Public Key Infrastructure Study

Final Report

April, 1994

National Institute of Standards and Technology
Gaithersburg, MD

Public Key Infrastructure Study

MTR xxxxx

Final Report

Dr. Shimshon Berkovits
Dr. Santosh Chokhani
Judith A. Furlong
Jisoo A. Geiter
Jonathan C. Guild

CONTRACT SPONSOR: National Institute of Standards and Technology
CONTRACT NO.: 50SBNB1C6732
PROJECT NO.: 3357B/Y
DEPT.: G021,J024

MITRE
McLean, Virginia

ABSTRACT

The National Institute of Standards and Technology (NIST) has tasked The MITRE Corporation to study the alternatives for automated management of public keys and of the associated public key certificates for the Federal Government. The public keys are envisioned to be used for secure electronic commerce. This Public Key Infrastructure (PKI) study focuses on the United States Federal Government operations, but also addresses national and global issues in order to facilitate the interoperation of protected electronic commerce among the various levels of government in the U.S., private citizens, commercial organizations, and international organizations.

Under the PKI study, policy and legal issues related to the operation and the management of the PKI are identified. Architectural and implementation alternatives for the PKI are developed. In addition, a methodology to determine the cost of the PKI is presented. The results of the PKI study are documented in this report. With the information and techniques presented in this report, federal agencies will be able to determine which infrastructure alternative is appropriate to their needs. In addition, agencies may use the costing methodology presented in the paper for planning and budgeting purposes.

EXECUTIVE SUMMARY

Use of electronic messaging and electronic commerce is becoming more widespread as information technology becomes cheaper and telecommunications become more advanced. However, increased user interconnectivity and reliance upon electronic communications means that more information is being carried electronically so that more information is becoming vulnerable to attacks such as eavesdropping, modification, and masquerade. Public key cryptography, which includes digital signature technology, can play an integral role in countering these attacks by providing end-to-end security of information in terms of confidentiality, integrity and proof of origin. The strength of these security services are dependent upon the security of the underlying cryptographic keys. Specifically, this requires protection of the confidentiality of the private keys and the integrity of the public keys in delivery and storage.

In a small community, the integrity of the public keys can be insured by manual delivery of the keys. However, manual delivery of keys is infeasible in a national or an international electronic messaging and commerce environment where thousands or millions of keys are involved. Therefore, in order to facilitate the use of public key cryptography in such an environment, automatic key management is necessary. This study addresses the issues related to a Public Key Infrastructure (PKI), which will automatically manage public keys through the use of public key certificates. Each certificate certifies the association between a user's identity and his public key.

The National Institute of Standards and Technology (NIST) has tasked The MITRE Corporation to study the alternatives for automated management of public keys and of the associated public key certificates in both a national and an international environment. This study focuses on the United States Federal Government operations. It also addresses national and global issues in order to facilitate the interoperation of protected electronic commerce among the various levels of government in the U.S., the states, private citizens, commercial organizations, and international organizations.

Under the PKI study, user and technical requirements for the PKI have been developed using information obtained through the examination of relevant standards, through discussions held among project members, and through interviews with personnel at various federal agencies, standards committees, and commercial organizations. The identified requirements apply to the infrastructure as a whole as well as to specific components of the infrastructure. They include requirements that relate to the generation and distribution of keys, to the obtaining of public key certificates and to the distribution of "hot lists" of canceled certificates, commonly known as Certificate Revocation Lists (CRLs).

The user and technical requirements form the basis for the architecture and the implementation recommendations for the PKI. A number of entities are integral to the

functioning of the PKI. Some have policy responsibilities; others provide certification services; and a few do both. These entities are defined below.

> **PAA:** The PAA is the policy approving authority, which creates the overall guidelines for the entire PKI. It may also certify public keys belonging to PCAs.
>
> **PCA:** The PCA is the policy certification authority. Each PCA establishes policy for all certification authorities and users within its domain. It certifies CA public keys.
>
> **CA:** The CA is the certification authority with minimal policy making responsibilities. The CAs are expected to certify the public keys of users in a manner consistent with the cognizant PCA's and the PAA's policies.
>
> **ORA:** The ORA is the organizational registration authority. The ORA is an entity which acts as an intermediary between a CA and a user. Its sole purpose is to vouch for the identity and affiliation of the user and register that user with its CA.

These entities can be organized in either a hierarchical (tree) structure or a non-hierarchical (network) structure. In the former, the PAA issues certificates for the PCAs while in the latter it approves the PCA policies. Each PCA issues certificates for the CAs within its domain. In the non-hierarchical structure, roots of smaller trees must issue certificates for each other (cross certification) as circumstances require. As cross certifications create uncertainties about and ambiguities between the applicable certification policies, the PKI does not follow this alternative.

There are several approaches to how users are associated with PCAs and CAs within the PKI. At least three alternatives are possible. The Communities of Interest (COI) approach organizes users according to the tasks they perform most frequently. The organizational alternative parallels existing organizational hierarchy. The assurance level approach divides users according to the level of certification assurance they require. A small number of assurance levels, perhaps as few as three or four, may be sufficient.

Any determination of which implementation approach is most appropriate for the national PKI must consider the following qualitative attributes: robustness, scalability, flexibility and ease of use, trust, interoperation, implementation timeframe, management structure, and exposure to liability. These considerations lead to the following recommendation for the PKI:

- PAA, a national body, be created to establish overall PKI policy, to approve individual PCA policies, and to act as a root for the national certification infrastructure to be created.

- Each federal department implement its own PCA to establish its own policy. PCAs be established for sets of independent commissions and independent agencies.

- Each PCA be certified by the PAA.

- CAs be established by offices and bureaus within large federal departments and independent agencies, as determined by the PCA. ORAs be placed near individual facility security offices, as needed.

- Several assurance levels, with associated PCAs and CAs, be established for use by private corporations and citizens. ORAs be placed near corporate personnel offices, as needed.

- Each user caches (stores) all certificates he or she uses most frequently.

The following organizations are recommended for the various levels of the PKI:

PAA. Policy setting be done by a committee consisting of representatives from some or all of the following: Defense Information Systems Agency (DISA), Federal Reserve Board (FRB), General Accounting Office (GAO), General Services Administration (GSA), NIST, National Security Agency (NSA), Office of Management and Budget (OMB), United States Postal Service (USPS), and industry and trade organizations. The daily operation of the PAA's certification functions can be managed by GSA, the FRB or USPS.

PCA. Executive departments or major independent agencies are recommended for the management of the PCAs for the Federal Government. USPS, banks, and telecommunications service providers are among the qualified organizations for facilitating interoperation of the federal infrastructure with the rest of the national infrastructure.

CAs. Agencies below the executive department level are recommended for the CA management. USPS, banks, and telecommunications service providers are among the qualified organizations for facilitating interoperation of the federal infrastructure with the rest of the national infrastructure.

ORAs. For the federal agencies, local authorities such as badge-issuing offices or security offices are recommended to run ORAs. For the non-federal segment, ORAs may be established by private corporations for their own convenience.

Global interoperability is also of interest. There are two principle ways for achieving interoperability: (1) the creation of a single global root for certifying various national and transitional roots (PAAs); or (2) the cross certification by such roots (PAAs). While the

former alternative may create a neater infrastructure, the politics of international agreement and the natural flow of events may make the latter alternative the *de facto* choice.

The recommended PKI organization follows the lines of the existing Federal Government organization. The cost of the various PKI components will probably be divided along the same lines. Nonetheless, it is useful to estimate a PKI-wide price tag for deployment and for the annual operations budget. This effort requires the development of a cost model for the PKI. Such a model should also aid implementors with their planning and budgeting activities for their segments of the PKI. The model should include only quantitative impacts which can be specifically ascribed to the PKI as well as the cost of the additional demands that the PKI will place on the directory service and its supporting communications. Costs associated with developing and running an electronic transaction environment are specifically excluded. They are deemed prerequisites for the use of digital signatures.

It is not possible to estimate the cost of a large computer and communications system without some detailed concept of how that system will operate. A concept of operations for the PKI consists of thirteen distinct activities. Each of these PKI activities can be divided into multiple steps. The resources necessary to accomplish each step must be enumerated. These resources fall into four categories: storage, communications, processor and staff. The model consists of four modules, one for each of the four PKI entities: PCA, CA, ORA and key generator. There are additional modules for the user and for the directory. Within each module, costs and frequencies for all needed resources are computed and a total cost is derived. From these totals and the numbers of each type of PKI entities included in the costing, it is possible to estimate PKI costs. Start-up and the yearly running costs are estimated to be about $9.5M and $16.2M, respectively.

Policy makers must consider their legal obligations in establishing a PAA, PCAs and CAs. Although there are questions that must be answered, there appears to be no legal impediment to delay continued development and ultimate implementation of the PKI. There seems to be a consensus that properly supported digital signatures on electronic documents can and do meet "signed" and "in writing" laws and regulations. What exceptions exist to this principle can be remedied by new legislation or further regulations. Electronically executed contracts, especially those that are covered by written trading partner agreements, are becoming legally accepted. The electronic filing of certifications, reports and briefs, if supported by digital signatures, will be widely used and accepted. These considerations point to the recommendation that PKI prototype development and examination of PKI legal needs and implications should proceed in parallel.

One important legal consideration is the development of a structure for PKI liability. As an agency of the Federal Government, the PKI may be considered to have sovereign immunity. That would imply that the PKI and its managers cannot be sued for any losses resulting from their actions or from their inaction. While such status may be attractive, it undermines the usefulness of the PKI. Without reasonable assurances that potential losses due to malfeasance will be recoverable, a typical non-government user will shy away from

relying on the PKI. (Even many government users will balk at mandated PKI use unless some system is in place for fixing blame when the PKI fails in its mission.) Any set of laws and regulations must strike a balance between protection of the government from excessive claims and blocking users from any chance of reimbursement. The entities of the PKI should not be liable if they operate according to their established policies. Furthermore, their liability should be limited for losses caused by the entities not following the agreed policies and procedures. In addition, legal issues relating to personal privacy concerns must also be resolved.

This summary has highlighted some of the issues that must be considered in defining the PKI. Based on the recommended infrastructure, work must begin on its implementation. The infrastructure should be implemented incrementally to gain practical experience with PKI policies, user acceptability, usage statistics, resource requirements, costs and liability issues. The practical experience obtained through implementation will help to refine and alter the infrastructure design and operation as well as the full scale implementation.

The designers of the PKI should be aware that there are still issues that need resolution. They include some concerns that pertain to the PKI itself and others that are external to the PKI, in parallel systems with which PKI may have to interface in the future. Examples of the latter include user authorizations, user attributes, time and date stamps, interoperation with other infrastructures and algorithms, and directory services. PKI issues involve archiving, a unique naming system, and confidentiality. This last issue is of utmost importance and demands immediate attention. The overall trust that users place on the PKI derives, to a considerable extent, from the ready availability of CRLs. Questions of liability depend on the way in which the CRLs are handled. Yet, the frequent and extended distribution of CRLs creates a great financial drain on the system. CRLs are to be obtained from the appropriate directory elements as needed. The problem of liabilities associated with the use or non-use of the CRLs should be studied immediately and resolutions found quickly.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Carolyn Barnes, Dennis Branstad, Shirley Kawamoto, Lynn McNulty, Judith Messing, Cyril Murphy, and Miles Smid for participating in discussions regarding the Public Key Infrastructure and/or for reviewing and commenting on this paper. We are especially grateful to David Gill and Shari Galitzer for their contributions to this project. For his advice and guidance on the legal issues, we are indebted to Michael Baum and we acknowledge Peter Weiss's comments and corrections on the draft of that section. We are also indebted to the following experts for their peer-review comments: Richard Ankney, Fisher International Systems Corp.; James Bidzos, RSA Data Security Inc.; Stephen Crocker, Trusted Information Systems; Dorothy Denning, Georgetown University; M. Blake Greenlee, M. Blake Greenlee Associates, Ltd.; Stephen Kent, BBN Laboratories; Stuart Stubblebine, University of Southern California; and Frank Sudia, Bankers Trust. Finally, we must thank all who participated in the interview phase of this study and all who challenged and encouraged us at the periodic participating agency meetings.

TABLE OF CONTENTS

SECTION			PAGE
1	Introduction		1-1
	1.1	Background	1-1
	1.2	Purpose	1-2
	1.3	Scope	1-3
	1.4	Audience	1-3
	1.5	Overview of Report	1-4
2	Overview		2-1
	2.1	Project Methodology	2-3
		2.1.1 Developing Requirements	2-3
		2.1.2 Analyzing Requirements	2-3
		2.1.3 Developing Alternatives	2-3
		2.1.4 Developing operational Concepts	2-4
		2.1.5 Cost Model and Analysis	2-4
	2.2	Highlights of Key PKI Requirements	2-4
3	Architectural Alternatives for the Infrastructure		3-1
	3.1	Policy Management and Certificate Management	3-1
		3.1.1 The Policy Approval Authority	3-1
		3.1.2 Policy Certification Authority	3-1
		3.1.3 Certification Authorities	3-2
	3.2	Trees and Cross Certification	3-2
	3.3	Architecture Alternatives	3-6
		3.3.1 Cross Certified CAs	3-6
		3.3.2 Cross Certified PCAs	3-6
		3.3.3 A Root PAA	3-7
4	Implementation Alternatives for the Infrastructure		4-1
	4.1	Functions Performed by CA Entities	4-1
		4.1.1 PAA Functions	4-1
		4.1.2 PCA Functions	4-2
		4.1.3 CA Functions	4-3
		4.1.4 ORA Functions	4-4
	4.2	The Implementation Alternatives	4-5
		4.2.1 COI Alternative	4-6
		4.2.2 Organizational Alternative	4-7

SECTION **PAGE**

4.2.3 Assurance Level Alternative 4-8
4.2.4 Hybrid Alternative 4-9
4.3 Recommendation 4-10
4.3.1 Comparison of Implementation Alternatives 4-10
4.3.2 The Recommended Federal Government Alternative 4-17
4.3.3 Beyond the Federal Government 4-18
4.4 Management of the Recommended Infrastructure 4-19
4.4.1 Management of the PAA 4-19
4.4.2 Management of the PCAs 4-20
4.4.3 Management of the CAs 4-20
4.4.4 Management of the ORAs 4-20
4.4.5 Summary of Recommended Organizations for the Infrastructure Management 4-20
4.5 Towards Global Interoperability 4-21
4.5.1 Cross Certification with All Roots 4-21
4.5.2 One Global Root 4-22

5 Operational Concepts 5-1

5.1 Introduction 5-1
5.2 PKI Activities 5-1
5.2.1 Key Generation, Certifying, and Distributing Keys 5-1
5.2.2 Signature and Verification 5-2
5.2.3 Obtaining Certificates 5-3
5.2.4 Verifying Certificates 5-3
5.2.5 Caching Certificates 5-4
5.2.6 Obtaining Certificates from a Cache 5-4
5.2.7 Reporting Key Compromise or Severed Relations 5-4
5.2.8 Recovering from a Key Compromise 5-5
5.2.9 Obtaining CRLs 5-6
5.2.10 Rekeying and Recertifying 5-6
5.2.11 Auditing 5-7
5.2.12 Archiving 5-7

6 Infrastructure Cost Analysis Results 6-1

6.1 Operational Concept Used in Cost Model 6-1
6.1.1 User Activities 6-1
6.1.2 PKI Activities 6-2
6.1.3 Directory Activities 6-3
6.2 Scenarios 6-3
6.3 Results 6-3

SECTION **PAGE**

6.4 Analysis 6-5
6.4.1 Analysis of Start-up Costs 6-5
6.4.2 Analysis of Yearly Running Costs 6-7
6.4.3 Analysis of Cost per Message and Cost per User 6-10

7 Related Issues 7-1

7.1 Authorizations and Attributes 7-1
7.2 Time and Date Stamps 7-2
7.3 Archiving 7-2
7.4 Confidentiality 7-3
7.5 The Next Steps 7-3
7.5.1 General Planning 7-4
7.5.2 Cryptography Awareness 7-8
7.5.3 Beyond the Executive Branch 7-8
7.5.4 Forge Ahead 7-8
7.5.5 Develop Multidisciplinary Development Group 7-9
7.5.6 Liability 7-10

Appendix A: Terms and Definitions A-1

Appendix B: Digital Signature Standard B-1

Appendix C: Applications of Digital Signatures C-1

Appendix D: Summary of PKI Requirements D-1

Appendix E: Applicable Standards and Analysis E-1

Appendix F: Certificate Formats F-1

Appendix G: Certificate Revocation List Format G-1

Appendix H: Sample Elements of PCA Security Policy H-1

Appendix I: PKI Cost Analysis I-1

Appendix J: Legal Issues J-1

LIST OF FIGURES

FIGURE		PAGE
2-1	Typical Digital Signature Scheme	2-2
3-1	Users with a Common CA	3-2
3-2	Users with Different Cross Certified CAs	3-3
3-3	Users with Different CAs under a Single PCA	3-4
3-4	Users with Different Cross Certified PCAs	3-5
3-5	Users with Different PCAs under a PAA	3-5
4-1	Communities of Interest Alternative	4-7
4-2	Organizational Alternative	4-8
4-3	Assurance Level Alternative	4-9
4-4	Sample Hybrid of Certificate Management Infrastructure	4-10
4-5	Example of COI Alternative	4-11
4-6	Example of Organizational Alternative	4-12
4-7	Recommended National Certificate Management Organization	4-18
4-8	Interoperability through One Global Root	4-23
7-1	First Phase of Public Key Infrastructure	7-5
B-1	The DSS Signature Process	B-2
B-2	The DSS Signature Verification Process	B-3
E-1	PEM Key Management Infrastructure	E-3
F-1	Proposed 1992 CCITT X.509 Certificate Format	F-2
F-2	1988 CCITT X.509 and PEM Certificate Format	F-4
G-1	PEM CRL Format	G-2

FIGURE **PAGE**

G-2 CCITT X.509 CRL Format G-3

G-3 ANSI X9.30 CRL Format G-4

G-4 Recommended PKI CRL Format G-5

I-1 Entity Spreadsheet Layout I-26

LIST OF TABLES

TABLE		PAGE
4-1	Summary of Policy-Setting Body and Users for Each Alternative	4-5
4-2	Infrastructure Management	4-20
6-1	Total Start-Up Cost Estimates	6-4
6-2	Total Yearly Cost Estimates	6-5
6-3	Start-up Costs	6-6
6-4	Yearly Costs	6-8
6-4	PKI Yearly Costs (Concluded)	6-9
6-5	Cost per Message and Yearly Cost per User	6-12
I-1	Combined Translation Table	I-30

SECTION 1

INTRODUCTION

1.1 BACKGROUND

As information technology becomes cheaper and proliferates further in both office and home environments, use of electronic messaging and electronic commerce is becoming widespread. The transaction of business electronically is further spurred by the National Information Infrastructure (NII) and by the Defense Information Infrastructure (DII) initiatives. Advances by telecommunications service providers and by the cable industry in establishing a national information highway also play a major role in the growth of electronic commerce. Together, the information technology revolution and the communications infrastructure initiatives are changing the way we do business. They are bringing about a new interconnection of individuals and organizations both nationwide and worldwide. But this interconnectivity and reliance on electronic communications makes the information being carried more vulnerable to attacks – attacks to information confidentiality in the form of eavesdropping, to data integrity through message modification or substitution, to origin integrity by an impostor's submission of messages in another user's name, and to message dependability destroyed by the possible future repudiation of the message by its sender.

Public key cryptography can play an integral role in providing end-to-end security of information in terms of confidentiality, integrity, and proof-of-origin. This cryptography is based on asymmetric keys. Each user has two keys: one is called the public key and may be available to everyone, the other is called the private key and is known only to its owner. When two typical users, call them Alice and Bob, communicate, they can use their public key capability to keep their messages confidential. If Bob wishes to hide the contents of a message to Alice, he encrypts it using Alice's public key. Encryption, however, is not germane to this study. If Bob wishes to sign a document, he must use a key available only to him that is, his private key. When Alice receives a digitally signed message from Bob, she must verify his signature. She needs his public key for this verification. She should have high confidence in the integrity of that key. If she can be tricked into accepting an impostor's public key as Bob's, then she will accept the impostor's signature on documents as Bob's also. Digital signature technology offers some of the desired information security services, namely sender authentication, message integrity and sender non-repudiation, provided that private keys are kept secret and the integrity of public keys is preserved.

In a small community, the integrity of keys can be guaranteed by the manual delivery of public keys. This is impossible in national or international electronic messaging and commerce environments. Anyone, anywhere, may send a message to anyone, anywhere. There is no way to exchange thousands or millions of keys manually and their storage is also difficult. This study addresses the management of public keys which will facilitate digital signatures for the worldwide electronic transactions of the Federal Government. It is

assumed that readers of this report are familiar with public key cryptography in general and with the Digital Signature Standard (DSS) in particular. For those who are not, appendix B contains a brief description of the standard.

1.2 PURPOSE

The following agencies in the United States Federal Government wish to explore the role of digital signatures in electronic commerce within in the Federal Government: Advanced Research Project Agency (ARPA), Department of State, Federal Bureau of Investigation (FBI), General Services Administration (GSA), Internal Revenue Service (IRS), National Aeronautics and Space Administration (NASA), National Security Agency (NSA), and United States Postal Service (USPS). These agencies have asked the (NIST) to study the technical, policy, and legal issues associated with establishing an automated system to manage keys electronically and distribute public keys and associated public key certificates.

NIST has tasked The MITRE Corporation to study the alternatives for automated management of public keys in a national and international environment. The focus of the study is the Federal Government operations. However, the study also needs to address the issues on a global basis, since the Federal Government conducts business with both national and international entities (governments at various levels, private citizens and business, and other organizations).

The automated management of public keys is henceforth termed the Public Key Infrastructure (PKI). The purpose of PKI is fourfold:

1. Generate public key certificates that bind the identity of users and their public keys in a secure manner

2. Provide users, directly or indirectly, with easy access to the certificates of other users

3. Provide users with easy access to circumstances (security policy) under which the certificates were issued.

4. Provide users, directly or indirectly, timely announcements of certificate revocations.

The purpose of this study is fourfold:

1. Identify policy and legal issues in the use of digital signatures in the Federal Government operations

2. Identify policy, technical, and legal issues related to the operation and the management of the PKI

3. Develop PKI alternatives for federal agencies and techniques for selecting the most appropriate alternative based on an agency's needs

4. Provide a PKI costing methodology to calculate costs for planning and budgeting purposes on an agency-by-agency basis

1.3 SCOPE

The PKI will provide a secure binding of public keys and users. In that sense, its scope is limited to providing strong authentication of the users. This report examines the issues involved in such an undertaking. It does not set any standards for an infrastructure. While this study focuses on the PKI needs of the Federal Government, it must consider other national and international entities such as individuals, businesses and other organizations. A considerable portion of the Federal Government's electronic commerce is expected to involve people outside the Federal Government. Thus, the PKI needs to be flexible in terms of accommodating a wide variety of digital signature schemes (since different entities may use different digital signature algorithms), and the PKI must be flexible in allowing heterogeneous substructures to interoperate. However, the final architectural recommendation is based on the needs of the Federal Government and not on those of the external agencies and entities.

The United States is moving towards split-key, multiple trusted parties, key escrowing to allow law-enforcement agencies to decrypt court-authorized wiretaps. The multiple trusted parties, key-escrowing and associated key management is germane to both the secret keys in symmetric encryption schemes and to the private keys in asymmetric encryption, that is, in public key cryptography. The secret key and private key escrowing and management will complement the PKI. Thus, PKI issues can be studied independently of the secret and private key escrow schemes.

1.4 AUDIENCE

This report describes the results of the PKI study. The primary audience for this report is the decision makers at the General Accounting Office (GAO), Office of Management and Budget (OMB), GSA, and at other federal agencies with representatives in the following areas:

1. Policy associated with using digital signatures

2. Policy associated with setting up the PKI

3. Selection of management and operations alternatives for the PKI

4. Budgeting for the PKI

The audience also includes researchers and standards developers within such federal agencies as NIST, NSA, the ARPA, and the Defense Information Systems Agency (DISA).

It is assumed that the reader is knowledgeable in the following areas:

1. Electronic commerce and role of public key cryptography and digital signatures in secure electronic commerce

2. Public key cryptography and digital signatures technology

3. United States cryptography policy

1.5 OVERVIEW OF REPORT

In section 2 of this report, the project methodology is described. A summary of key requirements that the PKI must fulfill is also given. Architectural alternatives for the PKI are described in section 3. The next section discusses several approaches for organizing users within the different architectural choices. It concludes that one approach holds several advantages for the Federal Government portion of the PKI. It suggests how that approach might interact with the remainder of the national infrastructure and other, international infrastructures. It also suggests candidate government agencies to be policy and certification authorities in the PKI. Section 5 describes possible concepts of operations for the PKI. One of them is selected, in appendix I, as the basis for a cost analysis. The results of the cost analysis for the recommended PKI infrastructure operating under the chosen concept of operations are presented and discussed in section 6. The paper concludes with issues related to the use of DSS that do not fall within the scope of the PKI Study but which require further examination.

Ten appendices are included in this report. The first appendix contains a glossary of pertinent terms related to the PKI. The next appendix presents a brief tutorial on the DSS. An overview of the applications in which DSS may be used are presented in appendix C. The user and technical requirements for the PKI are summarized in ensuing appendix. There then follows brief overviews of the standards that were examined in support of this study. The formats for certificates and for a Certificate Revocation List (CRL) are described in the sixth and seventh appendices. Appendix H contains a sample security policy for a Policy Certification Authority (PCA). The elements of the cost model are presented in appendix I. Finally, federal laws and policies examined as part of the PKI Study and which were most pertinent to the PKI are described in the final appendix. To facilitate the use of DSS and allow for the establishment of the PKI, existing laws and policies may need to be modified or repealed and additional laws and policies may need to be enacted.

SECTION 2

OVERVIEW

The purpose of a digital signature scheme can be met only if each user who verifies a signature has confidence in the integrity of the public key he uses in the verification computation. Figure 2-1 depicts the general signature and verification processes. In the simplest variation, the signature function is a transformation of the message digest. The transformation depends on the signer's private key. The verification function uses the signer's public key to recover that message digest. The verify equation is a simple comparison between the recovered digest of the original message and the recomputed digest of the received message. Other signature schemes have more complicated signature and verification functions and a more complex verify equation. In all cases, however, the process depends on the use of associated private and public keys.

The verifier may trust the key he uses in the verification function as being the signer's public key because it was manually delivered to him by the signer, whom he knows personally. Failing that, he trusts the key because he obtained it from a certificate signed by an entity for which he holds a public key he trusts. He trusts that key either because it was manually delivered to him or because he received it in a certificate of its own. Obviously, his trust in the integrity of the key with which he verifies a signature on a document depends on his holding a chain of trusted keys. His trust in the first key in the chain derives from having it delivered to him in a trusted "out-of-band" manner. He trusts all the other keys because each key in the chain is contained in a signed certificate. He can verify the signature on each key certificate by using the key that immediately precedes it in the chain. Thus, he has established a chain of trust from a key that was handed to him by a known individual to the key he uses to verify the document's signature.

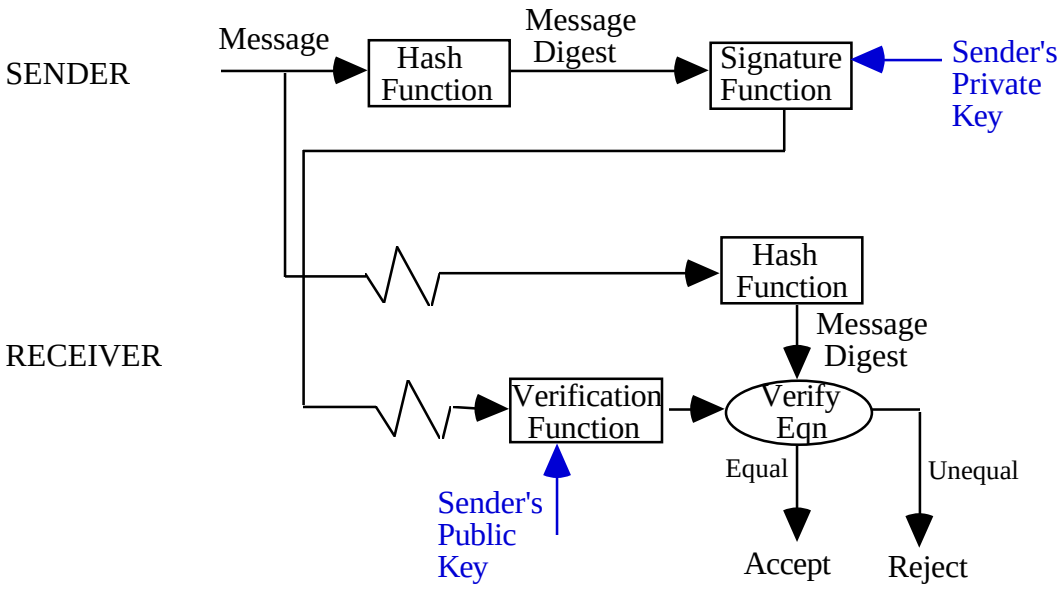


Figure 2-1: Typical Digital Signature Scheme

Our objective is to design an infrastructure that will allow users to establish chains of trust, commonly called "certification paths," which contain more than one key but which, in most cases, are no more than a few steps in length. Certification paths of length greater than one are important. Users will not personally know all other users with whom they must interact electronically. They will have to verify signatures of users with whom they have never communicated previously. The purpose of the public key infrastructure is to facilitate trusted electronic correspondence beyond those users with whom one has manually exchanged public keys.

Pivotal to the creation of chains of trust are Certification Authorities (CAs). These authorities certify the association of a PKI entity's identity and that entity's public key. The identity is contained in a unique name, that is a name which distinguishes this entity from any and all other PKI entities. Such a name is often formed by concatenating a sequence of locally unique names of some hierarchical substructure of the PKI. Each certification is contained in a certificate which is signed by the issuing CA. A certification path is a sequence of CAs, the first being a CA for which the verifier holds a trusted copy of the public key and the last being the CA that issued the certificate certifying the needed PKI entity's public key. The intermediate steps in the certification path are all CAs, each of which has certified the next.

Some CAs may have an additional, policy-setting responsibility. Such CAs are known as Policy Certification Authorities (PCAs). How CA's and PCA's functions and responsibilities are defined and how they are made to interact determines the nature of the infrastructure that is created. The majority of tasks associated with the PKI centered on researching and

developing these functions, responsibilities and interactions. The study includes the identification of the user, technical, and legal requirements for the PKI, development of architectural and implementation alternatives for the infrastructure, and recommendation of the preferred one. A cost analysis of the recommended infrastructure was also undertaken. This effort required the investigation of alternative concepts of operations, selection of a concept of operations, and development of the cost model for the selected concept of operations. The results of this effort are reported in the remainder of this paper. Much of the background information appears in the appendices.

2.1 PROJECT METHODOLOGY

2.1.1 Developing Requirements

The identification and development of the user, legal, and technical requirements for the PKI was an iterative process. An initial set of requirements was developed from project discussions and through the examination of the appropriate standards, literature, laws, policies, and federal regulations. The initial user, legal, and technical requirements were presented to NIST and the participating agencies in three draft papers [1, 2, 3], which were circulated to stimulate discussion and elicit comments.

A series of interviews was conducted with individuals at the participating agencies and at other federal agencies to obtain insight into the agencies' requirements for the PKI. Interviews with individuals from private user groups, standards bodies, and vendors of cryptographic hardware and software were also conducted. Using the comments provided by NIST and the participating agencies on the draft requirements papers, along with information gathered through the interviews, the requirements were refined to produce the final set of requirements for the PKI. Brief outlines of the relevant standards appears in appendix E while a detailed list of requirements can be found in appendix D.

2.1.2 Analyzing Requirements

Throughout the study, the requirements have been analyzed to determine which requirements are common to most operational environments and/or applications and which are specific to a particular environment and/or application. The results of the requirements analysis have influenced the development of the infrastructure alternatives. Each alternative infrastructure must be able to satisfy all the common requirements, and certain portions of each alternative must satisfy the environment and/or application dependent requirements. The requirements analysis has also helped the development of the infrastructure alternatives. The final set of common requirements is summarized below in section 2.2.

2.1.3 Developing Alternatives

From project discussions, the analysis of requirements, and the review of evolving standards, several architectural alternatives for the national infrastructure were developed. For the most promising architectures, infrastructure implementation alternatives, which

focus on how the users should be grouped and who should set the security policies, were also developed. The implementation alternatives were analyzed to see how well each met the list of requirements as expressed in a set of qualitative attributes. These included such things as ease of use, level of trust, flexibility, scalability, robustness, and interoperability. The results of this analysis identified an architecture and an implementation suitable for use in the PKI.

2.1.4 Developing operational Concepts

In order to form a baseline for the development of a cost model, a concept of operation for the PKI was developed. There are several ways to conduct the PKI activities which include such functions as generating and certifying keys, signature, and verification. Different operational choices for each PKI activity were discussed and one approach for each activity was recommended. This recommendation was the product of discussions among project members, a synthesis of the information gathered through the interview process, and an understanding of how similar systems operate. For the benefit of PKI implementors, a brief discussion of the alternatives is presented in section 5. The concept of operations selected for the cost model can be found in appendix I.

2.1.5 Cost Model and Analysis

The cost model was developed as an aid to PKI implementors. This model includes only the quantitative impacts that can be ascribed to the PKI specifically. Costs associated with instituting and running an electronic transaction environment were specifically excluded as they were deemed prerequisites to any interest in a use of digital signatures. Generation of the model began with the identification of each PKI activity. Each activity was divided into multiple steps, and the resources required for each step were estimated. The model was modularized and parameterized, so that cost estimates for different infrastructures can be done with a single cost model. Finally, the model was used to obtain the cost estimate for the recommended infrastructure operating under the selected concept of operations. The cost model is described in some detail in appendix I. The final cost numbers are presented and explained below in section 6.

2.2 HIGHLIGHTS OF KEY PKI REQUIREMENTS

The following are highlights of the user, technical and legal requirements of the PKI, as well as other observations obtained through the interviews and the analysis of the pertinent standards.

- **Ease of Use** – Certificate infrastructures should not make applications utilizing digital signature capabilities more difficult to use. In order to support the ease of use requirement, the infrastructure must provide a uniform way to obtain certificates in spite of the possible differences in certificate management policies employed by different segments of the infrastructure, i. e., established by different PCAs.

- **User Authentication** – To assure proper linkage of a public key with a specific user, the identity of that user must be authenticated. User authentication is usually conducted by the CA during the key certification process. User identity authentication is at least as thorough as specified by the applicable PCA policies.

- **Certification Policies**–If the existence of different certification policies is allowed, certification policies for both individual users and organization users must be clearly articulated. In addition, mechanisms must be provided to enable each user to be aware of the policies governing any certificate that he may encounter. In particular, a user should be able to establish how carefully and thoroughly the CA authenticated owner identity of the public key before certifying the association between the user and the key.

- **Trusted Certificate Authority** – Digital signatures are used for sender authentication, non-repudiation and message integrity purposes. In order for a user to trust these security services the user needs to be assured that the public key used to verify a signature is actually the key of the person who signed the transaction. This means that certificates should be generated by and obtained from trusted sources. This implies that mechanisms are needed to prevent any user from creating false certificates which he signs with his regular private key. Even though his signature can be verified using his properly certified public key, the false certificates must not be accepted as legitimate. Then the pretender cannot create signatures that will be accepted because they are verified using keys obtained from his false certificates. Since the CA performs user authentication at key certification time and is responsible for keeping the user's name and public key associated, each CA must be a trusted entity – at least to the extent defined in the pertinent, published PCA policies. This implies the provision of some security protection for each CA, specifically the private key of the CA, so that the CA cannot be modified or impersonated. Certification policies can specify the security measures a particular CA undertakes. Users must determine whether the CA is sufficiently trustworthy for their applications. The basic trust rests in the certification policies and security mechanisms established for the infrastructure.

- **User Affiliation** – To have a CA certify a public key, a user must provide a unique name in addition to the public key which is to be certified. His unique name usually contains his organizational affiliation. It is possible, however, that some private citizens may wish to have their keys certified independently of any organization. Therefore, provisions for certifying private citizens must also be made.

- **Privacy of User's Identity** – Some users may wish to remain anonymous but still register with a CA. This may require the establishment of certification agencies that would register users requesting nondisclosure of their identification information. Alternatively, each PCA may include or exclude anonymous certificates through its policies.

- **Multiple Certificates** – There are situations where a user may have several certificates, each issued by a different CA. This situation may occur if a user belongs to more than one organization and needs a certificate from each organization or if a user has a certificate as an employee and another certificate as a residential user. If the naming convention includes a user's organizational affiliation in his unique name, then a user can have several unique names with a different certificate associated with each. Multiple certificates assigned to a single unique name may be used to simplify recovery from CA private key compromise. The infrastructure may need to handle multiple certificates for a single user.

- **Certificate Revocation Lists** – When a private key is known to be compromised or even when its compromise is only suspected, it must be replaced immediately. The certificate containing the associated public key must be revoked immediately. To inform users of such a compromised key, thus allowing them to identify and reject possibly fraudulent transactions, the certificate is placed on a Certificate Revocation List (CRL). Placing a certificate on a CRL can also be used to announce the severing of a relationship between a signer and the organization with which he was once associated.

- **Services of CA** – CAs will need to certify pubic keys, create certificates, distribute certificates, generate CRLs, and distribute CRLs. Distribution of certificates and of CRLs will be accomplished by depositing them with a generally available directory service.

- **Security and Legal Efficacy** – There is an inherent linkage between security and legal efficacy. The security of electronic messages and records is not only a business requirement, but also is an underlying legal requirement. This linkage determines what is sufficiently secure by considering what presumptions apply to a particular message's or document's purpose(s) and by considering the risks it confronts. Legal requirements should clarify reasonable security procedures without sacrificing needed flexibility. The question is not whether "to have security" or "not to have security," Rather it is a question of the strength of the security mechanisms implemented for the degree of security offered by the digital signatures. The answer rests squarely on the strength of the infrastructure's security mechanisms.

- **Liability** – The extent of the infrastructure's liability must be founded on a balance between the interest of the government, which would limit it, and of the private sector, which would rather expand it. Consequently, it must be allowable to sue but there must also be a reasonable limit on the extent of the infrastructure's liability. Different levels of liability limitations can be offered. For a price, users might even be allowed to tailor the extent of protection to their needs.

  As an agency of the Federal Government, the infrastructure may be considered to

have sovereign immunity. That would imply that the infrastructure and its managers cannot be sued for any losses resulting from their actions or from their inaction. While such status may be attractive, it undermines the usefulness of the certification infrastructure. Without reasonable assurances that potential losses due to malfeasance will be recoverable, a typical non-government user will shy away from relying on the infrastructure. Any set of laws and regulations must strike a balance between protection of the government from excessive claims and blocking users from any chance of reimbursement. The following bullets summarize what may be considered reasonable limits on the extent of liability to which a CA at any level and ultimately the PKI as a whole should be exposed.

- A CA has no liability associated with the loss of the private keys of its children or with their generating weak private keys.

- A key generation facility has no liability associated with the compromise of the private keys it produces, unless it can be proved that the documented policies and procedures were not followed during the key generation process resulting in a weak private key that is more susceptible to compromise or the actual revelation of a private key.

- A key generation facility has limited liability for the compromise of a private key during the key distribution process, if its documented policies and procedures are not followed resulting in the revelation of the private key.

- A CA has no liability associated with forged signatures, unless the forgery results because the documented policies and procedures of the CA are not followed.

- A CA has no liability associated with the wrongful binding of an individual's identity with an associated public key, unless it can be proved that the documented policies and procedures for identification and authentication are not followed.

- A CA has limited liability for not revoking certificates according to its revocation policy.

- A CA has limited liability for revoking a certificate for a reason not specified in its revocation policy.

- A CA has limited liability if, having followed the published policies and procedures, a certificate in the database is modified or deleted.

• **Liability Policy** – The extent of the liability in the above situations is conceivably a part of the PCA policy under which the CA or key generation

facility operate. The policy must distinguish between direct liability on the one hand and indirect and consequential damages on the other.

SECTION 3

ARCHITECTURAL ALTERNATIVES FOR THE INFRASTRUCTURE

This section presents architectural alternatives for the certificate management infrastructure. It examines structures that will allow users to establish chains of trust which contain more than one key but which, in most cases, are no more than a few certificates in length. We begin with a brief look at the functions and responsibilities of the CAs and of the PCAs. This is followed by a discussion of how the CAs might be interconnected to permit the establishment of certification paths plus a brief look at the advantages and disadvantages for the various options.

3.1 POLICY AND CERTIFICATION ALTERNATIVES

Trust is based on three factors: appropriate policies at all points in the infrastructure; careful supervision of the application of those policies; and reliable management of programs and resources in support of those policies. It is envisioned that the supervision and management functions will be the responsibility of the staff. Policy will be overseen by personnel who may or may not have a CA associated with them. It should be noted that, to this point, the name "certification authority" has been used in a generic way to indicate a management entity at any level or position in the infrastructure. The next few paragraphs introduce specific names for CAs, names that depend on the role each CA is playing.

3.1.1 The Policy Approval Authority

Associated with the entire PKI is a policy establishing authority. This authority will create the overall guidelines that all users, associations of users, CAs, and subordinate policy making authorities must follow. It will establish the overall infrastructure security policy.

The Policy Approval Authority (PAA) will also have the responsibility of supervising other policy making authorities. It will approve policies established on behalf of subclasses of users or of communities of interest. It will allow these policies to extend its own but not to detract from them. Thus, it is called the Policy Approval Authority. The PAA may or may not be required to certify the public keys of the lower policy bodies.

3.1.2 Policy Certification Authority

Policy details that expand or extend the overall PAA policies will be created by Policy Certification Authorities (PCAs). This is the commonly accepted name for these entities. We feel this name implies that PCAs certify policies rather than public keys and would prefer the name "Policy Creation Authorities" because it emphasizes the PCA's role in the promulgation of policy. However, we bow to the popular usage and call the authorities Policy Certification Authorities. Each PCA will establish policy for a single organization or for a single Community of Interest (COI). Appendix H contains a sample PCA policy. The policy will specify such details as who will create keys, in what range of sizes the moduli

may lie, what size moduli will be used to sign certificates, how long will certificates be valid, and how will CRLs be handled. It is expected that each PCA will be required to certify the public keys of the lower certificate issuing authorities although this is not obvious *a priori*.

3.1.3 Certification Authorities

The PKI will include many CAs with little or no policy making responsibilities. They will be expected to certify the public keys of users or of other CAs in a manner consistent with the cognizant PCA and with the PAA policies. They may, conceivably, make the requirements of these policies more stringent. The CA will ensure that all key parameters are in the range specified by the PCA. Thus, the CA either creates key pairs using a modulus of a size that satisfies the PCA regulations or it examines user generated keys to ascertain that they satisfy the same range requirements.

The generalized CAs mentioned earlier in this report are the certification authorities at all levels. The majority are plain CAs. A few are CAs that are associated with PCAs. Each of these will certify public keys for the CAs under them. There can also be CA functions at the PAA level for certifying PCA keys.

3.2 TREES AND CROSS CERTIFICATION

Users can physically exchange public keys in a face to face meeting. Then, each is assured of the integrity of the key he has received. However, there are far too many users with whom an individual may need to communicate. These users may not be located nearby and he may not be previously acquainted with many of them. Thus, each user must rely on the PKI to supply him with the public keys he needs. He gets them out of the associated certificates. There are several different ways in which CAs and PCAs can be interconnected to afford him access to the certificates he needs to build a certification path he can trust.

A user's certificate is signed by a CA. That CA creates certificates for many other users. The CA becomes the root of a small tree structure; the users are the leaves. Figure 3-1 shows such a small tree. User 1 can obtain User 2's certificate from the directory in which the User 2 is listed. User 1 is given the CA's public key when the CA creates a certificate for User 1's public key. User 1 uses the CA's public key to verify the User 2 certificate.

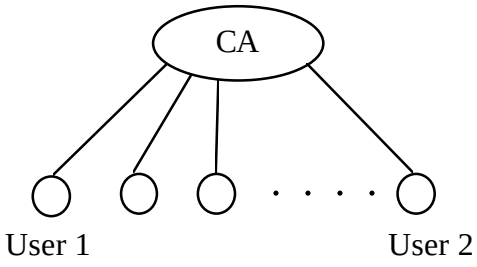


Figure 3-1. Users with a Common CA

Mangers of two CAs can meet and exchange CA public keys. Each will create a certificate binding the other CA's unique name to its public key and sign the certificate with its own private key. This process is called cross certification. Consider figure 3-2. The arrows indicate that CA1 and CA2 have cross certified. If the two CAs function under the same PCA and hence follow the same policies, cross certification is fairly straightforward but done for convenience only. Should the CAs fall under differing policies from different PCAs, cross certification may require more care or may not be permissible altogether.

When User 1 receives User 2's certificate, it is signed with CA2's key. User 1 obtains the CA1-signed certificate for CA2[1]. User 1 has CA1's public key and can verify the signature on the cross certification certificate. He can now trust the copy of CA2's public key contained in that certificate and can use it to verify the signature on the CA2-signed certificate for User 2.

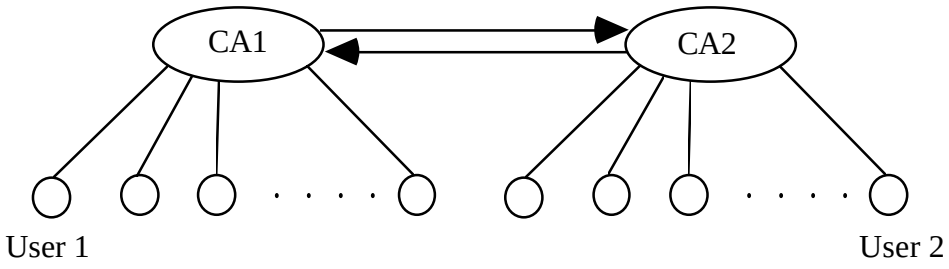


Figure 3-2. Users with Different Cross Certified CAs

---

[1] We do not discuss here where users obtain certificatres. It is expected that most certificates will be obtained from a directory service. However, some applications may call for sending some or all needed certificates with the message or document.

It is expected that the number of CAs will be quite large. In that case, it is still impractical for the CAs of every potential pair of communicating users to cross certify each other. Other difficulties with the cross-certification of CAs appears in subsection 3.3.1 below. There it becomes apparent that CAs should be certified by PCAs only. This structure is shown in figure 3-3.

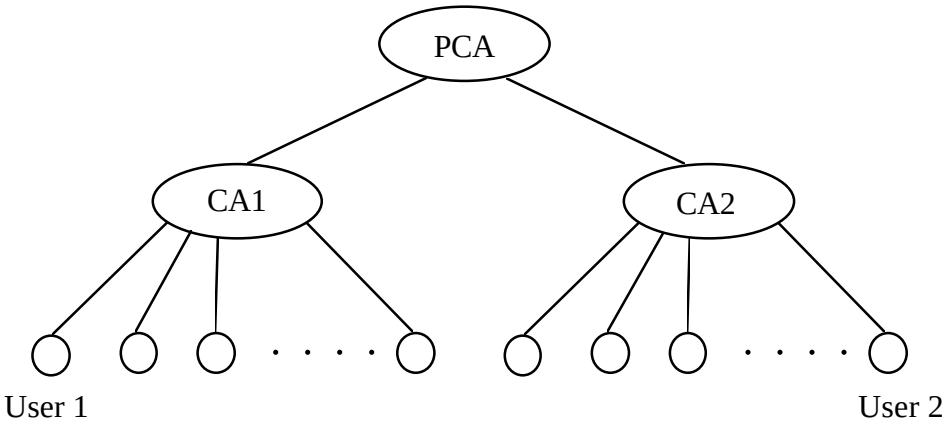


Figure 3-3. Users with Different CAs under a Single PCA

In this scenario, User 1 obtains CA2's public key from a certificate signed by the PCA rather than from a cross certification certificate signed by CA1. He verifies the signature on this certificate by using the PCA's public key. This key he receives from CA1 when his own key is certified there or on request as he needs it.

Similar architectural choices exist when User 1 and User 2 come under different PCAs. Suppose User 1 follows the policies of PCA1 and User 2 follows PCA2. The certificate chain that User 1 must verify before he trusts the key in User 2's certificate includes one for CA2 signed by PCA2. He can obtain that key either if PCA1 and PCA2 have cross certified or if both are certified by a single CA at the level of the PAA[2]. These two architectural alternatives are shown in figures 3-4 and 3-5.

---

[2] This report ultimately recommends the second alternative of establishing a certification authority at the PAA to certify all PCAs. However, the possibility of cross certified PCAs cannot be ignored. Until the PKI is fully operational, initial PCAs may have to cross certify each other.

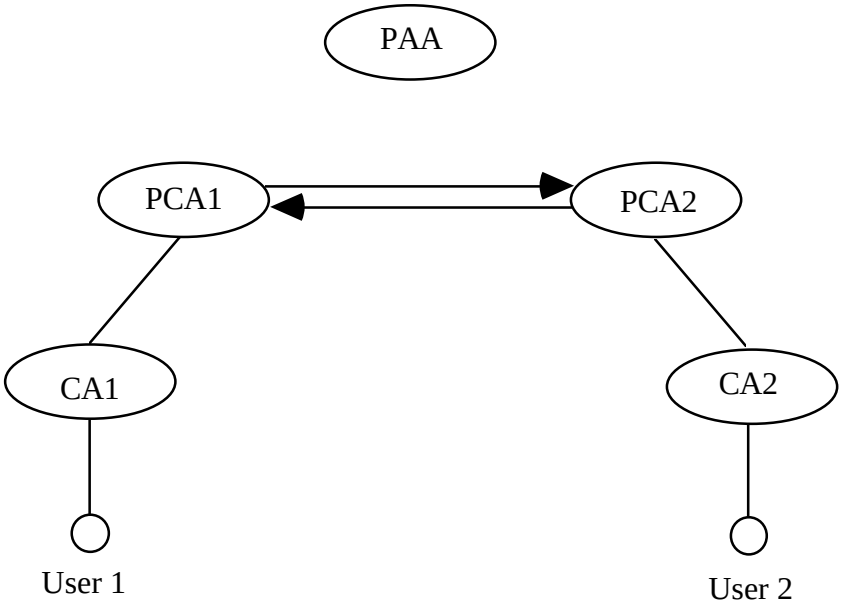


Figure 3-4. Users with Different Cross Certified PCAs

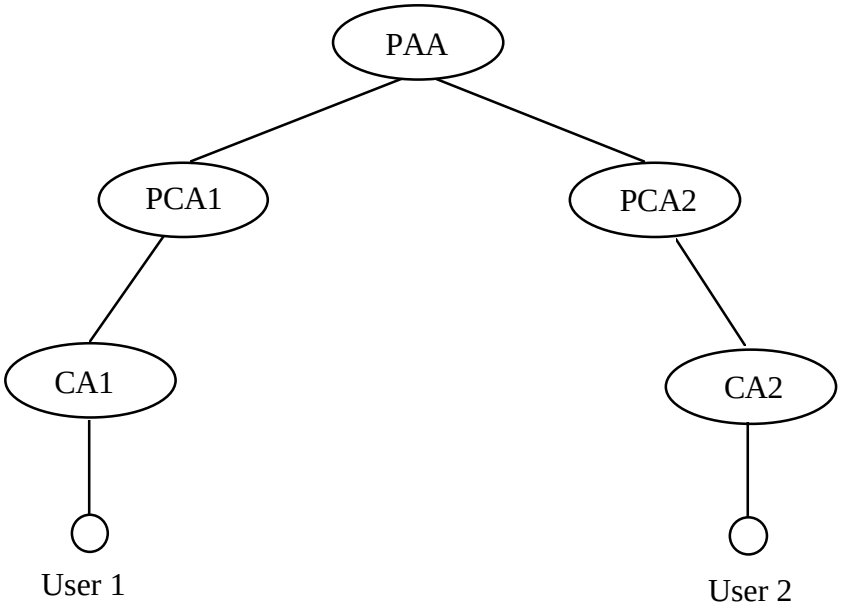


Figure 3-5. Users with Different PCAs under a PAA

3.3 ARCHITECTURE ALTERNATIVES

It is apparent from the previous discussion that the PKI has one of two general forms. It can be a single tree with a root at the PAA. Alternatively, it can consist of only the upper branches of that tree, removed at the CA level, at the PCA, or at some combination of these levels. The branches, which themselves have the structure of trees whose roots are CAs or PCAs, are then connected in a non-hierarchical way using cross certification. We examine several non-hierarchical architectures and the completely hierarchical tree.

3.3.1 Cross Certified CAs

At first glance, cross certification of CAs seems to allow the greatest flexibility and the shortest certification paths. A CA may serve a single government subdepartment or commercial enterprise. It cross certifies with all other similar CAs with whom its users conventionally transact their business. A comparison of the PCA policies of each CA can be undertaken at the time of cross certification to assure that interaction between their respective users violates neither policy. Furthermore, the certification paths are extremely short. They require only one certificate signature verification in addition to the user certificate verification that must always be done.

These advantages notwithstanding, cross certification of CAs does not yield a viable architecture for the PKI. The number of cross certifications required is far too great. Since each requires an agreement that must be negotiated "out-of-band," in total, they represent a significant hardship. Furthermore, the number of certificates they would produce is far too large. Managing them, especially if the associated private keys can be compromised or when it is time to rekey, is a burden which need not be tolerated. With cross-certified CAs, navigating certification paths can become a problem also. The verification software may have trouble deciding which CA certificate it needs in a given situation – that signed by the CA's PCA or that signed by some other CA. Additionally, CA cross certification alone does not allow for the occasional verification of a signature generated by a user outside the range of the existing cross certifications. Finally, difficulties in reconciling the differing certification policies may far outweigh any small benefit in shorter certification paths. In all, it seems impractical for the CAs of pairs of potentially communicating users to cross certify each other. All CAs should be certified by PCAs only.

3.3.2 Cross Certified PCAs

Cross certification of PCAs has most of the advantages of cross certification of CAs but few of the disadvantages. There is still some flexibility in determining who cross certifies with whom. Among the general public, it is normally the responsibility of an individual user to look up a PCA's policy and to judge whether he wishes to trust a certificate that was created under that policy. Government agencies may have their own view on this matter. It is possible for a PCA not to cross certify with another if the PCA policies are incompatible. The total number of cross certifications is greatly reduced. The need for "out-of-band" agreements is no longer as great a burden. In fact, it may be an advantage. If PCAs are associated with single, large government departments or with entire state governments or

with entire regulated industries, such agreements can specify all the legal and operational conditions of interaction and cooperation in a manner not unlike EDI Trading Partner Agreements. Policy issues can be resolved by the very people who create the policies. Furthermore, with a network of cross certified PCAs, the chances of a user receiving a signed document and not being able to construct a certification path are greatly diminished.

There is another, perhaps unexpected but significant, advantage to achieving nationwide connectivity through cross certification of PCAs. M. S. Baum points out in his report on liability and policy issues of a federal certification authority [4] that there are some constitutional issues that must be addressed. Briefly, he is apprehensive that harm potentially arising from a PAA abuse of power is potentially so significant that he urges further consideration of this issue. He sees possible violations of the separation of the three branches of the Federal Government, legislative, executive and judicial. He is further concerned about threats to federalism in which the PAA issues a CRL for a state government entity, thereby impairing or infringing upon the powers of that state's government.

The single, significant disadvantage to a PKI based on the cross certification of PCAs concerns itself not so much with the PKI but with how the PKI will interoperate with other certificate infrastructures on the national and the global scene. Requiring each PKI PCA to cross certify with similar entities in national, in foreign or in international infrastructures can, conceivably, become a burden. While certification of PCA by a PCA in another hierarchy leaves any decision concerning compatibility of policies squarely where the policy was created – the PCA, some hierarchies specifically disallow cross certification. It is felt that cross-certification disrupts the hierarchical naming convention, causes the possibility of circular certification paths and makes certificate revocation even more difficult than it normally is. Additionally, if PCAs are allowed to cross certify with the root of a tree for those countries and international organizations that have created a complete hierarchy, problems in transversal of the hierarchical, national, rooted trees can occur. It seems advisable to root all national subtrees at the same level and to cross-certify their roots. For the PKI, this means there is a PAA which certifies all PCAs and there is no PCA cross-certification.

3.3.3 A Root PAA

Building the PKI with a root certification authority associated with the PAA assures connectivity between any pair of users. For the most part, that is an advantage. Moreover, it leaves all decisions about the compatibility of policies in the hands of each end user, where it actually belongs. With a CA at the PAA, it is possible for any user to obtain a certificate chain for any other user. This chain is unique and verifiable from anywhere within the PKI. However, the user is responsible for determining the degree of trust which he can place in each certificate in that chain.

It is worth noting that a full certification chain with a root PAA is not longer or shorter than a similar chain when PCAs cross certify. The former has the PAA's certification of PCA2's key while the latter has PCA1's certification of the same key. It is assumed that

User 1 has equal access to the PAA public key in one case and PCA1's public key in the other.

3.4 ARCHITECTURE RECOMMENDATION

The above discussion suggests that there are only two viable alternatives for the architecture of the PKI. They are the purely hierarchical organization with a certifying authority at a PAA root and a non-hierarchical one with cross certification of PCAs. We recommend the former, purely hierarchical architecture. Considerations of assured existence of certification paths, absence of certification loops, interoperation with all other infrastructures, etc. lead us to this recommendation.

To allow for the gradual development and deployment of the PKI, however, we suggest the immediate establishment of the PAA, even without an associated certification authority. PCAs are temporarily allowed to cross certify, if they desire. Thus, any of the first established PCAs may elect whether its prototype PKI subtree will be completely open to all, some or none of the other early PCA and their prototype subtrees.

SECTION 4

IMPLEMENTATION ALTERNATIVES FOR THE INFRASTRUCTURE

This section presents implementation alternatives that are based on the certificate management infrastructure architectures that were recommended in the previous section. The objective of the implementation alternatives study is to develop an alternative that is sufficiently flexible to allow the infrastructure to serve a wide range of operations and to interoperate with other certificate management infrastructures (e.g., PEM certificate management). Although the implementation alternatives presented in this section are primarily for Federal Government activities, similar considerations apply when determining the infrastructure outside the Federal Government.

4.1 FUNCTIONS PERFORMED BY CA ENTITIES

This section presents the functions that are carried out by CA entities at all levels. It describes what the PAA, PCAs, and CAs perform. It also describes the role of an Organization Registration Authority (ORA).

4.1.1 PAA Functions

The PAA is the root of a national certificate management infrastructure. The public key of the PAA is the keystone in providing connectivity within the infrastructure. This key will be known to all entities in the PKI. It can be hand delivered to each user, to each CA, and to other entities at the time of certification of the user's or entity's public key. Further discussion of the distribution of this root key to all entities in the infrastructure can be found in section 5 which discusses operational concepts for the PKI.

The PAA performs the following functions:

- Publishes the PAA's public key
- Sets the general policies and procedures that all entities (PCAs, CAs, ORAs, and end-users) of the infrastructure must follow.
- Sets the policies and procedures that determine when and how a new PCA can join the PKI.
- Carries out identification and authentication of each of its subordinate PCAs and of national, international or multinational infrastructure roots it deems appropriate to cross certify.
- Generates certificates of subordinate PCAs and of national, international or multinational infrastructure roots it deems appropriate to cross certify.

- Publishes identification and locality information of subordinate PCAs (e.g., directory name, email address, postal address, phone number, fax number, etc.).

- Receives and publishes policies of all subordinate PCAs'.

- Specifies information required from subordinate PCAs for a request of the revocation of the PCA's certificate.

- Receives and authenticates revocation requests concerning certificates it has generated.

- Generates CRLs for all the certificates it has issued.

- Archives certificates, CRLs, audit files, and PCAs' policies.

- Deposits the certificates and the CRLs it generates at the directory.

4.1.2 PCA Functions

PCAs form the second tier in the infrastructure. They have both policy and certification responsibilities. It is important to note that, when a certificate issued by a PCA is verified, the receiver must decide whether the security policy associated with that particular PCA is the one he or she should accept. All PCA security policies are published and can be stored on an end-user's local database. The PCA policy can be obtained in similar manner as for the certificates, either via email or from the directory. If a PCA's policy is not already on the end-user's local database, the end-user can request that policy. The PCA then sends its policy to the end-user.[3]

Each PCA performs the following functions:

- Publishes its identification and locality information (e.g., directory name, email address, postal address, phone number, fax number, etc.).

- Publishes the identification and locality information of the CAs it has certified.

- Publishes who it plans to serve.

- Publishes, by making them available at the PAA or at appropriate directory elements, its security policy and procedures which specify the following:
  - Who generates key variables *p, q, g, x,* and *y* (cf. appendix B).
  - The ranges of allowed sizes of *p* for itself, its CAs, and end-users.

---

[3] It has been suggested that automatic processing of signature verification is enhanced by encoding the policy under which each certificate is generated directly within that certificate [6].

  - Identification and authentication requirements for the PCA, CAs, ORAs, and end-users.
  - Security controls at the PCA and CA systems that generate certificates and CRLs.
  - Security controls at ORA systems.
  - Security controls for every user's private key.
  - The frequency of CRL issuance.
  - Constraints it imposes on naming scheme (e.g., name subordination).
  - Audit procedures (e.g., scheduled and impromptu audits).

- Carries out identification and authentication of each of its subordinates.

- Generates and manages certificates of subordinate CAs.

- Delivers its own public key and that of PAA to its subordinates.

- Specifies procedures and information required to validate certificate revocation requests.

- Receives and authenticates revocation requests concerning certificates it has generated.

- Generates CRLs for all the certificates it has issued.

- Archives certificates, CRLs, audit files, and its signed policy if changed.

- Delivers the certificates and the CRLs it generates to the directory.

4.1.3 CA Functions

CAs form the next level below the PCAs. A CA may have any combination of users and ORAs whom it certifies. A CA performs the following functions:

- Publishes local CA augmentations of the PCA policy.

- Carries out identification and authentication of each of its subordinates.

- Generates and manages certificates of subordinates.

- Delivers its own public key and its ancestors' public keys.

- Verifies ORA certification requests.

- Returns certificate creation confirmations or new certificates to requesting ORA.

- Receives and authenticates revocation requests concerning certificates it has generated.

- Generates CRLs for the all the certificates it has issued.

- Archives certificates, CRLs and audit files.

- Delivers the certificates and the CRLs it generates to the directory.

4.1.4 ORA Functions

An ORA is an entity whose sole task is to help a user who is physically far from the user's CA to register with that CA and to obtain a public key certificate. The ORA performs identification and authentication of the end-user and then vouches for his or her authenticity in a signed message that it sends to the CA. If the user has created his or her own key pair, the ORA includes the new public key in its signed message. Either a signed confirmation of the creation of a certificate or the certificate itself is returned from the CA. In either case, the ORA must verify the signature to be sure it was really created by the CA. It must also examine the confirmation or the certificate to verify it is the response to its request. Further, if the certificate is returned, the ORA must deliver it to the user for whom it was created. This may require that the ORA load the certificate onto a smart card or floppy disk. The ORA has no authority to generate certificates on its own. It merely performs the identification and authentication function on behalf of a CA and delivers the CA-generated certificate to the end-user.

An ORA can be instrumental in the revocation of a certificate when such action is necessary. A lost token, suspected stolen private key, or severed relationship must be reported "out of band" to the CA that generated the associated certificate. Severed relations such as when an employee leaves his employer also need to be reported to the CA. The CA must authenticate the validity of any revocation reports. This can be accomplished by the user appearing in person at his local ORA to report the problem. The ORA uses a signed message to inform the CA of the need to revoke the certificate and to issue a new one.

An ORA performs the following functions:

- Carries out identification and authentication of users.

- Sends user identification information and, possibly, his public key to the CA in a signed message.

- Receives and verifies certificate creation confirmations or new certificates from the CA.

- Delivers the CA's public key and its ancestors' public keys as well as the certificate, if returned, to the user.

- Receives certificate revocation request, verifies the validity of the request, and if valid, sends the request to the CA.

4.2 THE IMPLEMENTATION ALTERNATIVES

The implementation alternatives are derived from two factors: (1) who sets the security policies, and (2) how are the users organized relative to those policies. Subdividing the users can be considered based upon:

- COI
- Organization
- Assurance level
- Hybrid of the above

The COI alternative is derived from the idea that users should be grouped by functions that they perform. The other users with whom they communicate most frequently in carrying out their daily tasks should be close to them in the infrastructure even if they are physically some distance away. This is a matter of convenience. Users are expected to cache those certificates they use most frequently, including some certificates along the certification paths encountered. The COI alternative insures that most paths will have many certificates in common and hence will minimize the number of certificates to be cached. The organizational approach is built upon the principle that the infrastructure should follow the current federal agency management structure. Assurance levels certificate management is founded on the realization that, in practice, only a small number of perhaps three or four security policies may be sufficient to satisfy all federal agency operations and that users should be organized according to their assurance needs. Table 4-1 shows the two deciding factors in play for each alternative user subdivision.

Table 4-1. Summary of Policy-Setting Body and Users for Each Alternative

Alternatives	Who sets the Security Policy	User Groupings
COI	Each COI sets its policy, by committee decision; there can be as many policies as there are COIs	Each community is defined by the commonality of the functions users perform
Organizational	Each organization's management sets its policy; there can be as many policies as there are different organizations	Employees of a single organization
Assurance Levels	The number of policies may only be a few ranging from basic to stringent assurance; a committee will specify all levels of policies	Anyone who needs a certain degree of assurance

4.2.1 COI Alternative

The COI approach organizes users according to the tasks that they perform most frequently. They should be subdivided according to how and with whom they communicate most frequently. Since federal agencies perform many different functions, there will naturally be multiple COIs in the infrastructure. The users in each COI are those who need to perform functions related to the specific community. The members of a community are those government workers who regularly perform the functions the community was established to support. It may also include corporate employees and private citizens. Each COI is further divided along organizational lines. For example, within the Law Enforcement COI there is a CA for the FBI within the Department of Justice and another for the Bureau of Alcohol, Tobacco, and Firearms (ATF) in the Department of Treasury.

The security policy of a COI is determined by the community itself. For example, one can imagine a community for the national law enforcement. The members of the law enforcement community would include some subset, but not all, of the employees from the FBI, ATF, IRS, Drug Enforcement Agency (DEA), Immigration and Naturalization Service (INS), Secret Service, Customs, etc. It might also include state and local law enforcement officers. Its policy might be set by the FBI or by representatives from all of these agencies. Figure 4-1 illustrates a case where there are two COIs under the PAA.

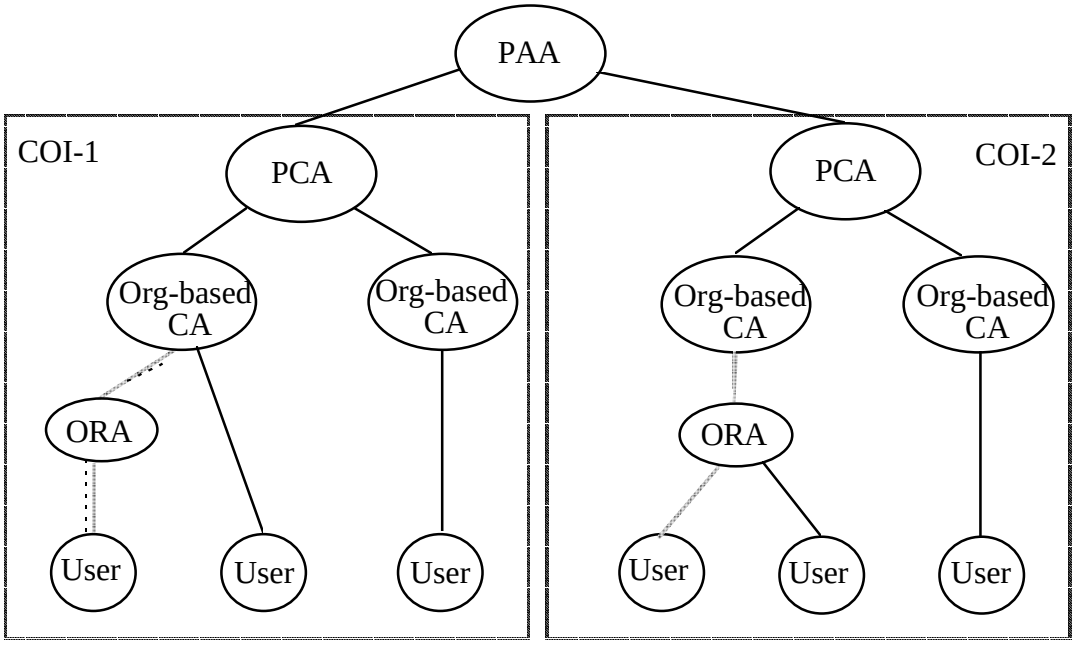


Figure 4-1. Communities of Interest Alternative

4.2.2 Organizational Alternative

The organizational alternative for the Federal Government parallels the Federal Government organization hierarchy. The security policy will be set by each large government department or agency. CAs will be located at the agencies and bureaus under the departments. Figure 4-2 illustrates the organizational concept.

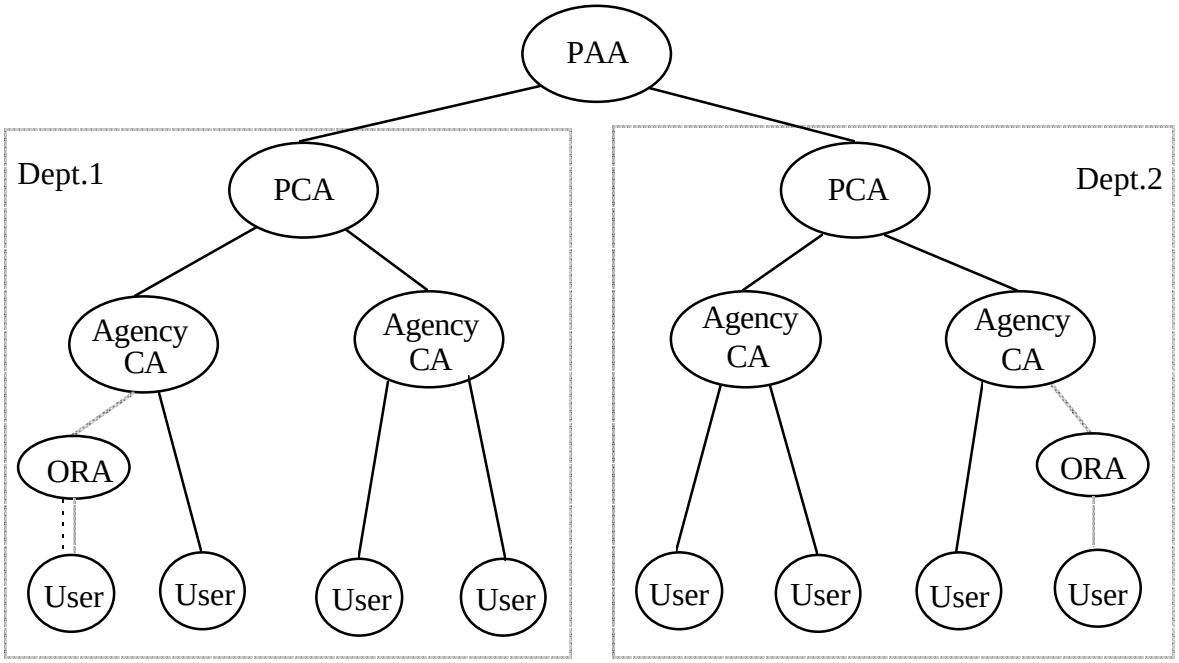


Figure 4-2. Organizational Alternative

4.2.3 Assurance Level Alternative

The assurance level approach is based upon the idea that a small number, perhaps three or four, of security policies may be sufficient to satisfy all federal agency requirements. Each government organization can adopt the policy that best fits its requirements. This alternative does not restrict users either by community of interest or by organizational affiliation. Regardless of functions performed or organizational affiliation, a PCA issues certificates to CAs that require the same level of assurance. Figures 4-3 shows several possible assurance levels: stringent, basic and persona. This alternative can play an important role in providing a bridge for non-federal organizations such as state and local governments, private corporations, and individuals. These entities may only be interested in general service at a certain level of assurance and not in specific communities or agencies. For example, if an individual who is not a part of the law enforcement COI needs to transfer data to someone in that COI, the individual can get a certificate from a CA whose security policy is at least as stringent as the law enforcement COI.

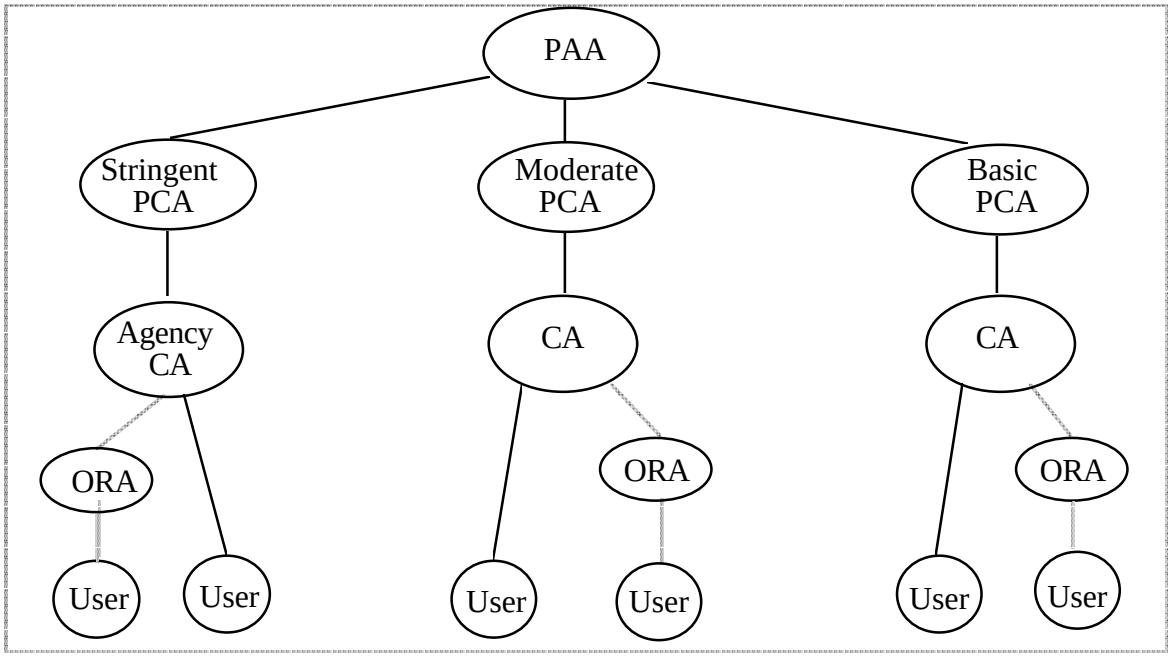


Figure 4-3. Assurance Level Alternative

While it is expected that there will only be a few different assurance level policies in total, it may be difficult to identify the exact number at first. It is conceivable that one security policy will satisfy most of the federal operations. The large number of users subscribing to a single policy may dictate the incorporation of several PCAs that operate under that policy. At what point there needs to be another assurance level policy and the degree of difference between the security features of two assurance policies are questions beyond the scope of this study.

4.2.4 Hybrid Alternative

It is important that the national infrastructure be flexible and accommodate a wide range of organizations. It must also be scalable to accommodate the addition of many new users and organizations. Each alternative has its strong points, but one alternative alone may not satisfy all infrastructure needs. An infrastructure that utilizes only one alternative will likely serve one group or one type of activity well, but may not serve another group of people or another type of activity well. The infrastructure, therefore, will reflect the real world better if the it allows segments which follow each of the three implementation alternatives as shown in figure 4-4.

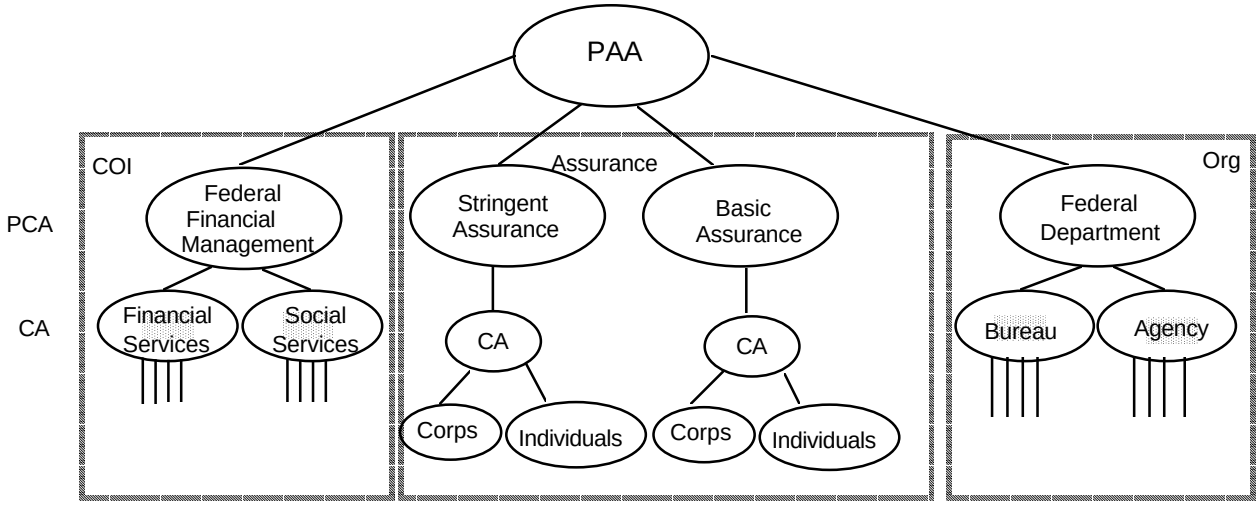


Figure 4-4. Sample Hybrid of Certificate Management Infrastructure

4.3 RECOMMENDATION

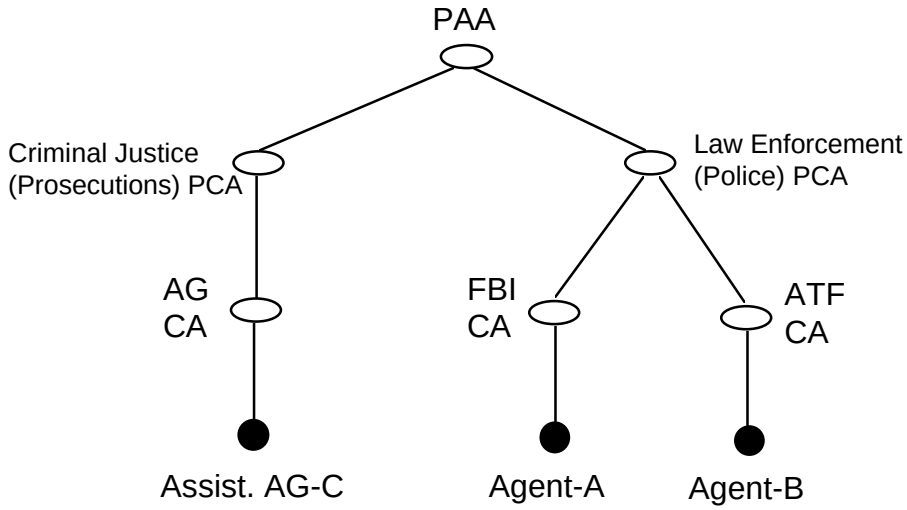
4.3.1 Comparison of Implementation Alternatives

It is difficult, purely from an operational viewpoint, to choose between the three basic implementation alternatives. There are functions within the Federal Government that fit naturally into a COI approach. Yet, the individuals involved in performing these functions daily engage in other tasks that are better undertaken with an organizational approach. There are also functions where the reverse is true. They are more naturally executed in an organizational configuration, but the individuals involved perform many duties that fit more readily into a COI approach.

For example, it would make sense if there is a community composed of employees from many different federal agencies that transfer a large volume of information within the community to form a COI with its own PCA. At the same time, these same users do a fair amount of communication with other users in their own organization. They write memoranda, submit time reports, file travel requests, and initiate purchase requests. If, on the other hand, there is an organization where information transfer occurs mainly within the organization, then the organizational alternative would serve this group better than any other alternatives. But, by the same token, there may be employees among these users who regularly transact much business within a certain wide-spread community.

To illustrate why the COI alternative would be advisable for a group whose members are from different organizations, yet need to communicate often, consider the following example illustrated in figure 4-5. Suppose that FBI Agent-A sends some information to ATF Agent-B. All Agent-B needs in order to verify Agent-A's signature is his/her certificate, which was issued by the FBI CA. In the law enforcement (i. e., police) COI, as shown, both agents are under the same law enforcement PCA. Thus, Agent-B does not need the PCA's

certificate because, it is assumed, he already has the PCA's public key. On the other hand, in the organizational approach of figure 4-6, the FBI and the ATF are under the Department of Justice and the Department of Treasury, respectively. Their PCAs are different. Therefore, Agent-B will need the entire chain of certificates, including the Department of Justice's PCA certificate that was issued by the PAA.



AG: Attorney General

Figure 4-5. Example of COI Alternative

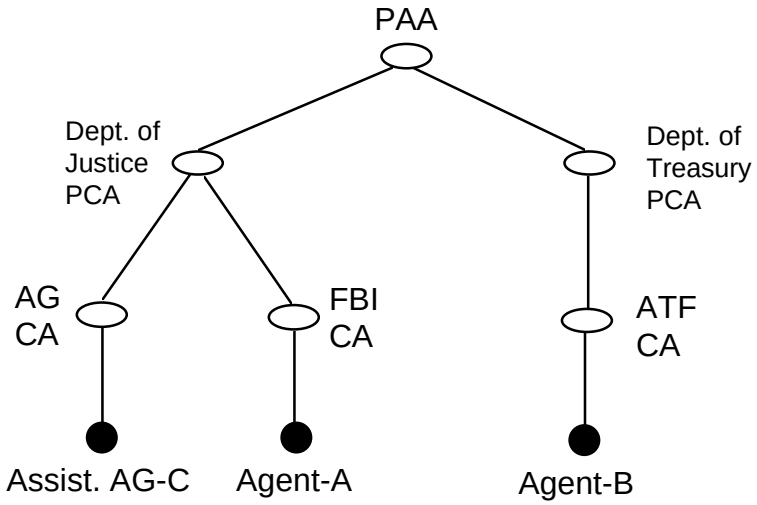


Figure 4-6. Example of Organizational Alternative

At the same time, one can envision a situation where the relative merits of the two approaches are reversed. Suppose FBI Agent-A must send evidence to Assistant Attorney General-C. In the COI approach of figure 4-5, the assistant attorney general is in the Criminal Justice (US court system) COI. He must obtain the certificates along the entire certification chain from Agent-A to the PAA's certificate for the Law Enforcement PCA. On the other hand, in the organizational alternative shown in figure 4-6, the chain is one link shorter. It needs only two certificates, the FBI CA's certificate and Agent-A's certificate.

These two examples show there is no technical reason for preferring either the COI alternative or the organizational alternative to the other. With users caching most frequently needed certificates, differences between the approaches are further reduced. Additionally, it is worth noting that, in actual practice, the security policies chosen by PCAs in either the organizational or the COI approach will probably consist of no more than a selection of an assurance level. The choice of one of the alternatives over the others, especially for the Federal Government portion of the PKI, must, therefore, rest on other considerations.

A decision on which approach to employ in structuring the Federal Government's portion of the national infrastructure requires the consideration of a number of points. They include the following:

System Reliability. Are there differences in the reliability of the various alternatives?

System Scalability. How easily scalable is each PKI alternative?

Flexibility and Ease of Use. Is one alternative easier to use than another? Does one alternative offer more flexibility than the others in how it is used? Does their flexibility to expand allow for flexibility to incorporate new technology?

Trust. Do the common users, the CAs or the PCAs experience qualitative differences between the various implementation alternatives in the level of trust that can be placed on the system?

Interoperation. How readily can each PKI alternative interoperate with another infrastructure and another signature algorithm?

Implementation Timeframe. How long does it take to get a new user, a new CA, or a new PCA "up and running?"

Management Structure. Does a structure already exist to manage each alternative or must that structure be created? Are budget mechanisms in place to fund the establishment and the operation of the alternative?

Liability. Is user liability or CA liability greater in one alternative than in another? Are the lines of responsibility more clearly defined in one alternative than in another?

The discussion of the COI and the organizational alternatives that accompanied figures 4-5 and 4-6 suggests that, in many of these areas of consideration there is very little to distinguish between the alternatives. This is especially true with the inclusion of caching of certificates most frequently used and permitting entities to register with more than one certification authority. Nevertheless, we examine each of the above questions in turn.

4.3.1.1 System Reliability

System reliability involves the availability of certificates to create certification paths. Thus, it depends on a ubiquitous and robust directory service. With no directory service, the CAs take on the role of the directory servers. They maintain a database of certificates and CRLs. They must be available to respond to requests for those certificates and CRLs. More importantly, CA key compromise can be a serious hindrance to robust functioning of a whole segment of the infrastructure. A truly robust system is able to continue to supply the needed certificates in the unlikely event of a CA compromise.

There are several methods to achieve system reliability. Any one or any combination of them can be made sufficient. They include the following. Each CA has a dual CA with a different key. Every entity which is certified by one CA must also be certified by its dual, thereby obtaining two distinct certifications of the binding of its unique name to its public key. Either certificate is sufficient to verify that binding. Each CA deposits its certificates with a different directory server. If one server is unavailable because it is malfunctioning, its communication support is inoperative, or if the first CA's key has been compromised, the needed certification can be obtained from the second server. Of course, one can save the cost of dual CAs by making the second directory server simply a hot backup holding exactly the same information as the primary server. Then, if one directory server malfunctions or is inaccessible because its communications are down, it is possible to switch requests for certificates or for CRLs to the second server. Patently, if the CA's key has been compromised, diverting to the backup CA would not solve the problem. In the absence of a dual or a backup directory server, anyone needing service from that server must wait until it is again accessible.

The only time the unavailability of a CA itself has any effect is when an entity needs to have its public key certified by that CA or needs to report a key compromise or a severed relationship. Certification does not appear to be an extremely time–sensitive event. However, the need for certificate revocation could well be more urgent.

These robustness considerations apply equally to all implementation alternatives. This conclusion rests on the assumption that there are roughly the same number of PCAs and of CAs under each alternative. That implies, under the assurance level approach, that the workload for each level is divided between several PCAs with identical security policies. If, on the other hand, there is only one PCA for all the users at one assurance level, then that PCA is a serious single point of failure.

4.3.1.2 System Scalability

A similar situation applies to the question of scalability. A server is more likely to be overloaded when performing directory services than a CA is in issuing certificates. In fact, certificate generation is, for the most part, an "off-line" task. When the requester appears in person at the CA, no infrastructure network service are required. The completed certificate need only be sent to the appropriate directory server at some convenient time. Even when the certification request comes through an ORA, only one signed message is required. The CA already holds the ORA's public key to verify the signature. The certificate again must be placed with the appropriate directory server and, possibly, sent to the ORA to be loaded onto the user's disk, smart card or other token.

Adding another PCA, adding a new CA when the certification load or the length of the CRL for a given CA becomes burdensome, or adding a very large number of new users appears to have equal impact on the organization and the COI implementation alternatives. It has the same effect in the assurance level alternative as well, again provided that there is more than one PCA for each level of assurance offered. Obviously, the creation of a new government department or agency requires a new PCA and several new CAs under it in the organizational alternative. This may be slightly more straightforward to accomplish than what is required in the COI or assurance levels alternatives.

On the other hand, a new COI is created when enough current users express the need for one or when an existing community decides to adopt the use of digital signatures. When the latter happens, the establishment of a new COI is analogous to the creation of a new department involving the transfer of many employees. It will require the establishment of a new PCA as well as several new CAs. When current users move to a new COI, old relationships may be severed with the rescinding of certificates which are placed on appropriate CRLs. New relationships are established with the issuing of new certificates.

It should be noted that moving of current users' public keys from one CA to another causes some confusion if the PKI adopts the convention that each user's unique name includes that of his certifying CA[4]. Then, when a user moves, his name changes. If the move is undertaken because the user actually changed his affiliation, then a change in his unique name is appropriate. If the move is brought about solely because of PKI expansion, a change of unique name becomes an inconvenience. However, the inconvenience is the same in all organizational alternatives.

---

[4] The entities in the PKI must have unique names. The naming scheme is not discussed in this report but one attractive approach is a naming scheme which parallels the certification hierarchy.

4.3.1.3 Flexibility

There is no apparent difference in the ease with which new technology can be incorporated into any of the three implementation approaches. The provision of distinct, role-based certificates for a single user and his capability to cache certificates of many other users provides some degree of flexibility in the use of the PKI under all implementation approaches.

Flexibility in defining security policies exists where every large scale organizational unit has its own PCA and can fine-tune its policy to suit its exact needs. That occurs in the COI and the organizational approaches but not under assurance levels.

4.3.1.4 Trust

Trust derives from the details of PKI; that is, from the specifics of the architecture, the implementation, the system management, and the concept of operations combined. In none of these areas has this report suggested any alternative that does not support the ultimate goal of trust. In fact, only in the area of system management is it possible to discern a slight advantage for the organizational alternative. Having each PCA in a position to supervise all its CAs directly through existing managerial channels within a single department or independent agency seems to suggest that trust can be placed on a CA-issued certificate in an organizational implementation. New COI PCAs will be established only in response to need. Presumedly, entities belonging to the new COI will trust its operation but it is less clear how much entities beyond that COI will be willing to trust it. This last point applies also to an assurance level approach, where the PCAs and the CAs are essentially service suppliers only. This statement notwithstanding, certificates at each of the various assurance levels may well be trusted to the specified degree. This is because the laxness may not change the assurance level of the certificate any appreciable amount and, in fact, may not exist at all.

4.3.1.5 Interoperation

For a PKI user who signs with DSS, interoperation with other signature systems means interoperation with other signature algorithms and interoperation with other key certification infrastructures. All the currently available, public key signature schemes (DSS, RSA, El Gamal, and Schnorr) are based on modular exponentiation. They all use the same set of fundamental computational routines. Any user who needs to verify signatures produced in some non-PKI signature infrastructure must obtain additional software both for hashing information and for creating actual signatures. He must also have access to the certificates required to verify the certification path in the other infrastructure. To this end, he must hold a trusted copy of the public key of the root of the other infrastructure. He can hold this key because he has it in a non-PKI certificate created by some national or international certifying authority at a level higher than the PKI or one of the PKI certifying authorities above him must be able to supply him that key in a signed PKI certificate.

4.3.1.6 Implementation Timeframe

This issue is tied to the next one, the existence of a management structure. Once the policy authorities have been identified, the CA managers have been appointed and needed funds have been allocated, the time to install the CAs and to write and install their software does not differ from one implementation alternative to the other. Any difference that may arise in the time to get the federal portion of the PKI running is a consequence of the presence or absence of the management structure.

4.3.1.7 Management Structure

By definition, the organization alternative has a management structure already in place. The various government departments, commissions, and agencies already have defined hierarchical structures within them. In most cases, people working in the same lower level organizational unit are collocated. The PKI can easily follow the same organizational structure, assigning all the employees in such a lower level unit to the same ORA or CA. The budgetary vehicle for financing this effort is already in place. However, if some department wishes to organize its workers' certificate infrastructure differently, this can readily be done, too. Each department has its own PCA and can make its own policy on this matter.

For the COI alternative, some decisions must be made before anything can be implemented. How many COIs should there be? Is there one agency that everyone within a COI will accept as the policy maker for the COI PCA or will there be a committee to set policy? If so, who will be on the committee and how long will it take that committee to establish the COI security policy? Who will actually run each COI PCA? Will the policy and certificate management functions of the PCA be budgeted as part of each participating department's regular appropriation or will there be a separate appropriation for the COI? Lower levels of a COI infrastructure may well follow the existing departmental tree structure. Management and budgeting of these levels may be straight forward. It is apparent that, in the area of management structure, the COI alternative lags well behind the organizational one.

If that is so, the assurance level alternative may lag even further behind. There is currently no natural constituency from which to draw the policy setting committee that should establish the security policy for each level of assurance. Nor is there a natural choice for who should run the several PCAs at each assurance level and all the CAs unless it is some existing, easily accessible service organization. The phone companies and the USPS both fit this bill. So does the banking system and any number of other service groups although USPS's loose ties to the government may make it the primary choice for government needs. (For a further discussion of PKI management, see section 4.4 below.)

4.3.1.8 Liability

While issues of liability relating to the PKI are discussed elsewhere [4], it is possible to make one statement here. When PCA policy is established by a department, independent

commission or independent agency for all its employees, it is easy to establish a chain of responsibility. It is a little less simplistic when the policies are set by representatives from another department or from several other departments. Under the organizational alternative, workers within a single department or independent agency will operate under a single set of rules established by that department or agency. The department or agency will have the responsibility to ensure that its security policies are enforced. However, each COI or each assurance level is spread among several departments and agencies. There is no single set of security rules being followed within each department or agency and no person or persons to enforce them. Establishing liability when something goes wrong may be far more difficult.

4.3.2 The Recommended Federal Government Alternative

It is apparent from a consideration of the questions of management structure, implementation timeframe, trust, and liability that the organizational alternative holds some, admittedly slight, advantage over both the COI and the assurance level. The organizational approach allows each department, independent commission, and independent agency to establish its own security policy and its own level of assurance. (In practice, this may well reduce to the assurance level alternative with only a few actual levels. However, an organizational structure allows a certain amount of autonomy.) With certificate caching, none of the benefits of the COI approach need be lost in implementing the PKI along existing organizational lines. Indeed, departments willing to cooperate may create what amounts to their own COI. Thus, for the Federal Governments portion of the PKI, we recommend the following:

- PAA, a national body, be created to establish overall PKI policy, to approve individual PCA policies, and to act as a root for the national certification infrastructure to be created.

- Each federal department implement its own PCA to establish its own policy. PCAs be established for sets of independent commissions and independent agencies.

- Each PCA be certified by the PAA.

- CAs be established by offices and bureaus within large federal departments and independent agencies, as determined by the PCA. ORAs be placed near individual facility security offices, as needed.

- Several assurance levels, with associated PCAs and CAs, be established for use by private corporations and citizens. ORAs be placed near corporate personnel offices, as needed.

- Each user caches (stores) all certificates he or she uses most frequently.

4.3.3 Beyond the Federal Government

Beyond the government sector of the PKI, it is expected that all three alternatives will arise. Since the COI and organizational alternatives do not generally provide an inherent way of registering unaffiliated corporations and private citizens, the assurance level alternative may well be the preferred way to serve these populations. On the other hand, if a large corporation wants to have its own PCA, it should be expected to create one as long as it follows the guidelines of the PAA. Also, if a community, like the banking industry, wants to create its own community PCA, it too should be allowed.

Figure 4-7 illustrates the mixture of hierarchical certificate management alternatives to be encountered in the national infrastructure. The Federal Government sector follows the government's organizational structure. The non-federal portion of the national PKI has all three types of PCAs: assurance level, organization, and COI. The result is a hybrid infrastructure.

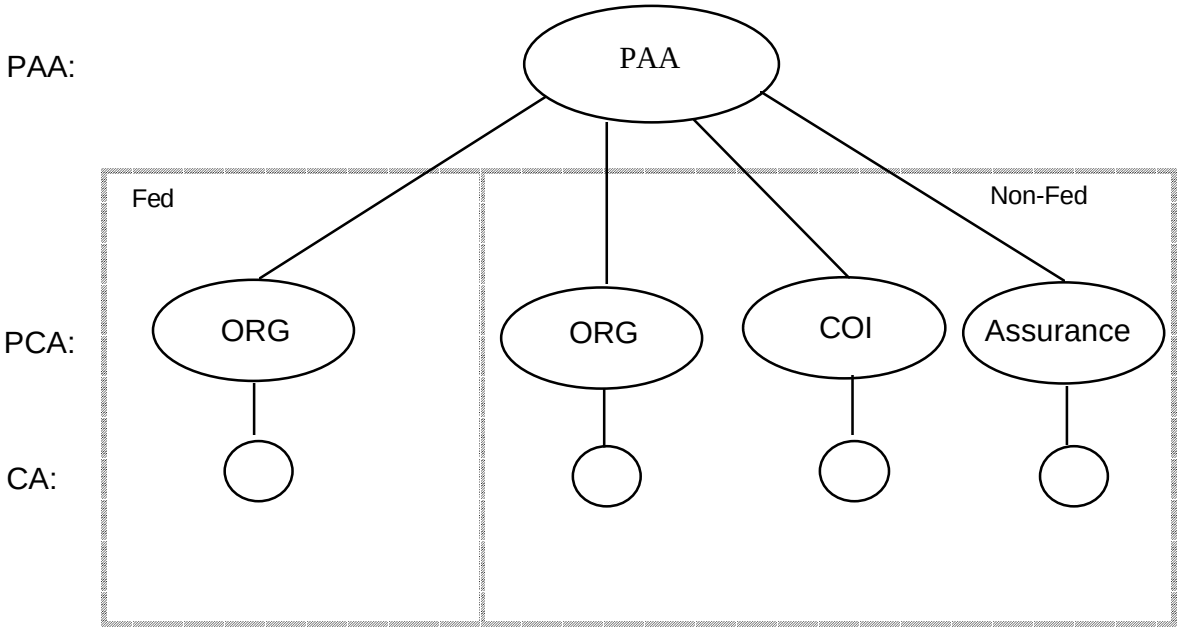


Figure 4-7. Recommended National Certificate Management Organization

It is worth noting that, individuals or entire offices within certain organizations on the non-federal side that transact large amounts of business with the Federal Government may wish to register within the federal sector. That registration can conceivably obtain the user's initial certificate but will, most probably, produce an additional certification of the same user key. If an end-user needs to exchange data with more than one organization that requires different levels of security, the user has the following two options:

- The user may obtain one certificate that satisfies the most stringent security requirements (e.g., largest *p* size, most strict Identification and Authentication (I&A)) he will encounter and he uses it with all organizations

- The user may obtain separate certificates, one per security requirement. This option allows optimum signature and verification process performance by not using a larger *p* than necessary. The logistics of keeping multiple keys and certificates may overshadow the benefit derived

4.4 MANAGEMENT OF THE RECOMMENDED INFRASTRUCTURE

This section suggests an organization or a group of organizations to be considered as candidates for managing the various level CAs of the recommended national infrastructure. The word management can mean different things to different people because management means different people with different tasks and obligations in different contexts. There are two aspects to management: (1) determining policies and procedures and (2) conducting the day-to-day operation. Of course, these two categories of work need not be performed by the same set of people or organizations. As an example, one group of people or of organizations may set the policy for a CA, but a different group performs the daily operation of that particular CA (e.g., an actual act of issuing certificates to CAs). While setting policies and procedures for the government operation will be the responsibility of agencies of the Federal Government, the day-to-day operation may be delegated by each agency to another organization within the government or to some service provider in the private sector.

4.4.1 Management of the PAA

Since the PAA will be responsible for the entire certificate management infrastructure, a committee consisting of representatives from government and industry would be an appropriate group to establish the policies and procedures for the entire infrastructure. The following organizations should be represented on the PAA: DISA, FRB, GAO, GSA, NIST, NSA, OMB, USPS, and appropriate industry and trade associations. While it is desirable that a group of organizations, rather than a single organization, participate in setting the policies and procedures, it is not necessary that the same group be responsible for the daily operation of the PAA. In fact, daily operation could be less effective if run by a committee.

The trust people put in organizations that perform the daily operation of certificate management can be an important factor in the acceptance and widespread use of the digital signature technology. It is desirable for organizations that are already established as being trustworthy to operate the PAA node in the certificate management tree. Three organizations, FRB, GSA and USPS seem to fit these criteria. Therefore, it is recommended that one of them be responsible for the daily operation of the PAA.

4.4.2 Management of the PCAs

The recommended infrastructure implementation states that each executive department, independent commission, and independent agency develop its own security policy. On the other hand, it may be more expedient that a committee of federal agency representatives, which may be the same committee that sets the general rules as the PAA, provides policies and procedures for all federal agencies. Each agency can then modify these policies and procedures to suit its particular needs. It is even conceivable that only one policy will be sufficient for all federal agency business. In any case, the daily operation of the PCAs would be under the charge of executive departments and certain other major independent executive agencies.

4.4.3 Management of the CAs

Once the policies and procedures are set by the PCAs, CAs augment those policies if appropriate, translate them into practical terms, and see them carried out. Divisions and agencies within executive departments, at the level of the Civil Rights Division within the Department of Justice or Consular Affairs at the Department of State (DOS), will manage their own CA. The actual running of the CA may be contracted out to an agency experienced in supplying such a service.

4.4.4 Management of the ORAs

ORAs play an important role in authenticating the identity of a user on behalf of a CA. In general, the office that issues employee badges or the office of human resources within an organization are good candidates for running ORAs. A new employees will have a certificate generated for them when he first reports for work or when he is issued a badge. Conversely, a departing employee will have his certificate revoked when he hands in his badge or when he checks out from the human resources office.

4.4.5 Summary of Recommended Organizations for the Infrastructure Management

Table 4-2 summarizes recommended organizations for the management of the PAA, PCAs, CAs, and ORAs. The specific organizations to operate the non-federal sector are preliminary suggestions. Their selections and recommendation is beyond the scope of this study.

Table 4-2. Infrastructure Management

Authority	Federal Sector Organizational Alternative	Non-Federal Sector All Alternatives
PAA	Committee of DISA, FRB, GAO, GSA, NIST, NSA, OMB,USPS and/or others for the federal sector; standards and/or commercial associations and/or others for policy approving. FRB, GSA, or USPS for daily operations	
PCA	Departments and/or major agencies	USPS, Banks, Telecommunications Service Providers
CA	Agencies under department level	USPS, Banks, Telecommunications Service Providers
ORA	Local organization authorities such as badge-issuing offices	Local organization authorities such as badge-issuing offices

4.5 TOWARDS GLOBAL INTEROPERABILITY

To this point, the discussion has concentrated on the U.S. infrastructure. Global interoperability is inevitable in a world of ever growing network connectivity. Obviously, there will be diverse certificate management infrastructures worldwide. Therefore, one of the most important requirements for the PKI is an ability to interoperate with other certificate management schemes. The diversity of algorithms used in signing certificates may cause interoperability complications. Nonetheless, the infrastructure should be responsive to the need for federal agencies to interact with state and local agencies, with the private sector, and with any international entities that may be utilizing alternate public key cryptographic algorithms with diverse infrastructures. In the end, the infrastructure should support global interoperability.

Interoperation also has to occur on different levels. First, data should be able to reach from one end of the globe to the other. To accomplish this, all the systems that are interconnected through the web of national and international networks must be open systems and must comply with the emerging U.S. and international standards. Second, once the data arrives at the receiving end, the receiver must know what public key cryptographic algorithm was used for the signature. X.500 Directory Service makes provision for this through a certificate. (The certificate format recommended in appendix F follows the 1992 version of the X.509 certificate. Other standards such as ANSI X9.30 [6], however, are moving towards certificate formats that allow the insertion of considerably more information to support automatic evaluation by signature verification software of certification policies and of signature authorizations.) All certificates carry information about the algorithms used for signature generation and any relevant parameters associated with the keys, both for the certificate holder and the issuer of the certificate.

There are two approaches for achieving global interoperability: cross certifying with other national or international infrastructure roots as long as they are following a hierarchical model, and establishing one global root.

4.5.1 Cross Certification with All Roots

This option achieves interoperability by cross certifying each root with all available and willing roots. The roots of each hierarchy (e.g., U.S. PAA and Internet's IPRA or U.S. PAA and Great Britain's (GB) equivalent of the PAA) cross certify each other. This option eliminates political complications. It adds only one link to each full certificate chain. However, each root may need to keep many certificates, one for each of the other international roots with which it cross certifies.

4.5.2 One Global Root

As illustrated in figure 4-8, this option ties the different certificate management hierarchies with a common global root administered by an international organization such as the United Nations (UN) or the Swiss based Bank for International Settlements. This mechanism only adds one link to the certificate chain. However, this option may not be politically feasible as many national and multi-national certificate infrastructure managers may be reluctant to relinquish complete autonomy.

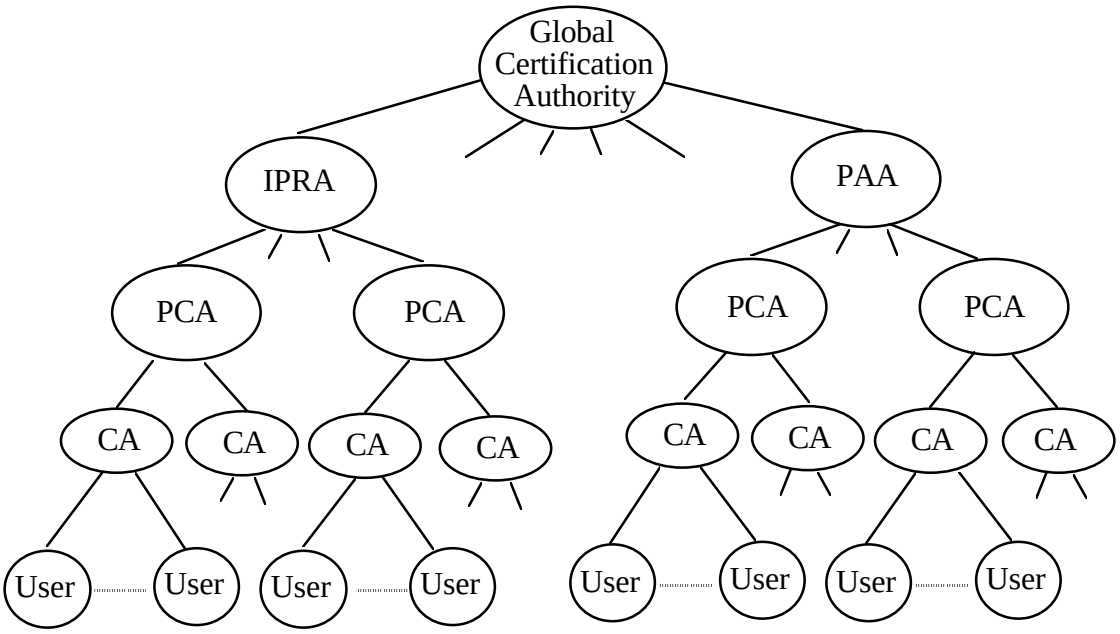


Figure 4-8. Interoperability through One Global Root

SECTION 5

OPERATIONAL CONCEPTS

5.1 INTRODUCTION

In order to develop a cost model for the PKI, one must understand how the PKI will operate. This section describes operational concepts for the PKI. Many activities associated with the operation of the PKI may be conducted in more that one manner. In addition, some of the activities are dependent upon one another. Therefore, deciding to perform one activity in one manner may limit the ways in which another activity can be conducted. Various combinations of how the activities are conducted result in different overall concepts of operation (CONOPS) for the PKI. In the remainder of this section different operational choices for each activity are presented. The section does not present an all inclusive list of operational choices, but it does discuss the more practical approaches.

5.2 PKI ACTIVITIES

There are twelve major activities associated with the operation of PKI: Generating, Certifying, and Distributing Keys; Signature and Verification; Obtaining Certificates; Verifying Certificates; Caching Certificates; Obtaining Cached Certificates; Reporting Key Compromise or Severed Relations; Recovering from a Key Compromise; Obtaining CRLs; Rekeying; Auditing; and Archiving. These activities associated with PKI operation are flagged according to the components that perform them: the PKI certifying authorities (indicated by P), the Directory (noted by D) and the users (signified by U).

Within this section, the term CA is used in the generic sense to include CAs, PCAs and the PAA. The term entity is used in this section to mean any component of the PKI be they human or machine and include: users, processes, CAs, ORAs, CA operators and ORA operators.

5.2.1 Key Generation, Certifying, and Distributing Keys (U, P, D)

There are two basic ways in which a user's public/private key pair, hereinafter referred to as the key pair, can be generated, certified and distributed. The choices are that the user generates his own key pair or that another entity generates the key pair for the user. The decision of which option to follow depends on the applicable PCA policy. In either case, a method that produces a good key pair should be employed. Appendices 2 - 4 of the DSS [5] describe how good key pairs can be generated.

If the user generates his own key pair, he is responsible for ensuring that he used a method for generating a good key pair. He must store the private key in a secure location so that it cannot easily be compromised, possibly on a smart card, a PCMCIA card or an encrypted diskette. He is also responsible for having his public key certified by a CA.

To have his public key certified by a CA, the user can present himself and his key to the CA. The CA will authenticate the user. If authentication is done in person, it may consist of the examination of several forms of identification which the user presents. The CA may also perform some tests on the key to determine its strength and to assure that it was truly generated by the user who is registering it. After the CA is sure of the identity of the user and the validity of the key, it will generate a certificate for the user that binds the user's identity to his public key. The CA will distribute the certificate to the user in person, through the mail, or electronically. The CA will also deposit the certificate with the appropriate directory server.

It is possible that the user will be located closer to an ORA than to a CA. In this case, the ORA will perform the authentication of the user and forward the user's credentials, which include his unique name and his public key, to the CA in a signed message. The CA can perform tests on the public key. The CA will then generate the certificate for the user. The CA will then send the certificate to the ORA, which will in turn deliver it to the user. Alternatively, the CA can deliver the certificate to the user electronically.

In the discussion presented above, it was assumed that the user had either software or hardware that was capable of generating the key pair for the user. It may not be practical from both cost and ease of use perspectives to give key generating capabilities to all users. Therefore, it may be necessary to have a central key generating system to which users in a particular location can go in order to generate key pairs for themselves. The key generating system can be either collocated with a CA or in a separate location.

To obtain a key pair from a key generating system, the user goes to the key generating system and requests that the system generate a key pair for him. The system generates the key pair and gives the public and private keys to the user, possibly on a smart card, a PCMCIA card or a diskette. The key generating system automatically destroys the copy of the user's private key once it was given to the user. It is insufficient to have the key generating system forward the public key to the CA for certification. It must give the copy of the public key to the user so that he can be properly identified during one of the certificate generating procedures described above.

CA keys are generated by the CA itself. Thus, the PAA, the PCAs and CAs all generate their own key pairs. An ORA can generate its own key pair or have it generated by a third party, as the PCA policy dictates. A PCA has its public key certified by the PAA. At that time, it can obtain the PAA's public key. A CA's public key is certified by the appropriate PCA. Along with its own certificate, the CA can obtain the PAA's public key and the PCA's public key. It is then in a position to pass either the PAA's public key or the PCA's public key or both to the entities whose keys it certifies.

5.2.2 Signature and Verification (U)

An activity that will be performed many times within the PKI is the signing of messages or files and the verification of the digital signatures on these messages or files. Actually, the PKI is being developed specifically to facilitate the verification process.

Entities within the PKI can use any one of several algorithms to compute and to verify digital signatures. For the purpose of this discussion, assume that the SHA and the DSA are used. A description of the DSS signature and verification process is presented in appendix B of this paper. The algorithms can be implemented in software, hardware, and firmware. Keys and certificates may be stored on computers, in smart cards, or on other types of media such as floppy disks.

5.2.3 Obtaining Certificates (D, U)

In order to verify a digital signature on a message, one must obtain the public key of the sender of the message. This public key is contained in the certificate of the sender; therefore, it is the sender's certificate that needs to be obtained. The recipient may need to obtain one or more additional certificates, such as those of CAs, in order to verify the signature on the sender's certificate.

There are a number of ways in which certificates may be obtained. The certificates can be sent with the signed message. They can be sent in a separate message. They can be obtained from the Directory, once it is universally available. They could also be obtained from the entity with which they are associated. One of these methods, or a combination of several of these methods can be used to obtain certificates within an infrastructure.

Along with the signed document, the receiver may have been sent the certification path directly from the sender. Thus, he will have obtained all the certificates within the path. If the receiver of a message is not sent all the certificates he needs to verify the sender's certificate, he will have to determine what other certificates he must obtain. He may do this in several ways. If he has the sender's certificate, he can examine the certificate and see which CA issued it. The CA's certificate can be obtained from the Directory. The user will examine this certificate and repeat this process until he finds a certificate signed by an entity for which he has the public key. In this manner he calculates the certification path of the sender. Alternatively, depending upon the naming scheme used within the infrastructure, the receiver may be able to calculate a certification path from the sender's unique name. For this each unique name must include the unique names of all the certifying authorities in a chain up to the one whose public key he holds. In either case, the directory service must aide him in locating the directory server to which he must and each certificate request and each associated CRL request.

5.2.4 Verifying Certificates (U, D)

The process of verifying a signature on a document includes iteratively determining the next step in the certification path and obtaining the pertinent CA certificate. Each certificate obtained must be checked against the appropriate CRL before it can be used. Starting with the key of the PKI entity where the certification path ends, the user verifies the signature on the certificate signed by that entity. Once this certificate is verified, the public key within it is extracted and is used to verify the next certificate within the path. The verification process

continues until the signature on the document signer's certificate is verified and the public key is extracted from the certificate.

5.2.5 Caching Certificates (U)

Caching is the process of storing certificates for later use. Caching is usually done to reduce the number of times a user actually has to contact the PKI for certificates and to make the verification process more efficient. A user can select to store all, none, or a certain set of certificates. This section describes several caching schemes.

The simplest, but probably the most inefficient, way of caching certificates is to store every certificate that the user receives. These can be certificates received from other users or retrieved by the user. It is recommended that the user verify each certificate before storing it.

Once the user has obtained and verified all the certificates within a certification path, he may choose to cache all or some of these certificates. A user can cache all the certificates in the path, the sender's certificate along with the certification path or just the sender's certificate.

There are several maintenance issues associated with certificate caches. If the cache becomes full, a procedure must be established for determining which certificates are to be removed from the cache. The most common procedure is to delete the certificate that was least recently used (LRU) from the cache to make room for a new certificate. Caches can be cleaned on a periodic basis, where all certificates that have been revoked or have not been used within a specified period of time are deleted from the cache. To determine whether a certificate within a cache has been revoked, it should be compared to the latest CRL issued by the CA which signed it. Any certificates that appear on a CRL are removed from the cache. Checking cached certificates against CRLs should be done regularly. However, certificates that are used to verify signatures in highly sensitive applications should always be checked against the most recent CRL before their use to obtain the greatest assurance of their validity.

5.2.6 Obtaining Certificates from a Cache (U)

Certificate caches are created by a user to reduce the number of times he has to contact the PKI for certificates and to make the verification process more efficient. When the user receives a signed message he will check his certificate cache to see if the certificate of the sender of the message is stored in the cache. If the certificate is in the cache, then the user obtains a copy of the certificate from the cache, extracts the public key of the sender from the certificate and uses the key to verify the sender's signature on the message. Prior to using the public key from the certificate retrieved from the cache, the user may choose to check the certificate against a CRL (see section 5.2.9, Obtaining CRLs, for a discussion of CRL distribution) to ensure that the certificate had not been revoked. Alternatively, he may periodically verify the validity of all certificates in his cache. If the user does not have the

sender's certificate within his cache, he will need to obtain and to verify the sender's certificate using the methods described in sections 5.2.3 and 5.2.4.

5.2.7 Reporting Key Compromise or Severed Relations (P, U)

It is expected that the private keys of some of the entities within the PKI will be compromised, thus requiring that the public key certificates associated with those entities be revoked. For CAs, compromise will most likely occur through deliberate attacks on the system. For users, compromise of their key will occur if the smart card on which their private key stored is stolen or lost.

In addition to compromise, public key certificates may be revoked because the relationship between the user and the organization specified in his unique name has been severed. For example, a user leaves his employer for a new job. His key pair, which is certified by the employer's CA, is no longer valid. Therefore, the public key certificate issued by the employer's CA should be revoked.

If either a compromise or a severed relation occurs, the CA that issued the certificate containing the public half of the affected key pair should be notified as soon as possible through some out-of-band method. Out-of-band methods include written notification and notification by phone. It is the responsibility of the compromised entity to notify the appropriate CA. In the case of severed relations, the employer's personnel or security office will notify the appropriate ORA or CA that an employee has left the company. If an ORA is notified, it will in turn notify the CA of the severed relation.

When a CA receives notification of a key compromise or a severed relation, it must verify the authenticity of that notification by the use of both an authentication and an authorization scheme. Once the notification is judged authentic, the CA will mark the certificate as revoked within its database. The CA will include the revoked certificate serial number and date of revocation on the next CRL it issued. The certificate remains on the CRL until its expiration date is reached.

If the compromised entity is a CA, a highly unlikely occurrence for a well-run CA, it will notify its parent CA. The parent CA will place the compromised certificate on a CRL. It is sometimes important that entities certified by the compromised CA be quickly alerted to potential problems with the acceptance of their signatures. Therefore, the CA may also notify its children of the compromise. It is important that this notification be done in a such a way that the children can trust its authenticity. One way this can be accomplished is by having the parent CA prepare and sign a compromise message which the compromised CA sends to all its children.

5.2.8 Recovering from a Key Compromise (P)

If the private key of an entity is compromised and the certificate has been revoked, a new key pair and certificate will have to be generated for the entity. This will allow the entity to

resume its business and be able to produce digital signatures again. A new key pair will be created and the entity's CA will generate a new certificate for the entity.

In the unlikely event that the compromised entity is a CA, it will need to reissue all the certificates it generated using the compromised key. Anyone who possess that private key can submit, to the Directory, bogus certificates containing false public keys of his own creation. He can then masquerade as any of the PKI entities for which he has generated the false keys. Since the integrity of the public keys contained within the certificates the compromised CA has issued is no longer guaranteed, new key pairs must be created for all the CA's users as well as new certificates for these new public pairs. The new certificates are signed in the CA's new private key. The old certificates issued under the compromised key are revoked and placed on a CRL.

Generating many new key pairs and reissuing many new certificates is a laborious undertaking. Recovery from CA key compromise is easier if dual CAs are used, that is, if every PKI entity has its public key certified by two CAs. If one of the two CA's key is compromised, its certificates can no longer be used in a certification path. However, another path exists through the dual CA and its certificate is still valid. Users need to be notified of the compromise so that they can switch to using the certificate issued by the non-compromised CA. The unauthorized holder of the CA private key will gain little in trying to submit false keys in bogus certificates which are ostensibly signed by a CA which is no longer in any certification path. Recovery of the compromised CA and reissuance of all the certificates it generated may take place over time, since system operation is not excessively affected .

5.2.9 Obtaining CRLs (U, D)

It is assumed that all CAs will generate CRLs. The CRLs may be generated on a periodic basis such as once a month or may be generated every time a certificate revocation occurs. These CRLs will encompass certificates, generated by the CA, that have been revoked because of key compromise, changes in a user's affiliation, etc.

As with obtaining certificates, there are also many approaches for obtaining the CRLs generated within the PKI. For example, when a CA generates a CRL it can automatically send the CRL to its children. Pushing CRLs may be desirable for applications that need immediate notification of a compromise. To keep querying the Directory may be impractical and, if the Directory is temporarily out of service, may be devastating.

In a more common approach, no CRLs are automatically distributed. Instead, the responsibility for obtaining CRLs is placed completely on the entities within the PKI. All entities are responsible for requesting the CRLs that they need from the Directory. Any CA which generates a CRL is responsible for sending its latest CRL to the Directory. All the users of the infrastructure can query the Directory for the CRLs they need.

5.2.10 Rekeying and Recertifying (U, P, D)

A new key pair and a certificate are generated whenever a private key is compromised. Even if a key is not compromised, it should be changed on a periodic basis such as once a year. Various options for key change are available. All entities within an infrastructure can change their keys on the same day. Alternatively, different days can be selected for each entity or each type of entity. For example, all CAs can change their keys on the same day. Each user within the system changes his keys on the first or second anniversary of when those keys were created or, perhaps, whenever the certifying CA changes its key pair.

Specifically, this means that all certificates carry an expiration date which coincides with the earlier of the expiration dates of the certified and certifier's keys. If the keys of all entities under a CA expire on the same day as the CA's key lapses, all certificates carry the same expiration date. When that day approaches, the CA and all the entities under it will select new key pairs. The CA will certify the new public keys in new certificates signed in its new private key. The situation is slightly different if the expiration dates do not coincide. Suppose all keys are replaced on the first year anniversary of their creation. A new public key for an entity will be certified by the CA, using its current private key to sign. The expiration date of the certificate will be the same as the day on which the CA is required to obtain a new key pair. When that day comes, the CA replaces the certificate with a new one which recertifies the same entity public key, but signed in the CA's new private key. The expiration date of the new certificate is now set at the anniversary of the entity's initial key generation. When that day approaches, the entity obtains a new key pair. The new public key is again certified by the CA using its current (i. e., latest) private key for signing.

No matter what type of changeover is selected, the entities who are changing their keys must have the new keys prior to the changeover date. This simply means that a new key pair and certificate must be generated and distributed to the entity prior to the changeover date. On the changeover date, which is indicated in the old certificate's validity interval, the entity begins to use the new private key to sign messages. The entity may archive the old key pair and certificate. These rekeying alternatives apply to each level of the infrastructure.

5.2.11 Auditing (P, U)

Auditing can be done by any entity within an infrastructure, including users. However, it is more likely that auditing will be performed by CAs. It is expected that each CA within the PKI audit security relevant events. Events which should be audited include: requests for key pairs and certificates; reports of compromised keys; and reports of severed relations.

5.2.12 Archiving (P, U, D)

For various purposes including legal requirements, satisfaction of Federal Government regulations, and system recovery needs, the certificates and the CRLs which a CA generates should be archived. The CA may need to archive other types of files and information, such as audit files, to satisfy statutory or regulatory requirements.

SECTION 6

INFRASTRUCTURE COST ANALYSIS RESULTS

To implement the PKI along the lines of the recommended architecture and for it to perform the functions described in the last section will obviously cost money. In an effort to estimate the financial impact of implementing the PKI, this report includes the results of a cost modeling effort. The details of the model are contained in appendix I. They include a detailed description of the assumed concept of operations, a listing of many of the values assumed for model variables and the details of the model itself. This section presents a summary of the assumed concept of operations, the results of the cost analysis and a few comments on those results.

6.1 OPERATIONAL CONCEPT USED IN COST MODEL

This section provides a brief summary of the concept of operations upon which the PKI cost model is based. In this summary, the twelve activities associated with the concept of operations are divided between the entities which perform them: the users, the PKI components (ORAs, CAs, PCAs and PAA) and the Directory.

6.1.1 User Activities

In order to have the ability to sign an electronic document digitally, a user will need to have a key pair and certificate. We assume that a trusted third party, referred to as a key generator (KG), will be used to generate the key pair for the user and place it on a token such as a smart card. The user is then responsible for taking his public key to his local CA or ORA to have it certified. In the event that the user loses the token on which the private key is stored or if he suspects his private key has been compromised in some other manner, he is responsible for sending a certificate revocation report to the appropriate PKI component, ORA or CA, in an out-of-band method. In the case of compromise, the user will need to generate a new key pair and to request a certificate be generated for him. Key pairs and certificates will have expiration dates associated with them; we will assume keys and certificates are valid for one year from issuance. On expiration, the user will need to obtain a new key pair and certificate.

To verify the digital signature of another, a user will need to obtain the public key of the signer, which is contained in the signer's public key certificate. Certificates will be stored in the Directory, so the user will query the Directory for the signer's certificate. In order, to verify the signer's certificate, the user will need to obtain the certificate of the signer's CA from the Directory, and so on until the user finds a certificate signed by a PKI entity for which he already has a trusted copy of the public key. The user can then iteratively verify the signatures on all the certificates until he reaches the signer's certificate. Once the signer's certificate is verified, the signer's public key can be extracted from the certificate and the signer's digital signature can be verified. During the verification process, the user or process running on his behalf should check all the certificates in the certification path against the

appropriate CRL(s), which are also obtained from the Directory, to ensure that none of the certificates have been revoked. If any of the certificates have been revoked, the verification process should stop, since the integrity of the signature cannot be assured.

Users will not want repeatedly to go through this detailed verification process for users with which they communicate often. In this case, we assume each user maintains a cache of certificates for the users with which he communicate often. The user verifies all the certificates within the certification path once and then stores the signer's certificate in the cache. A user or the process running on his behalf looks first in the cache for a signer's certificate before querying the Directory for the certificate.

There are several maintenance issues associated with a certificate cache. First, if the cache becomes full a least-recently-used (LRU) method should be used to remove certificates making room for new certificates. The second concern is cleansing the cache of revoked and expired certificates. Periodically the user or process working on behalf of the user should query the Directory for the appropriate CRL(s). The certificates in the cache should be checked against the CRL(s); any certificate which has been revoked should be removed from the cache. The expiration dates of certificates within the cache should also be examined; expired certificates should be removed from the cache.

6.1.2 PKI Activities

A third party key generating service will be provided to all users of the PKI by the KG. Users will go to the KG and use the system to generate a key pair for themselves and have the key pair placed on a smart card or similar token. The KG destroys all copies of the key pair once it is placed on the token. The user is responsible for taking the public key in person to the CA or ORA for certification. We assume that all PKI components (ORAs, CAs, PCAs and PAA) will generate their own key pairs. The operators of the ORAs, CAs and PCAs will travel to the appropriate parent to have their entity's public key certified.

CAs within the PKI are expected to generate certificates for users or subordinate CAs after properly authenticating the entities being certified. These CAs deliver the certificates they create to the Directory and to the users or subordinate CAs for which they were generated. Some users within the PKI are served through ORAs. These users go to their ORAs to have their public keys certified. The ORA authenticates the user and forwards the public key and other necessary information to the appropriate CA. The ORA receives the certificate back from the CA and forward the certificate to the user.

CAs within the PKI also receive certification revocation reports from users. The CA is responsible for authenticating these reports. If a report is authentic, the CA marks the associated certificate as revoked and includes this certificate on the next CRL that it issues. If a user interfaces with an ORA, he will send the revocation report to the ORA. The ORA is responsible for authenticating the report. If the report is authentic, the ORA sends a signed message to the associated CA which in turn revokes the certificate. We assume that CAs will generate CRLs on a biweekly basis. After generating a CRL, the CA delivers a signed copy of the CRL to the appropriate directory server.

All the PKI components are expected to perform some auditing activities, especially on actions which are considered security relevant. Most PKI components are likely to have archiving requirements imposed on them. It is also possible, although unlikely, that the private key of a PKI component will be compromised. In that case, the PKI component will need to go through a recovery procedure.

6.1.3 Directory Activities

Certificates and CRLs will be stored in the Directory. Therefore, the Directory will be responsible for responding to all requests for certificates and CRLs. These requests may be received from users or from PKI components. The Directory is also expected to receive certificates and CRLs from the PKI components which generate them. The Directory may remove expired or revoked certificates.

6.2 SCENARIOS

There are several variables in the cost model whose values will be determined only after the PKI is installed and running. It seems inadequate to assume a single value for each of these variables. Thus, cost estimates corresponding to a range of values for each of them are presented. These variables and the ranges of their values are as follows:

- The number of message and document signatures an average user verifies per day: 5, 25, 50, 100.

- The percent of federal civilian employees who use PKI: 50%, 75%, 100%.

- The percentages of an average user's corespondents who are siblings (under the same CA), who are first cousins (under a different CA but the same PCA), and who are second cousins (under a different CA and a different PCA): 30% - 40% - 30%, 50% - 25% - 25%, 70% - 20% - 10%.

6.3 RESULTS

Cost estimates for each scenario are given for both start-up expenses and yearly running expenditures. The 36 scenarios are grouped into three sets. Each set models the same number of employees using the PKI. The cost model results are summarized in tables 6-1 and 6-2. Table 6-1 presents the start-up costs while table 6-2 presents the yearly costs. All cost values shown are in millions of dollars ($M).

Table 6-1: Total Start-Up Cost Estimates

(amounts in millions)

# of Mess	% of Siblings	50% of Federal Employees	75% of Federal Employees	100% of Federal Employees
5	30%	$525.1	$782.9	$1040.7
	50%	525.1	782.9	1040.7
	70%	525.1	782.9	1040.7
25	30%	529.0	790.7	1048.5
	50%	529.0	786.8	1048.5
	70%	525.1	786.8	1048.5
50	30%	532.9	798.4	1060.1
	50%	532.9	794.5	1060.1
	70%	529.0	790.7	1056.2
100	30%	544.5	813.9	1083.3
	50%	540.6	810.0	1079.4
	70%	536.7	802.3	1071.7

Table 6-2: Total Yearly Cost Estimates

(amounts in millions)

# of Mess	% of Siblings	50% of Federal Employees	75% of Federal Employees	100% of Federal Employees
5	30%	$299.7	$583.5	$962.2
	50%	274.7	528.9	866.4
	70%	239.9	452.7	732.9
25	30%	1096.1	2323.4	4008.8
	50%	971.3	2050.0	3529.8
	70%	797.1	1669.2	2862.4
50	30%	2091.6	4498.2	7817.1
	50%	1841.9	3951.5	6859.2
	70%	1493.8	3189.8	5524.1
100	30%	4082.8	8847.8	15433.8
	50%	3583.3	7754.6	13517.7
	70%	2887.2	6231.1	10847.5

6.4 ANALYSIS

An examination of the spreadsheet results quickly identify the main sources of both the start-up costs and the yearly costs. They are discussed briefly below.

6.4.1 Analysis of Start-up Costs

In table 6-3 a breakdown of the start-up costs for the PKI for a selected set of scenarios is presented.

Table 6-3: Start-up Costs

Percent of Federal Civilian Employees: 50% (Amounts in millions)				
# Messages	5		100	
% of Siblings	30%	70%	30%	70%
User Smart Card & Reader	515.6	515.6	515.6	515.6
PKI Equipment	9.5	9.5	9.5	9.5
Directory Equipment	0.0	0.0	19.3	11.6
Total Start-up Cost	525.1	525.1	544.5	536.7

Percent of Federal Civilian Employees: 100% (Amounts in millions)				
# Messages	5		100	
% of Siblings	30%	70%	30%	70%
User Smart Card & Reader	1031.2	1031.2	1031.2	1031.2
PKI Equipment	9.5	9.5	9.5	9.5
Directory Equipment	0.0	0.0	42.5	30.9
Total Start-up Cost	1040.7	1040.7	1083.3	1071.7

By far, the single most expensive category in the cost of deploying the PKI is the user hardware cost. The CONOPS in section 6.1 describes an approach, which has all private keys and all signing capability resident on smart cards. This implies that every user must have an interface unit that allows his workstation to interact with the smart card. The cost of these units, with the associated software, is given as $337 each. The total cost of supplying an interface unit to each and every federal employee is then $1,031M. A smaller unit cost for the total number of units required is a possibility and will reduce this expense. Exploitation of newer technology which uses existing workstation interfaces without additional hardware will also reduce this cost. Obviously, a decision not to compute signatures in the protected environment of a smart card but in the more exposed software of the workstation will save considerable money. However, it will expose users' private keys to the threat of possibly malicious software. Furthermore, a move not to hold private keys on

memory cards to save the cost of card readers creates very serious key storage hazards. The best fall-back position, in that event, is to put both the private keys, in encrypted form, and the signing software on floppy disks, which must be locked away when not actually being used to sign documents. This places the onus of protecting the confidentiality of each private key on the owner of that key, no matter how inexperienced in security procedures he may be. By not requiring a smart card interface on each user's workstation, designers can reduce the start-up costs to under $60M, $45M or $30M for the 100 percent, 75 percent and 50 percent of federal civilian employee categories, respectively.

The decision of whether to use smart cards for both key storage and the signature computation, for key storage alone or for neither, of course, falls under the purview of each agency. Part of each PCA's security policy should delineate the circumstances and conditions under which each alternative for smart card use is appropriate and acceptable. Obviously, security and cost trade-offs are integral to the setting of this policy. It should be noted that prices for smart cards and for readers are decreasing. The model assumes a cost of $337 for a card and a reader. Current quoted prices for the PCMCIA card and parallel port reader, for example, are only $250. (However, with the SCSI interface reader, the PCMCIA cost is currently $375.)

6.4.2 Analysis of Yearly Running Costs

In table 6-4 a breakdown of the yearly costs for running the PKI for a selected set of scenarios is presented.

Table 6-4: Yearly Costs

Percent of Federal Civilian Employees: 50% (Amounts in millions)				
# of Messages	5		100	
% of Siblings	30%	70%	30%	70%
CRL Directory Comm	190.3	135.2	3656.7	2554.6
CRL User Comm	1.5	1.0	28.3	19.8
Total CRL Comm	191.8	136.2	3685.0	2574.3
Other Comm	15.2	11.0	304.0	219.5
Total Comm	207.0	147.2	3989.0	2793.8
PKI Staff	15.8	15.8	15.8	15.8
PKI Maintenance	0.4	0.4	0.4	0.4
User Maintenance	76.5	76.5	76.5	76.5
Directory Maintenance	0.0	0.0	1.2	0.7
Total Maintenance	76.9	76.9	78.0	77.6
PKI Yearly Cost	16.2	16.2	16.2	16.2
Yearly Cost Without CRL	107.9	103.7	397.8	312.9
Yearly Cost	299.7	239.9	4082.8	2887.2

Table 6-4: PKI Yearly Costs (Concluded)

Percent of Federal Civilian Employees: 100% (Amounts in millions)				
# of Messages	5		100	
% of Siblings	30%	70%	30%	70%
CRL Directory Comm	756.7	539.7	14597.6	10198.0
CRL User Comm	2.9	2.1	56.6	39.5
Total CRL Comm	759.6	541.7	14654.2	10237.5
Other Comm	33.4	22.0	607.9	439.0
Total Comm	793.0	563.7	15262.1	10676.5
PKI Staff	15.8	15.8	15.8	15.8
PKI Maintenance	0.4	0.4	0.4	0.4
User Maintenance	153.0	153.0	153.0	153.0
Directory Maintenance	0.0	0.0	2.5	1.9
Total Maintenance	153.4	153.4	155.9	155.2
PKI Yearly Cost	16.2	16.2	16.2	16.2
Yearly Cost Without CRL	202.6	191.2	779.6	610.0
Yearly Cost	962.2	732.9	15433.8	10847.5

The PKI's yearly running expenses derive mainly from the expense of transmitting CRLs from the directory. For example, with all the federal employees using the PKI and each of them verifying 100 messages on average, the yearly cost is estimated at between $10,848M and $15,434M. All this except about something between $610 and $780M are CRL communications costs, which are charged at about 2 cents per kilobyte. It can be argued that the directory service agents issue CRLs at night and the costs in using uncommitted LAN or WAN capacity at night is essentially free. At present, the CRL communication costs constitute roughly one twentieth of the yearly running estimate.

The CRL cost makes a significant difference between the two ways of dealing with fewer than 100 percent of employees using the PKI. For example, with only half the employees using the infrastructure and each verifying 100 messages on average, the annual cost is $4,083M and $2,887M for the 30 percent and the 70 percent sibling scenarios, respectively.

This is with the assumption that each directory service agent serves 15,000 users. If each one is forced to serve 30,000 subscribers but their number is halved, these estimates jump to $7,797M and $5,249M, respectively. That is, an increase of between 82% and 91% simply because, while the number of recipients of CRLs is the same, the size of each CRL has been doubled. Any PKI implementation planning must look very carefully at the cost of CRL distribution and must make a concerted effort to minimize that expense. It should even consider the total elimination of CRL distribution, replacing that function by requiring users to request each certificate anew whenever they wish to check their caches. Of course, each directory service agent must maintain a current CRL and must be trusted to compare each requested certificate against that CRL before sending it to a requester.

The remaining communications costs (listed as "Other Communications" in table 6-4) derive mainly from the expense of transmitting the signatures on the various messages that users send to each other. Thus, there is a substantial difference between the dollar amounts listed under 100 messages and under 5 messages. One might argue that these are not truly PKI costs. They derive from a user's desire to receive signed messages and documents. The decision to use digital signatures is one that the user takes before he approaches the PKI. The cost of transmitting those signatures should rightfully derive from that decision and not from the infrastructure. Similarly, the cost of the DSS smart card, reader, software and maintenance are traceable to the decision to use digital signatures and not directly to the PKI. However, all these costs have been included in the tables so that planners can have a realistic estimate for what the total cost of digital signatures will likely be.

After communications costs, the next major factor contributing to the PKI yearly costs are maintenance costs. The maintenance costs have been partitioned in table 6-4 into three categories: PKI Maintenance, User Maintenance and Directory Maintenance. The PKI Maintenance value includes the costs associated with maintaining the software and the hardware of the PCAs, the CAs and the ORAs within the PKI. The User Maintenance value is the cost associated with maintaining the DSS software for each user. As one may note from the table, the User Maintenance cost is the prime contributor to the Total Maintenance cost. However, as with other user-related expenses, it can be argued that this cost is not a direct PKI cost; instead, it is a user cost associated with the decision to use digital signatures.

There are some staff costs for running the CAs and the ORAs. The model has assumed that some must be hired to manage these facility though, in truth, the task may well be assign to existing staff in the office in which the CA is installed. Similarly, the model made the same kind of assumption for user and CA manager time. Some time is costed when, in reality, the task for which the time is allotted could be automated.

6.4.3 Analysis of Cost per Message and Cost per User

For comparison purposes the Cost per Message and the Yearly Cost per User associated with the PKI have been determined for some of the selected scenarios used above. The Cost per Message is determined by dividing the PKI Yearly Cost by the product of the number of users, the number of messages a user sends each day and the number of working days per

year. The Yearly Cost per User is computed by dividing the PKI Yearly Cost by the number of users. The results of these calculations are presented in table 6-5.

Table 6-5: Cost per Message and Yearly Cost per User

Percent of Federal Civilian Employees: 50%				
% of Siblings **# of Messages**	30% 5	70% 5	30% 100	70% 100
Yearly Cost	299.7M	239.9M	4082.8M	2887.2M
PKI Yearly Cost	16.2M	16.2M	16.2M	16.2M
Total Cost Per Message	0.16	0.13	0.11	0.08
PKI Cost Per Message	0.01	0.01	0.00	0.00
Total Yearly Cost Per User	195.80	156.80	2668.48	1887.07
PKI Yearly Cost Per User	10.59	10.59	10.59	10.59
Percent of Federal Civilian Employees: 100%				
% of Siblings **# of Messages**	30% 5	70% 5	30% 100	70% 100
Yearly Cost	962.2M	732.9M	15433.8M	10847.5M
PKI Yearly Cost	16.2M	16.2M	16.2M	16.2M
Total Cost Per Message	0.26	0.20	0.21	0.15
PKI Cost Per Message	0.00	0.00	0.00	0.00
Total Yearly Cost Per User	314.44	239.51	5043.73	3544.94
PKI Yearly Cost Per User	5.29	5.29	5.29	5.29

As one may note from the tables, the Total Cost per Message ranges from $.08 to $.26 per message. The low end cost per message is obtained when 100 messages are sent by a user per day. The high end cost is associated with a user sending out only 5 messages per day. Therefore, the cost per message decreases as users send out more messages. The PKI cost per message alone is always less than a penny.

The Total Yearly Cost per User ranges from a low of $156.80 up to a high of $5043.73. The lower value is obtained under the 50% of federal employees, 5 message, 70%-20%-10% scenario. The high value is obtained under the 100% of federal employees, 100 message, 30%-40%-30% scenario. These numbers show that the cost per user increases as the number of users increase. This cost increase is primarily due to the cost of CRL distribution. The larger number of users increase the size of the CRL and increase distribution list of the CRL, thus driving the yearly cost up. Cost efficient techniques for CRL distribution will need to be employed in order to keep the cost per user at a reasonable level. The ultimate goal is to have the cost per user decrease as the number of users increase. Similarly, the PKI Yearly Cost per User ranges from a low of $5.29 for 100% of the federal employees to a high of $10.59 for 50% of the federal employees.

SECTION 7

RELATED ISSUES

During the interview process, several user needs were identified that were related to certain applications of digital signatures. These needs do not necessarily impose requirements on the PKI and the services it provides. They may be satisfied by systems or infrastructures that run in parallel with the PKI. The designers and users of the PKI should be aware of these needs, because the PKI may have to interface with these parallel systems in the future. Several issues related to the PKI, including the next steps required in deploying the PKI, are discussed in this section.

7.1 AUTHORIZATIONS AND ATTRIBUTES

Authorization and handwritten signatures are closely tied in a number of applications, especially within the financial arena. Contracting officers are authorized to sign documents associated with a specific contract. Procurement officers are allowed to authorize payments up to a certain monetary limit. Corporate officers are authorized to sign documentation on behalf of the corporation. In a paper world, the signature authority is assigned through a specific procedure and is documented on a paper form or certificate. These forms are sent to, or may be obtained by, any entity that needs to know that the person is authorized to sign in a specific situation. Authorization is established independently from validation of any signature.

Whereas authorizations are granted to individuals to define the extent to which their actions can be accepted, attributes describe those individuals themselves. Often, on the basis of their attributes rather than their authorizations, one must make a decision on what they may or may not do. For example, a person's credit line determines whether or not he can charge a given purchase. A characterization as a user rather than a CA indicates that the certificates an entity signs cannot be trusted. The extent of a CA's liability in the event of a problem with a certificate is an attribute of that CA. It will determine if a certificate the CA issued should be used in a specific and special context. In the infrastructure built to support some public key systems, an entity's public key for another system employing a different algorithm is also considered an attribute.

In an electronic environment, a digital signature may be used to authorize a specific action, such as the payment of a certain contract. The PKI can prove the authenticity of the signature, but it cannot prove that the signer was authorized to sign in the specific situation. These authorizations must be proved by an independent system. The system could be paper based, but it would be preferable if it were electronically based. Authorization certificates can be created by an infrastructure similar to the PKI that associates a person's authorization with his/her identity. They will require digital signatures and hence will depend on the PKI for their verification. These authorization certificates can be sent along with the digitally signed document to prove the authority of the signature. Alternatively, the authorization certificates can be obtained from the authorization infrastructure whenever the authority of a

signature needs to be proved. In either case, authorization certificates are simply special signed documents that rely on the PKI for the trust to be placed in them. Some people have suggested that authorization be encoded directly into public key certificates to expedite the automatic handling of electronic commerce transactions [6, for example].

Authorizations can also be handled making use of the PKI itself. A user can be considered to acting in different roles if he is signing a purchase order for 5,000 DSA smart cards or committing his organization to perform the tasks listed in a contract. He might have different unique names for these roles with, possibly, different keys. Certification by the cognizant CA indicates that the user is authorized to make the commitment. This way of using the PKI to implement electronic authorizations requires a carefully constructed PKI very likely employing many more CAs in strategic positions. A department can elect to use the PKI in this way. The subject of authorizations, however, needs further examination[5].

7.2 TIME AND DATE STAMPS

In some applications, time and date stamps must be affixed to documentation to denote when the documentation was received or sent. Examples of such documentation include proposals from vendors and patent applications. If the documentation is generated by and sent via electronic means, the date and time stamp must also be generated and affixed to the document electronically. The systems that provide the time and date stamps need to be accurate and secure to ensure that the stamps reflect the true time and date the documentation was stamped. Income tax returns and many other mandated filings carry a specific filing deadline. When the filing is done electronically, it requires an accurate date and time stamp.

7.3 ARCHIVING

There are any number of regulations and statutes that state how long documents and reports must be kept in storage before they can be destroyed. The archival periods range from one year to seven years and up. When the documents in question exist only in digitally signed, electronic form, the actual storing becomes easier, for electronic archival is far superior to the paper or the fiche approaches. However, other complications do arise. First, the signature and signature application software version number must be saved with the document. Additionally, the certificate containing the needed public key as well as the certificate chain that established trust in the key must be archived, although not necessarily with the document itself. The pertinent CRLs existing at the time of the signing will be needed to check whether the private key was valid while later CRLs will be required to prevent later repudiation.

---

5 Authorization is currently a topic of standardization in the ANSI Accredited Standards Committee's Subcommittee on Data and Information Security Working Group X9F1.

There are government agencies whose task is to devise rules for the method and duration of electronic document archival. In support of this, someone must decide how to store and retrieve both certificates and CRLs. This will be a long-term storage whose duration must also be determined. There is actually some question as to whether the storing of all the certificates and pertinent CRLs is required when the archiving is done by a government archival service. With well defined procedures in place and followed, it may be sufficient for the archiving authority to verify the signatures at the time of receipt. Any signature whose verification fails will be so marked along with an indication of the nature of the failure and , perhaps, the relevant certificates and CRLs. The majority of signatures will be accepted and no further certificate or CRL storage will be required for them. The fact that the documents have been accepted for archiving will imply that the signatures have been revarified by the archiving agency. That agency can, additionally, affix its own signature to the stored file. It will need to archive all certificates and CRLs in its own certification path, but it only has to do this for one path and not for the certification of every signature it has archived.

7.4 CONFIDENTIALITY

Some agency applications require confidentiality in addition to authentication, integrity, and non-repudiation security services. Confidentiality services must be provided through additional mechanisms such as data encryption. Some public key cryptographic algorithms such as RSA and El Gamal provide encryption directly. Others, such as the DSA, are intended to provide authentication, integrity and non-repudiation services only[6].

Applications are being developed that incorporate both DSS and encryption technology. In these applications, encryption may be provided by using either symmetric algorithms such as DES or public key algorithms such as El Gamal and RSA. Most common is a combination of DES for data encryption and public key cryptography for key exchange. Signatures are based upon message digests computed from the encrypted data or from the non encrypted data.

There are also several alternatives for key exchange protocols. Some of them can use the DSS keys or similar keys. Any two users can agree on a conventional key using a variant of the original Diffie-Hellman Key Exchange Protocol [7] or the El Gamal encryption algorithm [8]. These issues are beyond the scope of this study.

---

6 It is interesting to note that a general purpose DSA program with the correct sequence of calls and corresponding parameters can actually perform both RSA and El Gamal encryption. [34]

7.5 THE NEXT STEPS

This report has addressed the high level issues that must be considered in defining the PKI. Once the broad outlines of the infrastructure have been established, work must begin on filling in more of the details. This effort includes several tasks.

7.5.1 General Planning

Work should begin immediately on the overall plans to deploy the PKI. The step in the deployment of the PKI is to begin a small scale implementation of the PKI in order to gain useful working experience with the structure. The overall planning of the PKI must be done with one eye on the experience being gained in this first small scale implementation which is described in some detail in section 7.9.1.1. Additional steps that must be included in the overall planning process are also described below.

7.5.1.1 First Phase of PKI Implementation

It is important to gain some practical experience with digital signatures, certificates, CRLs and all the other aspects of the PKI. To that end, it is appropriate to implement part of the infrastructure as soon as possible. This effort will not only yield the needed working experience but will afford an opportunity for fine tuning the policy guidelines, for refining the cost analysis; and for verifying system interoperability. The first implementation would take approximately six months to put into place. An additional six to twelve months should be spent utilizing this small scale PKI to gain the necessary working experience.

Several agencies are already experimenting with the digital signature technology. These agencies make a natural starting point for the implementation of the PKI. These agencies and their applications include:

- IRS for tax filing
- Patent and Trademark Office (PTO) for patent applications
- FAA for airman medical certifications
- Defense Logistics Agency (DLA) for electronic submission of bids
- NASA for financial transactions

Although none of these applications interacts with any of the others, each expects the participation of U.S. citizens and private organizations. The first phase implementation, therefore, requires a CA or several CAs for the general public and a certification path from each individual to the agency CAs. Figure 7-1 shows a PCA and a CA for each of the listed agencies as well as a USPS PCA and several USPS CAs to serve the public. This may imply a modification of individual agency plans. Each agency may have envisioned a stand-alone system in which any outside individual who wishes to transact electronically would obtain a certificate from the agency's CA. In fact, till now no agency has considered implementing a PCA to facilitate handling certificates not produced by their own CA although USPS is considering the offering of certification services to the general public .

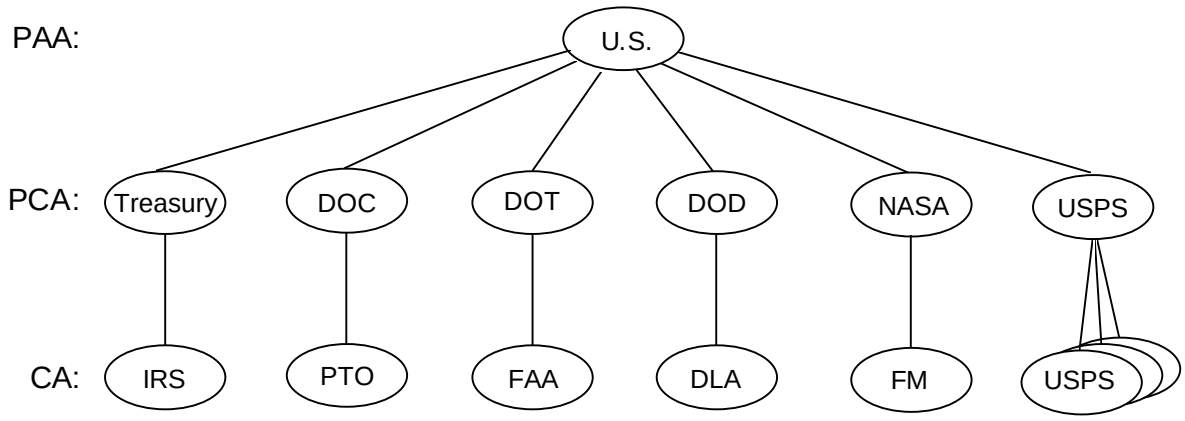


Figure 7-1. First Phase of Public Key Infrastructure

The implementation will be studied and analyzed on a ongoing basis to refine the results of this PKI study and to answer the following key questions.

- How much time and effort does it take to plan and implement the various aspects of the PKI?
- Is the PKI concept of operations workable in terms of user acceptance and operational implementation? Are there any legal issues associated with it?
- What kind of resources are required to operate the entities within the PKI? Does this validate the costing elements of this study?
- Are the PKI security policies complete? Are they implementable and enforceable? Do they provide necessary elements in dealing with legal issues?
- Are the user agreements in the areas of security and liability implementable and enforceable? Do they effectively deal with liability issues?

7.5.1.2 Security and Privacy

From the start, the PKI must have a unified security and privacy policy. Once established, the PAA must immediately commence work in developing this policy. At the same time, the several PCAs should begin developing their own policies. The PCA work must, of course, be coordinated with that at the PAA level, both so that the PAA has the benefit of input from the PCAs and so that the PCAs can remain within the broad security framework being established by the PAA. This nine month task will benefit from the experience gained in the first phase PKI implementation.

7.5.1.3 Statistics

Using information developed and experience gained in the small scale implementation, analysts can make a more accurate determination of the loads the PKI will have to endure. They will have to create an estimate of the number of users and to develop an approximation, based on the experience, to the volume of transactions those users will

generate. With this information and some targets on performance, they will be able to determine the size and number of CAs and PCAs the PKI will have to support. This task will be an ongoing endeavor.

7.5.1.4 Hardware Alternatives

Linked with the development of an approximation to the PKI size, is the selection of hardware for the needed computation and communication capabilities. What devices are available and appropriate as CAs and as PCAs? Which communication technologies will serve the PKI best? The hardware alternatives for the CAs and the PCAs need to be analyzed. Once some statistics on scaled-down PKI loads have been collected and extrapolations have been developed for the anticipated full-scale loads, a six month effort will produce a determination of the most appropriate hardware platforms for agencies of various sizes and differing workloads.

7.5.1.5 The Hardware vs. Software Choice

It is necessary to examine the alternatives for key generation, document signature, certificate generation, and CRL signature. What is the right balance between cost, security, and performance for each of these functions? Some "rules of thumb" should be developed to aid managers to choose between a hardware and a software implementation of each function in their own particular environment. This task will require about six months of study, once some experience has been gained with the small-scale implementation.

7.5.1.6 Availability

The immediate implementation of a directory service is of great importance. Additionally, the availability of the directory servers at all times becomes an issue as well. It is possible to implement a hot back-up for each server to assure that all certificates and all CRLs can always be obtained on request. Such an expedient is quite costly. It might be better to require each PKI entity to register each public key it holds with at least two CAs. Then if the directory server of one of the CAs becomes physically unavailable due to equipment failure or due to key compromise, the server associated with the other CA is still available. The question of CA/directory server availability needs further examination and would benefit from at least nine months of study.

7.5.1.7 Alternatives of CRL Distribution

Certificate revocation list distribution is by far the biggest cost driver associated with the operation of the PKI. Requiring that every request to the directory service for a certificate be accompanied by a similar request for the CRL on which that certificate may appear places an extremely heavy burden on the directory communications system. For the cost scenarios which most heavily load the PKI – large numbers of users and of signed messages – there are even extra equipment costs at each directory service agent. Other ways of dealing with the CRLs must be considered. CAs certifying many fewer users results in much smaller

CRLs. Thus, each CRL requested carries far less unwanted information. Perhaps few or no CRLs need be sent to users. If every directory service agent can be trusted to send out a certificate only if it does not appear on the pertinent CRL, much of these costs can be avoided. This requires that each service agent check its database whenever it receives a new CRL in the same way that each users checks his cache of certificates against each new CRL. Alternatively, CAs can retain responsibility for CRL distribution, sending each new CRL to its subscribers either directly or through the ORAs. They can even accept subscriptions to CRLs from other PKI users.

The technique chosen for handling and communicating certificate revocations will directly affect PKI operating costs and overall trust in terms of a public key being valid (i.e., not revoked). The trust will drive the acceptability of the partnership agreements and of the liability clauses. Detailed study of all options for managing and distributing CRLs and their cost, trust, and liability implications will help government develop the most prudent CRL management scheme. This effort must be undertaken immediately and completed within, at most, one year.

7.5.1.8 Concept of Operations.

Section 5 described alternative concepts of operation for the PKI and its components. As the overall plan for the deployment of the PKI progresses, specific details of a CONOPS for the PKI must be developed. Each PCA should be involved in this effort by clarifying how it would prefer the PKI to function. At the same time, it will develop its own extension to the overall CONOPS with refinements that limit choices or that implement functions to meet additional PCA unique requirements. The CONOPS at all levels will include provisions that implement the security and privacy policy. CONOPS requires immediate consideration before the first-phase implementation proceeds but must await a decision on the best method of handling CRLs. Thereafter, a six month review of the adequacies and shortcomings of the selected CONOPS should be undertaken once sufficient experience has been obtained with the pilot program. This should be done before the CONOPS for the full scale PKI is determined.

7.5.1.9 Schedule

As the plan begins to take form, it will be necessary to develop a schedule for its implementation. PKI managers will have to choose between a full scale deployment of much of the PKI or a gradual "ramp-up" to a large, far flung system. The schedule must allow for the completion of some of the other task listed above.

7.5.1.10 Procurement

As PKI hardware alternatives are established, the offices and agencies involved in the PKI must begin the procurement process. Facilities to monitor the procurement process and the PKI implementation must be in place.

7.5.1.11 Long Term Monitoring

Once the PKI has been deployed and is operational, it will require constant monitoring. It will be necessary to examine computer and communications system utilization along with the overall PKI usage to ascertain whether the infrastructure will benefit from extensions or upgrades and whether there are areas in which cost saving are possible. A parallel monitoring of the security and privacy policies is also needed. The examination should investigate the suitability of the policies and the adequacy of their implementations.

7.5.2 Cryptography Awareness

It would not be an exaggeration to say that success in implementing the PKI will largely depend on the awareness of digital signature technology and the PKI among personnel in key positions. These are the managers and technical leaders who can bring the technology into their respective organizations. Therefore, the key telecommunication, information security, and information technology personnel at various federal agencies need to be trained in the PKI, its role in electronic transactions, and its relationship to the field of cryptography.

7.5.3 Beyond the Executive Branch

This report has emphasized the Executive Branch of the Federal Government. Except for a brief mention GAO in table 4-2, the Judicial and Legislative Branches have not been included in the discussion of the PKI. The question of how these segments of the government can best be served by the PKI needs to be examined. The issue of whether the constitutionally separate branches can all be served by a single PAA also must be answered.

7.5.4 Forge Ahead

The continued development and ultimate deployment of the PKI can proceed. There appears to be some consensus that digitally signed electronic documents do satisfy most requirements for "signed" or "in writing." There is also reasonable confidence that liability can be controlled, particularly with the enactment of specific legislation. Even before the legislation is enacted, written contracts can expressly establish appropriate liability and warranties. They can be used until the PKI policies are developed and are recognized as binding. The certificate authorities are to be trusted entities and careful consideration must be given to the advantages to be derived from imposing liability on them consistent with such a role.

The PKI will have to communicate with and rely on diverse federal organizations and commercial parties to support a viable infrastructure. It is necessary to develop mechanisms to ensure the flexible and productive cooperation and interface between the private and public sectors. These mechanisms must incorporate issues of authorization, delegation, restriction, levels of service, policy encoding and signature purpose. This requires consultation and cooperation between the PKI, many regulatory bodies, and several alternate certificate distribution infrastructures.

It is crucial that the Federal Government take an active role in the development of a global public key infrastructure. It should do this for its own benefit and for that of the U. S. business and trade interests. The task should not be left exclusively in the hands of private and international standards organizations. Such entities do not necessarily promote the technical and operating requirements of the Federal Government.

These efforts should be ongoing from the outset. They are an immediate consequence of any decision to press ahead with the deployment of the PKI. In a similar fashion, the remaining recommendations also derive from the initial one. They should all be addressed, aggressively and in parallel.

7.5.4.1 Study and Develop Needed Legislation

It is an enormous risk but the development and enactment of the statutes and related regulations will likely take several years. This effort must be begun at once. To ensure that the PKI becomes a viable resource, legislation will ultimately be required. A fully functional, ubiquitous PKI is not specifically supported by existing legislation. A number of questions must be resolved. There is a need to define government liability for all foreseeable damages deriving from its negligence. Appropriate presumptions concerning the nature and effectiveness of the PKI mechanisms must be developed. This can lead to possible relaxation of the amount and nature of evidence required to prove the adequacy and integrity of a digital signature that is supported by a PKI certificate. These presumptions extend to the use of smart card devices to protect the confidentiality of private keys. There is a serious lack of business practices, of laws, and of viable alternatives. Not only should the legal efficacy of smart card technology be studied but pilot programs should experiment with card technologies to prove their viability operationally, financially and legally. These moves may be strengthened by passing new statutes specifically relating to PKI functioning. Until now, computer crime laws have focused on access rather than on authentication and integrity. There is a need for a rigorous consideration of the strength and weaknesses of current computer crime laws to determine the need for their legislative reform. Consumers will ultimately demand PKI services. This demand may bring consumer protection laws to bear on the PKI. The legal implications of consumer use of the infrastructure require rigorous study.

In parallel with this effort, there is a need to examine the legal issues involved in the choice of which governmental or quasi-governmental agencies should manage the PKI. Can or should an agency of the Executive Branch, an independent government-sponsored agency or a private contractor be selected to perform this function. Attention must be given to issues of independence, disinterestedness, limited liability and openness of government records. Questions of scaling should also be addressed. In the temporary absence of legislation, written agreements are to be employed. They may well become an impediment to the stability of the PKI. The behavior of the individuals who are actually responsible for the day to day running of the PKI entities must be circumspect. During the pilot stage of the PKI deployment, rules that provide assurance of the trustworthiness of employees should be evaluated. Inconsistent and insufficient standards of conduct for both government employees and private citizens, as well as inadequate criminal sanctions, will require

corrective action. Ultimately, clearly articulated rules incorporating swift, consistent and certain punishment for breach of fiduciary obligations must be developed and implemented.

7.5.5 Develop Multidisciplinary Development Group

As a part of or in addition to the current technical and administrative coordinating group or groups, the PKI will benefit from an interdisciplinary task force to oversee legislative initiatives, insurance, and policies. The same people can provide advice and review of private sector requirements as well as of consumer interests.

Members of this group should cooperate or oversee a PKI risk analysis. The PKI is contemplated to support many differing types of transactions. Risk analysis should, henceforth, assume an interdisciplinary approach. A legal risk analysis should be included with the more common technical and security risk analyses for the three areas are closely related. However, a precise identification and quantification of risk will require a trial period to see what forms legal concerns and relief take.

It is advantageous that the interdisciplinary development committee organize the educational needs. Among government and private executives at all organizational levels, there is an almost universal ignorance of even the basic concepts certificate-based digital signature technology. It is recommended that the government not wait for more computer-literate executives but begin immediately the task of educating and training the current generation of managers and potential PKI users. The program should include audit, legal and administrative issues and solutions along with the technical issues and solutions as well as much needed introduction to digital signature concepts. Workshops to assist in the development of appropriate policies, regulations, and guidelines are needed. (Some of the legal issues are discussed in appendix J.)

7.5.6 Liability

In the apportionment of liability, federal and private insurance programs may be of assistance. Furthermore, the flexibility of the insurance example offered by the USPS, should be investigated for its applicability to the PKI. That model allows the user to determine how much risk he will endure and how much message-specific insurance he wishes to purchase.

Pilot infrastructures should include the satisfaction of various legal requirements to maximize legal experience. These may include: satisfaction of criteria intended to produce computer-based analogs of signatures, notorial acknowledgments and negotiability, other specific requirements of specialized business and government documents, and similar legal demands. Most importantly, in those pilot programs for which it is appropriate, issues relating to liability must be monitored and different schemes tried.

Questions concerning disclosures, notifications and warnings to and among users, especially among government users, await answers. Issues that are of concern include: choice of media (Federal Register, mutual agreements, etc.) and whether the medium will be

paper or electronic. The policies of various federal agencies are inconsistent. Additionally, potential consumer usage of the PKI exacerbate the uncertainty. It is important that rules that are explicit, comprehensive and authoritative be developed.

LIST OF REFERENCES

1. Furlong, Judith A., 5 October 1992, *User Requirements for the Public Key Infrastructure (Draft)*, WP-92W0000378, The MITRE Corporation, McLean, VA.

2. Berkovits, Dr. Shimshon, 2 November 1992, *Legal Requirements for the Public Key Infrastructure (Draft)*, WP-92W0000431, The MITRE Corporation, McLean, VA.

3. Galitzer, Shari B., 16 October 1992, *Technical Requirements for the Public Key Infrastructure (Draft)*, WP-92W0000407, The MITRE Corporation, McLean, VA.

4. Baum, Michael. S., July 1993, *Federal Certification Authority Liability and Policy Issues,* Technical Report, Independent Monitoring, Cambridge, MA.

5. *Federal Information Processing Standard Digital Signature Standard (DSS),* 1 May 1994, FIPS PUB 186, U. S. Department of Commerce/National Institute of Standards and Technology, Gaithersburg, MD.

6. Sudia, Frank, January 1994, private communication.

7. Diffie, Whitfield, and Martin Hellman, November 1976, "New Directions in Cryptography," *IEEE Transactions on Information Theory*, Vol IT-22, Number 6, pp. 644-654.

8. El Gamal, Taher, 1985, "A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms," *Advances in Cryptography: Proceedings of Crypto'84*, Springer-Verlag, Berlin, pp. 454-464.

9. *Federal Information Processing Standards Publication , Secure Hash Standard (SHS),* 11 May 1993, FIPS PUB 180, United States Department of Commerce/National Institute for Standards and Technology, Gaithersburg, MD.

10. CCITT, Geneva, 1989, *Data Communication Networks: Directory*, Recommendation X.500 - X.521, Blue Book, Volume VIII–Fascicle VIII.8.

11. ISO/IEC, 25 December 1991, *Information Technology - Open Systems Interconnection - The Directory: Authentication Framework*, ISO/IEC 9594-8 (CCITT Recommendation X.509).

12. Linn, J., February 1993, *Privacy Enhancement for Internet Electronic Mail: Part I: Message Encryption and Authentication Procedures*, RFC 1421, IAB IRTF PSRG, IETF PEM WG.

13. Kent, S., February 1993, *Privacy Enhancement for Internet Electronic Mail: Part II: Certificate-Based Key Management*, RFC 1422, IAB IRTF PSRG, IETF PEM WG.

14. Balenson, D., February 1993, *Privacy Enhancement for Internet Electronic Mail: Part III: Algorithms, Modes, and Identifiers*, RFC 1423, IAB IRTF PSRG, IETF PEM WG.

15. Chalks, B., February 1993, *Privacy Enhancement for Internet Electronic Mail: Part IV: Key Certification and Related Services*, RFC 1424-C, IAB IRTF PSRG, IETF PEM WG.

16. Lambent, Paul A., October 1988, *Architectural Model of the SDNS Key Management Protocol*, Proceedings of the National Computer Security Conference.

17. ISO/IEC, February 1993, *Information Technology - Telecommunications and Information Exchange Between Systems - Network Layer Security Protocol*, ISO/IEC 11577.

18. ISO/IEC, 13 January 1993, *Information Technology - Open Systems Interconnection - Transport Layer Security Protocol*, ISO/IEC 10736.

19. Secure Data Network System, *SDNS Message Security Protocol (MSP)*, 23 November 1993, SDN.701, Revision 2.1.

20. Postal, J. B., August 1982, Simple Mail Transfer Protocol, RFC 821.

21. Secure Data Network System, *SDNS Directory Specifications for Utilization with SDNS Message Security Protocol*, 23 November 1993, SDN.702, Revision 2.4.

22. Accredited Standards Committee X9, December 6 1992, *Public Key Cryptography Using Irreversible Algorithms for the Financial Services Industry: Part 3: Certificate Management for DSA*, Draft, American National Standard X9.30-199X, American Bankers Association.

23. *Department of Defense Trusted Computer System Evaluation Criteria (TCSEC)*, December 1985, Department of Defense Standard DOD 5200.28-STD.

24. *Federal Organization Service*, May/June, 1993, Carol Publishing Co., Washington, DC.

25. *The World Almanac*, 1993, Pharisee Books, St. Martin's Press, New York, NY, p. 158.

26. Peril, Henry H., Jr., "The Electronic Agency and the Traditional Paradigms of Administrative Law," 44 *Admin. Law Rev.* 79 (Winter 1992).

27. *Admissibility of Electronically Filed Federal Records as Evidence: A Guideline for Federal Managers and Counsel*, October 1990, Systems Policy Staff, Justice Management Division, U. S. Department of Justice.

28. Decision. of the Comptroller General. of the US., 13 December 1991, *Matter of National Institute of Standards and Technology-Use of Electronic Data Interchange Technology to Create Valid Obligations*, File 245714.

29. "Korean Act on Promotion of Trade Business Automation (1992) (Law Enacted December 31, 1991) Art. 2.8 (Definitions, “Electronic Signature”)," reprinted in *UN/ECE/TRADE/WP.4/R.872*, 4 August 1992.

30. Legal Issues Committee of the Acquisition Task Group, CALS/EC Industry Steering Group, 10 November 1991, *Report on Potential Legal Issues Arising from the Implementation of CALS by the DOD*.

31. Bureau of Justice Statistics, June 1992, *Report of the National Task Force on Criminal History Record*, NCJ-135836.

32. Legal and Business Controls Task Force, Accredited Standards Committee X12, 1990, *1990 Survey*, American National Standards Institute.

33. EDI and Information Technology Division, Section of Science and Technology, June 1992, *Model of Electronic Payments Agreement and Commentary*, §7 comment 5, American Bar Association.

34. Stonier, Bruce, 1994, *Applied Cryptography: Protocols, Algorithms, and Source Code in C*, John Wily and Sons, New York, NY, PP 310-311.

BIBLIOGRAPHY

Baum, Michael S., November 1992, "Linking Security and the Law of Computer-Based Commerce," Paper presented at the *Workshop on Security Procedures for the Interchange of Electronic Documents*, Gaithersburg, Maryland.

CCITT (The International Telegraph and Telephone Consultative Committee), Geneva 1989, *Data Communication Networks: Directory*, Recommendation X.500 - X.521, Blue Book, Volume VIII - Fascicle VIII.8.

Chokhani, Dr. Santosh, and David L. Gill, August 1992, *Public Key Infrastructure Study: Task Plan*, WP-92W0000308, The MITRE Corporation, McLean, VA.

IEEE 802.10 Reporter of the LAN Security Working Group, 12 September 1989, *Standard for Interoperable Local Area Network (LAN) Security (SILS)*, P802.10/D6, Institute of Electrical and Electronic Engineers, Inc.

Information Processing Systems - Open Systems Interconnection - Basic Reference Model - Part 2: Security Architecture, ISO 7498-2: 1989 (E), International Standards Organization.

Kaliski, B., 3 June 1993, *An Overview of the PKCS Standards*, RSA Data Security, Inc.

Standard for Interoperable Local Area Network Security (SILS) Part B-Secure Data Exchange, 16 November 1990, P802.10b/D6, Institute of Electrical and Electronic Engineers, Inc.

APPENDIX A

TERMS AND DEFINITIONS

In this appendix, the terminology associated with a public key infrastructure, and which was used in this report, is presented. This appendix should assist the readers of this report in defining any terms with which they are unfamiliar. Many of these definitions are borrowed from the standards that have been reviewed under this study. Reference to the standard where more details can be found is given in parenthesis at the end of the definition.

Certificate: The document that binds an entity's unique name and its public key, together with some other information, rendered unforgettable by digital signature of the certification authority that issued it. (CCITT X.509|ISO/IEC 9594-8)

Certificate Authority (CA): An authority trusted by one or more users to create and sign certificates. (CCITT X.509|ISO/IEC 9594-8)

Certification Path: An ordered sequence of certificates of objects in the which, together with the public key of the initial object in the path, can be processed to obtain the public key of the final object in the path. (CCITT X.509|ISO/IEC 9594-8)

Ciphertext: Data produced through the use of encipherment. The semantic content of the resulting data is not discernible.

Cryptography: The discipline that embodies principles, means, and methods for the transformation of data in order to hide its information content, prevent its undetected modification and/or prevent its unauthorized use. (ISO 7498-2)

Decipherment: The reversal of a corresponding encipherment.

Decryption: See decipherment.

Digital Signature: Data appended to, or a cryptographic transformation of, a data unit that allows a recipient of that data unit to prove the source and integrity of the data unit. It protects against forgery, even by the recipient. (ISO 7498-2)

The Directory: A repository of information about objects and which provides services to its users that allow access to the information. (CCITT X.501|ISO/IEC 9498-2)

Directory Information Base (DIB): The complete set of information to which the directory provides access and which includes all of the pieces of information that can be read or manipulated using the operations of the directory. (CCITT X.501|ISO/IEC 9498-2)

Encipherment: The cryptographic transformation of data to produce ciphertext. (ISO 7498-2)

Encryption: See encipherment.

Hash Function: A many to one (mathematical) function that maps values from a large (possibly very large) domain into a smaller range. A "good" hash function is such that the results of applying the function to a (large) set of values in the domain will be evenly distributed (and apparently at random) over the range. It should be computationally infeasible to find two elements of the domain with the same hash or to find one element with a given hash. (CCITT X.509|ISO/IEC 9594-8)

Key: A sequence of symbols that controls the operation of encipherment and decipherment. (ISO 7498-2)

Key Pair: In a public key cryptosystem, the set of keys which consists of a public key and a private key that are associated with an entity.

One-way Function: A (mathematical) function f that is easy to compute but which, for a general value of *y* in the range, is computationally difficult to find a value *x* in the domain such that f(*x*) = *y*. There may be a few values *y* where finding *x* is not computationally difficult. (CCITT X.509|ISO/IEC 9594-8)

Public Key: In a public key cryptosystem, the key of the entity's key pair that is publicly known. (CCITT X.509|ISO/IEC 9594-8)

Private Key: In a public key cryptosystem, the key of the entity's key pair that is known only to that entity. (CCITT X.509|ISO/IEC 9594-8)

Relative Unique Name (RUN): A set of locally unique, attribute value assertions concerning a particular entity used to identify that entity. (c.f., relative distinguished name, CCITT X.501|ISO/IEC 9498-2)

Security Policy: The set of rules laid down by the security authority governing the use and provision of security services and facilities. (CCITT X.509|ISO/IEC 9594-8)

Trust: Generally, an entity can be said to "trust" a second entity when it (the first entity) makes the assumption that the second entity will behave exactly as the first entity expects. This trust may apply only for some specific function. The key role of trust in the authentication framework is to describe the relationship between an authenticating entity and a certification authority; an authenticating entity shall be certain that it can trust the certification authority to create only valid and reliable certificates.
(CCITT X.509|ISO/IEC 9594-8)

Unique Name: One of the names of an object, formed from the sequence of the Relative Unique Names (RUNs) of the object entry and each of its superior entries. (c.f., distinguished name, CCITT X.501|ISO/IEC 9498-2)

APPENDIX B

DIGITAL SIGNATURE STANDARD

In accordance with the draft Digital Signature Standard (DSS), given a private key, a user creates a signature to a document in two steps. First, treating the digitized document simply as a sequence of bits, a message digest is produced by applying the Secure Hash Algorithm (SHA). This algorithm is found in FIPS 180, the Secure Hash Standard (SHS) [9]. Hashing folds a document or message of any length onto itself to create a 160 bit digest. Changing a single bit of the data modifies at least half of the resulting digest bits. Furthermore, it is computationally infeasible to find two meaningful messages that have the same digest. Similarly, given a random 160 bit sequence, it is just as computationally infeasible to find a meaningful message with that sequence as its digest.

In the second step, the signer treats the message digest bit sequence as a 160 bit number. The signing parameters consist of three numbers: *p*, *q*, and g. *p* is between 512 and 1024 bits in length, inclusive; its length determined by the degree of security the signer needs. *q* is a 160 bit prime number that divides evenly into *p*-1. *g* is chosen so that *q* is the smallest exponent to which *g* can be raised to yield 1 mod *p*. The private key *x* lies between 1 and *q* - 1 and is used for signing. The signer chooses a large random number less than *q*. He combines the random number, *x* and the message digest in a mathematical computation. He reduces this result by dividing by *p*. He throws away the quotient and divides the remainder by the 160 bit number *q*. The 160 bit remainder after the second division is half of his signature. The other half is computed from the random number he chose, with the result also reduced to 160 bits by the same process of dividing by *q* and keeping the remainder. He appends the 320 bits of signature to the original document or message. The process is shown graphically in figure B-1.

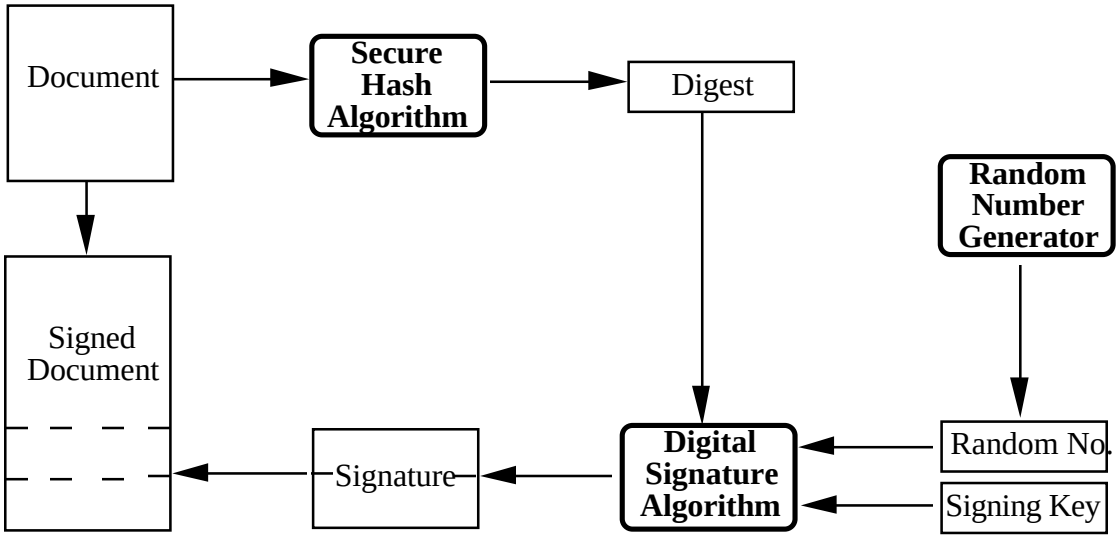


Figure B-1. The DSS Signature Process

Verification is accomplished using the supposed signer's public key. This is his public key and includes a number *y* derived from *x* and the three numbers *p*, *g*, and *q* that were used in the signing process. The verifier must first obtain these numbers. It is possible that the numbers *p*, *q*, and *g* are common to a group of users. In that case, the verifier may share the same parameters with the signer. If that is so, he already has them as part of his own signing process and he need only obtain *y*. Otherwise, he must obtain all four numbers *p*, *q*, *g*, and *y*.

When he obtains the public key, he must assure himself that he has the right key. He must be sure he has been given the public key that is associated with the claimed signer. If he is tricked into accepting a false verifying key, then he can be tricked into accepting a signature as belonging to the signer claimed, when in reality the signature was formed by an impostor. To assure that the public key he receives is the correct one, he accepts only keys that have been certified by a trusted Certificate Management Authority (CA). The key is contained in a certificate that also holds the identity of the individual with whom the key is associated. That certificate has the CA's digital signature, thus binding the user's public key to his identity. The verifier is given the CA's public key when he enrolls in the signature system. The role of the CA is similar to that of a notary. It verifies the identity of an individual and certifies his association with a key while a notary verifies his identity and certifies his association with a handwritten signature.

Once the verifier is satisfied he has the correct public key, he recomputes the message digest from the received document. Using the digest, the public key and the two halves of the received signature, he verifies whether they satisfy the verification equation given in the DSS. This computation requires two reductions using the same divisors as were used in the

signing process. If the remainders satisfy the given equation, the signature is accepted as valid; otherwise, it is rejected. This process is depicted in figure B-2.

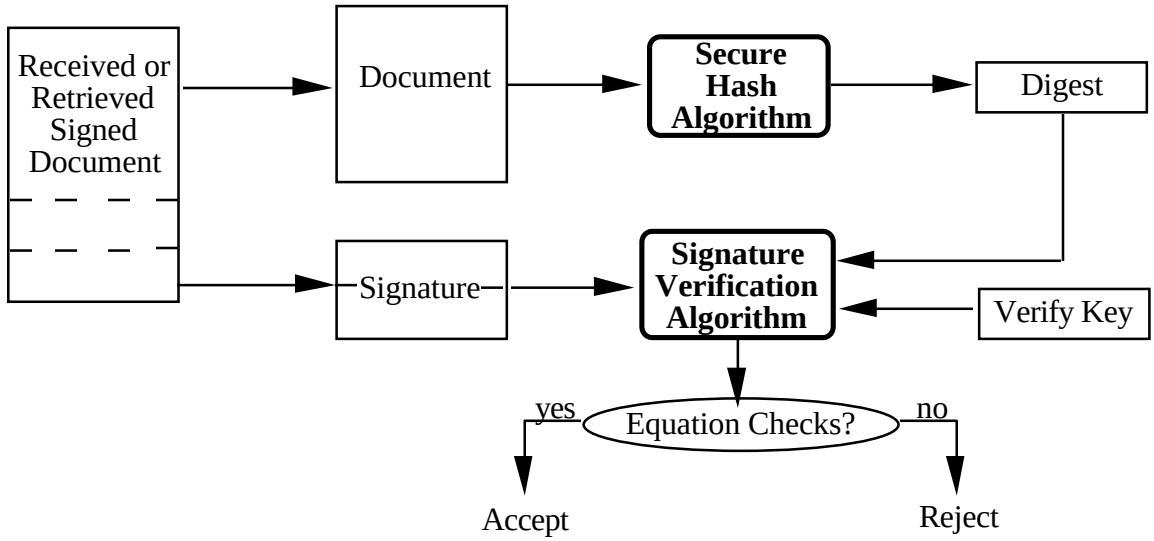


Figure B-2. The DSS Signature Verification Process

APPENDIX C

APPLICATIONS OF DIGITAL SIGNATURES

Digital signatures can be used in a number of applications, specifically in the automation of any paper process which required handwritten signatures. Digital signatures may also be used in a number of applications that previously did not require handwritten signatures. In general, any application that requires sender authentication, message integrity, and/or sender non-repudiation services is a potential application in which digital signatures may be used.

In this appendix, digital signature applications identified through the interview process and through attendance at several conferences are described. The applications described represent applications currently being used by agencies, applications being developed under prototype efforts, and applications that have been proposed for future implementation. Most of these applications are implementing the DSS. As digital signature technology becomes more widely understood and utilized, it is expected that other uses and applications will be developed.

The applications are presented in seven broad categories: Electronic Office, Financial Transactions, Electronic Filing, Software Protection, Authentication, Medical Information, and Other. The Electronic Office section describes applications used in the agency's day-to-day internal business. In the Financial Transactions category, applications involving the transfer of money or the authorization of the transfer of money are described. The Electronic Filing category discusses applications related to the filing of tax returns, the filing of patent applications, the reporting of large currency or stock transactions, etc. The Software Protection category describes applications related to the distribution and maintenance of software. The Authentication category encompasses applications such as cryptography-based authentication schemes. The Medical Information category contains information about medical industry applications. Under the Other category, applications that do not fall under the other categories are described.

C.1 ELECTRONIC OFFICE

Many of the day-to-day activities conducted within the Federal Government are being converted from paper processes to electronic processes in an effort to develop a "paperless office." Many of these applications required written signatures, and even in their electronic format many still require some type of signature. For some applications, documents can still be generated electronically and the hard copy may be signed by hand. This is often the case with memos and letters. However, other applications that have been transitioned to the electronic environment may be made more efficient if all paper is eliminated; therefore, the use of digital signatures within such applications is desirable.

Electronic applications utilizing digital signatures include time reporting, travel orders, travel expense vouchers, purchase requisitions, and purchase orders. Certain types of

contracts may be electronically generated and signed. Work requisitions and job assignments that are electronically generated could also benefit from digital signatures.

Electronic mail may also need to be digitally signed, especially in cases where sensitive information is being transmitted and security services such as authentication, integrity and non-repudiation are desired. For example, the DMS is using MOSAIC to provide security services to electronic mail messages. The DSA which is specified in DSS, has been incorporated into MOSAIC and is used to digitally sign mail messages as well as public key certificates.

C.2 FINANCIAL TRANSACTIONS

The financial transactions category encompasses a number of areas in which money is being transferred directly or in exchange for services and goods. Areas that fall under the umbrella of financial transactions include Electronic Funds Transfer (EFT), purchasing, billing, payments, and procurement. Many of these applications could benefit from the use of digital signatures.

Many EFTs require that security services such as authentication, integrity, and non-repudiation be provided for these transactions. Digitally signing EFTs would be a way to provide these services. Many documents authorizing payment require handwritten signatures. If the process that generates these documents is automated, then digital signatures may replace the handwritten signatures. Purchase requisitions, purchase orders, and travel expense vouchers are examples of this type of documents.

Finance–related applications exist within the procurement side of agencies. Vendor/Contractor proposals may be electronically signed. The agency may then verify the signature on the proposal to determine its authenticity. A prototype under development in one agency allows procurement officers to electronically create and sign payment authorizations. These authorizations are sent to the agency's accounting office, which verifies the signature and makes the payment.

C.3 ELECTRONIC FILING

Many federal agencies, especially regulatory agencies, require that individuals, corporations, etc., report certain information to them. Contracting officers expect certain mandated certificates from contractors. This information is often filed through the submission of a written form and usually requires a written signature. Some information reporting processes are being automated so that filings may be accomplished electronically. Digital signatures may be used in a number of these automated applications to replace written signatures and to provide authentication and integrity services.

One of the largest information submission processes of the Federal Government is the tax process. Many activities related to the payment of taxes and the request for tax-related

information require signatures. The IRS is converting many of these processes to the electronic world and is considering digital signatures for use within these processes. The IRS has several prototype efforts under development that utilize digital signatures generated by using DSA.

In one prototype Currency Transaction Reports (CTRs), which report transactions of over ten thousand dollars, are digitally signed and electronically filed with the IRS. The IRS verifies the digital signatures to determine the authenticity and integrity of the report. During the conduction of the prototype, the original paper process is run in parallel. The content of the electronic CTRs is compared with the paper version after the digital signature is verified.

TAXLINK is an IRS prototype that allows businesses to pay withholding taxes electronically. At present, businesses pay withholding taxes by going to the bank with a check that the bank forwards to the IRS. The prototype allows the business to prearrange with their bank an electronic transfer of money from their account to a Treasury Department account.

More and more tax returns are being electronically generated. Today, individuals may electronically file their tax returns with the IRS through tax preparers. At present, the electronic filing is followed by the submittal of a paper form that is signed by the taxpayer, along with any required attachments or schedules. In the future, the electronically generated tax returns may be digitally signed. The taxpayer may send the digitally signed electronic form to the IRS directly or through a tax preparer. At the present time, tax preparers send the individual tax forms to the IRS in bulk transactions. The IRS is considering requiring the preparers to sign the bulk transactions digitally to provide added assurances.

Information requests received by the IRS also require handwritten signatures. Individuals may request information on the status of their return. They may also request copies of old tax returns. State tax agencies also request information on taxpayers, such as address information, from the IRS. The Direct Electronic Access (DEA) prototype allows state tax agencies to dial up the IRS to request taxpayer information. The request is encrypted using the Data Encryption Standard (DES) and digitally signed using DSA. The digital signature allows the IRS to ensure that the request was sent from an authorized individual within the state tax agency.

Other examples of electronic filings with the Federal Government include: tariff information filed with the Federal Maritime Commission by shipping companies; financial reports filed with the Securities and Exchange Commission (SEC) by corporations; equal rights, minority subcontractor and other certifications in support of contract performance; and test results from clinical trials of new drugs submitted to the Food and Drug Administration (FDA). Both of these are potential digital signature applications. In addition, special patent applications that require notarized signatures could be digitally signed and electronically filed.

C.4 SOFTWARE PROTECTION

Digital signatures may be used to protect software. By signing the software, the integrity of the software may be assured. This method may be employed when distributing software. The signature may be verified when the software is installed to ensure the software was not modified during the distribution process. Digital signatures may be used on system software. When a machine is turned on, or when software is reloaded, the signature on the software may be verified to ensure the software has not been modified.

Use of digital signatures to ensure the integrity of software is also a means of virus protection. The introduction of a virus would cause modification to the software. If the software was digitally signed, the modification would be detected during the signature verification process.

C.5 AUTHENTICATION

Digital signatures can be used in cryptography-based authentication schemes to sign either the message being authenticated or the authentication challenge used in the scheme. X.509 strong authentication is an example of an authentication scheme that utilizes digital signatures.

Digital signatures may be used to generate one-time passwords. Digital signatures may even be used to replace passwords. NASA is considering replacing passwords on their computer systems with an access control system which utilizes DSS.

Strong authentication schemes are used to authenticate remote logins to agency computer systems from computers such as laptops. To provide additional assurance of the authenticity of the remote communications, each message or file sent from the remote computer to the agency computer could be digitally signed.

C.6 MEDICAL INFORMATION

The medical industry has a number of applications that could benefit from the use of digital signatures. Medical records being stored in computers and being transmitted could be digitally signed to provide integrity and authentication services. Prescriptions could be electronically generated and digitally signed to prove their authenticity.

The Aeromedical Certification Subsystem of the Federal Aviation Administration (FAA) is being developed to allow doctors, who certify that individuals are mentally and physically able to be airline pilots and airplane mechanics, to submit the certification forms to the FAA electronically. The prototype, which is currently being tested, uses DES encryption to maintain the confidentiality of the medical information on the form in transit. The paper version of the form is still submitted to the FAA. The FAA is considering adding DSS-

based signatures to the system to replace the handwritten signature and possibly eliminate the need for the separate submission of the paper certificate.

C.7 OTHER APPLICATIONS

The five categories of applications listed above encompass the major areas in which digital signature technology is being applied. However, many other application areas are considering using digital signatures to provide authentication, integrity, and/or non-repudiation services. This section provides a sampling of these other applications.

Digital signatures may be used to sign data being stored on computers in order to protect the integrity of the data. Bitmap images may be digitally signed to ensure their integrity. In the law enforcement community, digital signatures are being considered for sealing electronic evidence. NIST requires that its scientists keep notebooks. NIST is planning to allow these notebooks to be kept electronically and is considering using DSS signatures to ensure the integrity of the notebooks. Test results from various experiments can also be digitally signed to insure their integrity and authenticity. Permanent archive records may also be digitally signed to ensure their integrity.

Digital signatures can also be used in access control applications. Access control information may be digitally signed to ensure its authenticity and integrity.

Digital signatures may be used in date and time stamp applications to prove the authenticity and integrity of the stamp. The United States Postal Service (USPS) has two prototype systems, Electronic Postmark (EPM) and Electronic Postmark-Plus (EPM-PLUS), which prove digitally signed time and date stamp services.

Digital signatures are being considered for use in electronically locking and unlocking safety deposit boxes. These safety deposit boxes could be located in hotel rooms, banks, etc.

A completely electronic betting or electronic lottery system could be developed that uses digital signatures. The electronic bet and the authorization for the EFT to pay for the bet can both be digitally signed to provide authentication services.

Digital signatures could also be used to protect the intellectual property rights of electronic documents. A label that contains information such as the title, author, publisher, and date can be affixed to an electronic document and sealed with a digital signature to ensure its integrity.

APPENDIX D

SUMMARY OF PKI REQUIREMENTS

A number of user and technical requirements were identified during the interview phase of the PKI Study. Some of the requirements apply to the infrastructure as a whole, while others apply to specific parts of the infrastructure, such as the certificate authorities and the certificates.

In this appendix, the user and technical requirements for the PKI are identified. These requirements are presented in eight categories that reflect the nature of the requirements. The General Infrastructure category describes requirements that apply to the entire infrastructure. The Key Generation and Distribution category defines any requirements imposed on how the public/private key pair is generated and distributed. Within the Certification Authority (CA) category, requirements describing the types of services that the CAs within the PKI need to supply are identified. These requirements also apply to the PAA and PCAs within the PKI. In the Organizational Registration Authority (ORA) category, the requirements associated with the ORAs within the PKI are discussed. The Directory category contains the requirements imposed on a directory service in order to support the PKI. The Certificate category describes requirements for the infrastructure related to the different types of certificates identified by the users, as well as requirements for the format of the certificates. The Certificate Revocation List category describes requirements on the format of CRLs. In the user category, user types (e.g. non-human users) are discussed.

Within each category section, the user or technical need that led to the requirement on the infrastructure is described first. Then the actual infrastructure requirement is stated. The requirements are denoted by bullets and are in bold face for easy identification.

D.1 GENERAL INFRASTRUCTURE REQUIREMENTS

In this section, the general requirements for the PKI are specified. These requirements apply to the entire PKI as a whole. However, there may be implied requirements on the components of the PKI in order for the general requirement to be met.

Although not specified as a requirement, the cost of the PKI was considered in all aspects of the PKI development. The PKI was designed and requirements for the PKI were specified with cost in mind; the goal being to keep the cost of the PKI as low as possible. Further information on the cost of the PKI can be found in section 6 and appendix I of this document.

D.1.1 Trust

Digital signatures are used to provide authentication, integrity, and non-repudiation security services. Provision of these services hinges upon the proper association between the

users and their public/private key pairs. Entities verifying digital signatures need to be assured that the public/private key pair was generated in a secure manner and that the binding of the public key with the user identity was done properly. These entities also need to obtain certificates from a source they trust sufficiently and in a manner that is demonstrably secure enough for their application.

The PKI can prove its trustworthiness by establishing a security policy that describes procedures for identifying and authenticating users and generating and distributing key pairs and certificates in a secure and trusted manner. Components of the infrastructure may need to implement security precautions to assure users that the PKI can be trusted to perform these functions correctly. These precautions may be technical, procedural, etc., in nature. Variations in the trust levels associated with PKI components may exist and will depend upon the functions performed by the PKI component, but general PKI policy will establish a minimal level of trust that all components must meet.

- **The Public Key Infrastructure and its components must be trusted entities.**

D.1.2 Ease of Use

Agency users also noted that the incorporation of a digital signature capability within an application should not make applications more difficult to use. The use of digital signatures also must not be burdensome to the users. The concern is that, if an application becomes more difficult to use, users will not use that application. Alternatively, users may turn to others who understand how to use the application for assistance and give away their private keys. This event would violate the concept of having only the user knowing or possessing the private key and, thus, weaken the security provided by the utilization of digital signatures. It may also introduce the complications implied by the laws of agent and fiduciary. It should be noted that user friendly PKI and application program interfaces require sophisticated programs and designs.

Through the agency interviews, it became known that most users communicate with a specific set of persons on a frequent basis. Keeping this in mind, the use of digital signatures can be made more user friendly and faster for users, if applications utilizing digital signatures allow users to cache certificates and/or certification paths for users with whom they communicate often.

Federal agencies have established practices for conducting their business. These practices include assigning certain persons the authority to sign documents for certain contracts or for certain monetary amounts. If digital signatures replace handwritten signatures in these instances, then business practices may need to be modified to accommodate the use of digital signatures. The fewer changes made to established practices will make applications using digital signatures more acceptable to users.

- **The design and operation of the Public Key Infrastructure should not make applications which utilize digital signatures more difficult or more burdensome to use.**

D.1.3 Interoperability

Federal agencies communicate with a wide variety of entities outside the Federal Government. These entities include corporations, individual citizens, foreign nations, international organizations, etc. The communications include funds transfers, contract negotiations, mandated filings, and requests and applications. These external entities may use digital signature algorithms other than the algorithm used by federal agencies. Alternatively, they may utilize the same digital signature algorithm, but obtain their certificates from an infrastructure other than the PKI. In order for federal agencies to communicate with entities using different algorithms and infrastructures, the PKI will need to interoperate with these other infrastructures. Interoperation will require some type of cooperation between PKI and other infrastructures that allow entities in one infrastructure to verify signatures produced in the other infrastructure.

- **The Public Key Infrastructure should interoperate with different public key infrastructures.**

D.1.4 Naming Convention

To provide meaningful authentication services in a single CA domain, each entity in the domain needs to be uniquely identified. Obviously, it is impossible for an authentication scheme to differentiate between two entities with the same name. Bearing this in mind, all the users and trusted entities within the PKI must be uniquely identified by a unique name.

- **The Public Key Infrastructure must have an naming convention which ensures that the unique name of each entity is unique.**

D.1.5 Scalability

Most agencies were not able to specify how many of their employees would use digital signatures, since many applications that would use digital signatures are either in the concept or prototype phase. However, it is expected that the number of users of the PKI will grow as digital signature technology becomes more widely implemented.

- **The design of the Public Key Infrastructure must be scalable in order to accommodate a growing number of users and their associated certificates.**

D.1.6 Flexibility

The different alternatives which are being developed under the PKI Study may be implemented in a variety of ways using different technologies. If federal agencies elect to implement portions of the PKI in different manners, then steps must be taken to ensure that the different implementations will work together.

The PKI will be implemented in the next few years and will likely use the most advanced technology available. However, the PKI is expected to be in place for a number of years thereafter. Over this time period, technology will change and improve and the PKI will want to take advantage of the changes and improvements of technology.

- **The design of the Public Key Infrastructure must be flexible to allow different implementations to work together and to allow for changes and improvements in technology.**

D.1.7 Standards Compliance

Computer systems within the Federal Government must be compliant with a number of computer and networking standards. Some standards apply to the entire Federal Government. For example, all Federal Government computer systems must be compliant with the standards defined in the Government Open System Interconnection Profile (GOSIP). There are also other Federal Information Processing Standards (FIPS) that have broad application. Specific agencies may require compliance with additional standards. Depending upon the implementation of the PKI, parts of the infrastructure may be located at an agency. This portion of the PKI will need to comply with the agency's standards.

- **The components of the Public Key Infrastructure must be compliant with applicable Federal Government Standards.**

D.1.8 Archiving

Some of the electronic information that is digitally signed may need to be archived by agencies and possibly by the National Archives and Records Administration (NARA). In order to verify the signature on the information at a future date, the public key or the certificate associated with the signer will also need to be archived. CRLs produced by the CAs which generated the archived certificates should also be sent to the archive. The CRLs will indicate whether an archived certificate was ever revoked. The CRL or a notation specifying the time period for which the key/certificate was valid will need to be archived.

Items to be archived may be created using differing applications which produce documents in differing formats. To make archiving easier, information to be archived is often converted into a common format. Thus, keys, certificates, and CRLs to be archived may need to be converted into such a format.

As an alternative, the archiving application can verify the document after examining all the certificates in the certification path and all associated CRLs. It can then affix its own signature on the document or on a folio of documents stating that all signatures were verified at the time of archiving. The archival private key would be very long and would be archived indefinitely. This approach, however, requires legal investigation and acceptance.

- **The Public Key Infrastructure must provide support for the archiving of digitally signed documentation. In particular, the PKI may need to present certificates and CRLs in a specified format for archiving purposes.**

D.2 KEY GENERATION AND DISTRIBUTION

D.2.1 Key Generation

A user's public/private key pair may be generated by the user or by a trusted entity for the user, provided the key pair is generated using a strong method. Appendices 2 - 4 of the DSS standard [5] describe how to generate a good key pair. There are advantages and disadvantage to each type of generation. These advantages and disadvantages are enumerated in the following paragraphs.

The advantage of having the user generate his own key pair is that the user's private key is never released to another entity. This allows for the provision of true non-repudiation services. However, this method of key generation requires that the user has a certain level of competence and is trusted to adhere to defined security policies or that he be provided a tamper-proof key generation application. Otherwise, the key pair may not be very secure.

The benefit of having a trusted third party generate a key pair for a user is that the entity is likely to have the competence and trust to produce a good key pair. This method assumes that security measures are employed by the third party to prevent tampering. The disadvantage of the method is that the private key is known to an entity other than the user. The third party must be trusted to destroy all copies after it hands over the private key. This need for trust in a third party is a potential threat to a non-repudiation service in a publicly available infrastructure. Perhaps it is of less concern in an environment such as the Federal Government. The private key must, of course, be transmitted to the user in a secure manner such as on a token which might be a smart card, a PCMCIA card or an encrypted floppy diskette.

During the interviews, some agencies said that they would allow users to create their own public/private key pair, while other agencies said that the agency would have a device that creates the key pair for the user. Taking this into consideration, the PKI will deal with keys generated through both means. This is not a major concern for the PKI, assuming that the public key is presented to a CA for certification in a secure manner.

- **The Public Key Infrastructure must support public keys generated by a user or generated by a third party for a user.**

D.2.2 Secure Key Generation and Distribution

If key generation is conducted by a trusted third party on behalf of the user, it is necessary to assure the integrity of the public key and the confidentiality of the private key. Therefore, generation and distribution of key pairs must be done in a secure fashion.

- **Any activities of the Public Key Infrastructure related to the generation and distribution of the public/private key pairs must be done in a secure fashion.**

D.3 CA REQUIREMENTS

D.3.1 Trusted

In order to be assured of the authenticity and integrity of a certificate and public key contained within it, the users must have their certificates created by a trusted source. Therefore, the CAs which generate and manage the certificates must be trusted by the users.

- **To assure users the certificates they create can be trusted, CAs within the Public Key Infrastructure must function correctly, implement the specified security policy, and preserve the binding between the user and the user's public key.**

D.3.2 Availability

Users will require services from the CAs within the infrastructure at various times during the day; although, the major demand for service is expected during normal working hours.[7] However, in the case of key compromise, it is desirable to report the compromise as soon as it is suspected, thus preventing as much fraudulent activity as possible. It is possible that the compromise could occur outside of normal working hours and, if required by PCA policy, may need to be reported outside of normal working hours. Depending on the PCAs policy, a CA may need to provide a key compromise reporting mechanism at all times.

- **At a minimum, CAs within the Public Key Infrastructure should be available to provide all services during normal working hours.**

D.3.3 Services and Functions

The primary function of the CA is to generate and manage the public key certificates that bind the user's identity with the user's public key. In order to perform this function, each CA within the infrastructure will need to provide some basic services to its users. Services provided by CAs are identified in the following paragraphs.

D.3.3.1 User Identification and Authentication

Many of the future users of the infrastructure noted that CAs need to identify and authenticate users before generating certificates for the users. Some variations in the

---

[7] The definition of normal working hours is dependent upon the work environment and could also depend upon operation over several time zones.

strength of user authentication performed by the CAs are possible. Some CAs may utilize a weak authentication scheme which requires the user to produce some form of identification such as a driver's license. Other CAs may employ a stronger preexisting authentication scheme which requires the user to possess a token such as a smart card or picture badge. An even stronger preexisting scheme that utilizes biometric techniques could also be used. Users of highly sensitive applications would need to be authenticated by CAs providing strong authentication, while users of less sensitive applications would apply to CAs providing a weaker form of authentication. In either case, users must be authenticated in person.

- **The CAs within the Public Key Infrastructure must identify and authenticate their users. The Public Key Infrastructure should allow for different strength authentication schemes to be used by its CAs.**

D.3.3.2 Certificate Generation

One of the major services provided by CAs is the generation of certificates. It is through certificate generation that the binding of user's identity and a user's public key is made which, in turn, is based on the appropriate user identification policies and procedures. This binding is the key to providing authentication services through the use of digital signatures. Certificates are signed by the CA using its private key. This signature shows that the CA vouches for the authenticity of the information contained within the certificate.

- **The CAs within the Public Key Infrastructure must generate certificates.**

D.3.3.3 Certificate Distribution

CAs are responsible for sending copies of the certificates it generates to the appropriate directory server. Often, once a CA generates a certificate, it will provide a copy of the certificate to the user with whom it is associated. Some digital signature applications require users to forward their certificate along with the digitally signed document; thus the user needs a copy of his certificate to use such applications.

- **The CAs within the Public Key Infrastructure must distribute certificates to the Directory and may distribute certificates to the associated user.**

D.3.3.4 Certificate Storage and Retrieval

CAs perform management functions on the certificates that it generates such as notifying a user when a certificate is about to expire or revoking certificates. In order to provide these management service, the CA will need to store and retrieve the certificates it generates. The CA may also want to maintain a back-up file of certificates in case the certificates are needed by a directory server recovering from a failure or by the CA itself, in case of its own failure.

- **The CAs within the Public Key Infrastructure may store and retrieve certificates.**

D.3.3.5 Certificate Revocation Report

CAs will receive certificate revocation reports from its users in some out-of-band method, such as in person or via U.S. Mail. These notices report suspected compromises of a private keys or changes in a user's organizational affiliation. They may be made by the owner of the compromised private key, by his employer or by his sponsor. In the public sector, it may be his executor, his conservator or his legal guardian. To prevent a denial of service attack on the PKI, each report must be carefully authenticated by CA personnel.

- **The CAs within the Public Key Infrastructure must receive and authenticate certificate revocation reports.**

D.3.3.6 CRL Generation and Maintenance

The CA will need to generate CRLs that denote which certificates are no longer valid due to compromise or to employee severance. Each CA will generate CRLs for the certificates that it has generated. The CA will need to ensure that the information within the CRL is as current as possible. Therefore, the CA will need to update its CRLs periodically in order to incorporate new information.

- **The CAs within the Public Key Infrastructure must generate CRLs and maintain the CRLs so that they contain the most current information.**

D.3.3.7 CRL Distribution

Users must have up to date information about certificates which are no longer valid. In order to meet this need, a CA will periodically distribute a CRL to the appropriate directory server. When the CA distributes a CRL, it signs the CRL using its private key in order to prove to users that it generated the CRL. Users will be able to retrieve the CRLs they need from the Directory.

During an agency interview, it was noted that CRLs should be distributed in a timely manner and to as many of the concerned users as possible. This will prevent, or at least limit, fraudulent use of the private key. PCA policy may include a requirement that each CRL produced be sent to all users whose certificates were generated by the CRL issuing CA – even when the CRL is sent to the Directory as well.

- **The CAs within the Public Key Infrastructure must distribute CRLs to the Directory.**

D.3.3.8 CRL Storage and Retrieval.

Like certificates, CAs will need to store and retrieve CRLs. The CA will store CRLs that it created. It will also retrieve CRLs to update, to replace, or to distribute them.

- **The CAs within the Public Key Infrastructure may store and retrieve CRLs.**

D.3.3.9 Auditing

To provide additional assurance of the trusted nature of CAs and to provide information to agency personnel conducting internal audits, the actions of each CA should be auditable. Audit records and audit trails should be generated for events such as user registration, certificate generation, compromised key reports, etc.

- **CAs within the Public Key Infrastructure should have audit capabilities.**

D.3.3.10 Archiving

In order to verify the signature on the information at a future date, the public key or the certificate associated with the signer will also need to be archived. Certificates and CRLs produced by the CAs which generated the archived certificates should also be sent to the archive. It is unclear whether the user archiving a signed document is responsible for archiving the associated certificates and CRLs or whether CAs should be required to archive all certificates and CRLs they produce. In any event, CAs should archive such events as the creation or revocation of certificates.

- **CAs must archive logs of certificate generation and revocation. They may need to present certificates and CRLs for archiving purposes.**

D.4 ORA REQUIREMENTS

D.4.1 Trusted

In order to be assured of the authenticity and integrity of a certificate and public key contained within it, the users must have their certificates created by a trusted source. Since ORAs perform authentication functions for CAs, they must be trusted to follow the CA's user authentication policies and to pass the correct user identification information along with the associated public key to the CA. Similarly, ORAs must be trusted to pass certificate revocation reports to a CA in an accurate and timely fashion.

- **ORAs must be trusted to pass accurate certification requests and accurate certificate revocation requests to a CA.**

D.4.2 Availability

As previously noted in section D.3.2, CAs at a minimum will provide all services during normal working hours. Since the ORA is the interface between a user and a CA, it should adhere to the same availability requirements as the CA itself.

- **At a minimum, ORAs within the Public Key Infrastructure should be available to provide all services during normal working hours.**

D.4.3 Services and Functions

An ORA acts as an intermediary between users and a CA, specifically providing user authentication functions for the CA. The basic services provided by an ORA will be defined in the next few sections.

D.4.3.1 User Identification and Authentication

The prime function that an ORA performs is user identification and authentication. When an ORA performs this function on behalf of a CA, it must follow the same rules and method of authentication as the CA uses itself.

- **The ORAs within the Public Key Infrastructure must identify and authenticate their users by applying the same methods used by their parent CA.**

D.4.3.2 Certificate Request

After authenticating a user, an ORA will transmit a signed request for a certificate to the appropriate CA. The request will contain the user's unique name and his public key.

- **The ORAs within the Public Key Infrastructure will generate certificate requests.**

D.4.3.3 Certificate Receipt

In response to an ORA request for a key certification, the CA returns a certificate to the ORA.

- **The ORAs within the Public Key Infrastructure will receive new certificates from CAs.**

D.4.3.4 Delivery of New Certificate

The ORA passes the certificate on to the user. PCA policy may require the ORA first check the certificate by examining the unique name contained therein and by verifying a signature produced by the named user.

- **The ORAs within the Public Key Infrastructure will deliver new certificates to the users named therein.**

D.4.3.5 Certificate Revocation Report

ORAs can be instrumental in the handling of certificate revocation reports. The person making the reporting that a certificate 's key has been compromised or that its owner has severed his affiliation with an organization can do so at the ORA. It is then the responsibility of the ORA personnel to authenticate the report. If, by applying the same criteria the CA would have used they are satisfied that the report is authentic, the ORA sends a signed message to the CA containing certificate identification information and the reason for revoking that certificate.

- **The ORAs within the Public Key Infrastructure must receive and authenticate certificate revocation requests. They must forward the requests to the appropriate CA.**

D.4.3.6 Auditing

To provide additional assurance of the trusted nature of ORAs and to provide information to agency personnel conducting internal audits, the actions of each ORA should be auditable. Audit records and audit trails should be generated for events such as user registration, certificate request and receipt, compromised key reports, etc.

- **ORAs within the Public Key Infrastructure should have audit capabilities.**

D.4.3.7 Archiving

In order to verify the signature on the information at a future date, the public key or the certificate associated with the signer will also need to be archived. It may be important to know how a certificate was produced. ORAs should archive such events as the requests for the creation or revocation of certificates.

- **ORAs must archive logs of certificate generation and revocation requests.**

D.5 DIRECTORY SERVICES

The PKI needs the support of a directory service such as that defined in the CCITT X.500 series of standards [10]. In the initial stages of PKI development and deployment, small scale infrastructure prototypes may have to furnish this service themselves by including a Commercial Off-the-Shelf (COTS) directory server in their prototype, if a directory service is not already in place. The PKI requires that the directory display the following characteristics.

D.5.1 Expected Operation

It is important for the functioning of the PKI that the supporting directory service perform in the expected manner. It must respond to requests with the latest certificates and

CRLs. It must delete a certificate from its database if and only if its expiration date has passed. It cannot interfere with a PKI user's obtaining the latest certificate for an entity or with his ability to check the status of any certificate he holds. It might be beneficial for the directory to verify that all certificates and CRLs it receives do indeed originate with and were created by the same CA from which they appear to have been sent. However, COTS directory packages cannot be expected to do this and certainly cannot be trusted to do this. The onus of checking the validity of a certificate or of a CRL rests with the user.

- **The Directory should be perform its services in a manner which is expected.**

D.5.2 Speed of Signature Verification

Another need specified by agency users is that the computation of digital signatures and verification of these signatures must not significantly slow down applications that incorporate digital signatures. The time it takes to obtain certificates and CRLs will affect the signature verification time. Certificates and CRLs will be stored in the Directory, so it is the Directory's responsibility to provide these in a timely manner.

- **The Directory should supply certificates and CRLs to requesters in a timely manner; not significantly affecting the time it takes to verify a digital signature.**

D.5.3 Services and Functions

The following are the functions that the directory service must perform in support of the PKI.

D.5.3.1 Certificate Storage and Retrieval

A directory node must accept certificates from CAs and enter them into its database. The PKI entity's unique name as contained in the certificate is the key under which the certificate is stored. On request, the directory node sends this – and any other certificates listed for the same unique name – to the requester.

- **The directory must receive and store certificates. It must deliver them to any and all requesters.**

D.5.3.2 CRL Storage and Retrieval

A directory node must accept CRLs from CAs and enter them in its database. The CAs unique name as contained in the CRL is the key under which the CRL is stored. If CRLs are incremental, the directory adds the new CRL to any existing ones for that CA. Even if CRLs are complete listings, the directory adds the new CRL to any it already holds. Only if it can be trusted to verify that the new CRL comes from the named CA and supersedes all other CRLs from that CA can the directory be allowed to delete the older lists. On request, the

directory node sends this, the most recent CRL listed for the unique CA name, to the requester.

- **The directory must receive and store CRLs. It must deliver them to any and all requesters.**

D.5.4 Availability

User demand for certificates and for CRLs will be concentrated in, but not limited to, normal working hours. Thus, each local directory node should be available 24 hours a day with any maintenance down-time being scheduled outside normal working hours. During non-scheduled down-time, it is desirable that a back-up be available.

- **The directory service supporting the PKI should be available 24 hours per day, if possible. At a minimum, each node should be available during normal, local working hours.**

D.6 CERTIFICATES

D.6.1 Multiple Certificates

During the interview process several different types of public key certificates were proposed. Many of the proposed certificates were based upon the role the user is playing when digitally signing a document. For example, a residential certificate could be issued for a person to use when conducting private business, while an employee certificate could be issued for the same person to be used in conjunction with work related activities. Since each role a user assumes exists in a distinct domain, the user has a different unique name for each. Thus, private citizen Alice may have a unique name something like "c = us, s = ma, l = hertown, cn = alice" while employee Alice's unique name might be "c = us, s = ma, o = hercorp, cn = alice."

In creating role-based certificates, the PKI can issue a single private/public key pair to a user or one key pair for each role the user assumes. Thus, multiple certificates with the same key or with different keys are issued by different CAs. The system of unique names used within the PKI will need to allow the unique specification of any one of a single user's several certificates.

- **The Public Key Infrastructure must support multiple certificates for a single entity.**

D.6.2 Organizational Certificates

During the interview phase of the project, it was determined that some organizations may require a public/private key pair and, thus, an associated organizational certificate. The organization's private key could be used to sign documents on behalf of the organization, or

to sign certificates generated for the organization's employees. Care must be employed in applications where there exists the possibility of criminal misuse of the organizational key. Mechanisms must be in place to assure that any wrong-doer can be identified and can be associated with his actions.

- **The Public Key Infrastructure must support organizational certificates.**

D.6.3 Anonymous Certificates

For some applications, the identity of the individual computing the digital signature must be kept private; as for example, an undercover FBI agent filing a report from a remote computer. To meet this need, the concept of an anonymous certificate was introduced. This certificate would not bind a public key with the actual user's name. Instead the public key would be bound to an identity called anonymous, to a false name, etc.

- **The Public Key Infrastructure must support anonymous certificates.**

D.6.4 CA Certificates

The certificates of the CAs within the Public Key Infrastructure must be distinguishable from certificates associated with users. If these certificates are not distinguishable, it is possible for a user to masquerade as a CA. To eliminate this threat, a naming hierarchy should be used. The naming hierarchy requires that the name of the subject of the certificate to be subordinate to the name of the certificate issuer. Alternatively, each certification includes an indication that the entity whose name and whose key are in the certificate is a user, a device, an organization or a CA.

- **Within the Public Key Infrastructure, the certificates of the CAs must be distinguishable from certificates associated with users or organizations.**

D.6.5 Certificate Format

Throughout an infrastructure, the format of the certificate that binds the user's identity with his/her public key must be the same. This allows for ease of implementation. In addition to user identifiers and public keys, certificates usually contain the issuer's identifier, the version number, and the lifetime of the certificate. The entire certificate is signed using the private key of the issuer.

Further discussion on a certificate format for the PKI may be found in appendix F.

- **The Public Key Infrastructure should use a common format for its certificates.**

D.7 CERTIFICATE REVOCATION LISTS

D.7.1 Certificate Revocation List Format

Like the certificates, the CRLs within an infrastructure should all have the same format. CRLs that list the certificates of user and those that list the certificates of CAs should use the same format. CRLs contain information such as the CRL issuer's identifier, the serial numbers of the revoked certificates, and the date each certificate was revoked. Reasons for revocation may also be included in the CRL. The CRL is signed by the issuer using its private key.

Further details on a CRL format for the PKI may be found in appendix G.

- **The Public Key Infrastructure should use a common format for its CRLs.**

D.8 USERS

D.8.1 Applications as Users

During one agency interview, computer programs that automatically generate reports were briefly discussed. These automatic applications may require private keys to compute digital signatures and/or access to public keys in order to verify digital signatures. Some legal questions as to who is liable for the actions of such a process must be answered. If no general answers are available, the question of who is responsible and who is liable must be answered in each PCA's policy. The PCA should have the option to allow or to forbid applications from acting as users. Nonetheless,

- **The Public Key Infrastructure must be able to support non-human users.**

D.8.2 Unlisted Entities

Some users, especially those working in a closed environment, want to use digital signatures, but do not want to publicize their identity outside their environment and possibly inside their environment. To satisfy the needs of these users, CAs could offer a user anonymity to their registered users. A CA could provide such a service by not releasing the user's certificate to a directory service which would be accessible outside the closed environment.

- **The Public Key Infrastructure should be able to provide user anonymity services.**

APPENDIX E

APPLICABLE STANDARDS AND ANALYSIS

E.1 STANDARDS REVIEWED

Existing standards that involve public key cryptography, especially those with public key infrastructures, were examined as part of the PKI Study. Standards reviewed included the Comité Consultatif International Télégraphique et Téléphonique (CCITT)[8] Recommendation X.509–*Directory Authentication*, the Internet Activity Board (IAB) Privacy-Enhanced Mail (PEM) Request for Comments (RFCs), the .i Institute of Electrical and Electronic Engineers (IEEE) 802.10 *Standard for Interoperable Local Area Network (LAN) Standard for Interoperable LAN Security (SILS)*, RSA, Data Security's Public Key Cryptography Standards (PKCS), the Secure Data Network System (SDNS) standards, and the American National Standard Institute (ANSI) X9.30-199X *Public Key Cryptography Using Irreversible Algorithms for the Financial Services Industry* standards. These standards were examined for their insight into developing an infrastructure, particularly in the areas of user and technical requirements for public key infrastructures.

In the next two subsections, CCITT Recommendation X.509, the PEM RFCs, the SDNS Message Security Protocol (MSP) standard, and ANSI X9.30-199X, Part 3 are discussed, since they provide insight into the development of a public key infrastructure. The IEEE SILS and RSA PKCS standards are not discussed, since they both currently lack a section on key management. In section E.2, an overview of each of the standards is presented. The insights into user and technical requirements provided by these standards are discussed in section E.3.

E.2 OVERVIEW OF PERTINENT STANDARDS

E.2.1 X.509 Directory Authentication

The directory provides the means to determine where resources, such as public key certificates, are located in a distributed network. By providing the location of resources, the directory facilitates communication between, with, or about applications, people and terminals. The directory service standards are being developed in a cooperative effort between the International Standards Organization (ISO) and the CCITT. Due to this collaboration, the ISO 9594 series of standards are essentially the same as the CCITT X.500 Recommendation series.

The directory service standard relevant to public key infrastructures is Recommendation X.509–*The Directory–Authentication Framework* [10, 11]. X.509 describes two

---

8 The CCITT is now known as the International Telecommunications Union (ITU) Telecommunications Standards Section (TSS).

authentication methods: simple authentication and strong authentication. The simple authentication method is based on password usage. The strong authentication method is based on public key cryptography and is discussed further in the following paragraphs.

In the strong authentication method, each user is identified by the possession of the private key of a signature key pair. Confirmation that a user possesses a private key is obtained through the exchange of digitally signed authentication information. Party A who wishes to be authenticated signs an authentication request with his/her private key and sends this request to Party B. Party B obtains the public key of Party A from the directory and uses this key to verify the signature on the authentication request. If the signature is verified, the identity of user A is proven, since only Party A possesses the private key used to sign the message. Additional checks are made on the authentication information contained in the request to ensure that the request is fresh and intended for Party B.

An off-line Certification Authority (CA) encloses the public key in a *certificate*, signs the certificate, and places it in the directory. Users of this authentication scheme obtain a certificate from the directory and authenticate it using the public key of the CA that signed the certificate. The standard allows for the existence of more than one CA. A user has a copy of the public key of the local CA who signed his/her certificate. Sometimes the user will need to obtain copies of the public key of remote CAs. Mechanisms are in place within the directory to allow users to obtain the keys of remote CAs using a *certification path.* Such a path is a sequence of CAs from the verifier's CA to the signer's CA. Every CA in the path certifies the public key of the next CA.

E.2.2 PEM

PEM was developed by the Privacy and Security Research Group (PSRG) of the Internet Research Task Force (IRTF) and was refined based on discussion in the .i Privacy Enhanced Mail Working Group (PEM WG) of the Internet Engineering Task Force (IETF). A set of four RFCs were written to describe the privacy enhancement for electronic mail on the Internet. The original versions of these RFCs were published in draft form in August 1989 and have been periodically updated. The most recent versions of the PEM are RFCs 1421-1424 which were published in February 1993. RFC 1421[12] defines and describes message encipherment and authentication procedures used with PEM. RFC 1422[13] specifies the key management infrastructure to be used by the Internet community with this mail system. RFC 1423 [14] specifies the algorithms and related information relevant to RFC 1421 and RFC 1422. RFC 1424 [15] provides details on the paper and electronic formats for the key management infrastructure.

E.2.2.1 RFC 1422

RFC 1422–*Privacy Enhancement for Internet Electronic Mail: Part II: Certificate-Based Key Management* [13] describes the key management infrastructure, based on public key certificate techniques, which provides key information to users of the privacy enhanced mail system. This key management infrastructure is compatible with the certificate

infrastructure described in CCITT Recommendation X.509 (ISO 9594-8), but reflects a lower level of implementation detail than X.509.

The key management infrastructure defined by RFC 1422 establishes a single root, the Internet Policy Registration Authority (IPRA)for all certification taking place within the Internet. The IPRA establishes global policies that apply to the certification which takes place under the hierarchy. At the next level of the hierarchy are Policy Certification Authorities (PCAs). The PCAs publish the policies for registration within their authentication domain. Each PCA is certified by the IPRA. CAs are the entities found in the next level of the hierarchy and these entities certify users or subordinate organizational entities such as departments or subsidiaries that represent users. Three types of CAs are defined in the PEM RFCs: organizational, residential, and PERSONA. An organizational CA certifies entities associated with a specific organization. A residential CA generates certificates for users not associated with a specific organization. A PERSONA CA certifies users who wish to conceal their identities while making use of PEM security features. Each CA is certified by a PCA. Users and subordinate organizational entities are found at the lowest level of the hierarchy. The PEM key management hierarchy is shown in figure E-1.

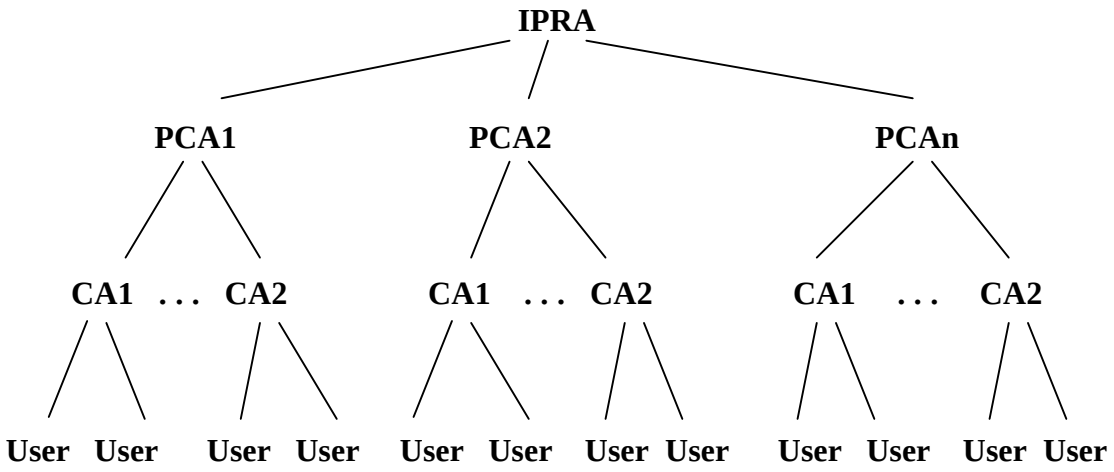


Figure E-1. PEM Key Management Infrastructure

The key management scheme described in RFC 1422 is specific to the use of the RSA algorithm. Algorithm identifiers, included in the key management protocol defined by this standard, facilitate the use of other algorithms such as DSA. This feature is intended to support interoperability with key management systems using algorithms other than RSA.

The key management scheme is based on the use of public key certificates. The certification authority representing an organization signs a collection of data consisting of the user's public key, information used to identify the user, and the identity of the organization whose signature is affixed. The certificate is signed with the private key of the

organization the certifying authority is representing. The organization is known as the *issuer* of the certificate. By signing a certificate, the certification authority vouches for the identity of the user and the user's affiliation to a specific organization.

When generated, certificates need to be made available to the users of the mail system. This can be achieved by storing them in a directory or electronically transmitting them to the user or to some location where the user can have access to them.

E.2.2.2 RFC 1424

RFC 1424–*Privacy Enhancement for Internet Electronic Mail: Part IV: Key Certification and Related Services* [15] describes three services (key certification, CRL storage, and CRL retrieval) which CAs provide in support of PEM. This RFC also specifies the format of the electronic mail messages used when communicating with the CAs that provide these services.

A CA provides the key certification services to subjects. A key certification request is generated by a subject that wants its public key certified by a CA. The request contains a self-signed certificate, a certificate generated and signed by the subject containing the subject's name and public key, and encapsulated text signed by the subject. When the CA receives the request, it verifies the subject's signature on both the certificate and the encapsulated text. If both signatures are valid, the CA creates a certificate for the user using the name and public key contained in the request. The CA signs this certificate using its private key and sends this signed certificate back to the subject.

The CRL storage service allows a CA to send its CRL to global PEM CRL database for storage. Access to this database is provided through a mailbox maintained by each PCA. When a PCA receives a CRL storage request, it verifies the signature on the request and the associated certification path to ensure that these are valid. If they are valid, the PCA stores on the global CRL database both the CRL that is provided to it in the request and with the associated certification path.

The CRL retrieval service allows a subject to obtain CRLs from the global database. When a CA receives a CRL retrieval request, it will provide the subject with the latest CRLs of the issuer(s) identified in the request. The CA also provides the requester with the certification paths and cross certificates associated with the CRLs which have been requested.

E.2.3 Secure Data Network System

The SDNS [6] project was a cooperative effort of government and industry to investigate various ways security could be implemented in distributed computer networks. The SDNS architecture and the associated specifications that resulted from the project provide a basis for standardization of security services in the Open Systems Interconnection (OSI) architecture.

The SDNS project was sponsored by the NSA and supported by NIST and the Defense Communications Agency (DCA)[9]. SDNS was developed by a committee of 12 United States computer and telecommunication companies (Analytics; AT&T; Bolt, Beranek, and Newman, Inc. (BBN); Digital Equipment Corporation, GTE; Honeywell; Hughes; IBM; Motorola; UNISYS; Wang; and Xerox [16].). These companies contributed to the project by providing both personnel and development resources.

The SDNS security protocols developed under the SDNS project were designed to provide security in an OSI environment. Four protocols were developed: Security Protocol 3 (SP3), Security Protocol 4 (SP4), Key Management Protocol (KMP), and MSP. The first two protocols provide security at the network and transport layers (OSI layers 3 and 4), respectively. KMP is used to generate, distribute, and update cryptographic keys required by the other three security protocols. MSP is a secure messaging protocol. Only MSP will be discussed further in this paper since it incorporates the use of certificates and directory services.

NSA turned over the SDNS specifications that resulted from phase 1 of the project, which completed in 1989, to NIST for possible standardization. NIST published these specifications in 1990 and requested comments on them. NIST also presented the SP3 and SP4 standards to ANSI who in turn introduced them to ISO. SP3 is the basis of the ISO 11577, *Network Layer Security Protocol (NLSP)* [17] and SP4 is the basis for ISO 10736, *Transport Layer Security Protocol (TLSP)* [18]. Updates to the SDNS standards have been made over the years since the completion of phase 1 of the project. These updates have brought the standards in alignment with the international versions of the standards as well as provide more functionality to the specified protocols.

E.2.3.1 Message Security Protocol

MSP, as described in SDN.701–*SDNS Secure Data Network System: Message Security Protocol* [19], is a security protocol that permits mail messages to be sent securely over the CCITT X.400 message handling system (MHS). Although the MSP specification is oriented toward the use of an X.400 MHS, the latest version of the specification (version 2.1) modifies the protocol so that it can provide secure message encapsulation in other messaging environments such as an environment using the Simple Mail Transport Protocol (SMTP) [20].

MSP provides writer-to-reader security services. These services include confidentiality, integrity, data origin authentication, access control, non-repudiation with proof of origin, and non-repudiation with proof of delivery. MSP is transparent to the X.400 message transport system since the mail message is encapsulated, and an MSP header is added before the message reaches the message transport system.

---

9 DCA is now known as the Defense Information Systems Agency (DISA).

Confidentiality, data origin authentication and integrity are provided to a MSP mail message through the encryption of the message contents and the associated key management mechanisms. Non-repudiation of origin is provided through the application of a digital signature to the mail message. Non-repudiation of delivery is provided through a digitally signed return receipt. The non-repudiation of delivery service may be requested only on signed mail messages, since the signed receipt is dependent upon the signed message. Two types of access control are applied to MSP message, rule based access control (RBAC) and identity based access control (IBAC). RBAC is based on the sensitivity of the mail message and the authorization of the message originator, the message recipient and the workstations associated with both. IBAC is the responsibility of the message originator and is supported by the strong authentication provided by MSP.

MSP supports three different X.509 certificates, each containing different public key information. One certificate contains only the signature public key information. Another certificate contains only the key management public key information. The third certificate contains both the signature and key management public key information. Which certificate is sent with the mail message is dependent upon the type of security services applied to the message. For example, if only non-repudiation of origin is selected, the certificate with only the signature public key information needs to be included with the message. However, if both confidentiality and non-repudiation are selected, then the certificates with the signature only and key management only public key certificates need to be sent with the message. Alternatively, the certificate with both the signature and key management public key information could be sent with the message.

The MSP specification states that CAs create and manage public key certificates and CRLs. However, the standard does not impose an infrastructure of CAs. Development of a certification hierarchy is left to the systems that utilize MSP such as the Defense Messaging System (DMS).

MSP utilizes the X.500 Directory to store information it needs to perform its processing. Information stored in the Directory includes: X.509 public key certificates; Auxiliary Vectors (AVs) which provide additional access control information; mailing lists; etc. The SDNS specification, SDN.702–*SDNS Secure Data Network System: SDNS Directory Specifications for Utilization with the SDNS Message Security Protocol* [21] defines object classes, object identifiers and attributes for directory entries to support the MSP requirements. The addition of these attributes and objects do not affect the operation of the directory service.

E.2.4 ANSI X9.30-199X

The ANSI Accredited Standards Committee (ASC) on Financial Services (ASC X9) has developed two sets of public key-based standards designed to protect financial information and support electronic commerce. The X9.30-199X *Public Key Cryptography Using Irreversible Algorithms for the Financial Service Industry* set of standards is based on the use of DSA. The X9.31-199X *Public Key Cryptography Using Reversible Algorithms for the*

Financial Service Industry set of standards is based on the use of the RSA public key algorithm. For the purposes of this paper, the X9.30-199X series will be examined in more detail.

The X9.30-199X series of standards consists of four parts. Part 1 describes DSA and is equivalent to the draft DSS FIPS [5] issued by NIST. Part 2 describes Secure Hash Algorithm (SHA) and is equivalent to the draft Secure Hash Standard (SHS) Federal Information Processing Standards (FIPS) [9] issued by NIST. Part 3 describes certificate management and an authentication framework to be used with DSA certificates. Part 4, which is in the early stages of development, will describe how public key algorithms may be used to manage secret keys used with symmetric algorithms. Part 3 of X9.30-199X is the most relevant to the development of the user requirements for the PKI and will be briefly summarized in the next section.

E.2.4.1 ANSI X9.30-199X, Part 3

ANSI X9.30-199X *Public Key Cryptography Using Irreversible Algorithms for the Financial Service Industry*: *Part 3: Certificate Management for DSA* [22] defines a certificate management and an authentication framework to be used by the financial service industry. The certificates generated and managed through the means specified in this standard are DSA public key certificates. These certificates will be used to validate DSA signatures applied to financial transactions. The use of DSA provides integrity, non-repudiation, and origin authentication services for these transactions.

ANSI X9.30-199X, Part 3 specifies the format and content of a public key certificate. The format of these certificates is based on the format specified in 1992 version of CCITT X.509. The 1992 X.509 certificate has two optional fields, Issuer Unique ID and Subject Unique ID. ANSI X9.30, Part 3, requires that the Issuer Unique ID be implemented and that it contain an identifier which will uniquely identify the private key used to sign the certificate. This standard also suggests the preferred contents of the Subject Unique ID field. This standard also specifies the content of the credentials required to obtain certificates.

Unlike the CA hierarchy specified in PEM, the hierarchy described in ANSI X9.30, Part 3 is a "bottom-up" hierarchy. Each CA certifies its subscribers and any adjacent CAs, be they superior or subordinate. Cross-certification of other CAs is also allowed. Entities are given the public key of their own CA rather than the public key of the root. Certification paths extend up a hierarchy from the verifier to a CA which is a common ancestor of the signer and the verifier and then down to the signer. This approach is especially useful in limiting the scope of a CA compromise, since the certificate path must go through the compromised CA. In addition, recovery from compromise is easier, since not all users would need new certificates.

The standard defines controls for CAs and other management requirements for CAs and their subscribers. Specifically, the standard describes how certificates are generated and revoked by the CAs, and how the certificates are validated by subscribers or other CAs.

The standard also presents an authentication framework that is defined as the structure that encompasses CAs and the entities that they certify. The framework is based on that which is provided in the CCITT X.509 standard. The framework described in this standard may be a hierarchical or a non-hierarchical structure and is intended to provide point-to-point connections. The authentication framework allows for the authentication of subscribers and their associated keys. It also facilitates obtaining subscriber certificates in a manner that assures the authenticity and integrity of the certificates.

APPENDIX F

CERTIFICATE FORMATS

F.1 PUBLIC KEY INFRASTRUCTURE (PKI) CERTIFICATE FORMAT

A common certificate format should be used throughout the PKI. This will allow for ease of implementation. The format should also facilitate interoperability with other signature algorithms. At a minimum, the PKI certificate should contain: serial number, subject's name, indication of the subject's signature algorithm, subject's public key and associated parameters, validity period of the certificate, issuer's name, and an indication of the issuer's signature algorithm. The entire certificate should be signed by the issuer, and this signature should be sent along with the certificate.

The certificate formats presented in several of the standards examined under this study meet the requirements specified above for the PKI certificate format. For several reasons, the format proposed in the 1992 version of Comité Consultative International Télégraphique et Téléphonique (CCITT) X.509 is considered the best selection for the format for the PKI certificate. First, X.509 is an internationally recognized standard, so use of the X.509 format should facilitate interoperability with the international community. Second, many of the other standards examined based their certificate formats on that presented in X.509; therefore, using the X.509 format will facilitate interoperability with other authentication schemes and infrastructures.

A description of the 1992 CCITT X.509 certificate format is presented in the next section. In the following sections, the Privacy-Enhanced Mail (PEM), American National Standards Institute (ANSI) X9.30, and Message Security Protocol (MSP) certificate formats are also presented. These formats are contrasted with the X.509 certificate; and thus, the PKI certificate format to show the differences between the formats. Additional considerations for the PKI certificate format are found in the last section of this appendix.

All the formats presented in this appendix show the elements contained within the certificate. In all instances, the entire certificate is signed by the issuer utilizing the algorithm specified within the certificate. This signature is always sent along with the certificate. In the figures, the issuer's digital signature is shaded to distinguish it from the contents of the certificate.

F.2 CCITT X.509 CERTIFICATE FORMAT

The 1992 version of the CCITT X.509 standard will be released in 1993. Figure F-1 shows the certificate format contained in the current draft 1992 version [11]. This standard is in the final draft stages and is not anticipated to change before it is released.

The 1992 CCITT X.509 certificate has nine fields: version, serial number, signature, issuer, validity, subject, subject public key information, issuer unique identifier, and subject

unique identifier. The last two fields were added to the certificate format in the 1992 version of the standard. They may be used to provide additional information about the subject and issuer within the certificate. The contents of each field are described in more detail below:

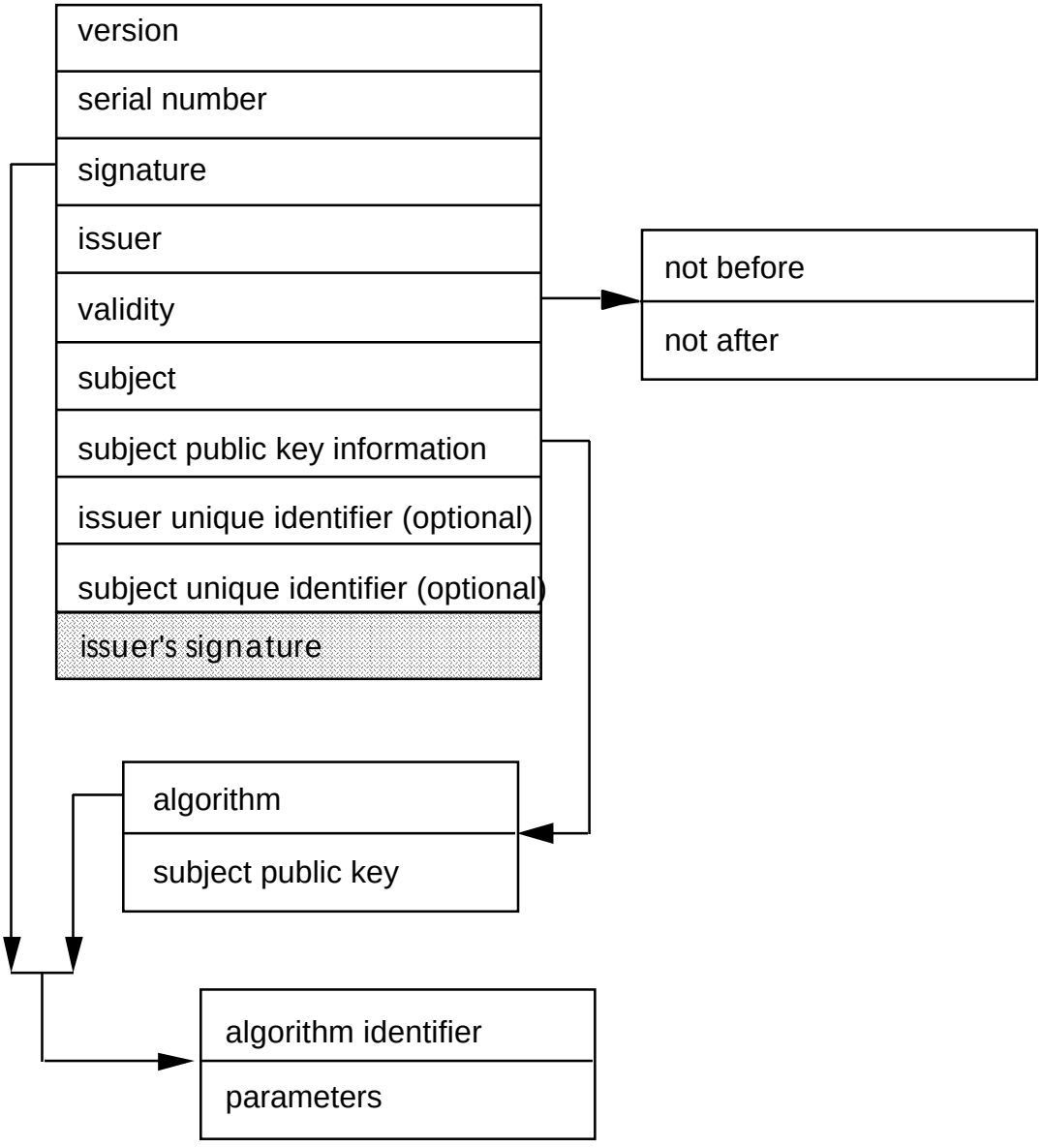


Figure F-1. Proposed 1992 CCITT X.509 Certificate Format

- **Version**: In this field the format of the certificate is identified. This field is included to facilitate orderly changes in certificate formats over time. The 1988 X.509 certificate format is assigned the value "0", and the 1992 X.509 certificate format is assigned the value "1".

- **Serial Number**: This field contains a unique identifier for each certificate generated by an issuer. The issuer must ensure that it never assigns the same serial number to two distinct certificates.

- **Signature**: In this field, the algorithm used by the issuer to sign the certificate, and any parameters associated with that algorithm, are specified.

- **Issuer**: This field contains the name of the entity that generated and signed the certificate.

- **Validity**: In this field, the time period for which the certificate is valid is denoted. This field contains two time and date indications that denote the start and the end of the time period for which the certificate is valid.

- **Subject**: This field contains the name of the entity for whom the certificate is being generated.

- **Subject Public Key Information**: This field contains the public key of the subject, an indication of the algorithm with which the public key will be used, and any parameters associated with the algorithm.

- **Issuer Unique Identifier**: This is an optional field and contains additional information about the issuer of the certificate.

- **Subject Unique Identifier**: This is an optional field that contains additional information about the entity for which the certificate is being generated.

F.3 PRIVACY-ENHANCED MAIL (PEM) CERTIFICATE FORMAT

The certificate format specified in the PEM standards [13] is the 1988 version of the X.509 certificate format. No changes to this certificate format are required by the PEM standards. The 1988 CCITT X.509 [10] certificate format does not contain the issuer unique identifier and the subject unique identifier fields which the 1992 version contains. The 1988 X.509 certificate format, and thus the PEM certificate format, is shown in figure F-2.

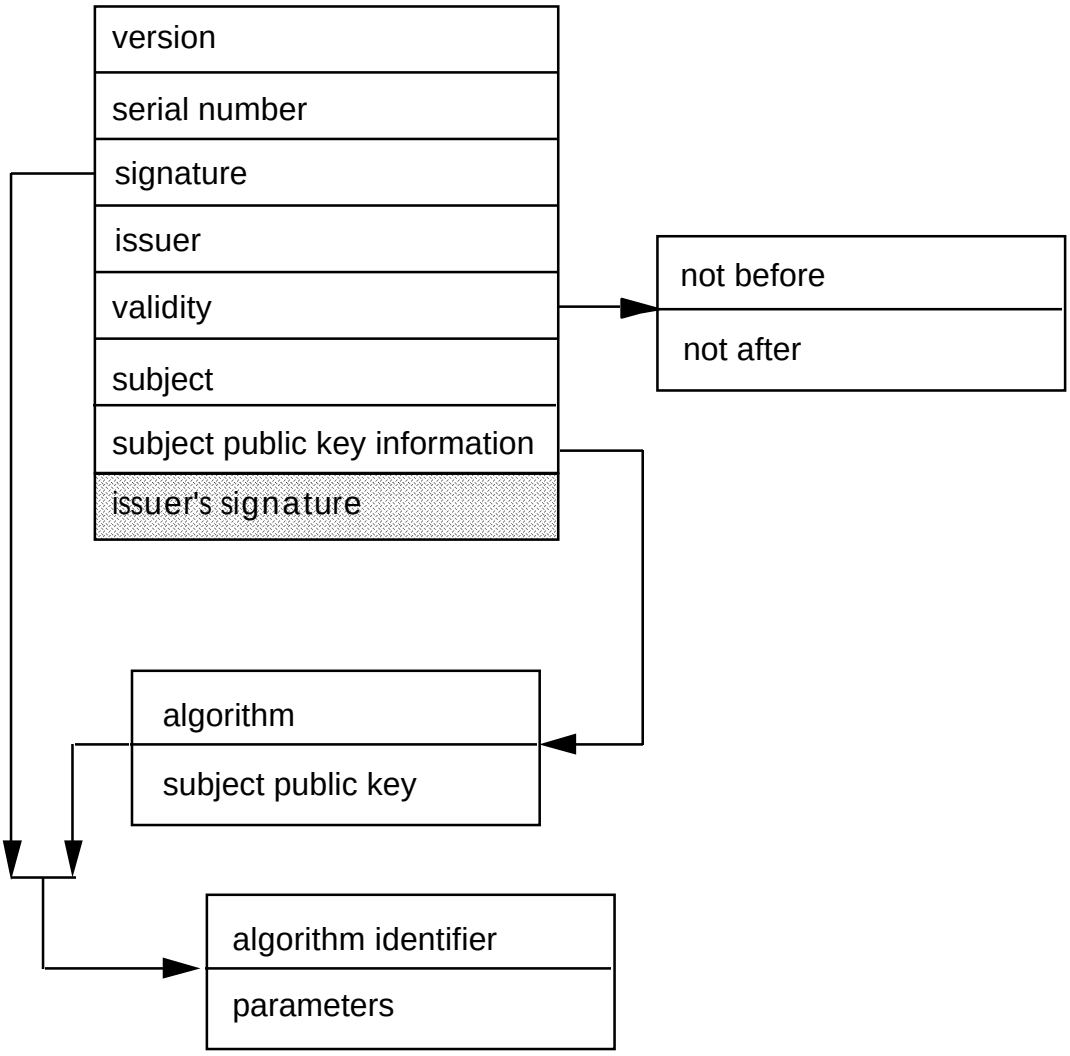


Figure F-2. 1988 CCITT X.509 and PEM Certificate Format

F.4 ANSI X9.30 CERTIFICATE FORMAT

The certificate format specified in the ANSI X9.30 standards [22] is based on the 1992 version of the X.509 certificate format. The ANSI X9.30 standard requires that the issuer unique identifier field be filled in. This field will contain information that allows the private key used to sign the certificate to be uniquely identified. The subject unique identifier field is optional.

F.5 MSP CERTIFICATE FORMAT

The certificate format used with MSP [19] is also based on the 1988 X.509 certificate format and therefore, does not include the issuer unique identifier or the subject unique identifier fields that are found in the 1992 X.509 format. MSP supports three X.509 certificates. Each of the three certificates contains different information in the subject public key information field of the certificate. One certificate contains only the signature public key information. Another certificate contains only the key management public key information. The third certificate contains both the signature and key management public key information. Which certificate is used is dependent upon the type of security services applied to the mail message. Further discussion of MSP and its use of X.509 certificates was presented in appendix E.

F.6 ADDITIONAL CONSIDERATIONS

In the PKI, Organizational Registration Authorities (ORAs) will exist. These entities will act as intermediaries between the users and the Certification Authority (CAs) by authenticating users for the CAs and by relaying credentials back and forth between the users and the CAs. Although users serviced by an ORA will have their certificates issued by a CA, it may be desirable to indicate within the certificate that the user interfaces with an ORA instead of directly with a CA. The extra field within the certificate would contain the identifier of the ORA. For users that do not interface with an ORA, this field in the certificate would remain blank.

The X.509 certificate format can be modified to accommodate an additional field that would contain the identifier for the ORA. However, the ORA identifier can be placed within the optional issuer unique identifier field. This may be the preferred approach since modification to the X.509 format is not required.

Some people [6, for example] would like the key certificate or an extended user attribute certificate to carry more information in order to facilitate total electronic processing. Some suggested entries are:

- **Entity type**: CA, person, device, or process.
- **Coded issuer policy**: issuer security assurance level, issuer liability limit, user identity checking method.
- **Coded user security data**: private key stored on smart card or other token, card or token is PIN protected, card or token has biometric activation.

APPENDIX G

CERTIFICATE REVOCATION LIST FORMAT

G.1 PUBLIC KEY INFRASTRUCTURE (PKI) CERTIFICATE REVOCATION LIST FORMAT

As with certificates, a common Certificate Revocation List (CRL) format should be used within the PKI. At a minimum, the PKI format should contain: issuer's name, indication of the issuer's signature algorithm, date the CRL was issued, date the previous CRL was issued by the Certification Authority (CA), date when the next CRL will be issued by the CA, and the list of revoked certificates. The certificate entries found in this list should include the serial number of each revoked certificate along with the date and time the certificate was revoked. This entire CRL should be signed by the issuer, and this signature should be sent along with the CRL.

Of the various CRL formats studied, the Privacy-Enhanced Mail (PEM) CRL format best meets the requirements for the PKI CRL format specified above. The PEM format is described in more detail within the next section. In the following sections, the Comité Consultative International Télégraphique et Téléphonique (CCITT) X.509 and American National Standards Institute (ANSI) X9.30 CRL formats are also presented. They are compared with the PEM CRL format to show where the formats differ. Additional considerations for the CRL format, including areas where the CRL format may need to be modified, will be discussed in the last section of this appendix.

All the formats presented in this appendix show the elements contained within the CRL. In all instances, the entire CRL is signed by the issuer utilizing the algorithm specified within the CRL. This signature is always sent along with the CRL. In the figures, the issuer's digital signature is shaded to distinguish it from the contents of the CRL.

G.2 PEM CERTIFICATE REVOCATION LIST FORMAT

The PEM CRL format [13] is illustrated in figure G-1. The PEM CRL format includes five fields: signature, issuer, last update, next update, and revoked certificates. These fields are briefly described as follows:

- **Signature**: In this field, the algorithm used by the issuer to sign the CRL, and any parameters associated with that algorithm, are specified.

- **Issuer**: This field contains the name of the entity that generated and signed the CRL.

- **Last update**: This field indicates the time and date when the CRL was issued.
- **Next update**: This field indicates the time and date when the next CRL will be issued.

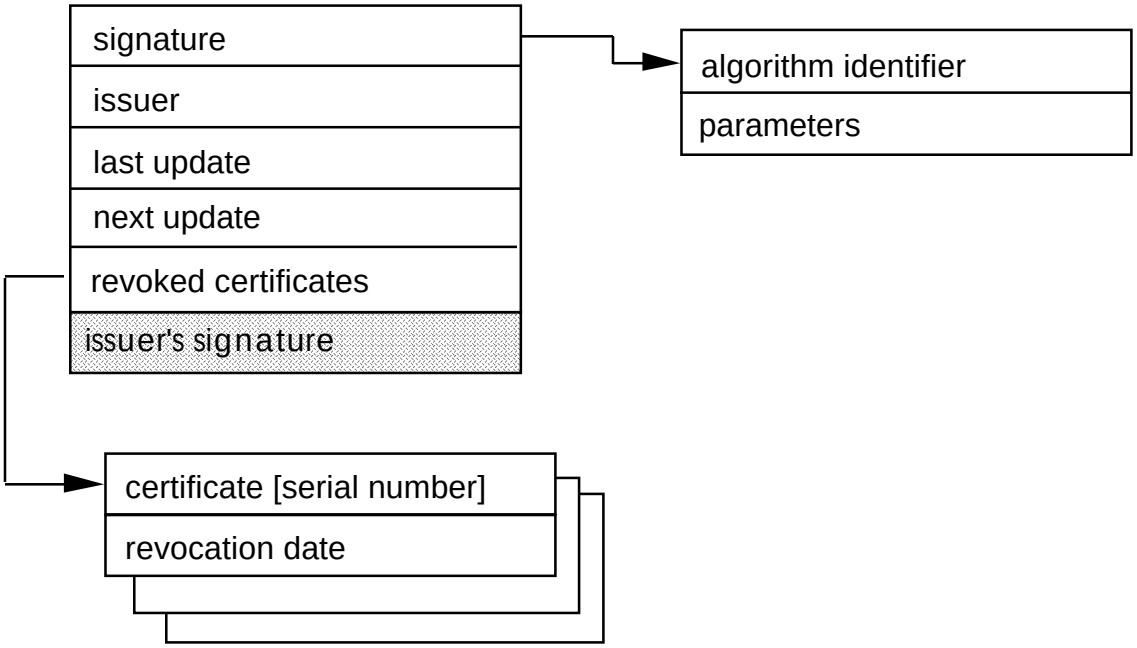


Figure G-1. PEM CRL Format

- **Revoked certificates**: This field contains the list of revoked certificates. Each certificate entry has two fields: certificate and revocation date.
  - **Certificate**: This field contains the serial number of the revoked certificate.
  - **Revocation date**: This field contains the time and when the certificate was revoked.

G.3 CCITT X.509 CERTIFICATE REVOCATION LIST FORMAT

The same CRL format is specified in both the 1988 and 1992 versions of the CCITT X.509 standard [10, 11]. This format is shown in figure G.2. The X.509 CRL format, unlike the PEM CRL format, only has four fields: signature, issuer, last update, and revoked certificates. The X.509 CRL format does not contain the next update field. There are also differences between the fields contained in the list of revoked certificates. The X.509 format

has a four-field entry for each certificate in the list. In addition to the serial number and the revocation date field, there is also a signature and an issuer field. These fields allow the X.509 CRL to contain revoked certificate issued by different CAs. The PEM CRL only contains the revoked certificates of one CA.

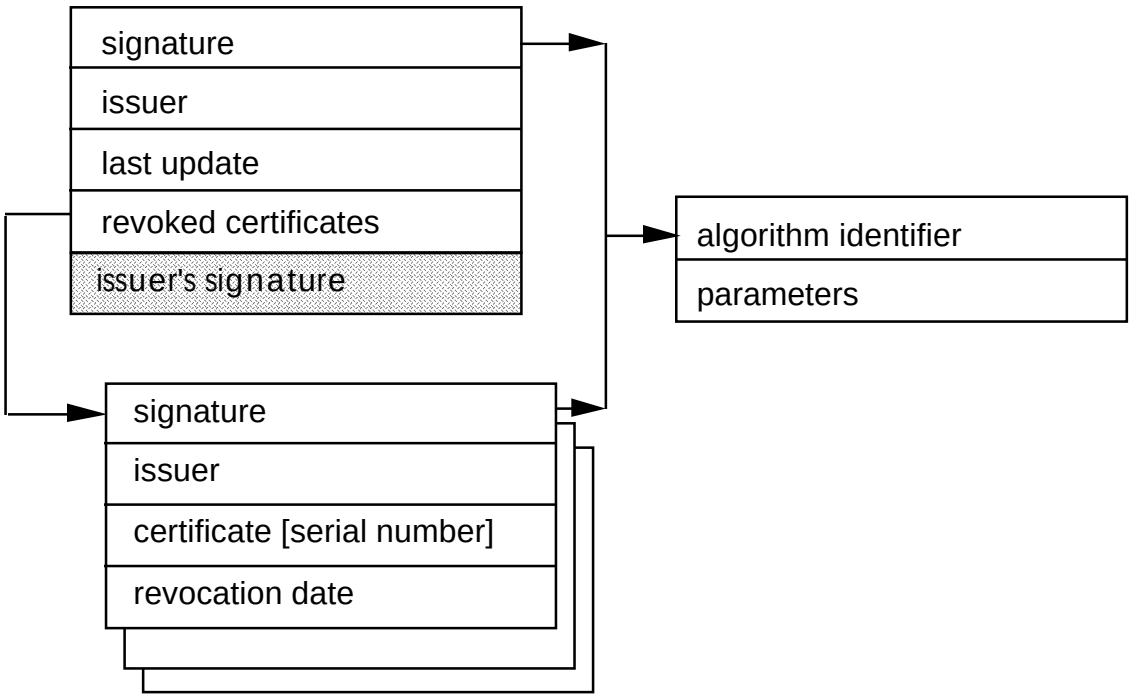


Figure G-2. CCITT X.509 CRL Format

G.4 ANSI X9.30 CERTIFICATE REVOCATION LIST FORMAT

The ANSI X9.30 CRL format [22] is based on the PEM format and is shown in figure G-3. The ANSI X9.30 format adds one field, reason code, to each certificate entry within the list of revoked certificates. This field is used to indicate the reason why the certificate was revoked. It was designed to allow automatic responses to occur when certain reason codes, such as key compromise, are indicated. Currently, there are six possible reasons (and codes) defined: key compromise (0), CA compromise (1), affiliation changed (2), certificate superseded (3), cessation of operation (4), and other (5). The X9.30 working group is considering making the reason code field optional in order to make the PEM and X9.30 CRLs compatible.

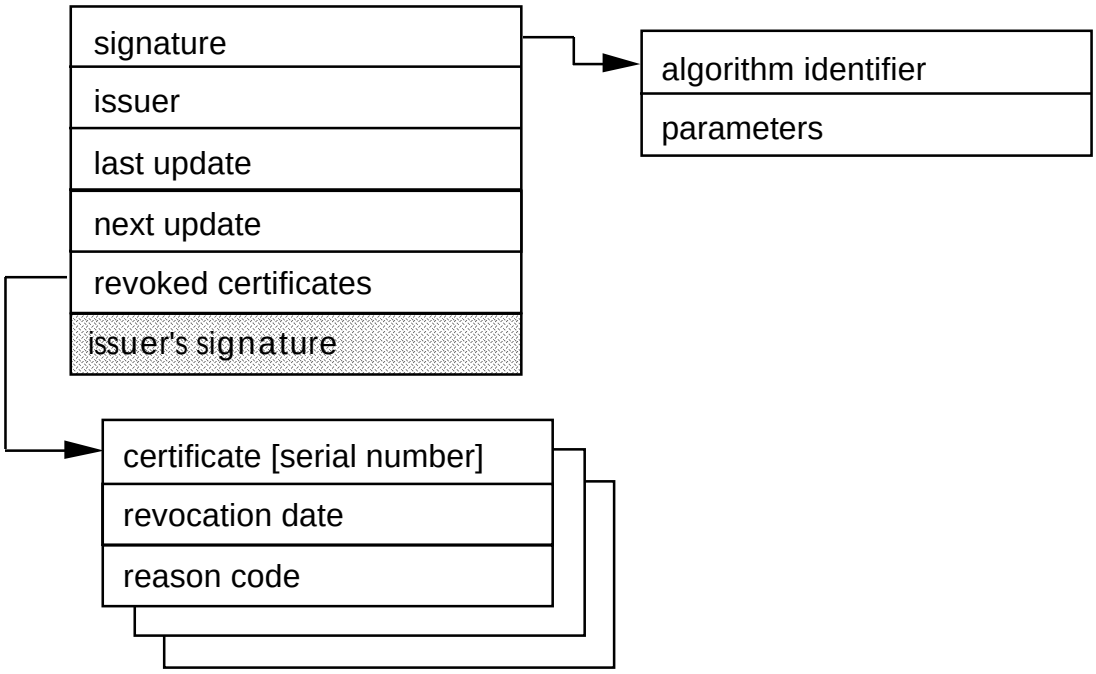


Figure G-3. ANSI X9.30 CRL Format

G.5 ADDITIONAL CONSIDERATIONS

In the proposed PKI CRL format, one field is unique and does not appear in the CRL format specified by PEM or by any other standard examined under this project. This field is the one that contains the date when the previous CRL was issued by the CA. Knowledge of the date of the last CRL will enable entities using the CRL to determine if they missed a CRL. Figure D-4 shows the proposed PKI CRL format. Fields which are unique to the PKI CRL format are shaded.

In the proposed CRL format, the last update field will contain the date on which the CA issued its last CRL. The current date field will contain the issue date of the current CRL. The next update field contains the date on which the CA will issue the next CRL. This field is unchanged from the PEM CRL format.

Consideration should also be given to including an optional reason code field for each certificate entry, as shown in figure D-4. There may be cases where the reason for revocation needs to be known and this field would allow the reason to be included in the CRL. The ANSI X9.30 CRL format contains a reason for revocation field, therefore, inclusion of this field in the PKI CRL would facilitate interoperability with the ANSI X9.30 standard. However, the ANSI X9.30 format uses a CRL number in place of the current data to allow users to check that they have received all previous CRLs.

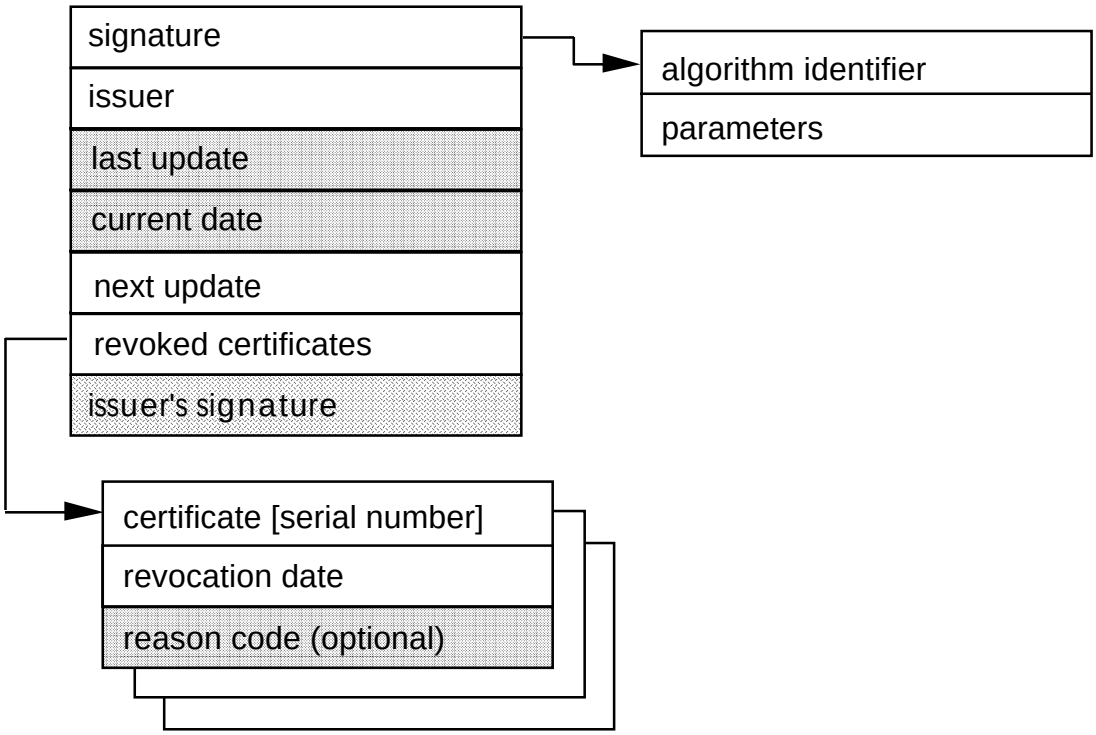


Figure G-4. Recommended PKI CRL Format

APPENDIX H

SAMPLE ELEMENTS OF PCA SECURITY POLICY

Policy Certification Authorities (PCAs) must publish their security policies, procedures, business related issues (such as any legal issues), any fees, or any other subjects that they deem necessary. This appendix discusses a minimal set of security policies that each PCA would need to define.

Who generates *p*, *q*, *g*, *x*, and *y*. The careful selection and appropriate protection of the prime numbers *p* and *q*, of the base number *g* and of the private and public components *x* and *y* of each key are the core of the security in digital signatures. Therefore, whoever generates these keys and their parameters are a security concern. Each PCA must define who will be allowed to generate these numbers.

How *p*, *q*, ,*g*, *x*, and *y* are generated. Even if the keys, prime numbers, and *g* are generated by acceptable sources, other factors that go into generating those numbers also have to be specified. For example, the algorithms used in generating the prime numbers may need to be specified, and the policy may require that the keys must be generated by hardware not by software. Each PCA would need to specify the acceptable algorithms in generating prime numbers and environments when generating the keys for itself and its Certification Authority (CAs). It may also specify them for its Organizational Registration Authority (ORAs) and end-users.

The range of *p*. The Digital Signature Standard (DSS) makes provision of having the size of prime number *p* from 512 bits to 1024 bits, in increments of 64 bits. Obviously, the larger *p* means more secure cryptographic function. But, larger *p* also means the signing and verification steps are more computing intensive. Thus, the size of *p* allows a trade-off between security and performance. Each PCA must specify the range of *p* for the PCA, its CAs, and its end-users. The range of *p* for the PCA, its CAs, and end-users may be different; the PCA's being the largest and the end-user's being the smallest.

Who gets the certificate. Each PCA must describe the users that it will serve. For one PCA, the users may be people who are affiliated with an organization; for another PCA, the users may be people who are part of a specific community; for yet another PCA, the users could be anyone (See table 8-1 for an example). Users do not necessarily have to be human; they can be a non human entity such as a specific role or an office in an organization.

Certificate renewal. Each PCA may describe procedures that it will employ for the renewal of the certificates that it issued. For example, it may require less stringent Identification and Authentication (I&A) requirements for the certificate renewal than when issuing the certificate for the first time. Each PCA also has to specify the maximum validity period for its CA certificates, as well as for its subordinate ORAs and end-users.

Identification and authentication requirements. When a user registers his or her certificate, the user must provide required I&A information to the certificate issuer in order to prove that he/she is indeed the claimed person. Each PCA must specify the I&A requirements that CAs must meet. It may also specify those that ORAs and the end-users must satisfy. The assurance level and types of information required may vary between CAs, ORAs, and end-users. Examples of the I&A requirements would be a driver's license, organization badge, passport, or birth certificate.

There are further I&A requirements that apply when a private key is reported as compromised or when an association is being severed. It is unwise to revoke a certificate unless the revocation request is established as authentic. Each PCA must institute policy delineating the requirements for determining the identity and the authority of the revocation requester.

CRL management. Each PCA must specify whether the CRLs will be pushed or pulled. In other words, each PCA must specify whether CRLs will be issued regularly or they will be issued whenever requested. In general, it is expected that the CRLs be issued regularly according to a schedule (e.g., every Monday or the first day of every month). In that case, each PCA must specify the frequency of the CRLs issuance. It should also specify how frequently a CA needs to receive CRLs from it subordinate CAs.

Validation of certificate revocation request. Each PCA must specify the procedures and information required to validate the legitimacy of a certificate revocation request.

Security controls. Each PCA must specify the security measures that it will employ for its hardware and software that are used for certificate generation and signing and maintaining CRLs. It also has to specify the physical securities, such as the storage for the archive of the certificates and the CRLs. It may also specify any security measures that it imposes on its CA's hardware and software, and how an end-user must safeguard his or her private key.

Audit procedures. Each PCA must specify the procedures for manual audits. The procedure may include a schedule of the manual audit and may also include that there may be impromptu audits.

Naming convention. A PCA may issue naming conventions that it imposes on its subordinates. If a PCA takes the policy that the UNs have to follow the hierarchy of the tree, it must so state. PCAs also have to specify procedures that have to be followed in case there are UN collisions.

APPENDIX I

PKI COST ANALYSIS

I.1 SCOPE OF THE ANALYSIS

In this appendix an overview of the analysis of the costs of developing the infrastructure recommended in section 4 is presented. The cost analysis is conducted at a high level and is based on simplified assumptions. It is hoped that the cost estimates will be helpful for planning and budgeting activities. In fact, it is intended that the cost model and its associated spreadsheet program can be used to estimate the cost of any subsection of the PKI. This should prove useful to any department or agency that is planning its own internal infrastructure.

It is important to note that the analysis includes only such quantitative impacts that can be ascribed to the PKI specifically. Obviously, creating a certificate infrastructure is for the express purpose of supporting the integrity of information and authenticity of sender in an electronic environment. Nonetheless, it is possible, though inadvisable, to deploy hardware and software in support of electronic business, reporting, and filing without including a signature capability. The costs associated with creating such an electronic business environment are explicitly excluded from this analysis. It is as if all installing, upgrading, and retrofitting to produce the electronic environment have been completed and now the PKI is being added to provide sender authentication and message integrity and to support sender non-repudiation. Thus, for example, costs for installing network servers are not included unless the PKI dictates the need for additional servers. The only software costs that are included are those related to the forming and verifying of digital signatures and to the creating and managing of certificates and CRLs. The software required for such applications as EDI or electronic filing of mandated reports is also explicitly excluded.

The cost analysis is based on the assumption that each CA within the infrastructure will at a minimum use a C2[10] level operating system (OS), a C2 level database management system (DBMS) and the appropriate associated hardware. The C2 level OS was selected as the minimally acceptable OS for CAs because of the assurance and the accountability services that such an OS provides. A C2 level OS provides accountability down to a individual level through login procedures and also provides auditing for security-relevant events. These features were deemed necessary for a CA to properly function and be consider trusted by its associated users. In addition to the security services provided by the trusted OS, public key cryptography and associated digital signature technology is used on the messages generated by the infrastructure, to ensure sender authentication and message integrity. Infrastructure messages include requests for certificates, requests for CRLs, delivery of certificates, and delivery of CRLs, etc. It is recommended that prior to the implementation of each CA, a thorough risk analysis be

---

[10] C2 is one of the levels within the Controlled Access Protection class described in the *Department of Defense Trusted Computer System Evaluation Criteria*. [26]

conducted to ensure the adequacy of the above security controls for that specific CA and to identify the need for and cost of additional security controls.

I.2 COST MODEL OVERVIEW

The PKI Cost Model uses a bottom-up methodology. The basic approach is to estimate the resources required for the PKI infrastructure and convert them into cost estimates. The cost model approach is flexible in that it may be adapted to any PKI organizational structure.

The PKI Cost Model starts with a selected set of operational activities described in section 6.1. The activities are discussed in detail, using lists of steps that show what actions are required to complete the activities. The cost model examines activities and steps at five levels: the user, the KG, the CA, the ORA and the Directory. For each step (or group of steps), a list of resources required to accomplish that step are estimated and used as inputs to the cost model. The resources are grouped into four categories: storage, communications, processor, staff. They are stated in terms of actions and items. For example, a communications need may be to send a message and a staff need may be travel to a CA.

After the resources have been estimated in terms of actions and items, the PKI Cost Model translates these needs into units applicable to the resource categories. Storage items are translated into megabytes (MBs), communications actions are translated into kilobits (Kbs), and processor and staff actions are translated into time. The translated quantities are multiplied by the number of times the respective activities will be performed annually, and then summed to arrive at the total capability required for each resource category.

Note that time was selected as the measure of processor usage, as opposed to a more traditional measure such as millions of instructions per second (MIPS), since the PKI functions could be more easily estimated in terms of processor time. For example, signing a message can be estimated in seconds more easily than in MIPS. However, the amount of time used is more dependent upon a particular processing system than is the number of MIPS.

Storage, processor, and staff resources are acquired in discrete amounts. These resources are fully paid for and can be utilized from 0 to 100 percent. Communications resources, on the other hand, are a continuous resource for which a user is billed only in proportion to the communications resources used. The cost model determines how many hard drives (storage resources), computers (processor resources), and staff (staff resources) are required, taking into account that the last unit of each resource will not be fully utilized. Based on the resources required, initial and annual costs are estimated. Depending upon the system life, one may calculate life cycle costs as the initial costs plus the annual costs times the number of PKI operational years.

I.3 RESOURCES ASSOCIATED WITH OPERATIONAL ACTIVITIES

This section describes the set of operational activities which are being used by this cost model. Each activity is broken down into functional steps which need to be conducted in order to complete the activity. Resources required for each functional step of these activities are defined. The resources that may be needed are hard disk storage (abbreviated below as "Storage"), communications ("Comm"), processor activities ("Processor"), and staff-time ("Staff-Time"). The resources required for an activity are presented in relation to the entity (user, KG, CA, ORA, or Directory) and in relation to whether the entity is initiating the activity or is a recipient of the activity.

In addition to estimating the resources required for each step within each activity, it is necessary to estimate the average number of times each activity is expected to be executed each year, or equivalently the number of occurrences per year. After the activity is discussed, an estimate of the number of occurrences is given as a constant, as an input variable, or as a function of other constants and variables. The variable names used are as follows:

# children:	The number of users or subordinate CAs reporting to a CA.
# siblings:	The number of equivalent entities reporting to the same CA.
# cousins:	The number of equivalent entities reporting to other CAs.
# PKI levels:	The number of administration levels in the PKI.
Revocation rate:	The probability that any single entity's key will be revoked.
CRL frequency:	The annual periodicity with which CRLs will be distributed.

Whenever possible, the steps associated with PKI activities that are described below apply to all PKI relationships: user-CA, user-ORA, ORA-CA, CA-PCA and PCA-PAA. In some instances different relationships require different steps. If this is the case, then the steps below discuss the activity in terms of the user-CA relationship and differences between the user-CA steps and other relationship steps are indicated.

I.3.1 Generating, Certifying, and Distributing Keys

For costing purpose, we assume that a user goes to a centralized Key Generator (KG) to generate his key pair. This KG will be either collocated with a CA or an ORA. We allow for the existence of ORAs between the PKI users, but not between other PKI levels. Once the user has generated a key pair, he needs to present himself and his public key in person to either a CA or an ORA. Authentication of a user by a CA or an ORA takes place in person. The user's PKI credentials are delivered to him on a smart card.

CAs within the PKI generate their own key pairs. To obtain certificates, CA operators go to the appropriate parent CA in person and represent their CA in the identification process. They receive the CA's certificate signed by the parent CA.

ORAs within the PKI can either generate their own key pairs or have their operator go to a KG to generate a key pair for the ORA. The ORA operator registers the ORA's public key with the parent CA. It is left to the discretion of the CA or the policy specified by the PCA

as to whether a certificate is issued for the ORA. A certificate is optional, since the ORA only communicates with a CA, the CA with which it registered its public key.

The functions associated with the generation, certification and distribution of keys along with the resources needed to perform the functions are summarized in this section. Separate sets of functions and resources are involved when a user presents himself to a CA or to an ORA.

I.3.1.1 Interfacing with a CA

Activities and Resources Needed:

1. The user goes to the location of the KG and the CA in person. Staff-time is allowed for the user to travel to this location. (CAs are assumed to be more distant from users than ORAs are from users.)

User	Storage	Comm	Processor	Staff-Time
Initiator				2 hours

2. The user uses the KG to create a key pair. KG software is run by the user, so the KG does not need an operator. Therefore, staff time is not included for a KG. The two hours allotted to the user for traveling to the KG/ORA or KG/CA location includes the brief time it takes to generate a key pair.

KG	Storage	Comm	Processor	Staff-Time
Recipient			Make key pair	

3. The KG presents the key pair to the user on a smart card and destroys all copies of the key pair that it possessed.

KG	Storage	Comm	Processor	Staff-Time
Recipient			Move files, destroy all copies of key pair	

4. The user takes his public key to the CA in person and requests that the CA generate a certificate for him.

5. The user is authenticated. Staff time is allowed for the CA Operator to authenticate the user.

User	Storage	Comm	Processor	Staff-Time
Initiator				1/4 hour

CA	Storage	Comm	Processor	Staff-Time
Recipient				1/4 hour

6. The CA generates a certificate for the user and stores a copy of the certificate in its database and sort its database.

CA	Storage	Comm	Processor	Staff-Time
Recipient	certificate		Make certificate, add record, sort database	= Processor time

7. The CA puts the entity's certificate, the CA's public key, the PAA's public key and the PCA's public key on a smart card and gives the smart card to the user in person. Staff-time is allowed for the users to return to their office.

User	Storage	Comm	Processor	Staff-Time
Initiator				2 hours

CA	Storage	Comm	Processor	Staff-Time
Recipient			Move files	=Processor time

8. The CA sends a copy of the user's certificate to the Directory.

   Communications costs are borne by the sender. Staff-time is equivalent to the time it takes to send the message.

CA	Storage	Comm	Processor	Staff-Time
Initiator		Certificate	Send certificate	= Comm time

9. The Directory receives the certificate and store it.

   Although the communications cost of sending a message is borne by the initiator (as mentioned above), the recipient's processor is occupied while receiving the message. Directory functions are automatic and do not require staff intervention on behalf of the Directory server; therefore staff-time is not included for any Directory functions.

Directory	Storage	Comm	Processor	Staff-Time
Recipient	Certificate		Receive certificate	

Number of Occurrences Per Year:

User[Initiator]: This function is performed for each new PKI subscriber. Thereafter, it is executed only as required by other functions.

KG[Recipient]: This function is performed for each new PKI subscriber at a KG. Thereafter, it is executed only as required by other functions.

CA[Recipient]: This function is performed for each new PKI subscriber at a CA. Thereafter, it is executed only as required by other functions.

CA[Initiator]: This function is performed for each new PKI subscriber at a CA. Thereafter, it is executed only as required by other functions.
Directory[Recipient]: This function is performed for each new PKI subscriber at a CA. Thereafter, it is executed only as required by other functions.

I.3.1.2 Interfacing with an ORA

Activities and Resources Needed:

1. The user goes to the location of the KG and the ORA in person. Staff-time is allowed for the user to travel to this location. (CAs are assumed to be more distant from users than ORAs are from users.)

User	Storage	Comm	Processor	Staff-Time
Initiator				1/2 hour

2. The user uses the KG to create a key pair.

KG	Storage	Comm	Processor	Staff-Time
Recipient			As in I.3.1.1, Step 2	= Processor time

3. The KG presents the key pair to the user on a smart card and destroys all copies of the key pair that it possessed.

KG	Storage	Comm	Processor	Staff-Time
Recipient			As in I.3.1.1, Step 3	=Processor time

4. The user takes his public key to the ORA in person and requests that a certificate be generated for him.

5. The user is authenticated. Staff time is allowed for the ORA Operator to authenticate the user.

User	Storage	Comm	Processor	Staff-Time
Initiator				1/4 hour

ORA	Storage	Comm	Processor	Staff-Time
Recipient				1/4 hour

6. The ORA sends user's request and user's credentials to the CA signed using the ORA's private key.

ORA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Steps 1-7			

7. The CA receives the message from the ORA and verifies the ORA's signature.

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.2, Steps 1-6			

8. The CA generates a certificate for the user and stores a copy of the certificate in its database and sort its database.

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Step 6			

9. The CA sends a copy of the user's certificate to the ORA in a signed message.

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Steps 1-7			

10. The ORA receives the message containing the certificate from the CA and verifies the signature on the message.

ORA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.2, Steps 1-6			

11. The ORA puts the user’s certificate, the CA's public key, the PAA's public key and the PCA's public key on a smart card and delivers the smart card to the user in person. Staff-time is allowed for the user to travel to back from the ORA's office.

User	Storage	Comm	Processor	Staff-Time
Initiator				1/2 hour

ORA	Storage	Comm	Processor	Staff-Time
Recipient			Move files	=Processor time

12. The CA sends a copy of the user's certificate to the Directory.

    Communications costs are borne by the sender. Staff-time is equivalent to the time it takes to send the message.

CA	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.1, Step 8			

13. The Directory receives the certificate and store it.

    Although the communications cost of sending a message is borne by the initiator (as mentioned above), the recipient's processor is occupied while receiving the message. Staff-time is equivalent to the processor time.

Directory	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Step 9			

Number of Occurrences Per Year:

User[Initiator]: This function is not separately executed.
KG[Recipient]: This quantity is dependent upon the number of times a user initiates this function and the number of ORA children.
ORA[Recipient]: This quantity is dependent upon the number of times a user initiates this function and the number of ORA children.
CA[Recipient]: This quantity is the same as the number of ORA[Recipient] occurrences.
CA[Initiator]: This quantity is the same as the number of ORA[Recipient] occurrences.
Directory[Recipient]: This quantity is the same as the number of ORA[Recipient] occurrences.

Key generation, certification and distribution steps for other PKI entities, CAs and ORAs will be similar to those described above for the user. The major difference are that CAs and ORAs can generate their own key pairs and that representatives for these entities will travel to the appropriate parent CA for registration and authentication purposes. CAs register directly with another CA and do not register via an ORA.

Number of Occurrences Per Year:

CA[Initiator]: # new CAs
ORA[Initiator]: # new ORAs
CA[Recipient]: # new children

I.3.2 Signature and Verification

For costing purposes, we assume that the digital signature algorithm being used is DSA and that the algorithm is being implemented on a smart card[11], the same smart card that is used to store the user's private key and certificate. Software on the user's workstation will provide the interface between applications that sign the messages and the smart card in the smart card reader. This software may also include an implementation of the SHA and DSA. Digital signatures will be computed on the smart card, so the private key never has to leave the card. Verification of the digital signatures may be performed by the software residing on the workstation or on the smart card if it has sufficient computing power and storage.

---

[11] We use "smart card" in a somewhat generic way to include cpu-equipped cards, PCMCIA cards, smart disks, and similar technologies.

To compute the user's digital signature, the software residing on the computer hashes the message being signed. The message digest is sent to the smart card. Of course, the software must be protected so that it cannot be altered to substitute a different document for signature. The private key is retrieved from the part of the smart card on which it is stored and the digital signature is computed on the smart card by the DSA implementation. The signature is transferred from the smart card to the computer, where it is appended to the message or file by the software running on the computer. The signed message is then sent using the preferred communications method.

To verify a digital signature on a message, the software residing on the computer will take the received signed message and strip off the signature. The software will also take the message and compute the message digest using the SHA. The software will then obtain the public key certificate of the message's sender and any other certificates needed to verify the sender's certificate. The signature on the sender's public key certificate will be verified, this may include verifying signatures on other certificates as well. Once the sender's certificate is verified, the sender's public key can be retrieved from the certificate and used to verify the digital signature on the message.

The functions and resources associated with the execution of the signature and verification process within the PKI are summarized below.

I.3.2.1 Signature

Activities and Resources Needed:

1. Message to be signed is hashed by the SHA software on the workstation to produce a message digest.

2. The message digest is sent to the smart card.

3. Digital signature is computed on the smart card.

4. Digital signature is sent from the smart card to the workstation.

5. Software on workstation appends the digital signature to the message.

Steps 1 through 5 are unified into single resource tables, one for the user and one for the CA. Staff-time is equivalent to the processor time it takes to sign the message.

User	Storage	Comm	Processor	Staff-Time
Initiator			Sign message	= Processor time

CA	Storage	Comm	Processor	Staff-Time
Initiator			Sign message	= Processor time

6. Signed message is sent. For the entity, the only resources that are of concern are the incremental resources required for PKI activities. In sending a message, the signature is the incremental element. For the CA, since the CA will be a new entity, the entire message is accounted for in the resource estimation.

   The communications costs are borne by the sender. The staff-time is equivalent to the time it takes to send the signature or the message.

User	Storage	Comm	Processor	Staff-Time
Initiator		Signature	Send signature	= Comm time

CA	Storage	Comm	Processor	Staff-Time
Initiator		Message	Send message	= Comm time

7. Signed message is stored. The signature or message is stored for future reference.

User	Storage	Comm	Processor	Staff-Time
Initiator	Signature			

CA	Storage	Comm	Processor	Staff-Time
Initiator	Message			

Number of Occurrences Per Year:

User[Initiator]: # signed messages
CA[Initiator]: This function is executed by CAs and ORAs only for PKI messages required in other functions.

For the user, the number of occurrences per year equals the number of PKI messages sent above (to its CA), plus the number of signed messages (signed messages are defined as non-PKI signed messages). For the CA, the number of occurrences per year equals the number of PKI messages sent above (to its PCA), plus the number of PKI messages sent below (perhaps to its users). All PKI messages are signed.

I.3.2.2 Verification of User Signature

Activities and Resources Needed:

1. Signed message is received.

   Although the communications cost of sending a signature or message is borne by the

initiator (as mentioned above), the recipient's processor is occupied while receiving. Staff-time is equivalent to processor time, i.e., the user or CA waits for the signature or message.

User	Storage	Comm	Processor	Staff-Time
Recipient			Receive signature	= Processor time

CA	Storage	Comm	Processor	Staff-Time
Recipient			Receive message	= Processor time

2. Sender's certificate is retrieved either from the cache (See section I.3.5.3) or from the Directory (See section I.3.3.1)

3. Sender's public key is retrieved from a certificate.

4. Digital signature is stripped from message.

5. The message is hashed to reproduce the message digest.

6. Digital signature is verified.

Steps 3 through 6 are combined in the resource tables below. Staff-time is equivalent to the time it takes the processor to verify the signature.

User	Storage	Comm	Processor	Staff-Time
Recipient			Verify signature	= Processor

CA	Storage	Comm	Processor	Staff-Time
Recipient			Verify signature	= Processor

Number of Occurrences Per Year:

User[Recipient]: # signed messages received + # PKI messages received from CAs in performing other functions

CA[Recipient]: # PKI messages, in support of other functions, received from all sources

It is assumed that every message has one sender and one recipient. Thus, the number of verifications is equal to the number of signatures in section I.3.2.1.

I.3.3 Obtaining Certificates

For costing purposes, we assume that certificates and certification paths are not sent with signed messages or in separate messages transmitted prior to signed communications between entities. We presume that, at a minimum, a sender will include both his PKI unique name and the PKI unique name of his CA with the signed message. Each application must, in general, decide how it will handle certificates and certification paths.

For our model, it is the responsibility of the receiver of a message to calculate the certificate path of the sender by examining the unique name of sender and determining from the name the location of the sender within the infrastructure. The receiver requests all the certificates in the path from the sender to an ancestor of the receiver for which he has the public key. Within the PKI, entities will have the public keys of their ancestors: PAA, PCA and CA. Requests for certificates are sent to the Directory, which stores the certificates for all entities in the PKI. Each certificate is sent to the Directory by the CA which produces it. It is the CA's responsibility to keep the Directory listing current and accurate.

The functions and resources associated with obtaining certificates from the Directory under the modeled PKI CONOPS are presented in this section.

I.3.3.1 Obtaining Certificates from the Directory

Activities and Resources Needed:

1. Receiver determines which certificate(s) it needs.

2. For each certificate that it needs, the initiator requests the entity's certificate from the Directory. A message is sent by a user or by a CA to request a certificate. The Directory receives the message.

   The communications costs are borne by the sender. The staff-time is equivalent to the time it takes to send the message.

User	Storage	Comm	Processor	Staff-Time
Initiator		Message	Send message	= Comm time

3. The Directory receives the message requesting a certificate.

   Although the communications cost of sending a message is borne by the initiator (as mentioned above), the recipient’s processor is occupied while receiving.

Directory	Storage	Comm	Processor	Staff-Time
Recipient			Receive Message	

4. The Directory retrieves the certificate(s) it has for the entity.

Directory	Storage	Comm	Processor	Staff-Time
Recipient			Find record	

5. a. The Directory sends the certificate(s) to the requesting entity.

   b. If the Directory does not have the requested certificate, the Directory sends a message to the requesting entity stating that the certificate does not exist.

   As the recipient of the original request, the Directory sends the message containing the certificate. The initiator of the original request receives the message containing the certificate and stores the certificate for future use. If the Directory sends more than one certificate for an entity, the requester needs to determine which certificate is associated with the digital signature it is trying to verify. Staff-time is equivalent to the processor time.

Directory	Storage	Comm	Processor	Staff-Time
Recipient		Certificate(s)	Send certificate(s)	

User	Storage	Comm	Processor	Staff-Time
Initiator	Certificate		Receive certificate	Processor

Number of Occurrences Per Year:

User[Initiator]: # certificates retrieved by a user
Directory[Recipient]: # certificate requests x # children

The number of certificates per user is a function of the number of siblings and the number of cousins as well as the number of PKI levels. The number of certificate requests is dependent upon the number of PKI entities and the number of certificates per entity.

I.3.4 Verifying Certificates

The functions and resources associated with verifying certificates within the PKI model are described in the following.

Activities and Resources Needed:

1. Entity has obtained all the certificates that it needs. (See section I.3.3, Obtaining Certificates)
2. The entity begins the verification process by taking the public key of common ancestor and verifying the signature on the certificate the common ancestor signed.

User	Storage	Comm	Processor	Staff-Time
Initiator			Verify signature	= Processor time

3. Entity extracts public key from verified certificate.

User	Storage	Comm	Processor	Staff-Time
Initiator			Extract key	= Processor time

4. Entity repeats steps 2 and 3 until sender's certificate is verified and sender's public key is extracted from certificate.

   The resources in steps 2 and 3 are multiplied by the number of PKI levels above the initiator in order to estimate the resources needed to verify signatures up to a common ancestor.

Number of Occurrences Per Year:

User[Initiator]: # certificates retrieved by a user

I.3.5 Caching Certificates

For costing purposes, we assume that a user cache certificates of the users with whom he communicates frequently. The user should verify the certification path once and then store the user's certificate within the cache. If the cache becomes full, an LRU-based deletion method should be employed. It is also assumed that any user that maintains a certificate cache will periodically check the certificates within the cache against the appropriate CRL(s); removing any revoked certificates from the cache.

In cost model, we assume that users maintain certificate caches and assumed that CAs would also maintain certificate caches. We expect users to always check their cache first for any necessary certificates before requesting them from the Directory. A user may choose to check the certificate retrieved from the cache against the latest CRL before using it. However, this specific function will not be included in the cost of this activity.

Functions and resources associated with this approach for caching certificates are outlined in the following.

I.3.5.1 Putting Certificates into the Cache

Activities and Resources Needed:

1. The user has verified the sender's certificate (See section I.3.4, Verifying Certificates)
2. Entity discards all certificates except that of the sender.
3. Entity stores sender's certificate in cache.

The certificate is stored for future use. A record of the certificate is kept in the entity's certificate database which making a new record and sorting the database.

User	Storage	Comm	Processor	Staff-Time
Initiator	Certificate		Add record, sort database	= Processor time

4. If the cache is full, an LRU-based method is used to delete certificates from the cache to make room for new certificates.

User	Storage	Comm	Processor	Staff-Time
Initiator			Find LRU certificate, delete certificate	= Processor time

For the cost model, it is assumed that step 4 is not executed.

Number of Occurrences Per Year:

User[Initiator]: # certificates in user's cache

I.3.5.2 Checking Certificate Cache Against CRLs

Activities and Resources Needed:

1. User obtains CRL(s) either automatically or through a request. The user has obtained CRLs using a method as described in section I.3.9. The cost of obtaining CRLs has been counted in that section.

2. User compares certificate in cache with those on the CRL(s).

User	Storage	Comm	Processor	Staff-Time
Initiator			Find record	= Processor time

3. Any certificate on the CRL(s) is removed from the cache.

User	Storage	Comm	Processor	Staff-Time
Initiator			Delete record	= Processor time

Number of Occurrences Per Year:

User[Initiator]: # certificates in user's cache x CRL frequency

I.3.5.3 Obtaining Certificates from a Cache

Activities and Resources Needed:

1. The user checks the certificate cache for a copy of the sender's certificate.

User	Storage	Comm	Processor	Staff-Time
Initiator			Find file	= Processor time

2. If the certificate is in the cache, the user obtains a copy of the certificate from the cache and extracts the public key from the certificate.

User	Storage	Comm	Processor	Staff-Time
Initiator			Extract key	= Processor time

3. If the certificate is not in the cache, the entity obtains and verifies the certificate in the normal manner (See sections I.3.3, Obtaining Certificates and I.3.4, Verifying Certificates).

Number of Occurrences Per Year:

User[Initiator]: # certificates retrieved by a user

I.3.6 Reporting Key Compromise or Severed Relations

The functions and activities associated with reporting a key compromise or severed relation are shown below. The steps for reporting to a CA and to an ORA are presented below.

I.3.6.1 Notifying a CA

Activities and Resources Needed:

1. Entity's private key is compromised or entity is no longer associated with a CA.

2. a. The CA which issued the certificate is notified of the compromise or severed relations. Round-trip time is allowed for the user or CA to travel to the next higher CA for out of bands notification.

User	Storage	Comm	Processor	Staff-Time
Initiator				4 hours

CA	Storage	Comm	Processor	Staff-Time
Initiator				1 day

b. If the compromised entity is a CA, it will also announce the compromise to its children and to other entities with which it is cross certified. This step is repeated, based once for each child and once for each cross certified sibling.

CA	Storage	Comm	Processor	Staff-Time
Initiator		Message		= Comm time

3. CA receives compromise or severed relations notification and verifies it authenticity.

CA	Storage	Comm	Processor	Staff-Time
Recipient				1/4 hour

4. CA marks the certificate in its database as revoked.

CA	Storage	Comm	Processor	Staff-Time
Recipient			Find record, change record	= Processor time

5. CA places the certificate on the CRL in readiness for the next CRL issue. A record is created and the database is sorted.

CA	Storage	Comm	Processor	Staff-Time
Recipient	CRL		Add record, sort database	= Processor time

Number of Occurrences Per Year:

User[Initiator]: Revocation rate
CA[Initiator]: Revocation rate x (# children + # siblings)
CA[Recipient]: Revocation rate x # children

I.3.6.2 Notifying an ORA

Activities and Resources Needed:

1. User's private key is compromised or entity is no longer associated with a CA.

2. The ORA is notified of the compromise of severed relations and verifies it authenticity. Round-trip time is allowed for the user to travel to the ORA for out of bands notification.

User	Storage	Comm	Processor	Staff-Time
Initiator				1 hour

ORA	Storage	Comm	Processor	Staff-Time
Recipient				1/4 hour

3. ORA sends signed message to CA notify it of the compromise or severed relation.

ORA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Steps 1-7			

4. CA receives notification and verifies the ORA's signature.

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.2, Steps 1-6			

5. CA marks the certificate in its database as revoked.

CA	Storage	Comm	Processor	Staff-Time
Recipient			As in I.3.6.1, Step 4	= Processor time

6. CA places the certificate on the CRL in readiness for the next CRL issue. A record is created and the database is sorted.

CA	Storage	Comm	Processor	Staff-Time
Recipient			As in I.3.6.1, Step 5	= Processor time

Number of Occurrences Per Year:

User[Initiator]: Revocation rate
ORA[Recipient]: Revocation rate x # children
CA[Recipient]: Revocation rate x # children

The revocation rate is the probability that any single entity will report a key compromise or severe relations.

I.3.7 Recovering from a Key Compromise

For costing purposes we will assume that no duplicate CAs exist. Instead, compromise of a CAs key would interrupt service and the detailed compromise recovery procedure outlined in section 5.2.8 would need to be implemented. It should be noted that if the child of a compromised CA is also a CA, the child CA may also need to reissue its certificates. This would be the case if it received a new private key during its parent's recovery

procedure. The child CA would then follow the recovery procedure described in section 5.2.8.

The functional steps and associated resources required for a PKI entity to recover from a key compromise are as follows:

I.3.7.1 Users

Activities and Resources Needed:

1. The user generates a new key pair and its CA generates a new certificate for the user that was compromised. The key pair and certificate is distributed in the normal manner. If a user is recovering from a key compromise, the roles of the user as initiator and the KG and the CA or the ORA as recipient are the same as in Generating, Certifying, and Distributing Keys, section I.3.1.

Resource tables for users interfacing with a CA are as follows.

User	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.1, Steps 1, 4-5, 7			

KG	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Steps 2-3			

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Steps 4-8			

Directory	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Step 9			

Resource tables for users interfacing with an ORA are as follows.

User	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.2, Steps 1, 5, 11			

KG	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Steps 2-3			

ORA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Steps 5-6, 10-11			

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Steps 7-9, 12			

Directory	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.2.1, Step 13			

2. Entity receives new credentials and begins to use them.

Number of Occurrences Per Year:

User[Initiator]: Revocation rate
KG[Recipient]: Revocation rate x # children
CA[Recipient]: Revocation rate x # children
ORA[Recipient]: Revocation rate x # children
Directory[Recipient]: Revocation rate x # children

I.3.7.2 CAs

Activities and Resources Needed:

1. CA generates a new key pair and the parent of the CA generates a new certificate for the CA. All public keys and certificates are distributed in the normal manner. If a CA is recovering from a key compromise, the CA as initiator and next higher CA is the recipient in the format used in Generating, Certifying, and Distributing Keys, section I.3.1. No interaction between a CA operator and a KG is necessary, but time should be allotted for key generation.

CA	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.1, Steps 1-2, 4-5, 7			

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Steps 4-8			

Directory	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Step 9			

2. CA receives new credentials and begins to use them.

3. If the CA's certificate database was compromised, the CA contacts its users (or subordinate CAs) in an out of bands method and requires that they initiate the process to obtain new key pairs (See section I.3.1).

CA	Storage	Comm	Processor	Staff-Time
Initiator		Message	Send Message	= Comm time

User	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.1, Steps 1-3			

4. Whether new keys are generated by a CAs subordinates or not, the CA issues new certificates to replace those it issued under the compromised key. The CA creates new certificates by signing public keys within its database using its new private key. Children request new certificates in the normal manner (See Section I.3.1 Generating, Certifying and Distributing Keys).

The chart for children who deal directly with the CA and not through an ORA is:

User	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.1, Steps 1, 4-5, 7			

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Steps 4-8			

Directory	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Step 9			

The chart for children who deal indirectly with the CA through an ORA is:

User	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.1.2, Steps 1, 4-5, 11			

ORA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.2, Steps 4-6, 10-11			

CA	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.2, Steps 7-9, 12			

Directory	Storage	Comm	Processor	Staff-Time
Recipient	As in I.3.1.1, Step 13			

5. The CA revokes the old certificates which were signed with the compromised key and places them on the CRL.

CA	Storage	Comm	Processor	Staff-Time
Initiator			Find record, change record	= Processor time

The above step is repeated for each child (# children).

6. The CA issues a CRL and distributes the CRL in the normal manner. (See I.3.8 Obtaining CRLs)

CA	Storage	Comm	Processor	Staff-Time
Initiator	As in I.3.8.1			

Number of Occurrences Per Year:

User[Recipient]: Revocation rate
CA[Initiator]: Revocation rate
CA[Recipient]: Revocation rate
Directory[Recipient] Revocation rate

I.3.8 Obtaining CRLs

For the PKI cost model, it is assumed that each CA periodically generate and distribute a CRL to the Directory. Requests for CRLs are sent to the Directory. The functions and resources associated with this approach to obtaining CRLs within the PKI are described in this section.

I.3.8.1 Automatic Distribution of CRLs to the Directory

Activities and Resources Needed:

1. On a periodic basis, the CA generates a CRL and distributes the CRL to the Directory.

   The communications costs are borne by the initiator. The staff-time is equivalent to the time it takes to send the message.

CA	Storage	Comm	Processor	Staff-Time
Initiator		CRL	Send CRL	= Comm time

2. The Directory receives CRL and store it.

   Although the communications cost of sending a message is borne by the initiator, the recipient's processor is occupied while receiving the message. Staff-time is equivalent to processor time.

Directory	Storage	Comm	Processor	Staff-Time
Recipient	CRL		Receive CRL	

Number of Occurrences Per Year:

CA[Initiator]: CRL frequency
Directory[Recipient]: CRL frequency

I.3.8.2 Requests for CRLs

The steps required to request a CRL from the Directory along with the associated resource tables are included below.

Activities and Resources Needed:

1. User sends a request for a CRL to the Directory.

   The communications costs are borne by the initiator. The staff-time is equivalent to the time it takes to send the message.

User	Storage	Comm	Processor	Staff-Time
Initiator		Message	Send message	= Comm time

2. The Directory receives the message requesting the CRL.

   Although the communications cost of sending a message is borne by the initiator, the recipient's processor is occupied while receiving the message.

Directory	Storage	Comm	Processor	Staff-Time
Recipient			Receive message	

3. The Directory finds the latest copy of the CRL.

Directory	Storage	Comm	Processor	Staff-Time
Recipient			Find record	

4. The Directory sends the CRL to the user that requested it.

Directory	Storage	Comm	Processor	Staff-Time
Recipient		CRL	Send CRL	

5. User receives CRL and stores it.

User	Storage	Comm	Processor	Staff-Time
Initiator	CRL		Receive CRL	= Processor time

Number of Occurrences Per Year:

User[Initiator]: # user cousins
Directory[Recipient]: # CRL requests x # children

The user will request the CRL associated with every certificate it retrieves from the Directory. The user will also request CRLs periodically to check its cache of certificates. The Directory will receive all these CRL requests.

I.3.9 Archiving

In the cost model, we assume that CAs within the PKI will copy their certificate databases and their CRL databases to the archive location at least once a year. CAs will also archive their audit files. User archiving certificates and CRLs with documents may be an application requirement and is not included in the PKI cost model.

The functional steps and resources necessary to archive information are as follows:

Activities and Resources Needed:

1. CA will periodically (at least once a year at key changeover time) make a copy of its certificate set, the CRL database, and its audit file. Users may also archive certificates they have used.

2. The CA or user will move the copies to the archive location.

User	Storage	Comm	Processor	Staff-Time
Initiator	Certificate set, CRL set, audit file		Move files	= Processor time

CA	Storage	Comm	Processor	Staff-Time
Initiator	Certificate set, CRL set, audit file		Move files	= Processor time

The certificate set size is dependent upon the number of certificates the user or the CA possesses.

Number of Occurrences Per Year:

User[Initiator]: 1
CA[Initiator]:`1

It is assumed that archiving is conducted once a year.

I.3.10 Auditing

For the cost model, we assume that each CA within the PKI audits security relevant events. Examples of such events were listed in section 5.2.11.

The functions and resources associated with auditing within the PKI are outlined in the following list.

Activities and Resources Needed:

1. When a security relevant event occurs, the appropriate audit message is generated.

2. The audit message is logged in the audit file along with other relevant information such as the date and time of the event. A new record is created for each audit event and stored.

CA	Storage	Comm	Processor	Staff-Time
Initiator	Audit text		Add record	=Processor time

Number of Occurrences Per Year:

CA[Initiator]: = I.3.1 + I.3.6 + I.3.8

I.3.11 Rekeying and Recertifying

For costing, we assume that new key pairs and certificates should be issued for the entities within the PKI once a year and that all entities within the PKI change their key on the same date. We assume that the smart cards have enough space to store two private keys and two certificates, the current and the new set of keys and certificates.

The functions and resources associated with the recommended PKI rekeying procedure are described as follows:

Activities and Resources Needed:

1. Entity requests new key pairs and certificates from their parent CA prior to changeover date.

2. Users generate key pairs using the KG, CAs generate their own key pair. CAs generate certificate and distributes in the normal manner.

3. On the changeover date the entity begins to use the new private key to sign and stores the new certificate.

Steps 1 through 3 require the same resources as Generating, Certifying, and Distributing Keys (section I.3.1).

4. Old key and certificate may be archived by the entity.

User	Storage	Comm	Processor	Staff-Time
Initiator	Private key, certificate		Move files	= Processor time

CA	Storage	Comm	Processor	Staff-Time
Initiator	Private key, certificate		Move files	= Processor time

Number of Occurrences Per Year:

User[Initiator]: 1
CA[Initiator]: 1
CA[Recipient]: # children
ORA[Recipient]: # children
Directory[Recipient]: # children

Entities need to get new key pairs only once a year, hence the "1"s for the initiators. For the parent CAs, the number of annual rekeys, as a recipient, is equal to the number of its children.

I.4 MODEL STRUCTURE

The PKI Cost Model was implemented using Microsoft Excel for the Macintosh version 4.0 in a workbook. A workbook is a collection of Excel files (e.g., spreadsheets, charts, macro sheets). The PKI Cost Model workbook is a set of eight spreadsheets, each considered a separate module in the model. (A separate module for the PAA was not created since its functions would be similar to a PCA.) There is a module for each entity represented: user, KG, ORA, CA, PCA and Directory. There is also a module where the analyst enters the infrastructure configuration data and another that tabulates the costing results. Using a workbook structure ensures that the spreadsheets are maintained as a set and that cross-spreadsheet links are not broken.

The six entity modules have the same basic structure, shown in figure I-1. This figure is repeated at the beginning of each entity module in the PKI Cost Model. The figure shows five tables: Module Inputs, Translation Table, Function Needs, Capabilities Required, and Resources Required and Estimated Costs.

The Module Inputs are at the beginning of the module and are divided into analyst inputs and calculated inputs. The person using the model enters inputs into the Input Module and they are automatically copied into the entity modules in which they are used. Each module computes its calculated inputs from these entries. The calculated inputs are shown to provide feedback to the analyst.

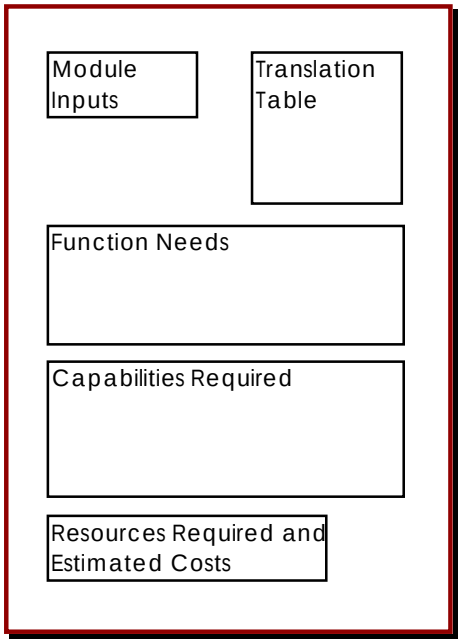


Figure I-1. Entity Spreadsheet Layout

The Function Needs table is a summation of the resource tables, by function, as presented in section I.3. The Translation Table is a table of values for converting the Functions Needs table into the Capabilities Required table (section I.2 discusses the translation process in more detail). Sums are calculated from the Capabilities Required table and converted into cost estimates in the Resources Required and Estimated Costs table. Initial and annual costs are estimated for a single entity. These results are automatically transferred to the Summary Results spreadsheet, the last page in the workbook.

I.5 MODULE INPUTS

Several inputs are required for the PKI Cost Model. The inputs fall into two categories: architectural and operational. Both types of inputs are discussed in the following paragraphs. It should be noted that several of the inputs require some hand calculations before being entered into the model.

Architectural:

- *# of PKI levels* . This input is the number of PKI levels in the PKI hierarchy.

- *# of children per CA.* The average number of users whose public keys are certified (directly and indirectly) to a CA. This input is related to the next two by

the equation
# of children per CA = (*# of ORAs per CA* + 1) * *# of users per ORA or CA.*

- *# of users per ORA or CA.* This input is an estimate of the average number of users reporting to either an ORA or to a CA directly.

- *# of ORAs per CA.* This input is an estimate of the average number of ORAs under a particular CA.

- *# of CAs per PCA.* This input is an estimate of the average number of CAs under a particular PCA.

- *# of PCAs per PAA.* This input is an estimate of the average number of PCAs under the PAA.

Operational:

- *Revocation rate for users*. This input, stated as a percentage, is an estimate of the proportion of users that will have their private keys revoked in a year for any reason except periodic rekeying. For example, a 10 percent revocation rate for users means that one out of ten users will have his private key revoked in a year.

- *Revocation rate for ORAs and CAs*. This input, stated as a percentage, is an estimate of the proportion of ORAs or CAs that will have their private keys revoked in a year for any reason except periodic rekeying. Since ORAs and CAs private keys have considerable importance, it is reasonable to expect that they will be safeguarded with maximum care and therefore their revocation rates will be less than the users' revocation rate.

- *CRL frequency*. This input is an estimate of how many times in a year each CRL is regularly distributed. A value of 52 would imply weekly distribution, while a value of 12 would imply monthly distribution.

- *cache size of user.* The average number of certificates that a user will cache in a year.

- *# of daily signed messages*. This input is an estimate of how many signed business messages (i. e., those not related to the administration or use of the PKI) an average user will receive and need to verify daily.

- *% of siblings*. Of all the certificates that an entity has cached, this input is an estimate of the percentage of those certificates that belong to siblings. Siblings are two entities who are certified by the same CA at the next level of the infrastructure.

- *% of 1st cousins for a user.* Of all the certificates that a user has cached, this input is an estimate of the percentage of those certificates that belong to first cousins. First cousins are entities that are under the same PCA but under a different CA.

- *# of cousins per CA.* Of all the certificates that a user has cached, this input is an estimate of the number of those certificates which belong to cousins under the same remote CA.

- *# of certificates retrieved by user.* This value is an estimate of how many certificates one user will need to retrieve within a year.

- *# of CRLs retrieved by user.* This value is an estimate of how many CRLs one user will need to retrieve within a year.

I.6 TRANSLATION TABLES

A Translation Table was created for each module. A Translation Table is a table of values for converting a Functions Needs table (stated in objects or actions) into a Capabilities Required table (stated in units with which to estimate costs). For example, the Functions Needs table may estimate that 200 messages be sent every year. If the Translation Table estimates that each message will be 100 bytes, then the communications load with regard to messages is 19.5 KB (200 messages times 100 bytes per message divided by 1,024 bytes per kilobyte [KB]). A combined Translation Table (i.e., all the model Translation Tables combined) is given in table I-1. An explanation for how each translation value was obtained follows.

Table I-1. Combined Translation Table

Storage	
audit file	# audit records * audit record size
audit record size	100 bytes
certificate	500 bytes
certificate set	# certificates * certificate
user revocation rate	10% certificates
CA/ORA revocation rate	2% certificates
CRL header	51 bytes
CRL record size	9 bytes
CRL set	CRL header + revocation rate * # children * CRL record size
private key	296 bytes
signature	40 bytes
Communications	
certificate	See Storage
CRL set	See Storage
message	100 bytes
signature	See Storage
Processor	
add record	5 seconds
change record	60 seconds
extract key	5 seconds
find file	10 seconds
find record	3 seconds
make certificate	30 seconds
make key pair	60 seconds
move files	10 seconds
receive message	0 seconds
receive signature	0 seconds
receive certificate	0 seconds
receive CRL	1 seconds
send certificate	0 seconds
send CRL	1 second
send message	0 seconds
send signature	0 seconds
sign message	2 seconds
sort database	10 seconds
verify signature	3 seconds

I.6.1 Storage and Communications

audit record size 100 bytes

This is a short record of a significant event. It requires a time and date, a local unique name, and a code for the event. A time and date are each six characters for a total of 12 bytes A name is assumed to have 32 characters or 32 bytes. The code for the audited event is no more than one byte. This accounts for 45 bytes. The remainder of the 100 byte audit record is a conservative estimate of the space required to save other pertinent information.

certificate 500 bytes

A public key and its parameters require three numbers of up to 1024 bits each and one number of 160 bits. This is a total of 3232 bits or 404 bytes. There are two dates to bound the validity period. These require 12 bytes. The certificate also contains two names, one for the issuer and one for the subject, totaling of 64 bytes. The allowance for a certificate serial number is 20 bits. The subject's and the issuer's signature algorithm identifiers are one byte each. We do not include the issuer's public key parameters or auxiliary information about the subject.

CRL 10 percent certificates

It is estimated that about 5 percent of all user certificates will be revoked because of key compromise and roughly an equal number will be revoked because the user is no longer associated with the particular CA.

CRL set 51 bytes of header information plus 9 bytes per certificate listed

A CRL header contains the issuer's unique name (32 bytes), three dates (six characters each or 18 bytes), and a signature algorithm identifier (one byte). Each certificate listed requires the insertion of, at a minimum, a serial number (20 bits) and a date of revocation (6 bytes)

message 100 bytes

These are PKI messages requesting certificates or CRLs.

private key 296 bytes

A key and it parameters consist of two numbers of up to 1024 bits and two numbers of 160 bits making a total of 2368 bits.

I.6.2 Processor and Communications

Processor times are derived in one of four different ways. Some come from DBMS function benchmarks; others, from advertised Information Security Corporation DSA chip benchmarks. Some times are determined by dividing the number of bits processed by the internal bus speed. That speed is taken to be 125KB/sec. The remaining times are the result of analyst assessment. As a conservative approximation to the costs, it is assumed that

processes that will normally run at night, with no real impact on the running expenses, are executed during normal working hours.

Communication times are derived by assuming that the speed of the LAN to which each workstation is connected is the controlling factor. The LAN speed is taken to be 1Mb/sec.

I.6.2.1 DBMS Benchmarks[12]

add record	5 seconds
change record	60 seconds
find record	3 seconds
sort database	10 seconds

I.6.2.2 DSA Chip Benchmarks

sign message	2 seconds
verify signature	3 seconds

I.6.2.3 Communications Times[13]

receive certificate	0 seconds
receive CRL	1 seconds
receive message	0 seconds
receive signature	0 seconds
send certificate	0 seconds
send CRL	1 second
send message	0 seconds
send signature	0 seconds

I.6.2.4 Analyst Assessment

extract key	5 seconds
find file	10 seconds
make certificate	30 seconds
make key pair	60 seconds
move files	10 seconds

---

12 Like all other assumptions in this cost estimate, these numbers are conservative. DBMS benchmarks are for a full-fledged relational DBMS. All that the PKI may need is some type of searching and sorting software.

13 Here the conservative assumption is that the CPU is inactive for the duration of entire communication operations. This, of course, is usually untrue.

I.7 ASSUMPTIONS

There are a number of assumptions that cause the model to produce a conservative estimate of the total cost of installing the federal portion of the PKI and of running it each year.

I.7.1 General Assumptions

Any cost estimate depends on some data that are unknown without any practical experience with full-blown digital signature operations. The missing numbers include such values as: the number of different users whose signature an average user verifies annually; the number of different CAs whose CRLs are required to revalidate all the certificates in an average user's cache; the time a workstation is unavailable because it is involved in telecommunication functions. In the absence of any empirical data, it is necessary to make the following worst case assumptions.

- There are 240 working days in a year.
- Every user workstation, CA and ORA is connected to a LAN.
- Effective LAN throughput is 1Mb/sec; LAN use is free.
- VAN costs are approximately $0.02/KB (based on average FTS2000 pricing).
- The Central Processing Unit (CPU) is inactive during telecommunication functions. This is not necessarily true, but is assumed for cost purposes.
- On average, a user verifies 80% of his received signed messages with public keys already in his cache. 20% of the time, he must request a new certificate.
- There are some unforeseen circumstances that may occur at a CA which increase its load by 10%.
- CRLs are distributed to subordinate and subscribing entities bi-weekly.

I.7.2 Number and Size of CAs

It is assumed that each major executive department and several groups of independent administrative agencies have a PCA/CA. There are about 20 of these. All bureaus, agencies and subdepartments that appear at the second level of their department's or agency's management tree (see [24]) are assumed to have either a CA or an ORA. There are about 510 of these organizational units.

The model begins with a belief that each CA can service about 30,000 users. The total civilian workforce in the executive branch is roughly 3,048,000 (see, for example, page 158

of [25])[14]. This number implies that a total of 102 CAs are required. Thus, the remaining 408 of the 510 organizational units will be served by ORAs. On the average, there are four ORAs for each CA and approximately five CAs per PCA[15]. At each CA, approximately /510 or 20 percent of the users are registered by the CA directly. This amounts to some 1026,000 users. The remaining 24,000 users register through one of the four ORAs–an average 6,000 users per ORA.

I.7.3 Sizes of Caches and CRLs

The size of a user's certificate cache will vary with the number of corespondents whose signatures he verifies. However, without any concrete experience, it is difficult to quantify that relationship. For the purpose of establishing cost estimates, cache size is fixed at 30 certificates.

As indicated in section I.6.1, it is assumed that about 5 percent of all certificates issued by a CA are revoked because the private key corresponding to the public key they hold has been compromised. A similar 5 percent of certificates are withdrawn because the holder of the certificate has severed his connection with the organization with which the CA is associated. At the same time, it is expected that an equal number of new associations will be made. Thus, about 5 percent of certificates held by a CA during a year are newly issued to replace those with compromised keys and another 5 percent are issued to new employees. This makes the revoked certificates ten percent of all certificates issued. It is a slightly more conservative estimate than the experience gained at the STU-III CA. There the fraction of certificates that are revoked is around 6.5 percent.

I.7.4 Equipment

The pricing estimates are based on a number of assumptions concerning equipment. Each CA has one or more SPARCstation 10 Model 30 workstations, with a corresponding trusted operating system and a relational database management system. Each PCA and each ORA runs on a single IBM PC, also with a corresponding trusted operating system and DBMS. Users need interface units so that their workstations and their smart cards can communicate. However, user workstations are already available as is access to wide area networks from the LANs to which these various entities are connected.

---

[14] It seems reasonable to base the analysis on this number. Obviously, not every employee in the Executive Branch will need to sign documents. On the other hand, there are many in the Armed Services who will use the PKI for non-classified electronic transactions. It is assumed that these numbers roughly balance out.

[15] This number seems extremely low, suggesting that the entire executive branch can manage quite well with a single PCA. Alternatively, if each department retains its own PCA, the number suggests that users register with the PCA rather than with a CA under the PCA. This is true for all users, whether they register directly or through an ORA.

I.7.5 Unit Costs

Various unit costs are used in the PKI Cost Model to reflect initial and annual costs. These unit costs are obtained via vendor quotes or catalog prices. These costs are discussed below by category: communications, commercial off-the-shelf (COTS) items, and staffing. The items chosen for the PKI Cost Model are chosen solely for cost purposes and may not be part of an optimal technical solution.

I.7.5.1 Communications

All the communications costs appear as annual costs only, never as initial costs. They are based on FTS 2000 costs which ranged from $0.45 to $2.00 per 64 KB. An estimated cost of $1.00 per 64 KB is used as an average. It is assumed that all PKI entities have network connections.

I.7.5.2 COTS Items

There are four COTS items in the PKI Cost Model: computers, operating systems, hard drives, and smart card systems.

Computer costs are both initial and annual. For the KGs, the ORAs and the PCAs, IBM PC clones are chosen, specifically the Dell Computer 433/L which uses an Intel 80486DX processor at 33 MHz. The 433/L computer comes with 4 MB RAM and an internal 230 MB drive and runs at about 1.4 Mflops. The 433/L purchase cost, about $2,000, was obtained from a Dell Computer advertisement. An annual maintenance cost of $99 is an analogy based on IBM's on-site maintenance cost for IBM PS/1 computers, obtained from a *PC Magazine* article.

For the CA and the Directory, which have much greater workloads than the other entities within the model, Sun SPARCstation 10 Model 30 workstations were chosen. The workstation comes with 32 MB RAM and a 424 MB internal drive and runs at about 10.6 Mflops. The *Combat Air Forces Workstation* (CAF-WS) contract, a contract that provides special Sun pricing to the Air Force, gives the workstation cost as $22,495. Annual maintenance costs are not readily available but are assumed to be 10% of the original purchase price. Recall, it is assumed that the users already have personal computers with active maintenance agreements.

For the KGs, the ORAs, the CAs, and the PCAs, a trusted operating system (minimum C2 class of protection) is required. An examination of the prices for existing C2 trusted operating systems for various platforms led to estimate of $1500 as the cost of a C2 trusted OS. This value is used within the KG, the ORA, the CA and the PCA components of the PKI cost model.

The hard drive costs are initial costs only. Although hard drives are available in many capacities, a 120 MB drive is chosen in order to obtain a drive of sufficient but not excess capacity. The drive chosen is the MASS Microsystems Diamond Drive 120 which retails for

about $450 each in small numbers through mail order companies. This drive is used in all the modules except the CA module. The CA module uses a 424 MB drive from Sun Microsystems, the minimum sized drive listed for the CA's Sun workstation. The $1,900 pricing of the Sun hard drive is obtained from CAF-WS contract.

A necessary part of any PKI architecture is to store and to use the private keys in a secure and convenient manner. For the PKI Cost Model, the same smart card system is chosen for all the modules, specifically the Datakey's SignaSURE System. The system includes a reader/writer (often just called a smart card reader), a smart card, and DSS software. Signatures are computed on the card so that the private key need never leave its protection. The Smart Card system's initial cost of $450 per unit for small quantities is obtained from Datakey literature. In quantities from 501 to 1000, the unit price is reduced to $337, with further reductions possible for larger quantities[16]. An annual cost of $50 is a very conservative estimate of the cost of procuring software updates, obtained from a Datakey representative.

I.7.5.3 Staffing Costs

For the ORAs, CAs, and PCAs, staff are required to provide PKI operations. A Bureau of Labor Statistics Engineering Technician III was chosen as representative of the type of skills required to perform PKI administrative functions. The annual cost of this position is $29,852, which includes both salary and benefits. A partial description of this job classification is provide below:

> Provides semiprofessional technical support for engineers. Work pertains to electrical, electronic, or mechanical components or equipment. Required to have some practical knowledge of science or engineering. Performs assignments that are not completely standardized or prescribed. Selects or adapts initial standard procedures or equipment, using fully applicable precedents.

---

16 Some estimates of future large quantity purchases of smart card interfaces, especially in the event of burgeoning smart card use for "pre-paid charge cards" and U. S. Medical Insurance cards, bring the price down to the $100 to $150 range.

APPENDIX J

LEGAL ISSUES

J.1 INTRODUCTION

The accelerating movement from paper-based transactions and records to their electronic replacements, and the resulting benefits from this movement, are well documented. Yet in many cases, the shift from conventional to electronic mechanisms has not enjoyed sufficient legal consideration and treatment. Real and perceived security weaknesses of electronic transactions and records present legal and practical barriers to their effective widespread use.

Arguably, perceived security weaknesses could be reduced or eliminated by accepting commercially reasonable security practices. The failure to do so causes perceived weaknesses to become unnecessary barriers. This section considers the legal efficacy and expanded use of electronic transactions and records in modern commerce, government, and other environments for undertaking commitments and other important purposes. Legal efficacy here denotes wide legislative and judicial recognition that properly secured electronic transactions and records satisfy traditional legal indicia of reliability.

These indicia include, where appropriate, transactions or communications considered to be in writing, signed, verified, or acknowledged. Such legal requirements often differ considerably among states and nations, as well as by application. Some attributes of conventional paper-based media are difficult to reproduce by electronic media, such as their singularity (uniqueness), which is an attribute of critical importance to negotiable instruments and comparable legal instruments.[17]

The goal is to arrive at a reasonable level of security for various classes of transactions and records to assure legal requirements are satisfied. This section, however, focuses on the legal implications of authentication, integrity, non-repudiation, and availability, rather than on those of confidentiality. This focus is not intended, however, to underplay the criticality of responsive private and government treatment of confidentiality issues—indeed, confidentiality is a most important requirement in some applications.

Of primary concern, of course, are electronic documents that have been signed. Such signatures indicate authorization, affirmation or approval, commitment, concurrence or certification. The legal efficacy of these documents is closely tied to the security of the signature and certificate infrastructure. However, more is known and more precedents exist for electronic transactions in general. Therefore, in order to understand the security/legal efficacy issues of the infrastructure, we examine these issues in the more general environment of electronic transactions.

---

[17] See, for example, section 4.4.2 which introduces trusted electronic notaries and trusted identities that can assure singularity.

It should be noted that this section discusses the law's treatment of security and reliability. Specifically, the discussion is limited to the legal standing of digitally-signed transactions and information rather than to a comprehensive examination of certificate infrastructures. It only briefly touches on the crucial issue of liability–what is the extent of a user's liability in losing a key, how far is the PKI liable for a compromised certificate, how can liabilities be divided between the several participants in an electronic transaction to keep the exposure within reasonable bounds? These questions, for which few answers have been attempted in the open literature, are addressed in a separate report [4].

J.2 TREATMENT OF SECURITY AND RELIABILITY IN THE LAW

Electronic data interchange (EDI) and transactions, certifications and records in electronic form are not yet accorded the extent of the legal efficacy enjoyed by the corresponding paper documents. Before these electronic forms can earn this legal efficacy, they must establish customs and practices, or they must at least be judged legally equivalent to their manual counterparts. There is gradual movement towards complete legal recognition of computer-based information.

Neither current technical and security standards for EDI nor similar future standards for digital signatures serve as substitutes for responsive legal consideration. Such standards are purposefully drafted to provide options and alternatives to accommodate use by diverse agencies and industries. They do not necessarily provide the guidance necessary to assure the creation of unequivocal legal acts. Technical standards developers cannot properly analyze and resolve complex legal issues.

This problem of legal efficacy arises in the following areas of law: contracts, evidence, government procurement and regulation, real property, and the judicial process. We touch on each in turn.

J.2.1 Contracts

Seeking to satisfy requirements for electronic transactions and records under the Uniform Commercial Code (U.C.C.) raises certain fundamental issues. For example, although the definition of signed in U.C.C. § 1-201(39) “includes any symbol executed or adopted by a party with the present intention to authenticate a writing,” the word authenticate is not defined in U.C.C. Articles 1 or 2 (although Official Comment 39 to § 1-201 includes mention of a thumbprint, a particularly forensically intensive type of authentication). This lack of definition has created confusion in the legal community. While the case law considering electronic writings and signatures is sparse and inconsistent, some of those cases addressing the issue confirm the importance of the probative value of signatures. On the other hand, "there is growing agreement among lawyers knowledgeable about electronic contracting that authentication and signature concerns can be addressed by existing legal concepts in conjunction with adequate audit and record keeping controls." [26]

J.2.2 Evidence

The Federal Rules of Evidence do not address electronic data security mechanisms specifically. The scope of proof of trustworthiness (and, arguably, security) as an evidentiary foundation requires closer scrutiny. According to a U.S. Department of Justice guideline, “[B]ecause electronic files are particularly susceptible to purposeful or accidental alteration, or incorrect processing, laying a foundation for their admission must be done with particular care. Proper control over creation and maintenance of these files can be crucial in overcoming inevitable objections that will be raised in the courtroom.” [27] The implications of burgeoning, open, interconnected, and highly diverse computer systems utilizing expert system components, which may change frequently and considerably, may call for strong evidentiary foundations. There is some case law supporting the notion that proof of reliability (and implicitly, of security) is recognized as appropriate and necessary in evaluating the admissibility of computer-based evidence. In any event, the extent to which a rigorous foundational requirement for computer-based evidence will ultimately result is still in question.

The Systems Policy Staff of the Justice Management Division, Department of Justice, has produced guidelines for the admissibility of electronically filed evidence. "Any tangible thing offered as evidence is subject to challenge regarding its genuineness. Computer printouts are no exception. Federal Rule of Evidence 901(a) states:

> The requirement of authentication or identification as a condition precedent to admissibility is satisfied by evidence sufficient to support a finding that the matter in question is what its proponent claims." [27, p. III-32]

Among the examples of authentication listed by way of illustration in Federal Rule of Evidence 901(b) is:

> *Process or system*. Evidence describing a process or system used to produce a result and showing that the process or system produces an accurate result. [27, p. III-32]

The Comptroller General states the conditions for obligations of federal funds. The "signature" that binds the funds should satisfy the following:

- Unique
- Verifiable
- Executed under control of the signee

It has been suggested that an additional attribute that the "signature" should exhibit is that it should be

- Linked or bound through accountability to the document being signed.

The Manual for Complex Litigation Second (1985) recognizes and addresses this problem of proof of reliability. It observes that “[n]ot withstanding the capacity of

computers to make tabulations and calculations involving enormous quantities of information . . . several sources of potential errors of great magnitude exist." The Manual further notes that "the proponent of computerized evidence has the burden of laying a proper foundation by establishing its accuracy," and "the existence or possibility of errors usually affects only the weight, not the admissibility of the evidence, except when the problems are so significant as to call for exclusion . . ." These concerns apply most particularly to public key certificates that are presented as evidence to support the authenticity of a digital signature. Proper foundation requires establishing the accuracy of the certificate from which the signature public key is extracted.

It is apparent that a digital signature algorithm, backed by an appropriate key distribution and protection system (such as the PKI, including its procedures and practices), will go far to prove authenticity to the trier of fact, but only after the laying of such foundation as each court requires. The System Policy staff, quoting *United States v. Briscoe* 896 F.2d 1494-5 (7th Cir., 1990) which in turn cites *United States v. Craft* 750 F.2d 1354 (C.A. Wis., 1984), summarize: "As long as the government provides sufficient fact to warrant a finding that records are trustworthy and the opposing party is afforded an opportunity to inquire into the accuracy thereof and how the records were maintained and produced, a proper foundation has been established." [26, p. III-35] Digitally signed documents should readily be admissible as evidence. What credence the trier of fact places on that evidence will depend on the foundation presented by expert testimony on the signature algorithm and on the key distribution scheme.

J.2.3 Government Procurement and Regulation

Interpretation and resolution of State, Federal, and foreign requirements such as those concerning signature requirements remains unsettled. Compare the following varied—arguably conflicting—signature definitions.

- Signature: "includes a mark when the person making the same intended it as such" (1U.S.C. Section 1)

- Signed: "includes any symbol executed or adopted by a party with the present intention to authenticate a writing" (U.C.C. § 1-201(39))

- Signed: "shall include the entry in the form of a magnetic impulse or other form of computer data compilation of any symbol or series of symbols executed, adopted or authorized as a signature" (17 C.F.R § 230)

- Signature: "in the case of an EDI transmission, means a discrete authenticating code intended to bind parties to the terms and conditions of a contract" (41 C F.R. § 101-41.002(d))

- Electronic signatures: "characters representing the nominated persons on documents, and signed or symbols identifying their writers" (Korean Act [28])

One working group who considered this issue apprehended the effect of such uncertainty when it concluded that "[t]he lack of adoption of an accepted electronic signature policy by the [Department of Defense] will prevent some contract transactions being conducted in digital form." [29] Independently, the Comptroller General has addressed uncertainty in electronic commerce with the following decision: "[c]ontracts formed using Electronic Data Interchange technologies may constitute valid obligations of the government for purposes of 31 U.S.C. § 1501, so long as the technology used provides the same degree of assurance and certainty as traditional "paper and ink" methods of contract formation." [30] Nevertheless, outside of the specific circumstances presented in that case, the decision begs for a closer definition of the indicia of assurance and certainty necessary to be deemed reliable.

J.2.4 Real Property

An example of how the problem of legal efficacy of electronic information could arise in the real property area involves the recording of deeds and related instruments where the recording statute mandates that "writings which are to be recorded or docketed in the clerk's office of courts of record in this Commonwealth shall be an original or first generation printed form, or legible copy thereof, pen and ink or typed ribbon copy. . . " (VA. Code § 5-108). Such a statute raises considerable barriers to computer-based commerce. Fortunately, there are very few Federal statutes and regulations with similar wording which can preclude electronic filing of mandated reports and certifications.

J.2.5 Judicial Process

The legal efficacy of information in electronic form also arises in judicial contexts. Despite the advance of computer automation in some aspects of the judicial process, electronic notice and service of process are not generally permitted by court rules. However, there are exceptions, and judicial reform is accelerating. A recent revision of the Federal Rules of Appellate Procedure 25(a) permits the filing of court papers electronically. The National Archive and Records Administration's *Electronic Records Management* regulations accommodate the judicial use of electronic records. Additionally, the U.S. Department of Justice has issued findings which "encourage the development of electronic data interchange technologies." [31]

J.2.6 Archiving

The preservation of important documents is mandated under several federal statutes and regulations. The period that most documents must be held is seven years. Some requirements are less; a few are more. The Federal Records Act includes some requirements for permanent archiving. Obviously, the signature or signatures on an archived document must be achieved as well. This is true of electronic documents as well. All the certificates necessary to verify those signatures must also be preserved. The period during which each certificate was valid must be carefully recorded. If a certificate was invalidated for any reason before its expiration date, that fact has to be preserved as well.

It is apparent that any digital signatures must accompany a document as it moves from one computer archival medium to another and that a careful audit trail be kept for the duration of the archival period. The audit trail must include information about the signatures as well. Again, it will be expert testimony about this entire procedure and about the audit data collection that will lay the foundation for the testimony should the documents be required as evidence. It is this testimony that will help the trier of fact determine the trustworthiness of the evidence the signed document presents. Thus, procedures must be in place to establish the validity of the signatures on archived electronic documents.

J.3 REASONABLE SECURITY PROCEDURES

Unlike conventional paper-based transactions and records, there is little jurisprudential guidance as to whether (and, if so, under what circumstances) a particular security technique, procedure, or practice will provide the requisite assurance of reliability in electronic form. Within the EDI community, for example, this lack of guidance concerning security is reflected in the multiplicity of current security and authentication practices. These practices, in many instances, appear to have been implemented in an ad hoc manner, with neither a clear understanding of the present state of law, nor the technical proof assurances of other chosen practices.

In fact, the EDI experience is quite enlightening. In a survey of EDI users, the mechanisms or procedures employed as legal signatures included the following: a "buyer code," a DUNS number and suffix, a password, a message authentication code, an account number, an ID/password combination, an electronic verification of symbol and codes, "and functional acknowledgments" [32]. The law should be flexible in permitting a variety of signatures in electronic form–most particularly, public key-based digital signatures. However, this survey reflects a lack of purposeful, consistent, and knowledgeable choices by the user community, as well as the law's lack of clarity. Where the law has responded, it has been arguably too vague—such as a requirement to implement *reasonable security procedures* (U.C.C. § 4A-201). This vagueness may be intentional for the requirement is intended to affect the apportionment of liability rather than precluding transactions in its absence.

On the other hand, consider the case of the Model Electronic Payments Agreement and Commentary. [33] It states the following: "While security procedures should certainly be reasonable, in certain situations a lack of specificity in defining "reasonable" security procedures may provide inadequate guidance causing such security procedures to fail in their intended purpose. . . . Specificity may help the parties implement and comply decisively and unambiguously with security procedures, reduce confusion and offer better expectations of reliability and certainty. Security procedures should be sufficiently precise so that they are not subject to discretionary, self-serving interpretation, in part, because: (i) few standard security procedures exist in the law. . . (ii) security technology is changing rapidly, and (iii) parties often hold particularly diverse opinions on appropriate solutions to security threats."

One difficulty in developing responsive laws involves deciding the extent to which law should detail and endorse particular security techniques, procedures or practices. Proponents of specificity argue that the migration from a paper to an electronic world needs greater guidance and that private agreements and legislation requiring only reasonable security procedures are vague and unworkable. Proponents of generality, on the other hand, argue that the endorsement of specific security procedures, practices or techniques leads to inflexibility and creates a presumption that the failure to implement such techniques, procedures and practices constitutes failure to exercise ordinary care. In recognition of these competing interests, a measured movement toward greater specificity on security procedures in the law may be needed. Such is most certainly the case with electronic signatures.

J.4 LEGAL AND POLICY RELIEF REQUIRED

J.4.1 Burden of Proof and Presumptions

Scant attention has been paid to burden of proof and presumption issues in electronic transactions. This is unfortunate because, after all, proof issues are at the heart of the meaningful resolution of disputes and the successful prosecution of wrongdoing. While undeniably a daunting task, and an issue worthy of further study, burdens of proof and presumptions must be examined and integrated into a workable legal framework for electronic transactions. In the area of commerce, maritime law is rich in presumptions because there are often no witnesses to events on the high seas. Another large size body of presumptions must be developed for application to electronic transactions and to the use of the PKI for there is a similar lack of witnesses to many computer and communications events.

The development of electronic commerce rules are intimately affected by burden of proof requirements that consist of both the *risk of nonpersuasion* and the *duty of producing evidence*. U.C.C. § 1-201(8) states that the burden of establishing "a fact means the burden of persuading the triers of fact that the existence of the fact is more probable than its non-existence." This would seem to be true in the non-commercial arena also. Burden of proof issues affect (1) electronic message reliability and genuineness, and (2) admissibility and enforceability of information in electronic form (for example, when substituted for paper-based documentation).

J.4.2 Trusted Entities

Despite the great benefits resulting from the use of digital signatures, they have some inherent limitations (as is true with any security mechanism), including an innate inability to provide "time-related" non-repudiation. Digital signatures and other cryptographic methods cannot, in the absence of a trusted entity, provide an unforgettable trusted time stamp[18].

---

[18] Bellcore has announced a timestamping technology which involves the publication, in specified leading newspapers, of digital signatures on the hash of all documents submitted for time stamping during the week immediately preceding publication.

Therefore, to achieve "full" non-repudiation, time stamping must be undertaken by a disinterested party beyond the control of the parties to a transaction or record—a *trusted entity*. Time stamping by a trusted entity is also necessary for certifications and submissions that must be filed by a specified time and date.

A trusted entity is an independent, unbiased entity capable of providing important security assurances that enhance the enforceability and reliability of electronic records. The key attributes of a trusted entity are that it is a *disinterested*, *unbiased*, *third party* trusted by the parties to the transaction and by the dispute resolution mechanism(s) relevant to a transaction or record.

Simply stated, a trusted entity's administrative, legal, operational, and technical infrastructure must be beyond reproach. Third Party Service Providers (TPSP) or value added networks (VANs), such as ATT or MCI, have arguably been inaccurately identified as trusted entities. VANs are not necessarily disinterested, since they may compete with each other, participate in the transfer or processing of information (and therefore have exposure), and introduce error, delay, unavailability, or misdelivery. However, it should be noted that there is no reason why a TPSP cannot be a trusted entity.

A trusted entity can time and date stamp, store (or forward) a "record copy" or hash of a transaction, keep an audited data log, or serve as an intermediary for other trust-based services between trading partners. The trusted entity's record copy of an electronic transaction would control in the event of a dispute regarding a document's authenticity or timeliness. The electronic notary offers unique solutions to one of the critical "missing links" of electronic transactions and records assurances: unforgettable trusted time stamping.

The electronic notary also may facilitate future TPSP and value added network service requirements by providing them with trusted-entity services. Notarizing data intended for record retention and archiving provides an unforgettable seal that may contain a time stamp and digital signature, together with additional audit, legal, and security information to enhance its legal efficacy. The electronic notary can even provide irrefutable proof of the time of the origination of the document.

The extent to which the PKI certificate authorities must be trusted will vary with their place in the infrastructure and with that infrastructure itself.

Included in the hash is the exact time and date at which Bellcore received each document.

LIST OF ACRONYMS

ANSI	American National Standards Institute
ARPA	Advanced Research Project Agency
ASC	Accredited Standards Committee
ATF	Bureau of Alcohol, Tobacco, and Firearms
BBN	Bolt, Beranek and Newman, Inc.
CA	Certification Authority
CCITT	Comité Consultative International Télégraphique et Téléphonique
CA	Certificate Management Authority
COI	Community of Interest
COTS	Commercial Off-The-Shelf
CRL	Certificate Revocation List
CTR	Currency Transaction Report
DBMS	Database Management System
DCA	Defense Communications Agency
DEA	Drug Enforcement Agency
DES	Data Encryption Standard
DII	Defense Information Infrastructure
DISA	Defense Information Systems Agency
DIT	Directory Information Tree
DLA	Defense Logistics Agency
DMS	Defense Messaging System
UN	Unique Name
DOJ	Department of Justice
DOS	Department of State
DSA	Digital Signature Algorithm
DSS	Digital Signature Standard
EDI	Electronic Data Interchange
EFT	Electronic Funds Transfer
EPM	Electronic Postmark
EPM-PLUS	Electronic Postmark-Plus
FAA	Federal Aviation Administration
FBI	Federal Bureau of Investigation
FEMA	Federal Emergency Management Agency
FIPS	Federal Information Processing Standards
FMS	Financial Management Service
FRB	Federal Reserve Board

GAO	General Accounting Office
GB	Great Britain
GEIS	General Electric Information Systems
GOSIP	Government Open Systems Interconnection
GSA	General Services Administration
HHS	Health and Human Services
I&A	Identification and Authentication
IAB	Internet Activity Board
IBAC	Identity Based Access Control
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
INS	Immigration and Naturalization Service
IPRA	Internet Policy Registration Authority
IRM	Information Resources Management
IRS	Internal Revenue Service
IRTF	Internet Research Task Force
ISO	International Standards Organization
ITU	International Telecommunications Union
KMP	Key Management Protocol
LAN	Local Area Network
MHS	Message Handling System
MSP	Message Security Protocol
NASA	National Aeronautics and Space Administration
NARA	National Archives and Records Administration
NIST	National Institute of Standards and Technology
NSA	National Security Agency
OFPP	Office of Federal Procurement Policy
ORA	Organizational Registration Authority
OSI	Open Systems Interconnection
PAA	Policy Approving Authority
PCA	Policy Certification Authority (or Policy Creation Authority)
PEM	Privacy-Enhanced Mail
PEM WG	Privacy-Enhanced Mail Working Group
PKCS	Public Key Cryptography Standards
PKI	Public Key Infrastructure

POSIX Portable Operating System Interface
PSRG Privacy and Security Research Group
PTO Patent and Trademark Office

RBAC Rule Based Access Control
RFC Request for Comment
RUN Relative Unique Names
RSA Rivest, Shamir, and Adleman

SEC Securities and Exchange Commission
SDNS Secure Data Network System
SHA Secure Hash Algorithm
SHS Secure Hash Standard
SILS Standard for Interoperable LAN Security
SP3 Security Protocol 3
SP4 Security Protocol 4
SSA Social Security Administrator

TEK Traffic Encryption Key
TPSP Third Party Service Providers
TSS Telecommunications Standards Section

UCC Uniform Commercial Code
UN United Nations
USPS United States Postal Service

VANs Value Added Networks