institute of Standards and	38	
	 Technology 7.2 FIRST System - Dennis Steinauer, National Institute of Standards and Technology 7.3 CERT Activities - Dain Gary, Software Engineering Institute, Carnegie-Mellon Uni- 	38 41	
	versity	43	
8	Session on Workshop Recommendations8.1 General Discussion8.2 Elaboration of Recommendations	45 45 49	
A	View Graphs for Section 4.1	52	
в	View Graphs for Section 5.1	56	
С	View Graphs for Section 5.2	59	

D	View Graphs for Section 6.2	62
E	View Graphs for Section 6.3	65
F	View Graphs for Section 7.1	69
G	View Graphs for Section 7.2	73
н	View Graphs from Xerox Special Information Systems - Russ Housely	78

1 Executive Summary

1.1 Purpose of the Workshop

Under the High Performance Computing Act of 1991, the National Science Foundation (NSF) is charged with investigating the security of the National Research and Education Network (NREN) and the National Institute of Standards and Technology (NIST) is charged with developing standards for the Federal information systems component of the NREN. In the discharge of these responsibilities, NSF and NIST organized this two-day security workshop.

The purpose of the workshop was to bring together experts, conversant in the technologies, to develop recommendations for enhancing the security of the NSFNET/NREN. The focus of attention was on those technologies that could be immediately deployed with the intention of first improving the security of the NSF-funded supercomputing centers. Scalability and exportability of the technologies were considered desirable attributes, but secondary in the sense that they were not to stand in the way of immediate deployment.

1.2 Structure of the Workshop

The workshop was divided into six sessions. A brief introductory session established the objectives and scope of the workshop. This was followed by four formal sessions, each of which was concerned with a major area of security in distributed networks – authentication, access control, application security, and security management. A final session was devoted to the formulation of workshop recommendations. Each of the formal sessions typically consisted of one or more prepared presentations followed by an open discussion period. The presentations were given by representatives of government agencies, national laboratories, and federally funded centers. Throughout the workshop, there were free exchanges of ideas among the participants, making it sometimes difficult to separate the formal presentations from the planned discussion periods. Also, because of the interdependence of the four specified areas of security, the discussion on any particular issue was not necessarily confined to its designated session.

These proceedings present an abridged and edited accounting of the various presentations and discussion periods, with the intent of capturing the substance of ideas and reporting on the essence of what transpired.

1.3 Workshop Recommendations

In formulating recommendations, the participants took into consideration a number of factors: objectives of the workshop, currently available off-the-shelf technologies - both hardware and software, interoperability of technologies, cost, ease of use, acceptability by the user, and also the secondary issues of scalability and exportability.

The recommendations brought forth by this workshop are purposely limited in their scope in recognition that it is currently impractical, if not impossible, to draw hard, security boundaries around the supercomputer sites. In view of these limitations, the recommendations contained in this report are considered to be initial steps that address only parts of the security problem. While not providing a complete solution, it is considered essential to the interests of progress that these steps be taken at this point in time. The recommendations, summarized below, are extracted from a more elaborate discussion presented in Section 8 and are restructured to reflect the formal session topics of the workshop. The recommendations do not compete with each other; rather they complement each other. They are prioritized into categories A and B to indicate a relative sense of urgency. However, all recommendations are considered important and should be implemented immediately.

- I. User Authentication (Priority A)
 - Avoid use of static (i.e. reusable) password authentication systems.
 - Use some form of challenge/response (C/R) authentication system if Kerberos is not implemented or not available.
 - Use a publicly available algorithm.
 - Use an exportable authentication algorithm.
 - Use a public key based authentication system when possible.
 - Allow for user access from multiple sites, including across international boundaries.
 - Implementation should allow for authentication via host software, smart tokens, PC software, special hardware and should be amenable to precomputation for printed C/R lists.
- II. Access Control/Authentication (Priority B)
 - Use Kerberos (Version 4) now.
 - Plan for transition to Kerberos (Version 5).
 - Review use of Distributed Authentication Security Service (DASS) as development proceeds.
 - Assure conformity with application access control (e.g. rlogin, telnet, ftp).
 - Support alternative ways of generating/providing Kerberos keys (e.g. smart cards, eeprom, software, user selected passwords/passphrases).
- III. Application Security (Priority B)
 - Implement and use Privacy Enhanced Mail (PEM).
 - Initial applications of PEM should include security administration/management.
 - PEM should be available for users to protect mail (integrity, confidentiality, signature).
 - Initially use existing suite of cryptographic algorithms (DES/RSA/MD2/MD5).
 - Explore later use of proposed NIST suite of cryptographic algorithm standards.
 - Support and use certificate registration authority infrastructure.
- IV. Security Management
 - Establish security responsibility (e.g. security officers) (Priority A).
 - Develop and maintain site security policy and procedures, including a policy for handling security incidents.
 - Support comprehensive security education programs for users.
 - Establish Forum of Incident Response Security Teams (FIRST) point of contacts.

- Establish configuration control for security purposes (Priority A).
 - Supercomputer sites are high priority.
 - Distribute/support use of automated security management tools (e.g. COPS).
- Establish a security perimeter protection capability (**Priority B**) (e.g. filters, control gateways, Simple Network Management Protocol (SNMP)).
- Establish security audit information collection and review capability (Priority B).

V. Follow-on Activities

- Maintain cognizance of new/developing security technology.
- Establish an agenda for follow-on activities such as additional workshops in specific security areas (e.g. signature certificate registration authority infrastructure).
- Establish security enhancement specification/implementation teams.



2 Workshop Participants

Aiken, Bob	National Science Foundation
Branstad, Dennis	National Institute of Standards and Technology
Cerf, Vint	Corporation for National Research Initiatives
Crocker, Steve	Trusted Information Systems
Dray, Jim	National Institute of Standards and Technology
Ellis, Jim	Computer Emergency Response Team
Ford, Peter	Los Alamos National Laboratory
Gary, Dain	Computer Emergency Response Team
Housley, Russ	Xerox Corporation
Jansen, Wayne	National Institute of Standards and Technology
Kawamoto, Shirley	Mitre Corporation
Kent, Steve	Bolt, Beranek, and Newman
Laws, John	British Defense Research Agency
Mansur, Doug	Lawrence Livermore National Laboratories
McNulty, Lynn	National Institute of Standards and Technology
Nessett, Dan	Lawrence Livermore National Laboratories
Newman, Gerard	San Diego Supercomputing Center
Oldehoeft, Arthur	Iowa State University
Rosenthal, Robert	National Institute of Standards and Technology
Schiller, Jeff	Massachusetts Institute of Technology
Smid, Miles	National Institute of Standards and Technology
Staton, Hal	National Security Agency
Steinauer, Dennis	National Institute of Standards and Technology
Wolff, Steve	National Science Foundation



3 Introductory Session

Session Leader: Dennis Branstad, National Institute of Standards and Technology

The purpose of the opening session was to articulate the objectives, scope and desired outcomes for the workshop. The participants from the sponsoring Federal agencies – the National Science Foundation (NSF) and the National Institute of Science and Technology (NIST) – expressed their views on scope and time-frame for implementation.

Lynn McNulty of NIST stated that the overall goals of the workshop were to assist NSF and other Federal agencies in improving the security of existing networks and to use this as a spring board to address similar problems in evolving high-performance networks.

Since NSF is charged under the High Performance Computing Act of 1991 with investigating the security of the network, Steve Wolff and Bob Aiken articulated their views on what NSF would like to see as guidelines for discussion and eventual recommendations. It was noted that there appear to be a lot of things that can be done in terms of authentication and securing resources on and off the network. By bringing together individuals who were conversant in the technologies, the hope was for some recommendations on how to begin to do this. The focus should be on technologies that can be implemented immediately with specific attention given to enhancing the security of the NSF supercomputer centers – recognized as valuable resources on the NSFNET. It was noted that each of these sites has thousands of users spread around the world.

Scalability, although a desirable feature, was considered to be a secondary issue in the sense that there are only four NSF-funded supercomputer centers. Scalability of up to 100 host computers will be sufficient for the near-term; scaling to 10,000 or more host computers is not necessary at this point in time. If the workshop recommends technologies that do scale to the larger scope of the NREN, then that would be considered a bonus.

A question was raised on the issue of exportability. The NSF position was that, although exportability of the technologies is a desirable attribute, export regulations should not stand in the way of getting started on the implementation of security measures. Dennis Branstad from NIST noted that the networks being discussed are international in scope and therefore international issues have to be considered; the only thing to be ruled out is a discussion of special cryptographic algorithms that may or may not be exportable.

Steve Crocker from Trusted Information Systems proposed for consideration a list of things that he felt could be implemented immediately:

- 1. use of stronger methods to tighten access controls, ranging from more rigorous administration of passwords to automated mechanisms such as token-based systems, one-time passwords, and perhaps smart cards;
- 2. use of front-end filters for cost-effective perimeter control, e.g. routers with access control lists, specialized gateways and front-end boxes;
- 3. use of audit trails with logs of relevant security events;
- 4. use of tools for storage of information in tamper-resistant ways;

- 5. use of privacy enhanced mail for sensitive communication for administration and security officers, Kerberos, Distributed System Security Architecture (DASS);
- 6. raising the level of attention with respect to configuration and operation of security relevant portions of the system within a particular site, particularly the domain name service; and
- 7. a potpourri of administrative and operational considerations development of security policies at each center, education of people, careful configuration of the cryptographic parts of each site, fire drills so that timely and effective action can take place when a intrusion is discovered to be in progress, and spending time keeping abreast of the technologies.

4 Session on Authentication in Distributed Networks

Session Leader: Steve Crocker, Trusted Information Systems

4.1 NIST Smart Card Authentication/Access Control - Jim Dray, National Institute of Standards and Technology

(Supporting view graphs are presented in Appendix A.)

DRAY: NIST is currently involved in a project called the Advanced Smart Card Access Control System (ASACS). Funded by DARPA, the overall objective is to combine cryptographic (password-based) authentication techniques with token or smart card technology to achieve better security than is normally possible using passwords alone. The desired end product would be a practical system, that is not prohibitively expensive, acceptable to the end users, and easy to manage.

Conforming to the ISO definition, the NIST smart card is the same size as the standard credit card. It has a single Hitachi H8/310 chip microprocessor with 10 Kbytes of ROM and 8 Kbytes of EEPROM. The card can perform cryptographic authentication, using DES, to multiple host computers in automated fashion. It currently supports the ANSI X9.17 key distribution protocol. The encryption rate (approximately 500 bytes/second) is considered adequate for authentication. One of the main goals of this project is to implement public key cryptography in the firmware and, at this point, the Digital Signature Algorithm (DSA) and one other algorithm have been successfully implemented in EEPROM. With EEPROM, the cards are flexible, allowing for reprogramming of the firmware to meet different requirements or fixing of bugs.

The current network testbed consists of an ethernet backbone with a number of computers – mostly Sun and Unix workstations and MS/DOS PCs. Each machine that can support an interactive login has a smart card reader/writer (R/W) attached through an RS232C serial interface. The authentication protocol consists of a three-way handshake for authentication of card to a host computer and also host computer to a card.

NIST is currently working with two different kinds of R/Ws. The R/W for the original model is simple and fairly inexpensive, providing communications and power for the card and an oscillator for a system clock. The more recent R/W developed specifically for this project is portable, providing the option of connecting through a serial port, but it also has a keypad and display allowing for a manual challenge/response (C/R) authentication when the first option is not available.

The firmware implements a set of commands. The host computer will issue commands in proper sequence to accomplish authentication and the card responds by sending back whatever data is appropriate to each command. Seven to nine commands have been added to support public key cryptography. In addition to the encryption algorithms, there are a lot of the other things in ROM, and space is tight at this point. With the next project in the series, something will have to be taken out in order to make room.

Initialization of a card is performed by a security officer or system administrator. A card is initialized for access to a particular system by issuing to it a series of commands. This amounts to loading onto the card some personal identification numbers (PINs) so that the security officer and user can authenticate themselves to the card, and at least one DES cryptographic key for authentication. The ASACS authentication sequence for logging onto a machine with a R/W can be implemented in a number of different ways. In the NIST approach, the host machine would first reset the card and then prompt the user for the PIN which has to be sent to the card. The user is authenticated to the card by entering the correct PIN, either through the keypad on the portable R/W or on a keyboard via the host machine. The former path is considered more secure. The card sends back an authentication number that the user can look at to determine whether it is the the right card. This is not considered a strong security mechanism – it is similar to writing your name on a card and having it embossed.

The important part of the process is the three-way handshake – the C/R protocol, for bidirectional authentication. The host computer directs the card to generate a random number (RN1). The host computer takes RN1 and encrypts it using the DES key that it shares with the card. It also generates a second random number RN2 and sends it along with encrypted RN1 back to the card. The card will decrypt the encrypted RN1 and verify that it matches the previously sent value. It then encrypts RN2 which is send it back to the host for decryption and verification. The protocol provides each side with assurance that the other side possesses the correct DES key. Currently, this handshake is performed over an ethernet between Sun workstations running TCP/IP. It is also running under DOS and NIST is looking at PC/NFS. The protocol is independent of the lower layers in the network.

An automated interface between the token and the host computer makes things more convenient for the user. If a user is logged on and and wants to connect to a remote machine, a program (written by NIST) will request from the card the list of host machines with which the card shares authentication keys. The program displays this in a menu format and the user can select the desired remote host. The card will then go through the same C/R protocol with the remote machine. This is transparent since the user is already authenticated to the card and the card can now act as an agent for the user. At one point, the card had enough available storage for up to 100 remote host authentication keys – they do not have to be different computers, they could be programs on various machines running at different security levels.

The DES keys, shared with different computers, can be entered manually or one can use an automated protocol. In the manual case, a security officer would load the key for whichever machine was needed, using a specific command on the card called "load key". In the automated case, one can use the X9.17 protocol and distribute it through a network. The very high-level protocols needed to support this have not yet been addressed. It is currently being done manually.

The biggest advance in this project is the public key capability and what it can do in terms of key management. The approach taken by NIST was to implement at a lower layer of the firmware those fundamental operations that are required by most public key algorithms. This allows for implementation of a variety of public key algorithms in the next higher layer.

Reprogramming the firmware, modifying it, or completely replacing its contents is accomplished by generating an executable image of the new firmware contents and downloading it to the card. There is a mechanism whereby one can lock the card so the firmware can no longer be modified, thereby preventing people from changing their cards (although there may be some cases where certain portions of it should be modifiable). To unlock the card would require tearing the chip apart and modifying individual memory locations, or playing with the system in some very sophisticated ways. ELLIS: Do you have a target cost factor for the card and an R/W?

DRAY: Right now, the portable R/Ws are selling for a little under \$500. The simpler R/Ws, in small quantities, cost \$275. Smart cards, in quantities of 10,000, cost \$10 or less. NIST costs are high, because it is buying small quantities and is paying for development. In small quantities, smart cards sell for around \$100 which is fairly expensive.

CROCKER: Another question deals with lead time. Suppose an immediate result of this workshop is the recommendation that this technology should be deployed in the supercomputer centers. How long will it take to go from signing the purchase order to an implementation of this in the field?

DRAY: A number of companies sell ISO standard R/Ws. The NIST group was able to get one to work with a very minor modification. There have been estimates from other persons of something like 12 weeks to manufacture cards.

CROCKER: Aside from the mechanics of getting R/Ws or cards run off the assembly line, there are details of development and all the ramifications of going through that cycle. How would we relate to and connect it to any specific system? I think if we pushed it hard, it would be a year from now, maybe two years at the most, before the technology is ready for on-the-street, wide distribution.

DRAY: One of the big problems with anything like this is that there are no standard interfaces to, for example, operating systems. So, in a lot of cases, it ends up being a custom port. We have it running under UNIX and DOS, but we may not have it running the way someone would want it to work in a particular application. A lot of the work that is going on with common authentication technology and things like that should help that situation. How long it is going to take someone to integrate something like this into a particular system is hard to guess without knowing more details. The place to start, as someone said earlier, is with a strong authentication mechanism and build on that. It would be nice if all that work had already been done but in most cases, it has not been done.

AIKEN: From the user's perspective, suppose I have accounts on a number of machines at different sites. I have one card and I need a key for each host that I am going to log onto. Does that mean I am going to have to deal with each system administrator of each host to get keys for those hosts? How do I know what keys to use as I migrate to different machines?

DRAY: From the card's perspective, each key is stored with a unique key identification. What we have done is make the key "ids" the same as the names of the host computers on our network so that when you tell the card to do an authentication with a particular host, it can look up the corresponding key by name. The user does not have to know about that. In our simple demonstration, the user simply runs a program that asks the card for the key names that correspond to host names.

CROCKER: The point that is being stressed here is that the smart card can hold multiple disjoint sets of information to allow independent authentication to different systems. That can be contrasted with current kinds of token-based systems where if you want to authenticate yourself to multiple systems, you are going to need coordination among those systems in terms of the single identity that is known to each of them.

SMID: Another thing to mention is that our next version is going to provide for a public-key approach to distribute secret keys.

STATON: Thinking of the number of users, 4000 to 5000 for each super computer center, do you have ideas on how to manage all of this?

DRAY: A lot of that is tied up in whatever higher level authentication protocols are used and how one manages public key certificates. There are a number of ways you can do it. We have not addressed that; we just put in the hooks. Some of the other standards and publications we are working on address things like that. In this particular project, we have not gone to that level. We may end up working with a number of other government agencies – that would be one of the things we would have to do for them.

KENT: Given the world-wide distribution of supercomputer users and the use of "guest" cards, have you explored the likelihood of getting a general export license for this technology since it has a more generic encryption capability?

SMID: We can either make it with the features we want or we could make it exportable. You can take out the commands or the calls to the commands from the EEPROM, or you can even put in new commands if you wish before you lock the card. You could take out the generic encryption capability and if you had an exportable algorithm and it fits in the remaining memory, you could install it in the EEPROM and make an exportable version of the card.

CERF: With regard to the requirements of the NREN, it seems to me that we cannot escape the likelihood that much use of the system is going to take place outside of the U.S. by citizens who are traveling or who are on leave or who are at various other institutions. It will happen because of collaborative activities that take place in NASA, DOE, and elsewhere. So it seems to me that it would be a terrible mistake to imagine that we should somehow carve up these cards into exportable and non-exportable versions because surely we will want to do the best we can when we are operating outside of the United States. I am concerned about making the card less functional for use outside the U.S. when that is exactly when you want it to function best.

CROCKER: The advice we have been given is on the conservative side. If the algorithm is for authentication, an easier set of rules apply. If it is for confidentiality, there is a very stringent set of rules and and it will not be easy to export it. And if there is ambiguity, that is, if the capability is not focused only on authentication and it is not clear what it is used for, the default is to treat is as general purpose confidentiality. With respect to taking a system that has the full capability and stripping out the parts that are offensive, it has to be done in such a way that it is not obvious how to put it back together again.

SCHILLER: Taking your card right now with all of its capabilities, I could probably use it in the Kerberos system because I could use the generic encryption capability to do the required cryptography. If you take the algorithms out and I have to figure out how I can do Kerberos transactions with what is left, then it could be difficult.

CERF: I am sorry to sound like a broken record, but imagine that you are traveling either temporarily or even for some appreciable length of time and suppose that you brought your own laptop or notebook computer with you. You want to have access to the NREN and pretend for a moment that happens by way of the existing or evolving Internet. The last thing in the world that you want, most likely, is to plug into your host institution's LAN and have to read your mail in the clear, and have to carry out various transactions unprotected. Once again, the period of time when the most significant interest in protection, both confidentiality and authenticity, arises is when you are

outside of your normal domain.

CROCKER: Let me propose a division here so that, when we talk about access control and talk about authentication, one way or another we can probably choose the technology so that export issues do not get in the way. This discussion we are having here is part of the presentation on the smart card project, but I intend for my statement to be outside of that context. If you want to talk about protecting confidentiality of the interactives internationally, you are in a whole new ball game with respect to export controls and that is independent of whether you can do it on a card or if you can do it in software or if you can do it in a different piece of hardware. The question of whether or not the powers that be want to protect the confidentiality and the mechanisms is sort of irrelevant at that point.

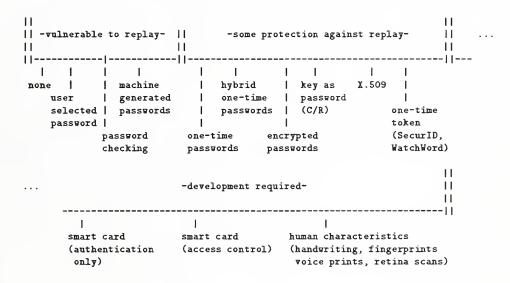
SCHILLER: Its more than export control. I suspect if I take a laptop with me overseas and do an encrypted connection back home to read my mail, I will probably have a number of people annoyed with me.

MANSUR: I am not sure that is true. My understanding is perhaps that an American citizen can take something overseas and use it and bring it back as long as it is not being sold over there.

BRANSTAD: There is at least one solution to it – there are a couple of algorithms for confidentiality purposes that have just been approved for export.

4.2 General Discussion on Authentication

Branstad constructed a chart depicting a spectrum of possible authentication technologies. During the course of discussion as reported below, this chart evolved into the following form:





It was noted that, in many systems, users are still allowed to select passwords of length four. Newman said SDSC imposes a eight-character minimum whenever possible. In the majority of these cases, the UNIX operating system creates the password and it is stored in encrypted form. There are some additional enhancements like password expiration, but there are currently no restrictions on reusability.

Rosenthal stated that NIST started a research project two and one-half years ago on machinegenerated passwords with the intent of publishing an algorithm that would be useful in automatically generating pronounceable passwords. They chose an algorithm that been available for more than a decade. The hope is to eventually issue it as a Federal standard.

Branstad defined a "one-time" password system as one in which the user is issued a list. On each login, the user must select the next password on the list. The system would generate the next password and, upon successful comparison with the user entry, would then expire that password. Schiller pointed out that a useful enhancement is to generate ahead of time a number of passwords (say 100) and allow the user the choice of selecting any one from the list.

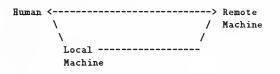
Branstad defined an "encrypted password" as one that is encrypted under some key as it passes through the network and is decrypted by the (remote) host and validated. It could not be compromised in transit and would presumably carry a time-stamp to counter attempted replays.

Branstad defined "key as a password" as a C/R system where one waits for a connection to a particular host instead of passing a password. Then one passes a value, encrypted with some authentication key, to the host which in turn decrypts it. Timestamps are used to guard against replay.

Schiller explained the Multics algorithm. In the Multics password generator, the host computer stores a seed to the pseudo-random number generator and a sequence number, i.e. the point to which one has advanced in the sequence of pseudo-random numbers. The user, upon logging onto the host, provides the current password and then advances the sequence number by one so that the subsequent login will require the next password in sequence. Passwords are converted into character strings that can be typed. A user at the home office machine can run the same program and generate as many "advance" passwords as desired. So, if the user is planing to travel, this program is used to generate in advance a list of passwords which is carried in person. Upon return to the home office, the user tells the program how many passwords were used while traveling. So this enables the traveling user to generate passwords without having to engage administrative overhead on the host. Schiller explained the hybrid one-time password system as one that eliminates the need the need to carry a list. Suppose one does "not" want to use a new password at every single point. As long as the same password is used, the host does not advance the position in the sequence. As soon as the next password in the sequence is used, the host will also advance to the next one.

Cerf wondered if the assumption that we are authenticating "people as users" might be too narrow. Software processes initiate transactions that access resources on the network requiring authentication without human intervention. Crocker echoed this concern with a picture of possible communications, noting the distinction between

- a direct connection between some human and a remote system in which you just type basically what we do today with remote terminals, and
- carrying around a notebook computer or PC or some other computationally significant device with the requirement of machine-to-machine communication.



Cerf felt that this picture still implied that a user is actively involved. He imagined a large number of connected machines on which there is software initiated by some user that subsequently runs independently of the user, although on behalf of the user. An example cited by Mansur was a transaction that requested information on a given topic, resulting in processes that reach out across a thousand systems. Cerf cited another example of initiating periodic probes, resulting in the operating system periodically creating processes, each of which has to be authenticated. Another example cited was an electronic cash problem where one wants to execute anonymous transactions. Steinauer noted that underlying all of this is the requirement of accountability back to some person or legal entity. So the question that needs to be asked is "Are there situations where you don't need or want that accountability?" Steinaurer did not think so!

Subsequent discussion led to the establishment of demarkation points on the above spectrum chart for distinguishing those authentication technologies that are less/more vulnerable to replay and those technologies that are presently available/unavailable. Crocker suggested that the focus of attention be given to those technologies that have some protection against replay, are reasonable to deploy (in terms of cost and difficulty), and meet the current needs (and beyond) for potential payoff in a broader part of the NREN community. Nessett felt that tokens are not practical in the short term because of the need for a lot of people to buy hardware and the difficulty of administrating the system. Ellis felt that there may be groups of organizations (and people) that will be able to afford the higher class options and we should encourage others to also consider the possibilities. It was noted that the discussion had led to the following set of possible candidates for authentication: onetime passwords, keys as passwords, and encrypted passwords (X.509 protected simple protected). Newman stated that things to the left of the first demarkation line (user selected passwords and machine-generated passwords) are unacceptable in the supercomputer centers because they do not provide sufficient security.

At this point, discussion turned to the criteria that should be used to evaluate the specific technologies. Cerf proposed the first six metrics in the table below for measuring what technologies can be practically and satisfactorily deployed at this point in time. The metrics were not intended to be any specific order. In the course of discussion, Crocker proposed adding metrics 7 and 11, Nessett proposed adding metrics 8 and 9, and Wolf proposed adding metric 10.

- 1. Costs
 - development
 - capital (acquisition cost)
 - operational (administrative costs)
- 2. Availability of the Technology
 - broadly available vs. only from certain places
 - available from numerous vendors
 - proprietary or open
- 3. Scale of Deployability
 - some of the ideas may not work over thousands or millions of users
 - broadening the scope to the NREN (not only supercomputers) requires thinking in the range of 10M users
- 4. Degree of Protection
 - different kinds of threats and vulnerabilities should be considered (in addition to the "replay" threat)
- 5. National Interests or Policies
 - must be taken into account
 - NSA problems
 - exportability issues
- 6. International Availability of the Technology
- 7. Consistency with Longer-Term Architectures
- 8. User Acceptability
- 9. Administrative Pain
- 10. Degree of Vendor Involvement
 - dependency on vendor
 - level of vendor support
- 11. Common Use Across Multiple Systems

Proposed Metrics for Measuring Current Deployability of Technologies Wolff reiterated that need for the workshop participants to focus on near-term implementation and the protection of a small number of resources on the network. "Small" was to be taken as probably less than 100 and certainly less than 1000. Anything larger constrains the things we are able to think about. If all things are equal, one should choose scalable solutions.

Kawamota suggested that the need to support a traveling (roving) user might influence what solutions would be acceptable. Crocker pointed out that there may be a distinct difference between someone who is traveling to Zurich carrying perhaps a PC and someone who is in Zurich with potentially powerful computing facilities (and potentially affected by export regulations).

Smid noted that, at some point, it would be necessary to identify the features that are felt to be absolutely essential and those that are felt to be less essential - for example, if one has a physical token, does it have to authenticate to any of a desired set of computers or is a different token needed for each computer. Discussion ensued on the need for multiple authentication algorithms. Wolff pointed out that there was presently some discussion taking place about a "National Machine", which will involve all the NSF supercomputer centers in an allegedly seamless distributed computing environment. Ford noted that the centers that are being formed under the HPC Act are also distributed. For example, the Los Alamos and Oak Ridge facilities are supposed to be multisite and NASA is in the process of bringing up sites that include both NSF and NASA principal investigators. Aiken expressed a desire for the workshop to address authentication in the context of a totally distributed environment with a large number of principal investigators who access multiple machines at multiple sites and teleconference between the centers.



5 Session on Access Control in Distributed Networks

Session Leader: Vint Cerf, Corporation for National Research Initiatives

5.1 Kerberos Authentication/Access Control System - Jeff Schiller, Massachusetts Institute of Technology

(Supporting view graphs are presented in Appendix B.)

SCHILLER: The traditional approach is to implement a security perimeter around a host with the requirement of authenticating the user to the host. However, in a UNIX network, the components of the system are distributed and there are several different perimeters that need to be protected. What is needed is authenticated access across the network as opposed to distributed network protocols that use the "trust me" approach. Some method is needed for authenticating things other than just human beings.

The Athena computing environment has client workstations placed in public places with untrusted software; the root password on all the Athena-based client stations is public knowledge. Servers are placed in locked rooms but are considered only moderately secure since numerous individuals are able to access these rooms; no secret information is stored on these servers; the compromise of any one server will not result in the compromise of another server. The three key distribution centers (KDCs), with unpublished alarms, are placed in physically secured areas; they use symmetric cryptography (DES algorithm) implemented in the software.

The goals addressed by Kerberos are detection of spurious association initiation (authenticity), detection of message stream modification (data integrity), prevention of release of message contents (confidentiality). Traffic analysis and denial of service have not been addressed.

Kerberos provides authentication, not authorization. The idea of a centralized server for access control decisions was rejected at the outset since managers of servers may not want to delegate that decision to some third party. Also, authorization is application specific. For example, a file server may allow operations "read/write/execute" and perhaps "delete". For a print server, the operation may be "print this document against a certain account". So rather than building a system that would be expandable or general enough for all possible authorizations, the simpler problem of authentication was addressed.

The design goals included such things as no cleartext passwords transmitted over the network, no client cleartext passwords on the servers, and minimal exposure of client/server keys (by not encrypting too much within one key). Compromises should affect only the current session – prevention is effected through session keys that have limited lifetimes. Passwords or derivatives of passwords are not stored on local workstations or servers. All of this is transparent to normal users as the interface projects the image of a standard UNIX login protocol.

It was assumed that the DES is good enough to protect the resources and would be widely available in the software. It turned out that not even the KDCs, implemented on the slowest machines, need a hardware-implemented DES. The RSA algorithm was not adopted due to the expense of computation and expense of licensing. The expense of computation is now questionable, but licensing is still a problem. Also, it was assumed that a global clock would be available.



The Kerberos Model is based on the Needham & Schroeder algorithm with the private keys of all principals stored at a trusted KDC. A session key is created and down-loaded to the two principals of communication along with assurances that they are who they claim to be. The exact nature of the communication is as follows. The client talks to the KDC, gets tokens, and sends them to the server (there is no direct communication between the KDC and the server). Rather than discuss the details of the protocol, which can be found in numerous papers, this discussion will focus on where Kerberos fits in, what is available today, and plans for the future.

Kerberos is a deployed technology with commercial support available for everything from installation help to total turn-key operation to end-user support. Kerberos, version 4 (v.4) is presently distributed by MIT within the U.S. free of charge. There are Kerberized versions of the Berkeley UNIX r-commands (e.g. rsh, rlogin) and there is an encrypted rlogin command and an encrypted rcp command that use Kerberos authentication. There is a Kerberos-authenticated implementation of the telnet command. Kerberos, version 5 (v.5) is currently in Beta test. A version of Kerberos is shipped with DEC's ULTRIX operating system, modified to not support confidentiality. TGB, Inc. has or will shortly have a version of Kerberos for DEC's VMS with the entire suite implemented which includes all of the utilities as well as the KDC. MIT has ported Kerberos to the MacIntosh

How to Kerberize the ftp command is not clear because of its two-connection nature and because it is not clear what people really want. For example, ftp might use Kerberos only for authenticating the control connection and do nothing with the data connection. If so, the Kerberos-authenticated rcp command is much easier to deal with because there is only one (secure) connection.

V.4 was designed to meet the needs of MIT. V.5 evolved from suggestions from various vendors on certain things that were incompatible with their versions of UNIX. In v.5, one can have multiple realms as separate independent entities that share keys with each other and interoperate. One can chain through multiple realms. The maximum lifetime of a ticket has been increased and renewable tickets have been added to handle lengthy computations and batch processing. A renewable ticket has both an immediate lifetime and an ultimate lifetime. The ultimate lifetime can be long but, when the immediate lifetime is close to expiring, the ticket has to be sent back to the KDC for reissuance. The reissued ticket will have a slightly longer immediate lifetime. There is no consensus among vendors on naming - everyone seems to want their own scheme. For acceptance, names are defined to be an array of strings. This could provide some future difficulties in v.5 because two implementations may be compliant with the specifications but might not interoperate because of incompatible naming schemes.

Plans for the immediate future include finishing v.5 (requires a detailed in-depth code audit), finishing the administration server, and working with DEC to combine with the DASS technology so there will be a unified system that meets the needs of both communities (X.500 distinguished names have been agreed upon). The idea is to allow an organization that does not want to use public key cryptography (for whatever reasons), but wants to use a Kerberos-like approach, to interoperate with another organization that uses public key cryptography much like the DASS approach.

CERF: Is the KDC replicated for availability and reliability?

SCHILLER: Special care is taken to keep the KDCs running. One subtlety of using symmetric key encryption is the tradeoff between replication (for availability and reliability) and vulnerability - brought about by providing multiple centers (to attack) that contain all the keys. MIT has only three replicas of the KDC, each with an uninterruptible power source, located in the best possible

places. They listen only to the Kerberos service port and not to anything else (such as finger, sendmail). All management must take place on the console or via a back port that requires an encrypted session. There are no user accounts – one must login as root.

When dealing with interoperability of multiple realms, the amount of trust one realm has to place in another is limited. Suppose one realm is less secure than another and a key is shared between them. If the first realm is compromised by a perpetrator, the perpetrator can masquerade as a user only of the first realm, but not as a user of the second realm.

CERF: What is needed in compatibility of naming conventions to make things work?

SCHILLER: Three different naming conventions that immediately come to mind are OSF, MIT version 4 Kerberos naming convention (which is of the form "user.instance@realm"), and X.500. Once could envision writing some translation software to glue it all together. Otherwise, there is no way of comparing names. Looking at an access control list on a file, one expects to see names of a specific syntax. Its not a matter of authentication, it is a matter of providing syntactically correct information to application programs (which assign meaning to names). For "n" naming schemes, the complexity of the translation problem is of order "n-squared".

KENT: Does the facility that allows one to choose a naming scheme provide for identifying different schemes?

SCHILLER: The facility allows identification of different naming schemes. At least, it can detect when it is dealing with a different naming scheme as opposed to a broken system.

KENT: In an integrated authentication system, the structure of the name has some significance because that determines the path of certification. In earlier discussions about a traveling user, if someone establishes name-based access control for access to some application, then it might be desirable for it to work whether accessing that application from home or while on the road. And that requires a certain degree of uniformity in expressing names.

SCHILLER: (continuing the presentation) If someone is interested in using Kerberos, v.4 is recommended today because it works and it is supported by several vendors. There are terminal server vendors working on Kerberos. MIT will come out with a transition plan from v.4 to v.5 to convert its own community of users. So an upgrade path will be available. Similarly, MIT will have an upgrade path from v.5 to the Internet authentication system, if and when they actually get around to building one.

Scalability is not a problem. MIT operates a database with 20,000 entries and has tested it with as many as 250,000 entries. On any given day, as many as 5,000 uniquely identifiable individuals log into the system. The KDCs run on MicroVAX IIs. When a KDC is needed, the master server is tried first; the slaves are tried if the first packet is lost. The slaves get hardly any traffic, which means the master MicroVax II is able to handle all of the traffic.

We have done quite a bit of testing with inter-realm communication. Not every application does it right. Kerberos does it right because for applications that are inter-realm aware (e.g. the MIT bulletin board system), authentication works just fine. The problem you get into again is translating names like NFS.

NESSETT: In applying this to NSFNET or NSF supercomputer centers, it may be the case that

each center would have its own KDC.

SCHILLER: Each supercomputer can operate in its own realm and realms share keys. Rlogin, telnet, rcp and rsh should work. With a little work, one can make NFS work too.

FORD: Supercomputer centers will have to deal with the startup and operational costs for typically 40 or 50 other institutions (if not more) that connect to them. At MIT, how much does it cost to continue to keep Kerberos going?

SCHILLER: That is where the the administration server comes into play. The administration server is the way to change passwords and also add users to the database (if you are appropriately privileged). Individuals can use this service to change their passwords. When they invoke the "passwd" command in UNIX, they are really invoking the administration server with a request to change their password. Password checking is now implemented, using a 150,000 word dictionary which also has some popular "nonwords".

FORD: Do you think that Kerberos is "the solution", "a solution", or the "beginning of a solution"?

SCHILLER: One of things we haven't done is determine the requirements of our user group. Depending on how that is done, I would say that this is "a solution" for networks. Kerberos is a great solution for a supercomputer center, with the KDC running on a workstation located in a secure place along with a staff to set up and administer a center-wide name space.

FORD: From the supercomputer center point of view, Kerberos seems like a good idea except when you start looking at the remote users in universities and laboratories. Does one ask the remote sites to adopt Kerberos? Or are remote users put in the supercomputer center's Kerberos database?

SCHILLER: I would recommend putting users in the supercomputer center database. I would distribute binary programs for all the major software platforms that are preconfigured. If you are willing to distribute binaries and there is no need for confidentiality, then exporting is not a problem.

NESSETT: In terms of inter-realm applications, you mentioned rlogin and some others. What about rcp, ftp and some of these other things?

SCHILLER: Inter-realm is not the problem in Kerberizing ftp. The issue with ftp is determining what the authentication service should guarantee. You wouldn't want to implement a Kerberosauthenticated ftp that provides no integrity on the data channel – especially since the encrypted rcp does provide this integrity (if you rcp a file with Kerberos, a check sum is sent).

ELLIS: Installing Kerberos at the supercomputer centers requires putting it on all the centers' workstations. How much work does that require?

SCHILLER: Kerberos runs under ULTRIX and we have it running under SUNOS.

KAWAMOTO: Can you give us your reasons for not going to a public key system?

SCHILLER: We may in fact do that in the successor to v.5.

5.2 Supercomputer Center Access Control Requirements, Dan Nessett, Lawrence Livermore National Laboratory

(Supporting view graphs are presented in Appendix C.)

NESSETT: I will talk about the requirements and threats for some applications in supercomputer centers. The first application deals with concurrent supercomputer computations. We have a CRAY and we are slated to get an MPP soon. There will be a job queue serviced by both machines and a batch scheduler. When a job is initiated by a user, it is placed in the queue for later execution. Another component of the system is a performance monitor and we are presently working on the communications platform. It would be possible and quite typical for each of these systems to be in different supercomputer centers.

CERF: In composing this kind of computation and the supporting system, it seems like the authentication and access control infrastructure will need to be accessible to a variety of protocols (not just the telnet and rlogin varieties) that reach beyond the boundary of a particular center.

NESSETT: We are currently looking at the infrastructure and will do performance analysis on a couple of protocols that operate on top of TCP and ISODE.

(continuing formal presentation) For managing software, there is a distributed "make" facility for users who want to maintain source files at their local workstation (or it might be a LAN of workstations sharing a source file). Users can build codes on both the MPP and on the Cray concurrently. There is a program that runs on the Cray - it is like a daemon except it does not need "root" privilege and there is an internal protocol between the client program on the workstation and the server on the Cray that is used to initiate other applications. The clients call the Cray server, known as Remoxe, which uses things similar to capabilities that are passed from the local workstation to the Cray. Inside of the capability is encrypted information that allows Remoxe to recover a user's password. The Remoxe server has the master key to decrypt this information. Remoxe obtains the password and uses it with the "su" program to initiate another process that executes the code.

CROCKER: How many clients have that password?

NESSETT: In registering with a remote server, a user must enter a system-specific password, different for each user and different for each system.

(continuing formal presentation) Let me cite some perhaps controversial threats. The hacker/anarchist threat exploits the philosophy that if you are not using your resources, then he/she should be able to use them. Another threat is a disgruntled/unstable employee who can disable the system, destroy critical files, etc. As NSF supercomputer centers support industrial collaboration, there is going to be a foreign intelligence threat. Another threat is industrial espionage where companies might want to know what their competitors are doing. A related threat is academic intelligence gathering where a university team is working on some problem and someone else would like to know what they are doing or find out about outstanding proposals. Finally, there is the threat of unauthorized use of the resources such as stock market analysis.

CERF: With respect to NREN, it seems to me that there are going to be a large number of government-owned resources that are connected to the network. There will also be private resources and user concern over who can use the resources or see the content of those resources. A major

issue is intellectual property rights and tools and methods for protecting them. Also, the integrity of databases in supercomputer centers need to be protected.

NESSETT: Given the threats, what are some of the requirements we have for access control? A major one is accommodating heterogeneity – especially for NSF networks. There are multiple centers and each will want to retain control of the resources rather than give it up to some higher authority. There will be multiple authentication mechanisms and multiple authorization mechanisms. There will be different physical security environments for the more powerful equipment in computer rooms and for the workstations in less protected areas. There will be different operating system vulnerabilities – even if everyone runs UNIX because there are different implementations. Some versions of UNIX still have the old problems and vulnerabilities, e.g. the sendmail vulnerabilities while others may have been corrected. The access control requirements require solving this heterogeneity problem.

We need to design for reality, not utopia. Not every operating system will be certified at level C2 or better. We want to be sure that the compromise of a system that is normally accessible to a user will not compromise the whole distributed system.

SMID: Do you assume the attacker has passive access or active access to the protocols, e.g. can an intruder see the user logon protocols?

NESSETT: A sophisticated intruder can compromise the system and observe the transmission of passwords on a LAN.

ELLIS: You have to assume that intruders have the protocols. The problem is not due to a lot of intruders being that sophisticated, but rather the fact that they are sharing their information. We are seeing very sophisticated attacks and once they get into the system, it is clear they don't know very much about UNIX. They simply have obtained programs and tools from others and are using them to break into a system. So you have to assume the attack will be sophisticated.

SMID: Can they also inject text to modify a protocol – an active attack like changing some of the parameters?

KENT: There is no motivation to do more sophisticated attacks if a less sophisticated attack will work and one of the things we have to be concerned about is the threshold that is established. Whatever we recommend here will clearly impose some amount of inconvenience on users and system managers. But if we don't establish a sufficient threshold, the intruders will shift.

NESSETT: I agree! The relevant categorizations are "difficulty" and "visibility". Suppose it is easy to do a particular intrusion and it is almost impossible to find out that it is being done. If you eliminate the possibility of that attack and make the intruder do something that's harder and more visible, then you have probably succeeded in terms of your security mechanisms.

(continuing the formal presentation) If one installs a distributed access control mechanism, then its execution should not require "root" privileges for several reasons. First of all, system administrators are apprehensive about installing this kind of software and it is long process to convince them that it really works and will not cause problems. The second reason is that root access opens up compromise possibilities because of the way UNIX works. If someone finds a missed vulnerability, it becomes a point of attack. In the short term, I think it is better to build on top of existing mechanisms. In the long term, it is probably better to integrate distributed system access control methods into the operating system. In any case, one should try to limit the amount of root privilege needed in access control mechanisms that are local to the system and then build distributed access on top of them. Even if you do that, in the long term, one will have to accommodate systems that do not have this integrated access control mechanism because of the way software is developed and promulgated in communities. Going back to the heterogeneity point, there will always be a mix of systems.

Access control mechanisms should be flexible and comprehensive. We talked earlier about needing some sort of delegation facility so that processes do not have to be directly connected to users and can still access those resources on behalf of the users. You need to support a wide diversity of applications such as login, file transfer and routing. Our Privacy Security Research Group (PSRG) made a long list of applications that are good candidates for adding security features. Care must be exercised in building gateways between similar applications like ftp, telnet, etc. If the two applications support different access control methods, a gateway can introduce hazards. The PSRG has a good start on identifying appropriate protocol layers for security services. In access control mechanisms, a way is needed to quickly revoke access to resources by misbehaving users.

MANSUR: The NSF centers are beginning to explore the possibility of tying in mass storage systems and whatever we do has to be directly applicable to that. Right now, mass storage systems are protected by the login protocols to the systems. If we make them NFS-accessible to the entire Internet community, we have problems – but that is what users want. There is this idea of establishing a place where people can contribute software via NFS.

Also, there are a lot of requests for privacy enhancement and privacy enhanced mail.

5.3 General Discussion on Access Control

Vint Cerf began the discussion by posing a question, "If the workshop participants felt it important to use Kerberos, which version should be targeted -v.4 or v.5?" He noted that Schiller had been cautious about v.5 since production was targeted for some time next year.

Crocker pointed out that security provided by Kerberos at MIT assumed that the user had control of a workstation within the Athena environment. If there was any doubt about the security of the local workstation, the user would reboot the system to get a correct copy of the software. He raised the more general question regarding "remote" authentication capabilities of Kerberos. It was noted by others that remote authentication implies that the workstation is in the Kerberos environment, e.g. it is supposed to have client software running right up to a trusted host. Physically, there is no limitation on how far the remote station might be from the host.

Crocker inquired about the memory requirements needed to run Kerberos. Schiller stated the binary file "rlogin" is required along with a few configuration files. This resides in about 50-60 Kbytes of memory on a VAX. The DES layer can be optimized to gain additional space. At one point, Kerberos was running on a 256K PC. It will presently run on a laptop computer.

The discussion turned momentarily to a consideration of how the need for timely deployment would influence the workshop recommendations. Aiken said that the target date for beginning the deployment in supercomputer centers was fall of this year. Ford noted that the objectives of near-term deployment mandated concrete recommendations based on immediately available technologies. Nessett felt that this would not be a problem for the supercomputer centers themselves, but that the issue would be getting the user community involved and up to speed.

Given the time constraints for deployment, Schiller recommended the use of Kerberos, v.4. Crocker felt that Kerberos was, among other things, a complete and sophisticated implementation of challengebased authentication, but the real question was whether or not Kerberos was widely enough distributed. Cerf elaborated on this thought by asking if current users of supercomputer centers could be made v.4 compatible with any reasonable degree of effort and whether there would be roadblocks in terms of export controls, in terms of operating systems, or anything else other than the hard work of getting it distributed.

Subsequent discussion noted that the use of UNIX (and the "rcp" command) by all users was not a valid assumption. A ubiquitous software product used in the centers is NCSA "telnet" for PCs and MacIntoshes, implying a need to implement a Kerberized version of NCSA telnet. Other applications will need "ftp" support. Schiller said that the "control" portion of ftp can be Kerberized with only three lines of codes, but that would not secure the data transmission part of ftp execution. Cerf suggested that perhaps the workshop recommendations should include statements to the effect that the proposed Kerberos v.4 solution will not initially encompass 100% all of the users because the software is not available for deployment everywhere and development might be necessary to incrementally cover an increasing number of the end users. Aiken suggested a switch-over date of one year from the start of deployment, but Crocker thought that might be a bit optimistic because a roving user might not be able to get to a workstation that was configured as a Kerberos client. Cerf noted that this identifies a policy question with respect to a supercomputer center operation – will it (with or without Kerberos running internal to the center) accept unprotected connections from the outside world.

Crocker suggested the use of a dual strategy in which the roving user (when traveling) would use a list of one-time passwords for authentication. Schiller believes that Kerberos is the most attractive solution today. Although smart cards are a good idea, card readers are not likely to become widely available in the near term future. The down side of using Kerberos is that, without a token-based system, the potential of poor password selection still exists and one is still subject to Trojan Horse attacks. There are cases where one-time passwords are still appropriate and useful with Kerberos. For example, a user can have the appropriate software on a personal workstation, containing Kerberos cryptographic tokens so that the workstation will claim to be the user. The user can telnet from elsewhere over an untrusted network to the workstation. A one-time password is used to login to the technologies of things like SecurID and WatchWord could be employed in these circumstances to telnet to a Kerberos client but that people would probably find that a complete Kerberos solution much more pleasant to use. Cerf suggested that the recommendations of the workshop might reflect that Kerberos should be used when possible, but these other techniques could be used until Kerberos could be installed.

A discussion followed on the trade-offs between Kerberos and other technologies based on one-time passwords or tokens. Smid said the Kerberos forms a good framework in which some of the other technologies could also be implemented to counter the problem of password vulnerability. With a one-time token, the password is changing and is attractive if a good crytographic generator is used. It was noted that the passwords selected by users of Kerberos at MIT were compared against entries in a large dictionary thereby eliminating many poor passwords, but the strength was not equivalent to use of randomly generated 56-bit key. Nessett said that some systems at LCC did not allow the user to select a password and in some systems, passphrase software was being used. Ellis said that Kerberos has one flaw that one-time passwords do not have – the trust one must have that the local machine has not been compromised and will steal a password when it is typed in. Cerf suggested that the evaluations of the workshop should distinguish what the technology is capable of doing, how people may typically use it, and what rules have to put into place to make sure it is used well.

In response to a question about the availability of systems that were equivalent to Kerberos, the discussion provided the opportunity to speculate on a number of things – the comparison of Kerberos with Distributed Authentication Security Service (DASS), the interoperability between public key and private key systems, Common Authentication Technology (CAT) as specified by ATI and working on top of DASS and Kerberos, the differences between Kerberos v.4 and v.5., and the upgrade path from v.4 to v5.

At this point, Cerf shifted attention of the workshop to a second set of question(s), "What kinds of audit capability are required to make a network-based system acceptable in terms of tracking down abuse? Are audit trail capabilities needed at all? And if so, is that already part of the Kerberos specification? Or is it orthogonal to it?"

Mansur considered audit facilities to be separate from Kerberos, noting that a number of groups were now building intrusion detection systems that used audit information. Kent noted that Kerberos could play a central role in the collection of audit information since the KDC could record information each time it granted a ticket. Ford thought that the fundamental questions were concerned with what aspects of activity to audit and how to make use of the information. With respect to meeting the audit needs of services, Kent pointed out that there was very little integration of the concept of audits and application packages. Crocker said that he considered auditing to be very important and he had included it on his list of immediately available technologies that should be deployed. Within the time-frame constraints imposed on implementing workshop recommendations, the only sensible thing would be to issue a statement that the supercomputer centers must have some audit capability. Standardization would be too much to ask for, but the centers should share their experiences of implementation and also the experience of analysis and ultimate incident discovery.

Mansur said that some places had identified a small set of choke points for the monitoring of traffic and the software watched for certain signatures (or patterns of activity) that would strongly suggest intrusive behavior. Cerf echoed a special need for observing activity that is initiated from the outside world. Ellis pointed out that auditing could run counter to rights of privacy and that it may be legally necessary to notify users when monitoring. Mansur said the Justice department is currently working on this very issue. A policy has been written, but not yet promulgated, that requires informing the user community of monitoring activities. He noted that this could be difficult to do for all possible remote connections. As yet, there have been no court challenges and so the similarities to wire-tapping activities have not been legally established. Cerf suggested that the workshop participants did not have the necessary background to argue the legalities, but that the minutes should show that the collection of audit information is an important issue that cannot be ignored since without it, there is no satisfactory way to identify abuse.

Branstad speculated that the need for audit information and the amount collected might be inversely proportional to the effectiveness of the system security. This provoked a number of counter points: no matter how secure a system is, if a compromise is possible, then auditing is necessary, e.g. the telephone companies periodically audit telephone calls for quality control and in some sense the issue here is of quality control of the security system; audit and intrusion technology are required when there does not exist a clear set of rules that are enforced 100% of the time, e.g. individuals may be allowed access to resources to perform their work, but auditing is required to detect abuse of these privileges; and computer systems are fundamentally flawed and will be for a long time to come and therefore monitoring is necessary, e.g. auditing is required in air traffic control systems.

Kent noted that compromises are also possible through exploitation of such things as weak passwords and that auditing is used by operating systems to determine what damage might have been caused by an unauthorized user posing as an authorized user. The difficulty is deciding when and how much of the auditing capability to turn on. Also, an "external" auditing capability is extremely important because the internal auditing mechanism might itself be compromised by an intruder. Mansur said that the main point was to have the ability to turn an auditing capability on if an intrusion is suspected. For example, a common intruder signature is to "finger" an account and then immediately login onto that account. While it may not be clear that an attack is actually in progress, if the activity is highly suspicious, it should trigger turning on the auditing facility. Cerf noted that if there were enough cases, a significant amount of information might need to be collected and a lot computation might be required for the analysis. Crocker suggested that an attacker who knows this can repeatedly use a signature to cause an overload of the auditing system – a form of denial of service.

Kent drew attention to the fact that the discussion up to this point had not addressed the designated topic of access control. He noted that Schiller had prefaced his discussion with the remark that Kerberos is an authentication system, not an access control system and that Schiller had also said that they could not figure out how to satisfy a wide range of users with a single access control service. Relevant questions for this session might be related to what granularity of access control is considered important or what the specific focus of access control should be for supercomputer centers.

Cerf said that one obvious access control issue is simply whether or not the remote user is able to run any programs on a machine. One might ask, for the current set of supercomputer resources that are available, what sort of current access controls are enforced? Do users have the ability to run arbitrary programs or are they restricted to only certain programs? The ensuing discussion concluded that, for the most part, access control is currently controlled by authorization schemes that come with and are resident in a host operating system. Cerf questioned what access control should look like in distributed applications where a user is not supposed to have an account on the machine in the classical sense, is not supposed to have full access to everything any user could do on that machine, and is only supposed to have access to a particular application that the system is willing to provide. Will systems have access control lists to control who can run applications associated with particular ports? It was noted that the current use of file transfer and electronic mail are very simple examples of distributed applications. Distributed computing, as might be implied by something like the National Machine Room Project, would require file sharing and cooperative efforts and would place heavy demands on remote files system interfaces (e.g. AFS, NFS or their successors) to properly control access. The more general case of applications that get requests for service transactions are not currently being run at the supercomputer centers with problems of access control. Medium-term, rather than short-term, solutions would be required. Crocker mentioned "certificates" as being an important concept in solving some of these problems. In bringing the focus back on short-term solutions, Aiken suggested attacking the problem of having telnet and ftp access to supercomputer centers without transmitting passwords in clear text across the network. In the grand scheme of things, this may seem miniscule, but it is a concrete step forward!

At this point the close-out discussion identified a a potpourri of issues that should be addressed (other than access control). If authentication mechanisms required the use of hardware, (e.g. smart cards and associated readers), an administrative issue of inventory control would be involved; disabling accounts would probably also mean reclaiming the smart card. Any short-term recommendation by the workshop should be evaluated in light of what the world is going to look like in 1993 and 1994; it would be unwise to put a lot of effort into a community of 20,000 users, only to have it undone in 12-18 months. It should be noted that it is not clear that Kerberos ought to survive the transition on cryptography. The issue of configuration control is of great importance and needs to be addressed.

6 Session on Application Security in Distributed Networks

Session Leader: Steve Kent, Bolt, Beranek and Newman

6.1 NSF Super Computer Center Security - Gerard Newman, San Diego Supercomputer Center

NEWMAN: I will start by briefly describing some of the current practices at the San Diego Super Computer Center (SDSC), especially on the systems with which I am most familiar. On those systems where password length is controlled, a minimum length of eight characters is imposed. On some systems, a dictionary search is used to check passwords; on other systems the use of a nonalphabetic character in the password is required. Unfortunately, the practices are not consistent across all of the operating systems and that is somewhat bothersome. There is no network filtering on routers. The idea of filtering outside traffic to prevent NFS access at the center has been considered, but it was felt that too many of the users would complain if, for example, the Berkeley r-commands were not available. At present, NFS access to the outside world from the MassPar is not allowed, but it will be necessary at some point in the future in response to significant pressure from users.

We have security policy that we are willing to share with others; it has already been shared it with NCSA. We have an incident response team (such that it is) of about three people. We maintain a contact in the local FBI office so they will not be taken by surprise if we call them for help on a computer security problem.

Configuration control for the Cray is handled by a set of rules that allow changes to be made only on Tuesdays and Saturdays to the Cray operating system software. The workstations (WSs) are more of a problem. Configuration changes are done somewhat haphazardly since we seem to have practically every kind of workstation.

Application security is a topic of keen interest among the center staff and director. Ever since we have demonstrated how easy it is to telnet to the SMDT port and make at least a passing attempt at forging mail, there has been concern about the authenticity of messages.

As for other applications of security, something is needed for mass storage systems. At present, the security is simply through difficulty of addressing the storage. We are under pressure to make the mass storage system more globally visible, but we are hesitant to do so because of the security concerns.

There are generally no restrictions on what programs can be run by users. There are a few proprietary packages that restrict access to subroutine libraries or restrict access to a specific set of users, e.g. academic researchers. This is accomplished through an access control list on the Cray. There are some things we do control, e.g. payroll which has to be the same mechanism every time it is run. But there are no other requirements.

CERF: Within the confines of SDSC center, I assume there are a lot of workstations in addition to the supercomputer centers (SCs). Even if you did a good job of controlling access to SCs resources from outside the SC center, is it true that the WSs are still a point of vulnerability because if one of these is penetrated, then its rlogin capability could be abused?

NEWMAN: The problem that SDSC has with security is not generally on the major resources (e.g. Cray, HyperCube) that are closely watched. In virtually every case of break-in, it is a WS that has been compromised.

ELLIS: It is not just the front-end WSs at the SC centers. There are a lot of users around the country with workstations in their homes or offices who have an account on an SC. It is the problem of break-ins on these WS where the intruder is able to get the password or login or whatever to get access to SC centers.

AIKEN: What is the extent of the responsibility for configuration of WSs for your principal investigators (PIs) at the different places?

NEWMAN: It varies from institution to institution. Some PIs manage their own WSs and won't let anyone else touch them. There are others who refuse to touch their WSs and rely on the campus or organization to take care of setting up and maintaining the configuration.

BRANSTAD: Can you differentiate the importance of three different types of security requirements that might be needed in electronic mail – authenticity of the sender, integrity of the message contents, and confidentiality of the message contents?

NEWMAN: We need all three. Authenticity is perhaps the most important for the center administration, but confidentiality is also important. Authenticity is also important for some of the researchers but some place a very high importance on confidentiality. For example, we have some pharmaceutical companies that are doing molecular modeling on the Cray and I am certain they would not want their data to fall into the hands of competitors. Users want assurance of security; they are less concerned about the mechanisms or algorithms that are used to get it done.

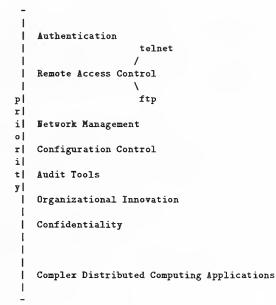
AIKEN: Do you filter at gateway levels through routers or do you set up filters on hosts by closing down certain ports?

NEWMAN: We filter source routing packets going out of the building because we have had some problems with student users, but none of the other hosts filter at the address or port level, and there is no filtering at the router level. We have talked about doing that.

(At this point, the workshop participants paused for reflection, with the purpose of determining how to budget the remaining time to insure that adequate attention would be given to the formulation of recommendations. The following summarizes the essence of this interlude.)

Aiken recalled that the major objectives were concerned with simple remote access to SC centers – telnet, ftp and possibly PEM as a third application, and also to make sure that passwords are not transmitted in clear text. That did not preclude recommendations for configuration control at sites and hosts, but the security of distributed computing as well the requirements of more complicated applications should perhaps be addressed in follow-on meetings. Cerf echoed these thoughts and noted the need to recommend technologies that can be deployed at SC centers without any development, those that will protect access to SC resources via NREN/Internet (modulo insider compromise), and whose deployment can begin in 1992. Crocker and Kent supported the idea of a coordinated effort among the SC centers to promote interoperability. Ellis said we should not ignore the unilateral actions that have already been taken.

Crocker felt that any plan should solve a prioritized set of requirements and to this end Cerf constructed a chart which through the course of discussion assumed the following form:



Cerf stated that authentication had to be high in priority because it is the one thing that provides a "leg up" on access control. Second in priority would be to apply some sort of access control to SC resources (telnet or ftp). Based on previous discussions in the workshop, he felt that confidentiality should be given a much lower priority. At the bottom are complex computing applications, the solutions for which are not of a short-term nature. Crocker felt that confidentiality should be significantly higher than complex distributed computing applications. In terms of impact on security, Ellis felt that configuration control should be given a high priority. He pointed out that the vendors have not picked up on it and sites do not have the time, expertise, or the money to configure systems for security. It is a matter of education, either at the user level or at the vendor level. Cerf recalled the major vulnerability of SC centers – individual users outside centers have WSs which, if not adequately configured and properly managed, provide a back door into the centers.

Ford said that, in centers with extensive networking, personnel cost is becoming the dominant factor. The cost of the technology has plummeted and the cost of the personnel has sky-rocketed and a radical switch has occurred at the SC centers. Yet centers are talking about increasing the number of individuals who are working on security issues. Mansur noted that there are tools like COPS to help with configuration management that might serve to ease the personnel requirements. Ford said the notion of registered security officers might be something that is needed. Crocker said that, despite the fear of getting into labor-intensive solutions, there should be some level of professional staffing focused explicitly and professionally on security at each of the sites. And there is a range of things that can be done from configuration of the networks to responding to incidents.

6.2 An Overview of Internet Privacy Enhanced Mail - Steve Kent, Bolt, Beranek, and Newman

(Supporting view graphs are presented in Appendix D.)

KENT: Privacy Enhanced Mail (PEM) is defined by a set of four documents which are presently available in Internet draft form. They are not yet available as Internet Request for Comments (RFCs). The latest issue to come up has been the compatibility of PEM with the Multipurpose Internet Mail Extensions (MIME). This has resulted in a major delay in proposing an Internet standard.

Two ways of implementing PEM will be illustrated here. They are considered equivalent from a standards perspective and they are interoperable. The differences should be invisible to the outside world. The first implementation is modular, but not especially "user friendly". One starts with an editor to produce a formatted message (exclusive of address information, etc.). This is passed through a PEM filter to produce another file. Then one goes to a regular mail user agent, addresses the message, inserts the PEM processed file into the message, and ships it off to the system. The second approach is the one that people are tending toward (because it is more reasonable to use) is to implement PEM as part of the user agent so that the user needs to specify the recipient address(es) only once. It is a more natural part of saying "send -s recipient" where the "s" flag designates secure. This approach requires taking the basic PEM capabilities and integrating them separately into each user agent. So there is a trade-off between ease of use and generality in terms of a particular set of software.

From the outside, these two implementations look equivalent; only the user sees the difference. The two approaches are interoperable in the sense that users can communicate with each other regardless of which implementation is used in a particular system. An important point is that PEM will transparently pass through intermediate mail relays (SMTP agents) or even through agents that go from SMTP to other mail systems. One can (at least in theory) take a PEM-processed message and insert it as a body part in X.400. To make use of that at the other end, there would have to something other than a standard piece of X.400 software, where (after X.400 message is received) the PEM body part would be put through a PEM filter and then processed. The off-the-shelf X.400 user agent does not support this. But it can be transported that way.

The security features are as follows. Data origin authentication is always provided for a high degree of confidence in the identity of the sender. Connection-less integrity (since messages are individual items rather than part of a connection) provides a high degree of confidence that the message has not been tampered with after it has been PEM-processed, or while it is in route, or while it is waiting in the recipient's mailbox. PEM provides a "basis" for nonrepudiation of the sender through the use of digital signature facilities. It is not correct to say that nonrepudiation is provided because there is a lot more to it (in an infrastructure sense) than that provided in PEM. But the tools are there so that one can set up a server on the network in conjunction with PEM to provide the required third-party time-stamping facilities. Finally, confidentiality is an optional facility in PEM. The originator of the message may optionally elect to use it. PEM versions can be produced without the confidentiality features.

Compatibility is a major concern in the design of PEM. It is compatible with SMTP mail relays and we believe it to be compatible with the vast majority of other mail relays, even the over-the-counter

systems of which we are aware. This is because a PEM message (when fully processed) uses a very restricted character set.

In the future, PEM will be extended to work with MIME. At this point, the plans are to install place holders to lay the groundwork for this future compatibility. To do a thorough job requires the nontrivial effort of making it possible for PEM to carry multimedia components that are processable by MIME and to make MIME processing work nicely with PEM.

In terms of producing an Internet standard, there is a lot of interest in PEM in the Internet Engineering Security Group (IESG) of the Internet Engineering Task Force (IETF). If there is agreement on our proposed way to deal with the MIME issue, it could be proposed to the Internet Activities Board (IAB) very soon. Since the PEM specifications were published three years ago, we gained a lot of experience with a number of independent implementations. Trusted Information Systems has been working on a reference implementation which they will release as soon as the remaining questions can be resolved. There are other implementations as well, e.g. MIT has one for MacIntoshes.

FORD: How do you see the future of PEM in the context of the massive explosion of mail facilities offered by the many vendors, e.g. Lotus and Microsoft and Apple? Are you getting positive feedback from them or are there things we should be looking out for in the future in order to achieve interoperability?

KENT: At the outset, we viewed PEM as an interim facility because we assumed that in the long run people would move to the X.400 standard that is widely touted in the vendor community. Although there are a relatively small numbers of native X.400 users today, it was assumed that they would convert their mail systems to X.400 in order to be able to send it to someone else's mail system. With the advent of MIME, which holds the promise of providing many of the facilities that X.400 does, the confidence in X.400 may be undermined.

CERF: The commercial electronic mail service as suppliers and the products suppliers are still very much oriented toward X.400 interface. All of the products I know about, especially LAN mail systems, use an X.400 interface as a way of getting out to external service providers. I would expect that, if some combination of MIME and PEM becomes the basic backbone for Internet, there will be a motivation to develop interfaces for them despite the conventional wisdom that X.400 is the primal standard interface.

KENT: As long as organizations have local proprietary mail systems that go through some interface to the rest of the world, a system like PEM is attractive because they can simply insert message inside of other things to get end-to-end mail security.

(continuing formal presentation) The processing steps in PEM are as follows. We start with a plain text message and perform an SMTP canonicalization on it at this point because it can not be done after the message has been protected. This is one of the steps that will probably be changed if we were to support MIME because it turns out to not be the right canonicalization in all cases. Next, the message integrity checks (MICs) are calculated for proper authentication, integrity, and nonrepudiation technologies. Optionally, the message is encrypted and the result is 6-bit encoded (except for one version of PEM messages) so we can ship it through anything and have it remain invariant. This last step would be something that would change in the MIME environment.

A graphical view of how it works is as follows. The originator of the message provides the addressing

and other information that appear in the RFC-822 header fields. Then the user provides the same information needed to perform encryption (if this option is exercised) and the encapsulated header is produced. Then the plain text from the user message is inserted into the PEM-protected portion of the message and it is optionally encrypted and placed as an encapsulated message. There is a significant amount of PEM processing that has to do with encapsulated header, the integrity check calculations, and the encryption operation that will be consistent with MIME, but some of the conventions or delimiting boundaries and some of the previously mentioned processing steps will have to be broadened for compatibility with MIME.

PEM is designed with the objective of making it suitable for a broad community and for standardization. Some of the simpler and less general mail systems do not perform the canonicalization and, as such, are designed for a community where everyone understands the same byte ordering, same character set, etc.

CERF: What about the broadcast case where someone is trying to provide confidentiality to a distribution list? Also, what if one includes the PEM-processed message in the middle of something?

KENT: PEM is designed to let you receive a signed and encrypted message. The encryption can be stripped away and the signature can be retained and put inside of another message that you in turn sign (and optionally encrypt) and send it to another person. The recipient of your message will be able to tell what you added to the message and signed, and what the originator had sent to you in the first place. That is why the bounds and encapsulation techniques are important. This also works correctly for non-repudiation.

(continuing formal presentation) PEM is parameterized in that it carries identifiers for the various algorithms that can be used for each of its processing stages. Because we use potentially six different algorithms, we are currently considering using suites (matched sets of algorithms) in order to reduce the number of combinatorial possibilities and insure interoperability. For encrypting the message, DES is the only choice at the moment. For message integrity, there are three choices. The DES mapping is included but it is not recommended if the message is sent to more than one person; also, it has the worst performance in terms of computation time. MD2 and MD5 are the approved options for message integrity. We started off with only MD2 but also included MD5 because its performance is much better (maybe by an order of magnitude). This can be significant in sending large files. MD2 was not eliminated because it is a more traditional type of hashing algorithm and is going to be used for certificate integrity anyway and, since certificates are small, performance is not an issue. For message signatures, the RSA algorithm is the only option. For key distribution, which is relevant only if you are talking about confidentiality, the RSA algorithm is the only option at the moment. Additional suites will be added to the list of algorithms that can be used with PEM and will interoperate. We hope to have a NIST suite that would include all the functionality that we need for the six kinds of algorithms we described. The Digital Signature Standard (DSS) and the secure hash standard together is a subset of what is needed to make up a complete suite.

CROCKER: In considering how communication can take place between government users who have the NIST suite and non-government users who have the RSA suite (or whatever), you have two systems that do not interoperate. What is going to happen is exactly the worst kind of solution – gateways that take RSA mail and re-sign it under a general purpose signature for the purposes of getting it validated at the other end.

FORD: Worse yet, you can not find a significant mailing list in the scientific community that does not have international components. And there is no international thrust to accept NIST algorithms. Software vendors that build systems, for example, are going to have build systems that have dual security stacks.

CROCKER: To solve this problem is beyond the scope of this workshop!

KENT: (continuing formal presentation) Looking at a certificate (used with PEM) in decoded form, one can see the various fields. Every certificate contains a serial number that is unique relative to whoever issued that certificate. This is the way to manage the distributed issuance of certificates. The name of the issuer is included in the certificate – a common form would be "country, organization name, organization unit name". Additional allowable attributes are "state", "province", or "locality". A validity interval provides the time frame during which the message is applicable. The subject field identifies who the certificate belongs to and has the same kinds of available attributes that are found in the issuer field. In particular, one would typically find the common name in here to identify a person (or a mail list agent or mail list server or whatever). The subject key information specifies the algorithm itself and any parameters associated with it along with the public key of the subject who is identified in the previous field. Finally this whole collection of information contains a digital signature and a specification of the algorithm used for the signature – a combination of both a hash algorithm and a signature algorithm.

PEM certification is designed as a hierarchy. The name of the top node is now called Internet Policy Registration Authority (IPRA). The Internet Society as a nonprofit international organization will operate a facility which registers entities at the next level down - high assurance, residential, midlevel assurance, persona. This next level of certification authority represents different policies under which organizations and/or users are certified. The idea is to accommodate a variety of policies. Note that an organization can be registered under multiple policies, if it has different classes of users with different requirements.

The high assurance authority is presumed to require some paperwork – perhaps signed legal agreements binding the organizations that sign up underneath them to pledge to provide high quality in authenticating the users whose certificates they sign and in managing their private signature key for signing certificates.

Residential refers to individuals who are not claiming an affiliation with any organization, employer, professional society, etc. It is possible to have different kinds of policies. Basically, the difference is in the form of the names for the registered individuals. Signing up a residential user may require the user to go to a notary public for verification of identity. Since this is the high end, it costs you more; it involves more overhead, paperwork, etc.

The mid-level assurance policy certification authority was motivated by the need to deal with an appropriate level of assurance for students at MIT (Kerberos). In the registration process, one can set up accounts because the names of the students are known and they come in and take one – one presumes that individuals will take the right accounts. If there is a conflict where a student cannot login because the account has been taken, then there is a natural alert. So mid-level assurance tries to do a good job, but there are limits on just how good it can be. And that would be part of the published policy.

Finally, a persona policy certification authority is a distinguished type of authority to serve users

who want to enjoy the benefit of PEM, but do not wish for their true identities to be represented in the certificate. They may have to communicate by other means to know true identities. All the persona names are viewed globally, so if one is chosen, it is like a CB handle. The same one cannot be chosen again in order to avoid confusion. But all of the rest of the facilities of PEM are there to make use of as one sees fit.

AIKEN: Since the government is not the Internet, there may be some components that feel they should not be under the IPRA. If we use different hierarchies, how are they going to interact?

BRANSTAD: That is a question we are trying to address in a study that NIST is launching. We are not going to look only at the Federal government part. We will also be looking at the commercial part and at international cooperation.

CERF: At the moment, there is a proposal that the Internet society act as the registration point for organizations that will authorize certificate issuers as the level below the IPRA – classifications for different registration policies and procedures. None of these issuers generate public keys; that is done independently. They only deal with the certification of someone's distinguished name with a public key. There could be a number of organizations that operate under these different kinds of levels of assurance (not just four).

AIKEN: There will be agencies which will not be part of that!

CERF: This is very unfortunate and a technically incorrect conclusion to reach. All that is going on in here is registration for the purpose of allowing interoperability. There are no restrictions or limitations or rules of behavior. It is critical to have this registration capability, along with some databases that the system can maintain, to help these authorities (and their subsidiary authorities) avoid the accidental generation of duplicates. What I would like to debate is whether it is impossible (for whatever reasons) for any authority, whether it is NIST or some other part of the U.S. government (e.g. the Defense Department), to agree to have themselves registered in common so as to make the system work.

KENT: From a practical standpoint, if government agencies register under this model, there may be a question of sovereignty. However, there is another way of viewing this. If a government agency registered under it, issued certificates under it, and wanted to impose an identity-based access control policy or authentication policy that said any certificates not issued by those in the U.S. government will be ignored, that would be easy to implement. Anybody underneath this could arbitrarily and unilaterally decide to refuse to recognize certification anywhere else in the rest of the tree. They will only hurt themselves.

AIKEN: Given the current status of the IPRA, there are going to be major objections unless one can show that it has been designated as the true international standards body recognized by the U.S. State Department. Otherwise, you are not going to get Federal agencies to buy into it.

CROCKER: Consider this scenario. We are in the business of creating credentials/certificates, but we are not yet at the point where the government is issuing them for the populace. Private organizations will fill the breach until there is enough infrastructure and the government recognizes it as an appropriate function for it to be doing. This is analogous to where we were 100 years ago when one carried letters of credit from a bank. Drivers licenses, passports, and social security cards are issued by government agencies.

6.3 Security in Open Systems Technology Demonstrator Programme -John Laws, Defence Research Agency, UK

(Supporting view graphs are presented in Appendix E.)

LAWS: Prior to reorganization, the Ministry of Defense (MoD) contained a number of units. Under the Centre, the Command Information System (CIS) is responsible for policy and operational requirements (POL & OR). The Procurement Executive (PE) is the purchasing arm of the MoD. The units on the left side of the chart essentially define the operational requirements and obtain the necessary endorsement and funding, and having done that, the PE is authorized to make purchases (e.g. battleship or tank).

The PE consists of four subunits, the Controller Air (CA), Controller of the Navy (CN), Master General of the Ordinance (MGO), and Controller Establishments Research and Nuclear (CERN). The Research Establishments, aligned under CERN, are each loosely associated with a particular arm of the PE: Royal Aerospace Establishment (RAE) was associated with air systems, Admiralty Research Establishment (ARE) with the sea, and the Royal Armaments Research & Development (RARDE) with land systems. Royal Signals & Radar Establishment (RSRE) was a little different; its long history in radar and electronics has led it to be seen as a technology-based establishment as opposed to platform-based, e.g. land systems. There are other establishments of a smaller nature like the Chemical Defence Establishment. Underneath the RSRE is the Information Systems (IS) Department, Communications & Computing (CC) Group, and the Distributed & Secure Information Systems (D&S IS) Division.

There has been a significant reorganization in structure over the last two years, with ARE, RAE, RARDE and RSRE merging into a new Executive Agency of the MoD, the Defence Research Agency (DRA). The intention is that there should be a contractor/customer relationship between the two (effectively uncoupled) organizations. The DRA has two managing directors (MDs); the one of interest here has five business units (BUs). The CIS BU is one of the largest BUs in the organization and within that I am in the Architecture Division of the Systems Engineering and CIS Technology Department.

The security problem I am addressing here may not be precisely applicable in this particular workshop in that you are looking for immediate deployment. But I hope it is appropriate for the longerterm ambitions of the NREN and Internet because I see some commonality of issues. In the security problem, there are many diverse CIS assets and many diverse security products and systems, not necessarily in the commercial market place. For example, we have an X.25 end-to-end packet encryption device rated at UK Level 5 which I believe corresponds approximately to TCSEC level B3 (good enough to carry secret information over a public network). The significant thing is that these products have been developed in the absence of a security architecture. (One of the more significant architectures that has been developed and is now undergoing evaluation is the Distributed Secure System (DSS) which is a complete solution for security in a local area network environment). One consequence of this is that you have uncertainty and risk when procuring a product requiring some element of security; there is uncertainty about what is available in the marketplace. It is generally not the intention of a project to buy a command and control system and then create the security. The more desirable situation would be that the products exist in the marketplace. The resources should exist to put together any size system that one might want although it may be necessary to commission the creation of extra software. Another consequence of the absence of a security architecture is a lack of interoperability, resulting in high cost or reduced functionality. The cost of a specially built solution is not amortized over a large commercial base or one has to compromise at a reduced capability.

The Technology Demonstrator Programme (TDP) is a mechanism for overcoming those deficiencies. It is not an R&D programme. The intent is to set a goal which is believed to address an area, particularly one in which there is risk and uncertainty, and in the space of a few years by working with the industry, demonstrate that there do exist prototypes and potential products. These have to be understood in the sense of how do they fit together to make something actually work.

The purpose of the TDP is develop a MoD-wide secure distributed information systems architecture. Because of time and funding limitations, there is a thrust to retain a sharp focus on the issues. Some important aspects include compliance with international standards, vendor independence, use of open systems products and particular emphasis on low and medium levels of assurance. An agency, equivalent to the U.S. National Security Agency, currently exists that focuses on high-assurance security levels. What is missing in the market place and consequently not available to MoD are the low and medium assurance products.

Having created a security architecture, the next step is one of demonstration. This will require the selection of some representative application and layer services. Attention will be given to evaluation, which has to be there eventually for medium and high levels of assurance (current thoughts are to focus on the ITSEC criteria, since it is representative of a significant community of the UK). One important aspect will be the use of prototype products available in the civil marketplace. The emphasis is on prototype since it is not intended that this programme deliver a product – although industry is free to develop products. Supporting services such as authentication and access control will be included in the demonstration. It will be more than the mechanisms inside a narrow application – it will be the entire required infrastructure. Another important aspect is the reuse of security elements in various applications, e.g. key distribution should not be application-specific.

There are some technical areas that have been selected for review as usable components in implementing the TDP - X.400, X.500, ISO NLSP, security services such as certificate authorities, and trusted functionality in database management systems (DBMSs) and operating systems (OSs). At this point, the future directions of MoD are not clear with respect to DBMSs and OSs since the state of development in these areas has yet to be assessed.

CERF: I have a question about the network layers and security protocols. Do you have any interest in the transport layer security protocols?

LAWS: I believe Transport Security Layer Security Protocol (TLSP) has a number of problems. The people involved in its development should be challenged very carefully about the technical quality of what has been proposed so far. In addition, the security evaluation of this protocol will be a technical challenge for some years.

(continuing formal presentation) Regarding the purposes of TDP, there is no intent to produce an evaluated prototype product and there is no intent to produce a military message handling system.

The programme has a number of phases. Phase 0 will initiate the programme. Phase 1 will consist of a programme definition study to provide greater technical detail such as who will be involved, what it is actually going to do, what will the costs be, and what size of team has to be employed. This phase will also select and implement an application in order to gain an early appreciation for some of the issues. Phase 2 of the four year programme should deliver a generic security architecture for use in distributed CIS systems, a security policy model, a specific architecture for this TDP, and provide contributions to the areas of standards and product review and selection.

Participation by industry will be essential. The intent is to form a partnership with industry contributing the funds for equipment. Such contributions will be a indication of its interest in the programme – without it, there will not be a programme. There is also a desire to have international participation from such organizations as NATO, Task Force C3 of the Inter-European Programme Group (IEPG, the European members of NATO), and the ICB.

CERF: Actually you won't get that level of participation from industry unless they are expending product development dollars. It is not a market environment. So your sale pitch has to be carefully constructed to show that the corporate contribution is in product development investment.

LAWS: Agreed, and in addition I wish to have international participation; and will work to achieve this through my membership on a number of international committees, such as NATO TSGCE Ad Hoc Working Group on Security and IEPG TFC3.

Phase I requires an appreciation of the things that are taking place in the R&D and commercial worlds. In particular, I am bringing to their attention work being done in the European community, something called COSINE PARADISE in the area of networking and directory services (X.500), ESPRIT THORN in directory services, U.S. efforts in the NREN and the Internet, and work by the Canadian Department of National Defense and the U.S. Department of Defense in implementing military message handling systems.

The first-cut security architecture would potentially include such things as PEM (U.S.), VALUE PASSWORD (EC), Military Message Handling System (NATO), and things that are happening in EWOS and other regional bodies that are specifying protocols. For the technology to implement this, the concentration is on X.400, X.500, and NLSP. Other components for consideration include current developments in secure multi-level DBMS and trusted OS.



7 Session on Security Management in Distributed Networks

Session Leader: Vint Cerf, Corporation for National Research Initiatives

7.1 SDNS Security Management - Wayne Jansen, National Institute of Standards and Technology

(Supporting view graphs are presented in Appendix F.)

JANSEN: The Secure Data Network System (SDNS) security management project was an NSAsponsored effort conducted during the 1991 fiscal year as part of the SDNS upgrade program. It involved a number of organizations that had been previously involved with this program (DEC, Hughes, IBM, Motorola).

The investigation was limited to lower layer security protocols SP3 and SP4 and also the key management protocols. However, the SDNS protocols for secure messaging were not part of this study. The effort included identifying the elements of information that needed to be managed and anticipating the management operations that needed to be performed. We tried to accommodate the requirements for policy independence, since one of the goals of the SDNS program was to address both the commercial and military sectors. The one element of security management that was already in place was the key management protocol and that had to be incorporated into the results. We had to take a look at how the security of the management operations was envisioned by the original SDNS program to be sure that it fit in with our overall plans.

As input to the investigation, we had an SDNS architecture and we also had the SDNS protocols – the lower layers and key management. SDNS is based on the OSI Reference Model. So, if a choice had to be made about which management protocol to use, it was clear that we should make use of the OSI protocols. We decided to look at the common management information services and protocols available in OSI. Not only do the set of OSI standards for management look at the protocols and services, they also have a variety of built-in system management functions and we realized that some of these things would be quite applicable to what we were trying to do. So we also tried to make use of the functions and their associated information definitions for object classes, attributes of object classes, and specific types of operations. The standards are all based on an object-oriented philosophy, so our job became one of defining objects for security management, determining what their attributes are, looking at the types of notification and events reports that would be emitted by these objects, and looking at the operations that we would want to be able to perform on the objects.

A dual-stack model captures some of those ideas regarding the plans to protect the management operations within the SDNS architecture. The SDNS security protocols reside at either OSI layers 3 or 4, and are used by normal user applications at layers 4-7. There is second stack that is used for management purposes. In particular, the functions that already existed in SDNS had to deal with key management. Whenever a user application would require key material, the security protocol would somehow, either before hand or on demand, obtain the keys through this channel.

CERF: What are the reasons for having the security protocols in each of the stacks appear as separate security functions?

JANSEN: If you take the case of security protocol in layer 4, the management stack is a separate addressable entity that is different than the user stack. The two "instances" of the security protocol might be different.

KENT: Before installing the Common Management Information Protocol (CMIP) for management of the security protocols in layers 3 and 4, there was already the need for a dual-stack model because the Key Management Protocol (KMP), from a security standpoint, cannot depend on anything beneath it. It performs its key exchange directly with the other entity. In order to use SP3 or SP4, one has to have a key in place and if one is contacting the other end to get a key in place, one cannot already be using SP3 or SP4. So it was already a separate stack which did not have a separate security protocol. It seems a little inconsistent to have both KMP and CMIP in there because KMP does not use the underlying security protocol but CMIP does use it.

JANSEN: (continuing formal presentation) The KMP is a two-phase protocol. The first phase consists of an exchange of credentials and a key is established. In the second phase, there is a negotiation of security attributes that are associated with that key. The KMP is self-protecting and therefore, in this dual-stack model, it must be capable of bypassing the security protocol at the lower layers. However, key management is not the only security functionality that you need and so the CMIP is also part of the security management stack and it has to protect its exchanges. Therefore, there are two choices on how to do that. One way is to choose to have CMIP do some self-protection, providing its own security services much like the KMP does. Even though it not the way SDNS is planning to do it, it is under consideration by some OSI implementors workshops. I believe the Simple Secure Network Management Protocol (the TCP/IP version that has security in it) is also planning on taking charge of security at the application layer as opposed to one of the lower layer security protocols. The SDNS approach was to protect the management protocol in layer 3 or 4.

I want to describe how we defined the management information bases (MIBS). In looking at object classes for the SDNS protocols, we realized that there were other MIBS being developed for the network and transport subsystems and a choice had to be made as to whether we would define objects that would fit into these particular subsystems or if we were going to have a disjoint set of objects. The choice we made was to have the SDNS object class definition separate (so it has its own MIB), independent of the network and transport MIBS that come out in OSI. That turned out to be an advantage because, at the time, this latter work was evolving in parallel and it would have been very difficult to try and coordinate the two efforts. It also keeps security as an optional add-on and that seemed to have some benefits as well.

CERF: There could be an advantage in trying to keep the network security MIB distinct because it might allow you to use the same network security layer in more than one underlying protocol.

JANSEN: There was actually a provision in the latest network subsystem that would allow linking it in separately and distinctly within a framework that has already been established for the subsystems.

(continuing formal presentation) The intent was to try to identify major object classes for the SDNS MIB using familiar terms and to describe the relationships between these classes. The relationships are depicted by arcs: one-to-one by an arc with no heads, one-to-many by a single-headed arc, and many-to-many by a double-headed arc. At the SDNS subsystem level, one can have an associated revocation list which in turn must be associated with credentials; it also must have an associated cryptographic device. The credentials are used to form cryptographic associations. Another class

associated with SDNS subsystem is the security protocol (SP) entity, which is a generic object class defined for the purpose of refinement into the specific SP3 or SP4 object classes. This is an object-oriented feature where one can take an object class like an SP entity and through inheritance mechanisms specialize it into SP3 or SP4 objects. One can do that for everything that is labeled SP, i.e. an SP entity, an initial value (IV) SP association, and an SP association. So the idea is that an SP entity represents a generic protocol entity at the lower layer. The IV SP association consists of initial values and defaults for that particular SP entity and an instance of this is a list of security attributes that are negotiated during key management. There is a similar process for the key management (KM) protocol.

Through the SP associations with an SP entity, one can specialize specifically for the SP3 and SP4 protocols by adding specialized attributes. The SP3 protocol required the definition of some additional object classes. In particular, the version of SP3 that we modeled was a revision of the original SP3; it was closer to NLSP and it covered both connection-oriented and connection-less types of networks. And in addition, SP3 has some capabilities to protect end systems that do not have a security protocol located within them.

The larger picture can be depicted by a containment tree that shows how all the instances of objects could be accessed within a particular end system. For example, starting from the SDNS subsystem, one can access the credentials, revocation list, SP4 entity, SP3 entity, crypto device, and KM entity. So, from the SDNS subsystem, all of the remaining objects can be accessed and manipulated through security management. This provides some interesting capabilities when one considers that some of these things can be created and destroyed.

There were several areas of built-in system management functions in the common management information services and protocol area that we could borrow and use for SDNS. One area was event recording where objects can emit notifications that can be filtered and if necessary an event report can be issued. This could be used in conjunction with the area of log control to serve as a basis of distributed auditing. Another area of built-in system management functions was access control. We felt this particular feature was quite flexible, allowing for all sorts of access control information and the specification of an arbitrary policy.

Although the project was primarily a paper exercise with no implementations, we did identify a good set of object definitions that we can bring into the OSI work and also into other work. The project allowed us to gain an appreciation for management and realize that it may be a mistake not to address it at the outset rather than waiting to deal with it at the end. In particular, some of the auditing requirements that we felt strongly about were not captured by the number of prototype implementations of SDNS. In the objects we defined for SDNS, there was some feeling that having these management services in place would be a very helpful counterpart to key management. The final point is that configuration management can be combined with network management practices. By using normal network management procedures the area of configuration management may be improved quite a bit, and the security of a system vastly improved.

The area is still developing, standards are continuing, and prototyping experience is badly needed (at least in the SDNS world) to understand how well we did in our definitions.

7.2 FIRST System - Dennis Steinauer, National Institute of Standards and Technology

(Supporting view graphs are presented in Appendix G.)

STEINAUER: The Forum of Incident Response Security Teams (FIRST) is a cooperative effort involving a number of organizations. The concept of FIRST arose as the result of the "Internet Worm" incident, along with a number of things that were going on in that period of time. It was clear that the basic nature of security threats had substantially changed and there were new requirements.

Awareness and self-protection are significant parts of the solution to the problems of computer and information security. This is more true now than before because so many people now are directly involved as part of the problem and must also be involved as part of the solution. One can no longer claim that a small number of computer systems operators can take care of the problem. It is increasingly more difficult to distinguish between the ordinary user and a system manager when one unboxes a full-fledged UNIX system.

Closer cooperation and coordination of activities are necessary in dealing with the increasing number of cases that affect more than one organization. Even in those cases where only one organization is affected, others might be able to help in solving the problem. At the same time, it is recognized that not every organization can have all the necessary skills, levels of expertise, and resources to cope with the problems. It is necessary to nurture and rely on centers of expertise in certain areas and to find ways for these centers to cooperate.

Even before the Internet Worm incident, DARPA saw some need for a security response capability and, as a result, funded the Computer Emergency Response Team (CERT) at the Software Engineering Institute (SEI) at Carnegie-Mellon University or, more specifically the CERT Coordinating Center (CERT/CC). In approximately the same time period, the DOE also established a response team which was called the Computer Incident Advisory Capability (CIAC). As a result of technical differences (in the systems being used) or organizational differences, there are different ways of identifying the constituency that might be affected by security incidents. It was clearly not possible to assemble all of the necessary expertise in one place. This led to the idea of individual constituencies and constituency response teams where each team would have one or more constituencies. From the outset, the defined constituency for CERT was the Internet and (as a corollary) UNIX users since so many of the hosts were UNIX-based. CIAC has an official constituency of DOE laboratories, but it also serves the rest of the DOE.

There is some overlap among the constituencies which tends to strengthen the whole process. But there was still a need for central coordination and interchange of information. The idea of FIRST was to extend the concept of individual constituency-based response teams to a network of response teams that would interact and support each other. When a problem occurs in one constituency, the corresponding response team would send out an alert to other response teams, either as a draft on a potential problem or as a live alarm if an incident was in progress. They would in turn determine whether or not the problem was relevant to their constituency, respond in kind within their constituency, or in many cases offer some assistance. An electronic mail network or mailing list has been established for the purpose of communicating among these various teams.

CERF: I assume you do not limit yourself to electronic mail in the event that you are dealing with

an incident in progress.

STEINAUER: That is correct, but that also identifies one of our problems, namely, we do not yet have (except through rather torturous methods) secure electronic mail communications. One of the complicating factors is the fact that this activity is not limited to the U.S. government or for that matter to the U.S. One of the first participants in FIRST was a European Standards Group and we now have a number of non-U.S. participants.

CERF: Apart from the fact that FIRST needs quality secure communication among its various components, this is a prime example of why one does not want too much diversity in order to facilitate the deployment of PEM.

STEINAUER: (continuing formal presentation) There are two levels of participation in FIRST, active response teams which are called members, and liaisons which are legitimately involved but do not have responsibilities (e.g. the FBI and one university group). FIRST does not have an official charter, but there is an operational framework in place in order to address some of the problems that will have to faced sooner or later. The activity of FIRST is strictly cooperative at this point.

In addition to the almost twenty member organizations, including the liaisons, there is a steering committee made up of ten individuals elected by the membership. NIST is currently acting as the secretariat, handling administrative matters. The steering committee consists primarily of people who are also from the member community. But we can get key people on the steering committee who have something to contribute and are not associated with a FIRST organization.

AIKEN: Is there anybody on the steering committee who is a coordinator of an end site, i.e. is there any deliberate attempt to include such persons?

STEINAUER: At this point, no one comes to mind but it is certainly possible. The charter does not state that the membership has to be selected from certain types of constituencies.

Most of those involved in FIRST are closely tied to the Internet. The reason is that they are the ones who probably more clearly see the threats and the need for the cooperation among emergency response teams, and they also have the tools at hand to actively participate. For membership, a response team must satisfy certain minimal communication requirements which include Internet connectivity. We are attempting to get other parts of the Federal government involved.

AIKEN: In looking at the cooperative efforts of the response teams and liaisons, there have been occasions when an incident or suspected incident has occurred, but someone placed a restricted label on the information and other entities were not informed to warn them of possible problems. Has this difficulty been addressed?

STEINAUER: The primary purpose of FIRST is to combat that particular problem. There is nothing that FIRST does in its main line activity that is classified. Several of the participants have to deal with classified situations and they do it within their own structure. We have discussed this and will continue the operation subrosa when that is necessary. But we also recognize that very little information is going to willingly come out of the classified community.

The Computer Security Vulnerability Working Group of the National Security Telecommunications Insistence Security Subcommittee is involved in an activity to set up a reporting mechanism and possibly a response activity. They are dealing more directly with problems in the classified community. NIST is an observer on that group. One of the points we have tried to make with that group is that if the classified community wants to keep secrets to itself and not actively participate with the non-classified community, then they are only hurting themselves because the vast majority of the problems are known and dealt with in an unclassified manner.

7.3 CERT Activities - Dain Gary, Software Engineering Institute, Carnegie-Mellon University

GARY: The incident response capability works well, but it serves a constituent community and that is why there are several of these capabilities. When constituent communities begin to cooperate, they begin to get policy overlap and policy contention. The attempt of course is to deal with things from a technical point of view rather than a single system that has been compromised.

The question was raised earlier about the scalability of CERT to the magnitude of the Internet. In concept, we can no longer afford the staff nor the telephones required to answer the reports on incidents. One of the things we have done at Pittsburgh is to move away from a response orientation (e.g. reactive) to becoming more pro-active. We have a team of three people, under the heading of research, looking at what tools and technologies can be developed to help installations understand where they have configuration problems and where the vulnerabilities exist, and at how to make the tool set available. The other aspect of pro-active orientation is an educational awareness and training program for the user community through tutorials, handbooks, how-to-do-it books, guidelines, etc. This information for our constituent community may be somewhat different than perhaps CIAC and the Air Force would tailor for their communities.

CERF: What kind of response are you getting from the software and hardware vendors when you identify a problem?

GARY: At Pittsburgh, we have a primarily UNIX orientation and we are working with approximately 30 UNIX vendors. Three years ago, we were not getting good responses, but things have changed. Over the last three years, we have been able to demonstrate to the vendors that responding to a known security problem will serve to improve their product line by making it stronger and less expensive to maintain. So the participation of CERT with the vendors is now a very positive thing from our perspective. At times we come between the vendors and their clients. Some vendors are very nervous about this; others appreciate the help. We have a workshop scheduled where the vendors can come in and present their perspectives.

MANSUR: There is a spectrum of vendors, some of whom are very cooperative - you call them up and they respond very rapidly; others do very little. Some have different policies - they may keep things "closed" until they do a full-blown release or, there may not be a lot of time between the dates when security improvements are made available to the community.

ELLIS: It is still the case that vendors are reacting to the pressures of consumers. The improved relationship does not necessarily translate into the result that they are doing well and progressing in the field. They still address the bottom line "What will contribute to sales and how much am I willing to spend?"

STEINAUER: Its a double-edged sword for them. We can work with them and make a big deal about how they participated and cooperated to solve this one problem. And they say they are not sure they want to make a big deal out of the fact that a problem existed.

STEINAUER: Another result of this may be a process of confidence building that has be done among the vendors. Initially, all they could see was another threat to their relationships with the users and they imagined that all they would do is distribute patches similar to what happened in response to the Internet Worm. The vendors have to recognize that this process is a fairly deliberate one and can be helpful.

AIKEN: What about the Internet? Is there something you would recommend that is pro-active – like, saying this is how one checks systems when you get them and before you connect them to the network – something that is easy to follow?

GARY: At the risk of sounding a bit authoritarian, there is no one in charge of the Internet. That is the problem; there is no enforcement mechanism. We were excited about participating in this series of discussions to see how we can begin to move in that direction. One could specify a policy, for example, which would be a good start as to how to do a number of things – like specifying for the vendor community what acquisition requirements are going to look like, what security attributes are associated with equipment, etc. It would go a long way, but how would it operate? As we go into the NREN program, we have learned a lot from the Internet that we could try to incorporate.

CERF; It is true that no one runs the Internet and it may be the case that no one will run the NREN. But it occurs to me that remote attacks are something everyone is concerned about. Where there is a common threat or vulnerability and a significantly large probability that some among a group will be attacked, this satisfies the classical reason for purchasing an insurance policy. Has anyone consulted with the insurance industry to see if they would like to write policies which would fund FIRST kinds of activities?

STEINAUER: There have been computer security loss policies, not necessarily network-oriented, for probably two decades.

CERF: Those have to do with loss of functionality or business due to natural disasters, fires and things like that.

STEINAUER: Insurance companies would probably adjust the premium based on the security measures that are put into place. It might very well be that the enducement to participate through a reduction in premium, combined with a self-funding mechanism for FIRST, would in effect do the same thing.

AIKEN: What is the major fear of an organization? Is it the loss of data? Is it the loss of operation of an important machine for some period of time? Or, is it the fear really that the organization's name is going to appear on the front page of Time magazine?

GARY: This model holds up well for the financial services industry, but may not be a good Internet model. We are working on what is called a "security maturity" model which allows one to look at what security features are in place at a particular installation and then say something to that organization about where they fit on the scale and what needs to be done in order to improve their posture.

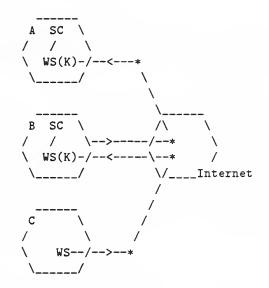
AIKEN: I am concerned about the communication between response centers, the infrastructure, and the need for a pro-active approach to security. There needs to be a recommendation from this workshop that supports these efforts.

8 Session on Workshop Recommendations

Session Chair: Vint Cerf, Corporation for National Research Initiatives

8.1 General Discussion

For the purpose of establishing a context for discussion, Vint Cerf began the session by drawing a reference diagram of a hypothetical network.



Network Reference Diagram

A, B, and C represent security boundaries of networks in the above diagram. These are connected by routers/gateways to the Internet. Each network has components like workstations (WSs) and supercomputers (SCs). The WSs with Kerberos implementations are denoted by WS(K). An intruder at WS in C is assumed to want to gain access to WS(K) in B with the ultimate goal of gaining access to SC in A. Cerf posed the following preliminary questions:

- 1. Which security boundaries can we actually do anything about?
- 2. If we configure the networks to use Kerberos within their environments, what can be said about remote access?
- 3. If authentication and access control are implemented by means of Kerberos, how will it look if there are several of these networks involved?
- 4. What happens if not all components of these networks are able to implement good configuration management? For example, suppose that some networks are Kerberos capable, but one of them is not and also it does not have anything else like useful access control or configuration management.

- 5. Given the presence of a random would-be-intruder, are we willing to accept the possibility of a weak spot in the network? (For example, suppose a WS in network B has weak defenses even though it is Kerberos capable in terms of access to SC resources elsewhere.)
- 6. What kind of security features can we assert in the policy we might propose or the recommendations we might make?

Cerf then presented a list of items for consideration in working toward the formulation of a set of recommendations.

- limited scope solution
- authentication
- access control
 - remote access (telnet, ftp)
- network management security
- configuration management
- audit tools
- organizational innovation
- confidentiality
- complex distributed computing applications
- scalability

Preliminary List of Constraints

The intent was for the participants to debate the list of items, suggesting additions and deletions and imposing a priority ordering. Cerf regarded the list as a set of constraints on the solutions that would be recommended. Through the course of discussion, several other participants felt that the list was actually a mix of objectives and constraints.

Extensive discussion ensued on the meaning of "limited scope solution". It was recognized that not all end systems at a site will be able to implement Kerberos. Furthermore, some Kerberos-capable end systems will be poorly managed. Therefore, it is not possible to draw hard, secure boundaries based merely on site adoption of Kerberos. It was considered important to do a good job within an SC itself in order to carefully control access. But even if SC sites use Kerberos, some client WSs may be compromised and the SCs accessed via this path. It will be hard (if not impossible) to put strong boundaries around the SCs as sites themselves because users will want increasing access to things like NFS and will also want direct access to the mass storage (not only through SCs). A reasonable, limited but effective, solution would be to first secure telnet and ftp access; perhaps provide for secure electronic mail; and implement authentication techniques that avoid resuable passwords in plain text.

The participants expressed the need for the workshop record to be up-front in noting that only initial steps were being recommended, addressing parts of the problem. Subsequent work will attempt to address other problems. And, in order to properly motivate those who might be reluctant to bother with an incomplete solution, the recommendations should also point out that it is essential for SC sites to take this initial step now in order to get anything done. The need for good configuration management was considered to be of high priority since, without it, it will be impossible to accomplish many of the other things. In the face of recommending only a partial solution, the use of audits was considered to be an important post hoc mechanism.

Echoing an objective previously articulated by NSF at the outset of the workshop, an important constraint was the ability to "immediately" begin to implement the final recommendations.

It was agreed that the recommendations should admit to either hardware or software implementations in an architecturally consistent and interoperable fashion.

The discussion on hardware vs. software raised the general issue of costs. For example, the recommendations should not require everyone to buy smart cards, but rather it should recognize that what is presently accomplished in software with passwords that provide user keys can also be accomplished in hardware with the aid of smart cards. The transition can take place on a piecemeal basis.

The discussion of cost raised the issue of scalability with concerns that recommendations for nonscalable technologies might be hard to reverse. Recalling NSF objectives, scalability was considered important, but of lower priority in the interest of effecting immediate enhancement of security.

Other issues raised were "ease of use" and "acceptance by users" noting that, if these constraints were not satisfied, the rest may be irrelevant. Ease of use (for end users) and ease of management (for sites) are especially critical since SC centers have paying customers other than NSF-sponsored researchers. Without reaching a resolution, there was some discussion on the nature of the SC center clientele, their insatiable appetite for processing power (with or without security measures), and the threshold of user discomfort deemed acceptable. It was noted that these issues were not independent of cost.

At this point in time, Cerf's preliminary list had been altered to appear as follows:

- 1. limited scope solution is a must
 - remote access to NSF supercomputers
 - telnet and ftp
 - no static plain text passwords
 - PEM if possible
- 2. must be able to start in 1992
- 3. solutions must allow hardware or software to be used (and interoperate)
- 4. ease of use
- $5. \cos t$

Modified List of Constraints

Steve Crocker then presented a list of recommendations which, through the course of discussion summarized below, evolved into the following form:

- 1. Use C/R User Authentication
 - public algorithm
 - exportable
 - S/W, token, PC implementation, pregeneration
 - multiple site
 - public key based
- 2. Distributed Authentication
 - Kerberos (version 4)
 - telnet, rlogin, ftp control connection
 - Kerberos version 5, DASS, etc.
- 3. Privacy Enchanced Mail (PEM)
 - for users, administrative, security
 - rationalize hash, signature, confidentiality algorithms (i.e. it has to work everywhere)
 - registration
- 4. Security Perimeter (for access control)
 - filters, gateways
- 5. Configuration Control/Management
 - at supercomputer centers
 - distribute/support COPS for remote users
- 6. Active Audit/Review
 - tools (e.g. COPS)
- 7. Security Officers/Team
 - implement policies and procedures
 - in charge of site education
 - serves as FIRST point of contact
- 8. Oversight
 - implementation and oversight group
- 9. Follow-on Developing Technology

List of Recommendations

8.2 Elaboration of Recommendations

Recommendations on Authentication

Crocker's list contained two methods of authentication (items 1 and 2). The stated philosophy was that, even if Kerberos is available, there is still the problem of initial authentication so that the two methods are complimentary rather than competitive.

After extensive discussion of these two methods, the consensus of the workshop participants was that:

- Kerberos (version 4) should be adopted now if possible since it provides a sound basis for securing a wide range of "applications" (e.g. telnet, ftp control connections, remote commands) and admits to a variety of assurance levels (e.g., it is possible to start with passwords as keys and move up to random keys, smart cards, etc. in an architecturally consistent and interoperable fashion). It can probably be used to support roving users across international networks, if the embedded DES is not used for confidentiality.
- The challenge-response authentication scheme should be used in those circumstances where Kerberos is not available (e.g. dialup terminal use, visiting a site and not wishing to expose a key). It is usable in software (and perhaps in hardware tokens) and is amenable to pre-computation for printed C/R lists.
 - As a candidate technology for C/R, a DES-based system has the advantage of being available and not requiring any development. The DES technology is exportable if it used only for authentication purposes and not for general data encryption; on the other hand, keys have to be stored in the clear at the end hosts.
 - Public key encryption for C/R avoids the problem of storage of clear text keys at hosts. It was also felt that public key systems will scale better and over time may be preferred in many (but not all) contexts.

Steve Kent recommended that, as part of the follow-on activities, we plan to track the work by the Internet Engineering Task Force as it provides an ability to interoperate between symmetric and public key systems.

Recommendations on Privacy-Enhanced Mail

The PEM technology will be available soon. The use of PEM is recommended, initially with its current suite of algorithms. Later, the use of the proposed NIST suite of algorithms should be explored, including a rationalization of the problems (hash, signature, confidentiality algorithms) of interoperability and management of multiple algorithm suites (i.e. it has to work everywhere).

PEM should be immediately employed by FIRST individuals and CERT for internal communication and for distribution of such things as security alerts and fixes.

Recommendations on Security Perimeters

The concept of a security perimeter incorporates the ideas of router and gateway filtering – access control by protocol and by source/destination. There was some concern expressed about the utility of this mechanism since addresses can be forged, the functionality may be unduly limited, and the management of port access on a fine-grained basis may be difficult.

While acknowledging limitations, there was substantial support for including security perimeters in the list of recommendations. The use of filters at this level helps to reduce the size of the problem. It advances the cause of configuration control, e.g it has clear utility in providing control for new machines brought on line (with pre-packaged network services); it provides a firewall which is not present in the software of each machine; it provides for logging of external connections; and it is an effective emergency response to attacks in progress from the outside.

Recommendations on Configuration Control

Configuration control/management consists of a potpourri of activities, some of which may be system specific. In addition to general password management and file protection, it might include such activities as initializing newly installed systems by immediately setting a root password, checking for proper settings of file access to various system files, closing off some functions like ftp or rlogin, closing down some logical ports, closing known security holes, etc. This area was considered to be of high importance to overall security. Without it, as previously noted, it may be impossible to accomplish many of the other things.

Automatic aids are available in support packages such as COPS and Security Profile Inspector (SPI) and their use is strongly recommended. Based on security configuration specifications, the latter package goes through a check of the system and constructs scripts that can be used as produced or can be modified to make the necessary configuration changes. These tools should be used by principal investigators and system administrators.

Recommendations on Audit/Review

A security audit information collection and review capability at each site was considered an important mechanisms for detection of security violations and points of vulnerability. COPS was again mentioned as one aid to help in the review process.

Recommendations on Security Officers/Teams

Steve Crocker recalled his earlier remarks and noted that the notion of a security officer or team actually referred to some professional staff time at each site devoted to such things as establishing and maintaining a security policy, monitoring the configuration, educating users, reviewing logs for breaches of security, having policies and procedures in place for handling security incidents, conducting fire drills to maintain a state of readiness, and self-education to keep up with changes in security technology and operations.

Each site should designate a security point of contact (POC), recognizing that the skill set and level of expertise required to do a good job of security audit and configuration management may vary among sites. At a minimum, the POC would serve as a postmaster who would contact appropriate individuals, especially in cases of emergency, e.g. messages from CERT or FIRST.

Recommendations on Oversight

(Although the workshop participants later voted to eliminate this recommendation, the discussion is included here so that the reader will appreciate the reason for its existence.)

It was noted that, once the workshop recommendations have been adopted, it will take some time to implement them. This particular recommendation was that each site have some person(s) responsible for implementing the recommendations (not necessarily the POC mentioned in the previous section). Furthermore, an oversight committee should be formed to promote cooperation and the exchange of information and tools. This committee should include representatives and implementors from each of the affected SCs along with the necessary security experts and sponsors.

Recommendations on Follow-On Activities

The final recommendation of the workshop was to establish an agenda for follow-on activities in order to emphasize that there are a number of things that cannot be done now, but should be done as new technology becomes available.

Prioritizing the Recommendations

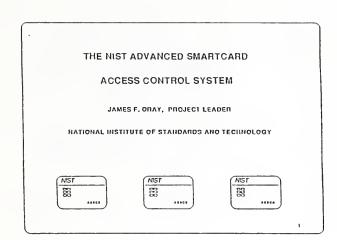
It was again emphasized that the items on the list of recommendations do not compete with each other; rather they complement each other. All items are considered sensible and important and can be accomplished now. The purpose for prioritizing the items was to identify those of highest urgency (Priority A), meaning that they "must be done now".

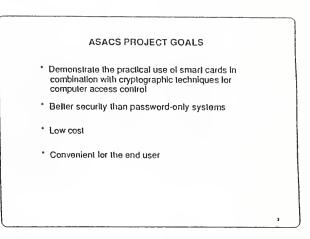
During the process of prioritization, it was decided to eliminate item 8 (Oversight). Item 9 (Followon Activities) was not assigned a priority because it is not an immediate implementation issue; rather it is an on-going activity expected of all. The priorities of the remaining items were established as:

Priority A:	C/R User Authentication Configuration Control/Management Security Officers/Team
Priority B:	Distributed Authentication PEM Security Perimeter Active Audit/Review

Cerf stated that an important follow-on activity would be to develop a security architecture for the NREN or the Internet. Aiken noted that the Federal Networking Council Security Working Group has some plans to work in that area. As part of follow-on activities, Branstad said that he would be providing the Federal Networking Council with a revised NREN Security Policy.

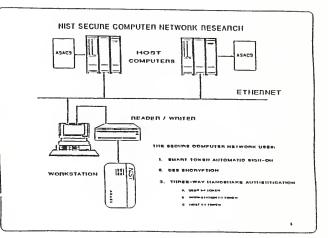
A View Graphs for Section 4.1





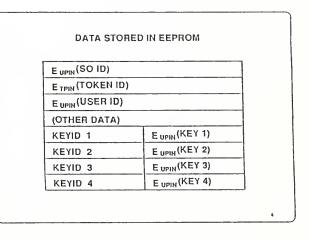
FEATURES OF ASACS

- Cryptographic authentication to multiple host computers
- Support for ANSI X9.17 automated key distrubution
- Low speed encryption
- Pseudorandom number generation
- Public key cryptographic capabilities
- * Reprogrammable lirmware



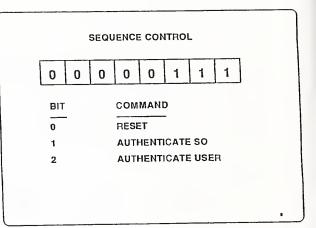
ASACS HARDWARE CONFIGURATION

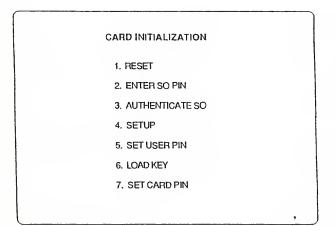
- * Hitachi H8/310 single chip smart card mlcroprocessor
- * ISO-compliant contact arrangement and form factor
- 10 kbytes ROM, 8 kbytes EEPROM
- Reader/writer with RS-232C serial Interface to host system

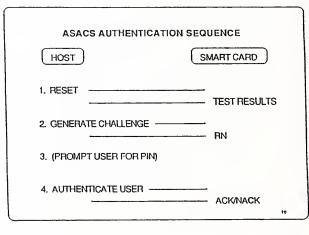


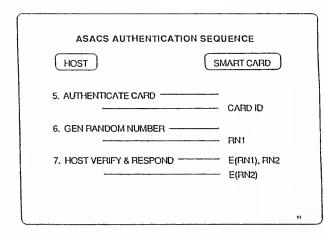
COMMAND SET INTERFACE

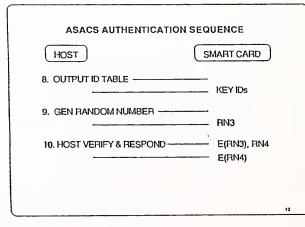
RESET ENTER SO PIN AUTHENTICATE SO SET USER PIN LOAD KEY AUTHENTICATE TOKEN GENERATE CHALLENGE GENERATE CHALLENGE GENERATE RANDOM NUMBER AUTHENTICATE USER CHANGE TOKEN PIN HOST VERIEY & RESPOND OUTPUT ID TABLE READ ZONE WRITE ZONE APPEND ZONE CALLDES TEST SETUP GENERATE KEY DELETE KEY EXPORT KEY MPORT KEY GET COUNT READ COUNT ACCEPT CHALLENGE SET ATTRIBUTE READ ATTRIBUTE READ GRAM











KEY MANAGEMENT

- 1. LOAD KEY
- 2. GENERATE KEY
- 3. DELETE KEY
- 4. EXPORT KEY
- 5. IMPORTIKEY
- 6. GET COUNT
- 7. SET COUNT

13

15

PUBLIC KEY APPROACH

Public key "primitives" will be Implemented In EEPROM firmware to support a variety of elgorithms. Primary goal Is to generate and verify digital signatures on the card.

А•в

A*(BMODM)

A ^ (B MOD M)

A + (B MOD M)

14

16

REPROGRAMMING THE ASACS FIRMWARE

- 1. Cross-compile source code for the H8/310
- 2. Issue the REPROGRAM command
- 3. Card erases all code and data stored in EEPROM
- 4. Download S-record lile to EEPROM
- 5. Firmware can be locked to prevent further changes

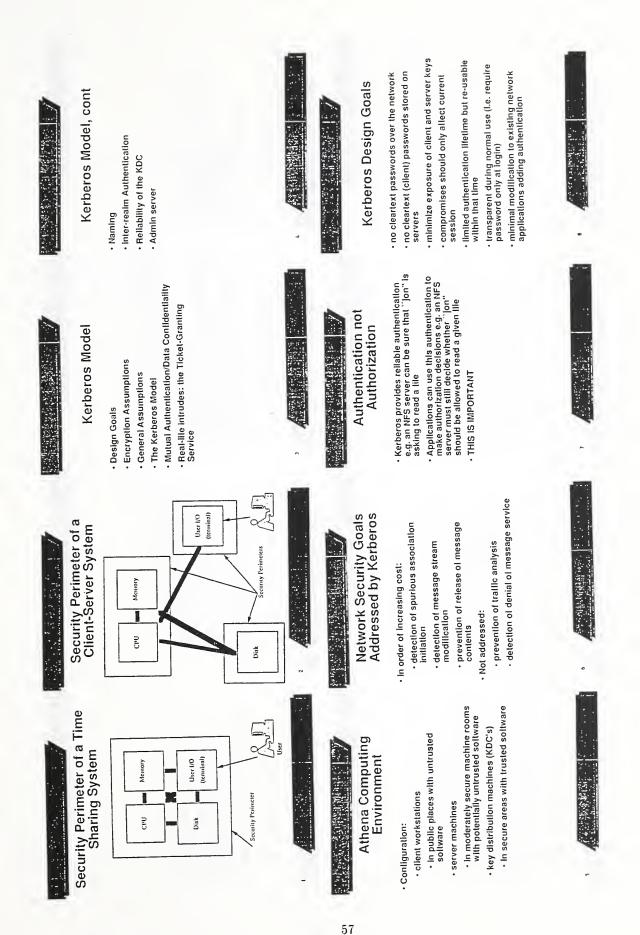
PROJECT STATUS

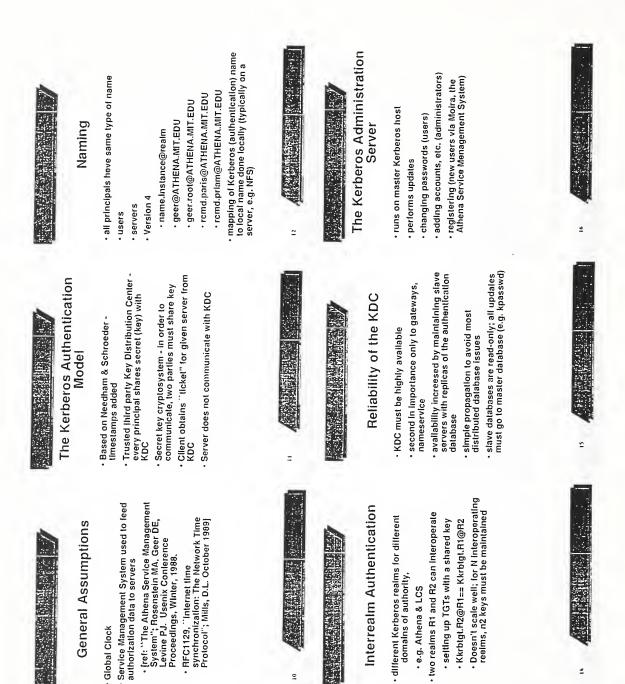
AUGUST 91 - Final delivery of secret key based product

4th quarter 92 - Completion of public key based smart card in conjunction with Trusted Information Systems

For lurther Information contact:

Jim Dray NIST (301) 975-3356 **B** View Graphs for Section 5.1







Encryption Assumptions

- DES is secure enough
- DES will only be widely available in software
- A few nodes (KDCs) may have herdware DES, though we don't feel It is necessary
- mechanism can be replaced (e.g. for export) Modularity is such that the encryption
 - RSA Is too expensive and not available lor domestic use without license
- This is actually changing... but was valid when Kerberos was designed



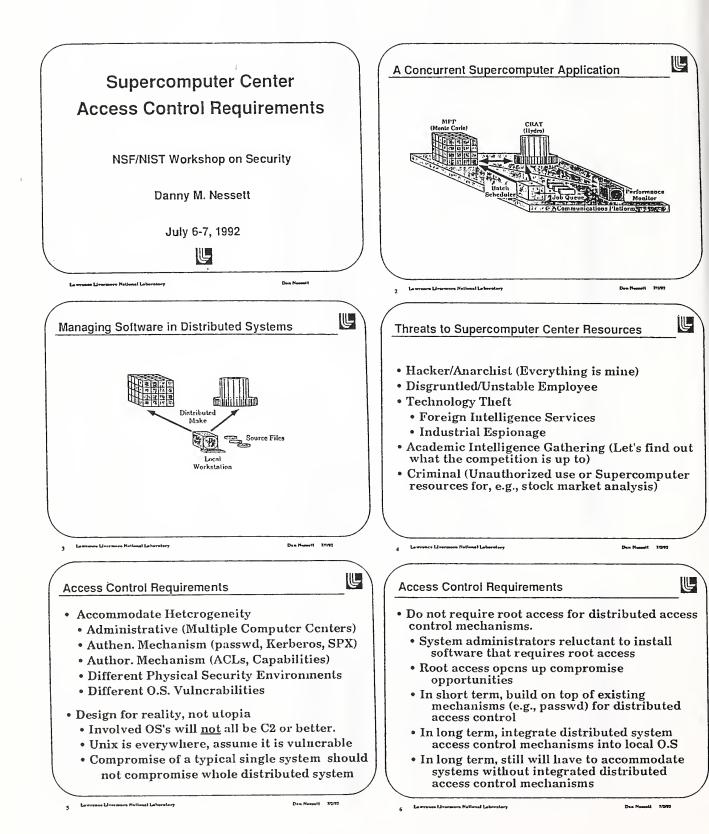
Version 5 Naming

- Some part of the name must include the A name is an arbitrary array of strings
 - realm identilier
- Designed to be general and capable of supporting X.500 names
- MIT V5 Implementation uses V4 style names
 - Still an open issue

2

2

C View Graphs for Section 5.2



Access Control Requirements

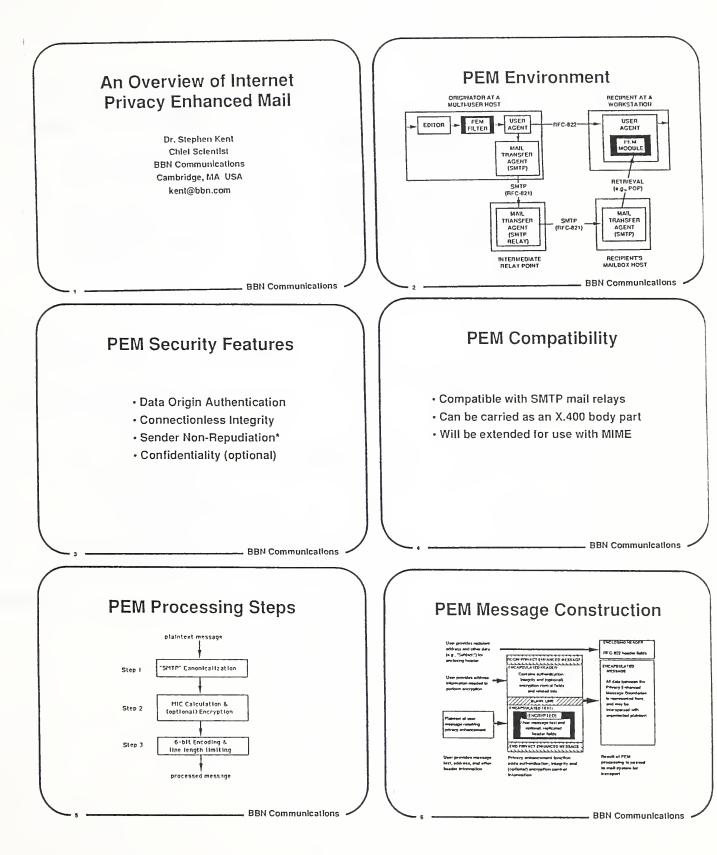
• Flexible and comprehensive approach

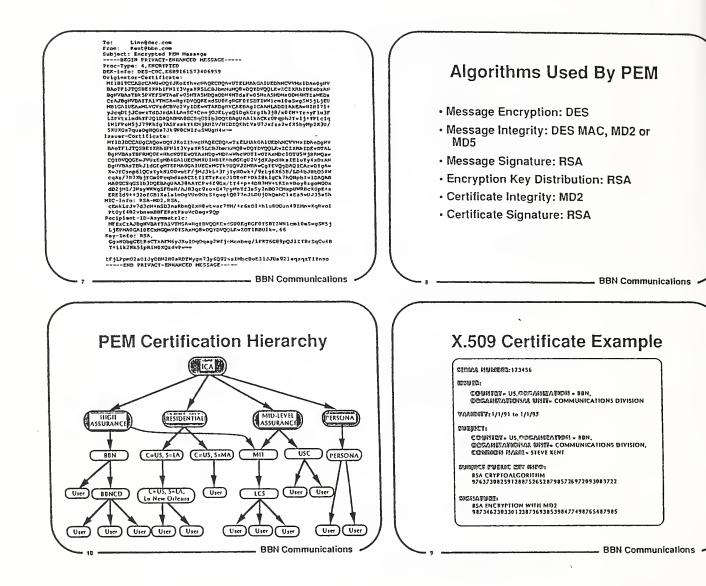
- Need delegation or equivalent functionality
- Support wide-diversity of applications (e.g., logon, file transfer, routing, etc. - see PSRG list of applications)
- Gateways between similar applications with different access control methods (design carefully to avoid hazards)
- Identify appropriate protocol layers for access control info passing – PSRG work on security architecture
- Need way to quickly revoke access to resources by misbehaving user

7 Lawrence Livermore National Laboratory

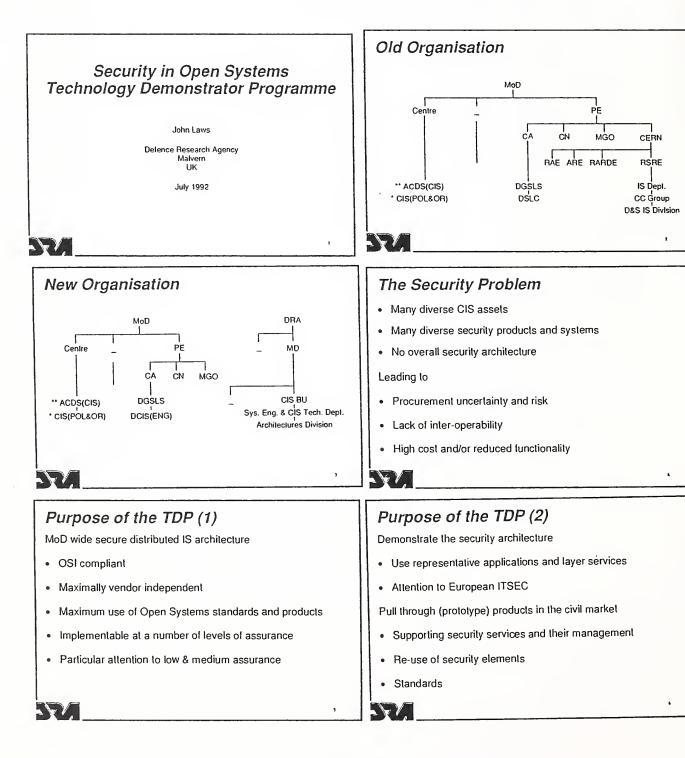
Den Nervell 7/2/92

D View Graphs for Section 6.2





E View Graphs for Section 6.3



NOT Purpose of the TDP (4) Purpose of the TDP(3) · Not to produce an evaluated prototype product Technical areas to be reviewed · Not to produce a MMHS CCITT/ISO X.400 inter-personal electronic messaging CCITT/ISO X.500 directory service ISO network layer security protocol Security services e.g. CA, authentication, access control Trusted functionality DBMS Trusted functionality distributed OS e.g. T-Mach & CHORUS - \- V Programme of work (1) Programme of work (2) Define and initiate the programme Phase II - 3 years Phase I - 1 year Generic security architecture in distributed CIS systems · Programme definition study to define Phase II System security policy model · Specific security architecture for TDP implementation · Selection and implementation of an early privacy electronic mail application Review, select and/or develop standards Review and select products supported by industry Implement upon extant MoD communications infrastructure Report, guide lines, recommendations 10 - Y Y ... Participation Phase I (1) Factor in extant R&D and commercial developments Industry, essential contribution EC COSINE PARADISE programme interworking national International directory services - NATO TSGCE SG9 Ad Hoc WG on Security EC ESPRIT THORN implementation of directory services - IEPG TFC3 US NSF NREN & Internet — ICB Canadian DND implementation of NATO TSGCE SG9 MMHS US DoD DMS 11 11

Phase I (2)

Security architecture

- US PEM
- EC VALUE PASSWORD programme
- NATO TSGCE SG9 MMHS
- EWOS and other regional OSI workshops

Application and layer services

- X.400
- X.500
- NLSP

-YY

Phase I (3)

DBMS

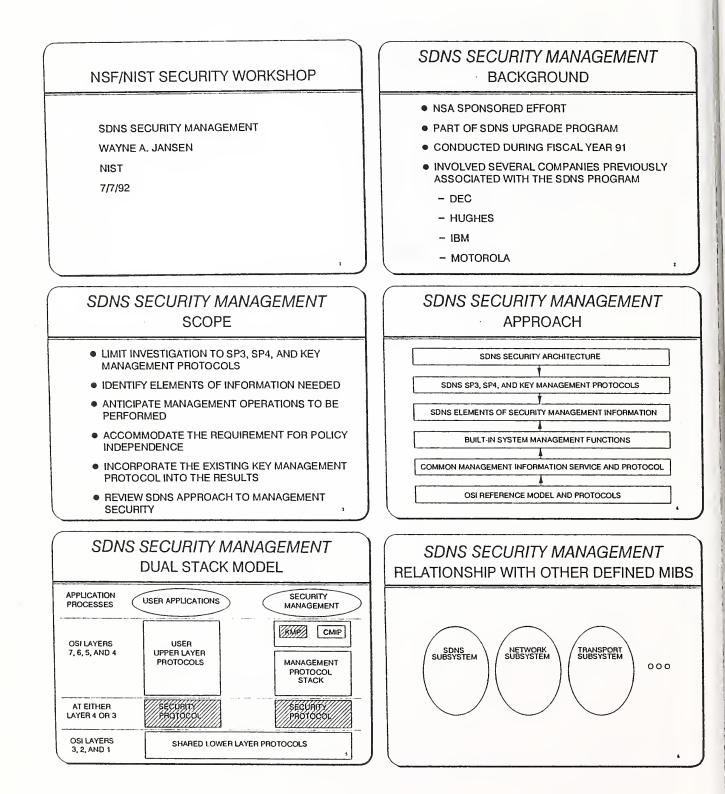
- Secure relational DBMS
- DRA SWORD DBMS security front end Distributed OS
- T-Mach
- CHORUS
- Privacy electronic mail
- US PEM

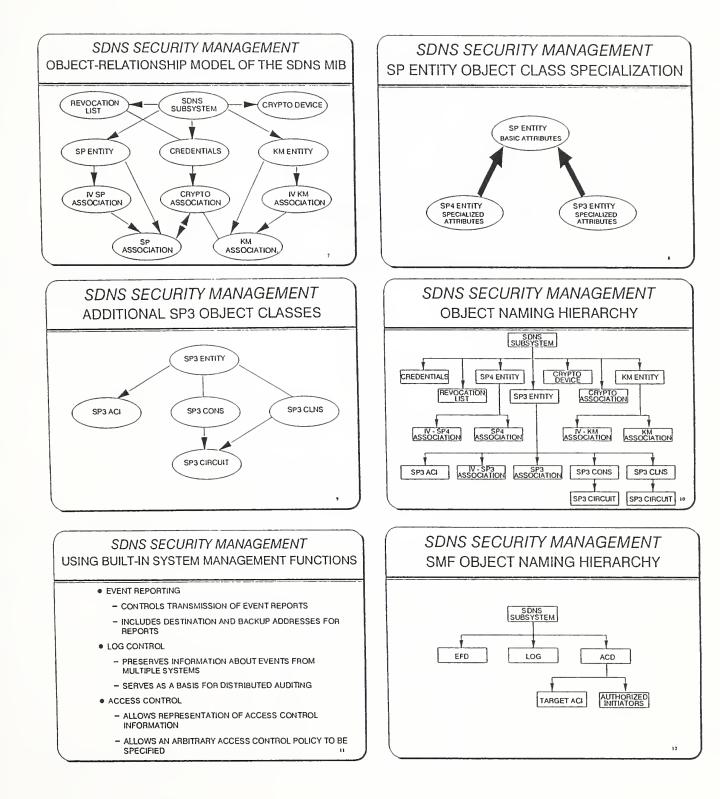
0

EC VALUE PASSWORD programme

84

F View Graphs for Section 7.1

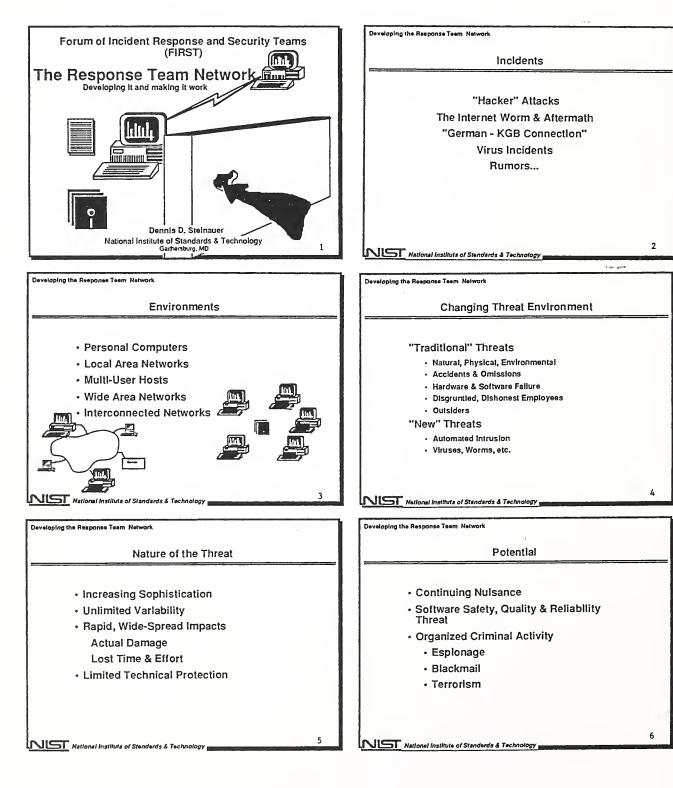


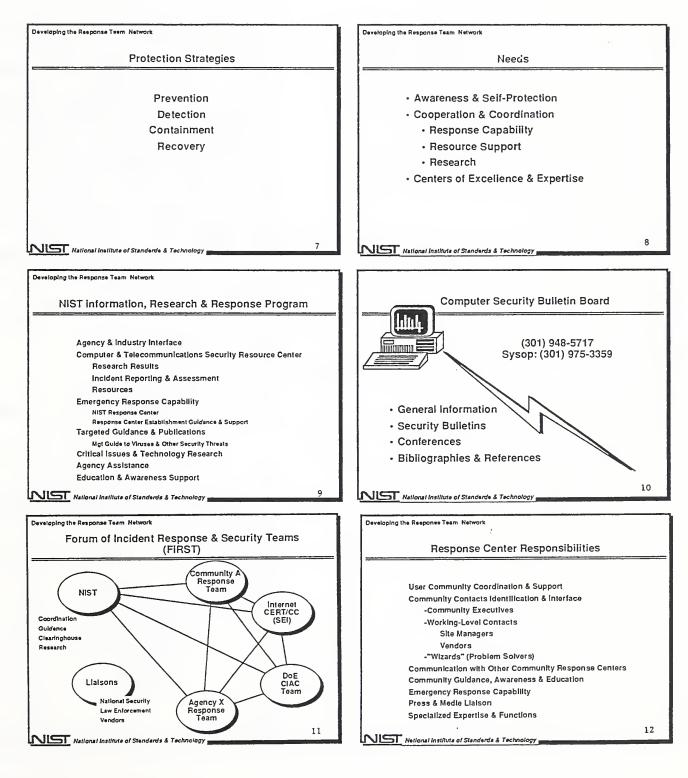


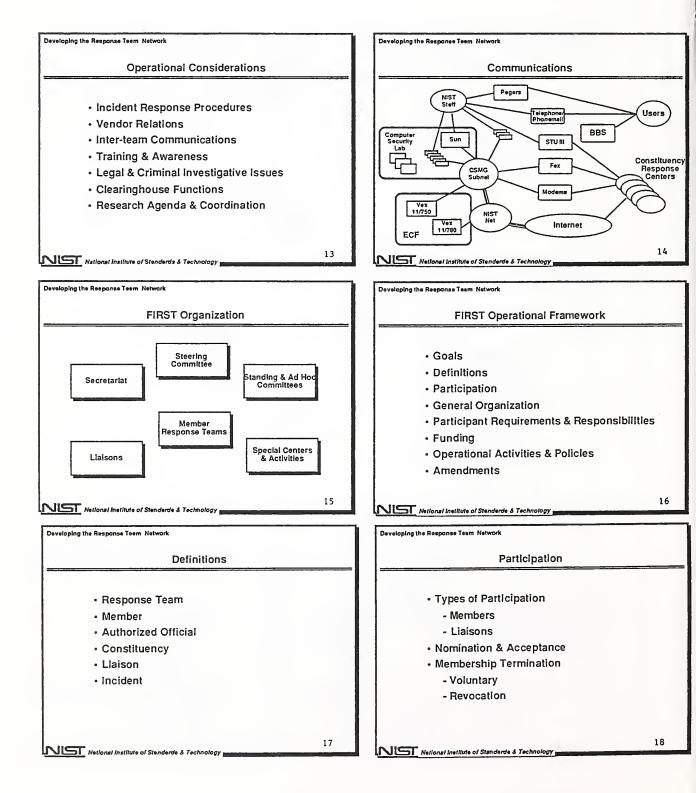
SDNS SECURITY MANAGEMENT OBSERVATIONS

- MANAGEMENT SHOULD NOT BE LEFT AS AN AFTERTHOUGHT
 - IT CUTS ACROSS ALL ASPECTS OF NETWORKING
 - IT DETERMINES THE COMFORT LEVEL OF DAY-TO-DAY OPERATIONS
 - NON-OSI SOLUTIONS ARE AVAILABLE TODAY
- SECURITY MANAGEMENT IS STILL A DEVELOPING AREA
 - STANDARDIZATION WORK IS CONTINUING
 - PROTOTYPING EXPERIENCE IS BADLY NEEDED 13.

G View Graphs for Section 7.2







Developing the Response Team Network	Developing the Response Teem Network
General Organization	Participant Requirements & Responsibilities
 Steering Committee Standing & Ad Hoc Committees Secretariat Meetings General Meetings Working Meetings 	 Participant Profile Communications Support Authorized Official
NUST Neilonal Inelliule of Stenderds & Technology	20
Developing the Response Team Network	Developing the Reepanse Teem Network
Operational Activities & Policies	FIRST Members & Liaisons
 Inter-Team Communications Information Handling & Dissemination Evidentiary Information Non-Disclosure Agreements Public Release of Information 	Members Air Force Computer Emergency Response Teem (AFCERT) Computer Emergency Response Team/Coordinaling Center (CERT/CC) Detense Information Systems Agency/Detense Deta Network (DISA/DDN) Department of the Army Response Team Department of Energy's Computer Incident Advisery Capability (CIAC) Digital Equipment Corporetion Software Security Incident Support Teem (ASSIST) Department of Energy's Computer Incident Advisery Capability (CIAC) Digital Equipment Corporetion Software Security Response Teem (SSRT) Godderd Spece Flight Center Micro-Bit Virue Center (MVC) - Germany NASA Ames Research Center Computer Network Response Teem (NASA ARC CNSRT) NASA Science Intermet Neval Computer Incident Response Teem (NAVCIRT) Netionel Institute of Stenderds end Technology (NIST) Space Physics Analysia Network France (SPA/NFrance) Purdue Computer Incident Response Committee for Unclassified Systems (CERCUS) Unsige CERT 22 NEST National Institute of Stenderds & Technology
Developing the Response Team Network	Developing the Response Team Network
FIRST Members & Liaisons (con'd)	Priorities & Issues
Liaisons Pann State University Defence Raseerch Agency, Royel Signels and Rader Establighment (RSRE) - UK DOW USA	 Operational Framework Operational System-Wide Procedures Vendor-Specific Constituencies Very Large Constituencies Cross-Agency Constituencies Alert/Notification Channels & Procedures Classified Information Issues
NIST National Institute of Standarde & Technology	NIST Neilonal Institute of Stendards & Technology

H View Graphs from Xerox Special Information Systems -Russ Housely

