PUBLIC KEY INFRASTRUCTURE - USING x.509 CERTIFICATES FOR DEVICE AUTHENICATION HERE A CERT, THERE A CERT, EVERYWHERE A CERT

Doug Jones
YAS Broadband Ventures, LLC.

Abstract

The Public Key Infrastructure (PKI) is a security standard designed to bring "trust" to Internet services. PKI is based on public key cryptography and uses digital certificates to authenticate entities to service providers. It is this authentication that provides trust for services offered over the Internet.

PKI presents a solution for several issues facing cable, and has been included in several CableLabs® projects. PKI is based not only on technology, but also policies. All of which will be discussed briefly in this paper.

INTRODUCTION

Cloning, false identities, theft of service etc., are all issues that Cable has faced before and will continue to face as it advances further into offering Internet-based services. And just like the Internet is a new technology, new security technologies will be needed. The Public Key Infrastructure (PKI) is a standard designed to bring trust to Internet-based services. Not only can PKI protect cable from cloned boxes, PKI will also protect consumers and service providers from fraud as they engage in E-commerce.

The purpose of any security technology is to protect value, whether a revenue stream or a purchasable information asset of some type. Threats to this revenue stream exist when a user of the network perceives the value, expends effort and money, and invents a technique to get around the necessary payments. Some users will go to extreme lengths to steal when they perceive extreme value. The addition of security technology to protect value has an associated cost; the more expended, the more secure the service can be. The proper engineering task is to design a reasonable costing security technology to force any user with the intent to steal or disrupt network services to spend an unreasonable amount of money to circumvent it. Security effectiveness is thus basic economics. Deploying PKI does require additional expense; however, the promise is are both a secure network and secure services that can be the basis for broadband interactive services going forward.

PKI protocols and technologies are developed within the Internet Engineering Task Force (IETF), an international standards body. PKI is a combination of encryption technologies and usage policies that enable the security of digital communications and business transactions on the Internet. PKIs integrate digital certificates, public key cryptography, and certificate authorities into a network security infrastructure.

PKI is included in most CableLabs® projects including PacketCable™, DOCSIS™, and CableHome™. CableLabs is taking an approach such that the PKI infrastructures for the projects will align.

The information in this paper is relevant to both engineers and managers. The material includes:

- introducing the concepts of PKI
- why it was developed
- what it protects against

The beginning discussion is an overview of how PKI works, and the later part of the paper goes into specific technologies.

Why PKI

For security purposes, it is important to both keep communications private and to know with whom you're communicating before exchanging any meaningful information. For electronic business to work, agencies and individuals must be convinced that transactions can be carried out both privately and securely and that both users and documents are authentic. The paper world relies on signatures. The computer world needed an electronic equivalent. PKI is a system for encrypting, decrypting, signing and verifying the authenticity of information that is transmitted over the Internet.

In a PKI, each entity on the network is issued a digital certificate. Hence the statement "Everywhere a cert." The information in certificates can be used for both encrypting and authenticating digital communications and transactions. An entity can be a device, a piece of software, or a user.

WHAT IS PKI

PKI is a group of protocols and techniques that, when put together with a group of policies, allow for secure, authenticated information exchange. The technical parts of PKI are standards and are widely available. The PKI is customized for specific applications based on administrative policies those enterprises set based on needs. It is the set of policies the truly defines how secure the PKI will be.

The center of the PKI is the Certificate Authority (CA). It is the CA that issues the digital certificates to parties within the PKI. The policy for issuing the certificates is what builds the trust in the PKI. Some CAs require rigid identification before issuing a key pair and certificate, other CAs may only require a phone call. While two CAs may use the same technologies, they may have completely different levels of trust based on how the certificates are issued.

PKI works by providing each user with two "keys" — one that is public and one that is private. The keys are represented as binary numbers, i.e., long strings of 1's and 0's. The user must keep the private key secret. The public key should be made available for anyone to use. A document encrypted with a public key can only be decrypted with the corresponding private key. The key generation algorithm is based in very complex mathematics such that the private key should not be able to be derived from analyzing the public key.

The digital certificate, issued by the CA, contains the public key of a member of the PKI. These certificates can be kept in any public place or directory as use of the public key is encouraged for encrypting data.

When a document intended to remain private is transmitted, the sender encrypts it with the public key of the recipient. Once encrypted with the public key, it can only be decrypted with the private key, which should be stored securely by the intended recipient of the message. The integrity of the PKI depends on keeping the private key secure.

In the event a private key is compromised, the certificate for the corresponding public key will be place on what's called a Certificate Revocation List (CRL). The CA will maintain the CRL, and it's the responsibility of the user of the public key to verify that key is not listed on the CRL.

PKI also includes functions that allow recipients to ensure the original message has not been tampered with. This feature uses a digital signature, based on the sender's private key, to "watermark" the message. Using the sender's public key (available in the public certificate) to decrypt the watermark, the recipient can verify that the received message

is authentic. Also, by using the sender's private key to create the watermark, the sender of the message is positively identified. Thus, a PKI can ensure that messages are both authentic and that the sender is who they say they are.

There is no single PKI today, rather, various enterprises have created PKIs using the base technologies and incorporating specific policies for their enterprise. As the use of PKI continues to grow, the enterprise PKIs will most likely gradually grow together, thereby increasing interoperability. Policy decisions to be made for the PKI include:

- how certificates are issued,
- who certificates are issued to,
- how certificates are revoked,
- core technologies used,
- how keys are generated, etc.

While base technologies are the same, the policies are what define the PKI. For instance, an informal PKI could be used for the recreational securing of email. A certificate could be issued based only on asking for one. A more controlled PKI would be created for use with e-commerce applications, requiring the users to identify themselves perhaps with birth certificates, drivers license, or credit card information with rigorous, regular ongoing audits of the certificates and their holders.

PKI certificates have a standard format that will be described later in the paper. The important contents of the certificate include material that can be used to uniquely identify the owner of that certificate, including the public key associated with the owner. It is this operation that provides the "trust" associated with the certificate. When a digital certificate is received, it can be used to verify the identification of the sender.

PKI Architecture

The PKI is a layered hierarchy of digital certificates, with the root certificate at the top. The top-most certificate belongs to what is called the root Certificate Authority (CA). The root CA private key is the absolute trusted source of information within the PKI and should be stored using considerable physical security. Generally the root private keys are stored in locked bunkers with many layers of physical security, including voice prints, retinal scans, and security cameras. The physical security of this key is critical because members of the hierarchy use the trust in the CA to authenticate transactions with other members. In economic sense, the physical security of the root CA private key should be configured based on the dollar amount of business that that the CA is designed to protect. This is serious business.

The CA issues key pairs to members of the PKI. The public key is contained in a certificate that is digitally signed by the root CA private key. Being "signed" means the certificate contains a construct that is encrypted with the root CAs private key. Anyone can use the root CAs public key, which is freely available, to verify the authenticity of the certificate, and hence the public key that it contains. As mentioned, revoked certificates will be listed on a CRL maintained by the CA.

For practical purposes of management, a single root CA may be divided into several subordinate root CAs. This is the case for how the Cable CAs are being managed. Figure 1 represents a possible CA hierarchy. The root CA issues certificates for three additional subordinate CAs: 1) software, 2) devices, and 3) service provider systems. The first two subordinate CAs issue certificates to authenticate software loads and the particular devices those software loads go into. The third subordinate CA is used to issue certificates to devices within the service provider's networks, such as various servers.

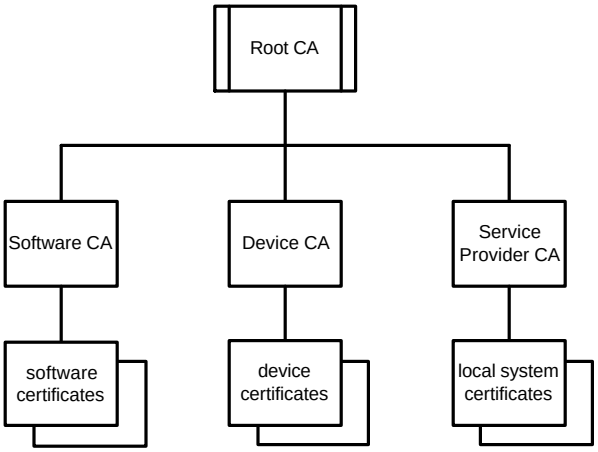

Figure 1

Note, in the specific hierarchy shown in Figure 1 no certificates are issued to users. Users might get certificates based on the e-commerce activities they would be engaged with. That is, the cable operator would issue certificates to ensure their network is trusted, and users would get certificates from their banks or brokerage houses to engage in secure authenticated e-commerce communications. There can be multiple dimensions of CA going on concurrently.

Based on Figure 1 is should be clear it is possible to have a chain of multiple certificates comprising a certificate of the public key owner (the end entity) signed by one CA, and zero or more additional certificates of CAs signed by other CAs. Such chains, also called certification paths, are required because a public key user is only initialized with a limited number of assured CA public keys. These chains are used to segment a CA into manageable entities.

THE X.509 DIGITAL CERTIFICATE

A CA issues digital certificates, and based on the policies of that CA, the certificate will have a certain level of trust associated with it. This is the basis of the operation of a PKI. The digital certificate is simply a computer file that contains information both to identify the holder and some of the policies of the CA. Most PKIs use digital certificates based on the X.509 standard, however, two CAs could have different polices and have incompatible certificates, even though all the certificates are based on X.509.

Important fields on the certificate include the public key of the issuing entity, how long that certificate is valid, and the technologies used by the CA. There can be many more additional fields in the certificate and these are defined in [1]. Because certificates contain only public keys, they can be distributed via untrusted communications systems, and can be stored in unsecured areas.

Another important field in a certificate is the digital signature placed on it by the issuing CA. This signature verifies the entire contents of the certificate. The technology associated with digital signatures is discussed later in the paper, but this is a method to verify the certificate is authentic. A person wanting secure communication with a PKI user need only verify the authenticity of that user's public key by checking both the digital signature on the certificate and that the certificate has not been revoked by the issuing CA. If both of these steps check, then trust is based on the policies of the underlying CA.

PUBLIC KEY CRYPTOGRAPHY

Cryptography, in general, has been with cable for years and is central to the Conditional Access (CA) systems used for digital video. A "key" is a string of binary bits that is used along with a core cipher to encrypt digital content. In the case of digital video, a core cipher is used to encrypt video at the headend and to decrypt it at the set top box. The core cipher is generally based on the Data Encryption Standard (DES), though there are several ciphers in use. The interesting part of a conditional access system deals with the keys and how they are shared between the headend

(to encrypt) and the set top box (to decrypt). There are CA systems based on both symmetric and public key cryptography, as described in the following paragraphs.

One type of cryptographic system uses what's called symmetric or "secret" keys. In this case, the message is encrypted and decrypted using the same key. While this certainly works, there are issues with how to manage and distribute the key that is used. Both parties involved in the transaction must have the same key, and that key must always be kept secret. But a symmetric key has to be distributed to both the encrypter and the decrypter, which means there is an opportunity to steal the key while it is being shared. Anyone stealing that one key can not only decrypt messages, but they can also change the message and forward it on as if it were the real thing. Symmetric key systems are not viable for Internet use because of the issues with key distribution.

A second type of cryptographic system uses asymmetric or "public" keys, and this is what the PKI is based on. Public key cryptography was introduced in 1976 by Whitfield Diffie and Martin Hellman, and it addresses several of the issues associated with symmetric key cryptography. Public key cryptosystems use two keys, one to encrypt and one to decrypt. The key used to encrypt data is called a public key, and as the name suggests, this key is distributed freely, in the form of a digital certificate, available for anyone to use. The key used to decrypt data is called the private key, and as the name implies, this key is to be kept secret by the holder.

Public Key Cryptography for Encryption

Figure 2 shows how public key cryptography for message privacy is used. The sender locates the receiver's public key (freely available in the form of a digital certificate) and uses it to encrypt the message. This message is sent to the receiver who uses their private key to decrypt it.

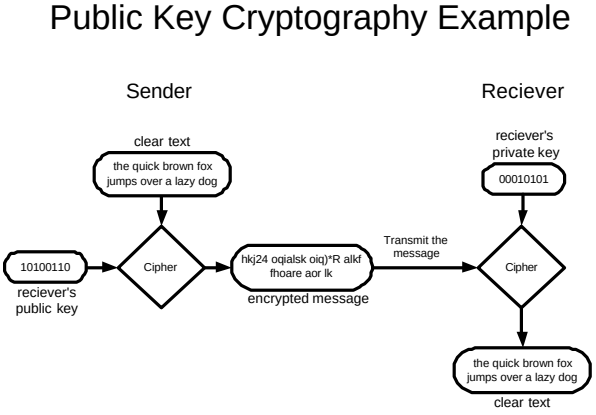

Figure 2

Public Key Cryptography for Authentication

In many cases of secure communication it is critical to not only encrypt the message, but also to verify the originator of a communication (or transaction). This is known as authentication, where the receiver of a message can verify and trust the source of a message. In some cases it's even permissible to send the message unencrypted, as long as there is a mechanism included with that message to verify its authenticity.

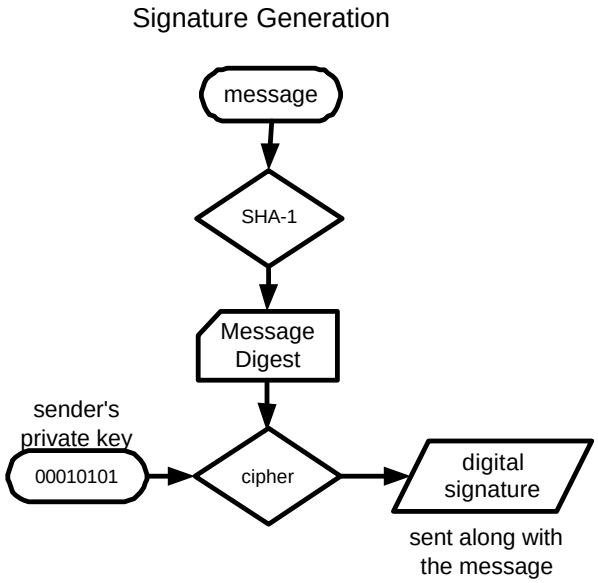

Figure 3

Adding authentication to a message is referred to as "signing" the message. To sign a

message, the sender does a computation over that message and the result is called a message digest. The message digest is encrypted with the sender's private key and the result is called a digital signature. The process is shown in Figure 3. The digital signature can be included with the message when it is sent as a form of authentication.

To verify the signature, the recipient first uses the sender's freely available public key to decrypt the digital signature. The result is the message digest over the original message. The recipient then computes a message digest over the received message and compares it to the digest received with the message. If the two digests match, the signature is verified to be genuine and the message is authentic; otherwise, the message and/or the sender is fraudulent. Signature verification is shown in Figure 4.

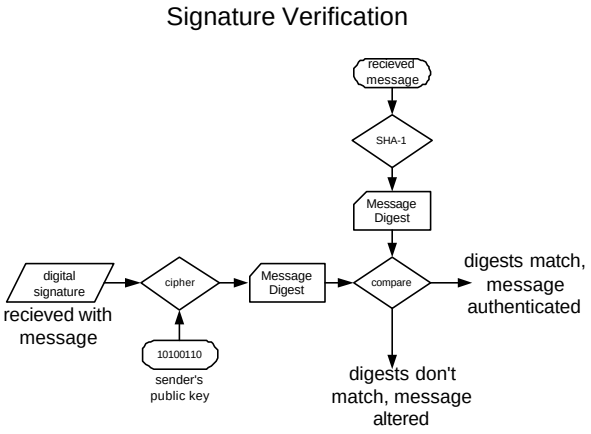

Figure 4

Message Digests

Authentication is accomplished using digital signatures. A digital signature is created by running a message digest algorithm, also known as a hashing algorithm, over the message and then encrypting the resultant message digest with the sender's private key. The input to the message digest algorithm is the original digital file. The output, the message digest, is a binary string, generally of fixed length, that is unique based on the original file. The message digest can be thought of as a mathematical summary of the original file.

A common algorithm used to create message digests is SHA-1 [4], the Secure Hash Algorithm. With SHA-1, the input file size can be up to 2^{64} bits long (a very, very large file, 2 million million megabytes). The output of SHA-1 is always a fixed-length, 160 bit (20 byte), string of digits that is unique for that file.

The SHA-1 is called secure for two reasons. First because it is mathematically infeasible to find two different messages that would produce the same message digest. Secondly, any change to a message will, with very high probability, result in a different message digest. Hence, the message digest provides a method to determine if the original message has been received with integrity. Rather than verifying the entire file, which can be millions of megabytes long, the receiver needs to only recomputed the digest over the message and compare that relatively short byte string with the original digest to determine if the message has changed during transit.

To provide authentication, the message digest is usually encrypted with the sender's private key and the result is a digital signature. Since the sender's public key is freely available, the recipient of the message can decrypt the digital signature and compare the resulting message digest to one computed locally. If the two message digests match, then the message has been received without being altered.

SUMMARY

PKI is an infrastructure that provides encrypted, authenticated communications over unsecured channels. Such a system was designed with the Internet in mind.

The trust in the PKI is based not only in the technology used by that PKI, but also in the policies and procedures instituted. PKIs with

similar technologies but different policies can have very different levels of trust.

PKI technologies are fairly straightforward, and are based on public key cryptography, which allows both encryption and authentication. PKIs can be used to secure hardware, software, and e-commerce users.

REFERENCES

1. RFC-2459, Internet X.509 Public Key Infrastructure Certificate and CRL Profile, IETF, January 1999.

2. RFC-2527, Internet X.509 Public Key Infrastructure Certificate Policy and Certification Practices Framework, IETF, March 1999.

3. Frequently Asked Questions about Today's Cryptography, RSA Laboratories, April 2000.

4. RFC-3174, US Secure Hash Algorithm 1 (SHA1), IETF, September 2001.

5. How PKI works, William Mathews, Federal Computer Week, June 2000.

Author Contact:
Doug Jones, Chief Architect
YAS Broadband Ventures
(303) 661-3823
doug@yas.com