Information Security Technical Report, Vol 7, No. 3 (2002) 81-94

Secure Access Management: Trends, Drivers and Solutions

Dr Stefan Brands

Credentica, Inc., Montreal, Quebec, Canada

As information systems move from closed environments to increasingly open environments, physical trust domains are rapidly vanishing and the demand for secure access management solutions is growing dramatically. This paper examines the trends and growth drivers in a variety of markets, and presents two different solutions. The first solution, X.509-style PKI, is being pushed forward by the PKI industry as the solution to secure access management, but nothing could be further from the truth. X.509-style PKI was invented in 1978 for the purpose of facilitating message encryption, with entity authentication serving to prevent man-in-the-middle attacks. Authentication as used for message encryption, however, is a far cry from managing the access of authorized parties to identity-related information, with all its privacy, security, and performance sensitivities. We show how X.509-style PKI fails in the context of access management, and present a solution based on Digital Credentials.

1. Introduction

Organizations are accumulating and storing more information than ever before. There is an increasing pressure on them to cost-effectively share much of this information across traditional trust domains. As a result, many organizations are currently adjusting their business processes and information management infrastructures to facilitate the sharing of information in electronic form, over both fixed and mobile networks. The ability to electronically share information is also fundamental to address an increasing number of business opportunities that cannot be addressed at all by non-electronic methods. Much of this is driven by the growing popularity of the Internet (the number of Internet users is expected to exceed 1 billion in 2005), as well as by the rapid increase in information appliances such as Web-enabled cell phones, PDAs, handheld computers, and personal access devices.

With the rise in data sensitivity and the sheer increase in information transfers, the demand for information security is growing dramatically. An April 2002 survey by the Computer Security Institute and the FBI reported that 90% of respondents in large corporations and government agencies detected a security breach within the last 12 months. Internet attacks in particular are dramatically on the rise. A November 2001 report by Goldman Sachs Global Equity Research lists the following drivers for companies re-evaluating their information security needs: growing numbers of devices and people are accessing the Web; corporations want to connect their networks to the Internet to facilitate strategic and economic business objectives; the rise of electronic commerce; security incidents are reaching critical proportions; evolving and emerging technologies create new security challenges; government regulations mandate further spending on security and privacy; mass market acceptance drives best-of-breed integration of security solutions into a single appliance; and, insufficient security resources drive the growth of managed services.

2. The need for access management solutions

The information security market can be broken down into five groups: access management (authentication and authorization); encryption; perimeter security (including VPN and firewall); network security (including intrusion detection, security management, Web and email monitoring, and vulnerability assessment); and, anti-virus. As information systems continue to move from closed environments to increasingly open environments, physical trust domains are rapidly vanishing. Firewalls, anti-virus software, intrusion detection applications, and vulnerability assessment products and services are good at defending closed corporate networks from attacks, but they provide very little security in the context of information sharing over open networks. With trust domains becoming logical rather than physical, security must be tied directly to the information itself rather than to the perimeter of its repository. This is especially urgent for personal information and for information that may only be accessed by authorized entities.

Correspondingly, the demand for secure access management solutions is growing fast. Wedbush Morgan Securities in a February 2002 industry report list the following four demand drivers for access management products: the increasing value of computer resources and their exposure to attacks; the growth in the number of points of access to most networks; government regulations requiring information privacy and security; and, increased awareness since September 11. IDC predicts that the market for authentication and authorization will grow from $2.8 billion in 2000 to $9.4 billion by 2005, making access management the fastest growing segment in information security software.

3. Information security and privacy trends

To get a better appreciation for the need for secure access management, we now briefly examine the information security and privacy trends in five example markets in different verticals.

3.1. Extended and collaborative enterprise

Traditional enterprise access management systems are good at allowing administrators to manage information sources within the physical boundaries of the company. They enable administrators to create and remove employee accounts, administer employee profile information, and deal with forgotten employee passwords. These days, however, enterprises increasingly need to share information across traditional trust domains. Secure access management for remote workers alone is becoming a huge problem. The US Department of Labor estimates that 19 million Americans work from home or outside the office, and the Yankee Group projects this number to exceed 50 million by 2004. More fundamentally, enterprises can cut costs and increase revenue by sharing their Customer Relationship Management (CRM) information with contractors, disperse business units, affiliate and partner organizations, providers of outsourced services, remote and mobile workers, supplier organizations, project networks, E-business communities, and so on.

Consequently, enterprises are investing in new Web-based applications that allow CRM information to be shared over the Web through portals. According to a January 2002 META report, "By 2003, most packaged business applications and newer custom applications will have fairly complete event-driven interfaces that will be exposed via the Web-services model, and Web services creation capability will be an automated part

of virtually all development environments." Web Services technology aims to fulfill the promises of cost-effective information sharing across disparate systems, by enabling software components to interact with one another dynamically in standardized ways. Leading industry players are actively driving the introduction of Web Services with heavily funded initiatives, such as Microsoft's .NET, Sun's SunONE, IBM's WebSphere, Oracle's Dynamic Services, and BEA Systems' WebLogic.

With the opening up of enterprise information resources, the need for security and privacy goes up dramatically. Enterprises seek to protect company secrets, to keep competitive information out of the hands of competing enterprises, and to prevent hackers and others from damaging or deleting information. In addition, data protection regulations force enterprises to take special measures to protect their customer information. Failure to protect the security and privacy of CRM data and other personal information can result not only in consumer loss and reputation damage, but also in legal liabilities. The security and privacy problems worsen in the context of enterprise collaboration, where information is managed jointly by individuals who are not working for the same enterprise.

According to a January 2001 executive white paper by the Aberdeen Group, "Authentication systems represent one of the largest security expenditures for the enterprise in terms of procurement and operating cost, dwarfing nearly all other security expenses." In the words of the CEO of the Burton Group in January 2002, "Bottom line: As an industry best practice, enterprises should develop an identity and access management strategy as soon as possible." The need for security is particularly strong in Web Services applications; in the words of the Hurwitz Group in August 2002, "There is no way Web Services can be deployed successfully without enhanced security."

3.2 Critical information infrastructures

Critical information infrastructures, such as those used by criminal justice systems and by national defense, are increasingly being made accessible over open networks to facilitate information sharing. The need for highly secure access management systems is enormous in this environment: lives could be lost due to attacks by terrorists, organized crime groups, foreign governments, hackers, and malicious insiders. In particular, since most of the information exchanges take place on the basis of privileges, the need to prevent access right lending and cloning of privileges is of utmost importance. Furthermore, information infrastructures that strongly rely on the real-time availability of central parties can be taken down by denial-of-service or physical attacks on these central parties.

The need for privacy is also high in critical information infrastructures, primarily driven by the threat of criminal elements actively monitoring the network to find out who is communicating with whom about what, and who is accessing which resources for what purpose. Infrastructures that systemically rely on uniquely identifiable serial numbers for every action are subject to a variety of attacks. For example, targeted key holders can be denied access to resources, communication attempts of targeted parties can be blocked in real time, and transaction-generated data conducted with target public keys can be filtered out by surveillance tools.

According to the US Federal Chief Information Officers Council in June 2001, "There are certain undeniable premises or trends [...] First is the fundamental need for strong authentication over a wide range of

applications that use electronic transactions. […] public key technology meets this need better than any other single technology […] Second, Federal agency use of public key technology is growing quickly both vertically (within organizations) and horizontally (across organizations)."

An information security infrastructure that revolves around the distribution and management of public keys and digital certificates is called a Public Key Infrastructure (PKI). PKI-based smart card infrastructures are deemed crucial for secure access management in critical information infrastructures; it is unacceptable to manage secret keys on personal computers or other devices that are not tamper-resistant. Dozens of government organizations around the world are implementing or have committed to PKI-based smartcards, not only for their own employees but also for the purpose of replacing driver's licenses, airport security documents, passports, and so on. Multi-application smart cards are believed to be the way to go.

The US Department of Defense is about to roll out more than 4 million PKI-based multi-application access smart cards, with the goal of securely managing access to the Department's computer networks, buildings, and controlled spaces. Special care has to be taken to ensure that a corrupt organization in the smart card supply chain does not build a backdoor into some or all of the cards to enable them to subliminally leak critical information (such as secret keys) to selected parties. Of course, citizens have strong reasons to be worried about their privacy as well when subjected to smart cards. According to Gartner in March 2002: "Gartner's survey shows that the public supports a national ID only for very specific purposes, and people are quite suspicious of what governmental agencies might do with it."

3.3 E-government

E-government refers to the electronic delivery of government services to citizens, particularly over the Internet. Other methods, such as proprietary networks and electronic self-service kiosks, are being explored as well. The primary objective of E-government is to simplify the interaction with citizens, enterprises and other institutions. In the words of UK Prime Minister Tony Blair in April 2002, "The opportunities are clear: better, more personalized, more efficient public services which handle personal information in a way that commands public trust." In the past two years, most governments around the world have established some degree of online presence. Among the leading countries to bring government online are the United Kingdom, Canada, and the United States. The United States alone is conducting over 1400 E-government initiatives, and has a law (the Government Paperwork Elimination Act) that requires federal agencies by October 2003 to accept electronic forms with electronic signatures when practicable.

In a June 2002 report, Giga Information Group distinguishes between five phases of E-government: providing information via websites; electronic service delivery; improving operations through Web interfaces and electronic data exchanges; moving toward more personalized electronic service delivery ('e-CRM'); and, introducing Web-based collaborative technologies. They state: "Implementing CRM initiatives, at least those appropriate for government, is a key priority. Federal agencies in particular as well as some states are using CRM to provide personalized citizen self-service, to track interactions with citizens and ensure responsiveness, to manage citizen contacts across multiple channels and to measure and assess citizen satisfaction with government services."

Concerns by individuals about information security and privacy are among the main barriers to the adoption of E-government services. In an August 2000 survey by Hart-Teeter about US citizens' view of E-government services, 53% of respondents indicated that they were extremely concerned with the potential loss of privacy. According to the survey, "The public's leading concern is identity theft, as 65% are extremely concerned about someone obtaining government-stored personal information and using it to steal their identity." In a 2001 Gartner survey, nearly 70% of consumers cited privacy concerns as one reason that could make them stop using E-government services. According to the UK Cabinet Office in 2002, "There is a lack of public trust in the way that the public sector handles personal information and the security of that information, and some concern about the risks to personal privacy from technological change. […] this could carry the risk of seriously undermining society's trust in public services and lead to significant and long-lasting harm to the effective delivery of services, including implementation of integrated E-government services."

It is becoming increasingly clear that electronic authentication is not the key security issue. As Giga Information Group observed in June 2002: "As governments allow more and more external citizens, suppliers and partners access to information via the Internet via internal systems, and as more government agencies provide access to internal databases and systems to employees via intranets, the old notion of a single enterprise security perimeter has become obsolete. […] Authentication — generally by user name and password for citizens, but increasingly by digital certificates combined with smart cards and/or biometrics for employees and government contractors — is the current visible battle. However, the real war is authorization through the use of directories that describe what access privileges an authenticated user will have to which systems and databases."

3.4 E-health

E-health is revolutionizing healthcare across the globe. It has been pegged by Forrester Research to be a $370 billion industry by 2004. A fundamental goal of E-health is to allow electronic access to health record information via the Internet and wireless networks. This is expected to greatly improve the quality of patient care, enhance the productivity of healthcare professionals, and reduce the costs of healthcare delivery and financing. According to a 2001 Harris Interactive/ARiA Marketing survey, patients are interested in communicating with their doctors, obtaining personalized medical alerts specific to their medical histories from their doctors, and participating in online communities to get information on diseases, hear about alternative therapies, and chat with others. Major efforts in Web connectivity are underway among hospitals, labs, pharmaceutical manufacturers, medical device companies, insurers, and managed care providers. Also, physicians are increasingly using handheld devices offering convenience, flexibility, and mobility.

Security and privacy are major impediments to E-health. According to a Health Canada report of January 2001, "Privacy is the most important policy area that needs to be addressed in relation to EHR's. Without public support on how privacy will be addressed, EHR systems will not be able to proceed." In the words of Bill Clinton in relation to HIPAA: "Nothing is more private than someone's medical or psychiatric records." The US industry costs for implementing HIPAA, which covers information security and privacy for virtually every organization involved in

health care, have been estimated at $43 billion until 2005. HIPAA requires organizations to comply with a variety of security and privacy codes starting in October 2002. Although HIPAA does not require a specific security technology, PKI is generally believed to be the only way to meet its stringent requirements. In a 2001 report, Gartner estimates that 98% of all healthcare providers will be PKI-enabled by 2003. In its 2002 top-ten list for healthcare, Gartner states that it a key component of E-health initiative success is "good security and privacy programmes, including public key infrastructure and two- and three-factor authentication services. [...] through 2004, the healthcare industry will focus on authentication technology-agnostic privilege management infrastructure, access logging and infrastructure to ensure healthcare delivery operations continuity. […] By mid-decade, mobile devices (likely palm-sized or slightly larger tablets) will have made strong inroads into breaking healthcare's input data capture/automation bottleneck. […] They will contain digital certificates, private keys and biometric capabilities to ensure strong user authentication and digital signatures."

3.5 Consumer space

In order to transact directly with consumers over networks on the basis of electronic information, enterprises must be able to authenticate consumers independent of their location, network, or device. This requires some notion of 'digital identity'. Digital identity management refers to the management of identity-related information in digital form. Although the notion has been much-hyped in the past year as the next big thing, digital identity management in essence is simply the digital authentication and certification of identity-related information, and its biggest use is in access management.

The digital identity management space has been heating up in the past year, with big-picture roadmaps from industry leaders. Most industry leaders have united in the Liberty Alliance project, while Microsoft is pursuing its own solution in the form of .NET Passport. The goal is to simplify the interaction between consumers and service providers by enabling consumers to use single sign-on for all their authentications. Single sign-on avoids the duplication of passwords by giving each user a single password for all resources.

However, single sign-on is only a user convenience, and not a security solution. In fact, with a single password giving access to everything about a user, single sign-on is a security catastrophe waiting to happen. While single sign-on may be acceptable for small companies for sharing non-critical information with employees over their enterprise VPN, it is totally inappropriate for sharing sensitive information over open networks. Hardly surprising, a Gartner study published in April 2002 indicates that consumers do not trust single sign-on services on the Internet for making purchases or using Web services. In the words of Gartner in February 2002: "The battle over control of the technology behind […] single sign-on will be entertaining, but will not be a major factor affecting enterprise IT deployments in 2002." In its November 2001 Information Security Hype Cycle, Gartner even lists single sign-on as a technology that will never reach market acceptance. Both Microsoft and the Liberty Alliance have already stated the importance of stronger authentication and authorization mechanisms in their roadmaps, and are emphasizing the importance of giving the consumer control over their own information.

Secure access management is all about modern cryptography. According to Gartner in March 2000, digital signatures for electronic authentication and authorization purposes are

"the killer application for PKI". Wedbush Morgan Securities in their February 2002 report state that "We do not believe it is far fetched to expect that in the future nearly every commerce- and Internet-enabled piece of hardware and software will posses or support digital certificates to ensure the privacy of sensitive transactions."

4. Why X.509-style PKI is not a solution

Contrary to popular belief, the PKI-based solutions of the leading vendors do much more damage than good in the context of access management. At the root of the problem is the fact PKI was invented in 1978, at the dawn of modern cryptography. The 1978 approach, which has been standardized through X.509 and related standards, was never invented for the purpose of access management: it was invented for the purpose of facilitating message encryption without the sender having to share in advance a secret key with the recipient. To send an encrypted message, the sender applies a public key of the designated recipient, which (to prevent obvious man-in-the-middle attacks) has previously been bound to the recipient's identity by means of a digital signature of a trusted central authority, called the Certificate Authority. The resulting digital identity certificate is the electronic equivalent of a passport: it authenticates the recipient, enabling the sender to convince himself that he is applying the right public key. In line with the 1978 goal of PKI, the leading PKI vendors continue today to define authentication as having to do with message encryption. In the words of one of the leading PKI vendors: "Encryption hopes to ensure: confidentiality; integrity; and, authentication (so that no one is sending false messages)."

Authentication as used in message encryption, however, is a very far cry from managing the access of authorized parties to identity-related information, with all its privacy, security, and performance sensitivities. This mismatch of the access management offerings of the leading PKI vendors with today's information security needs is not surprising: in 1978, open networks were hardly existent, let alone organizations seeking to share their information over such networks. Now that it is becoming increasingly clear that message encryption is not where the information security market is heading, PKI vendors are reinventing themselves to address the problem of access management. As an unfortunate consequence, the usage of X.509-style PKI is currently being distorted and stretched all the way into the realm of access management for sensitive information. There is only so much distorting and stretching that can be done, however. Applying X.509-style PKI to access management in increasingly open environments is like placing a car engine in a passenger plane; it simply will not fly.

Specifically, the X.509-style PKI approach to digital access management is to store identity-related information in central databases, and to use digital identity certificates as authenticated pointers to the database entries of their data subjects. This approach fundamentally suffers from bad performance, provides poor security, and violates many of the privacy principles codified in law in most countries.

4.1 Access right cloning and lending

X.509-style PKI does not provide software-only protection to discourage certificate holders from transferring (copies of) their access rights and entitlements to other parties: a user's secret key is simply a random number, and so revealing it to someone else has no direct negative consequences for that certificate holder. This defeats the entire purpose of PKI in the context of access

management. Even when secret keys are stored in smart cards, the break of a single smart card suffices to bypass the security of the system. Hackers around the world are already cloning pay-TV smart cards, high-value phone cards, and debit cards, causing hundreds of millions of dollars in damages. In May 2002, two University of Cambridge computer security researchers announced an inexpensive attack that employs a $30 camera flashgun and a microscope to extract secret keys from widely used smart cards. What is needed is a software-only protection mechanism; when implemented in smart cards, this would result in two layers of defence instead of just one.

4.2 Non-scalable

The approach of using an X.509 certificate as an authenticated pointer does not scale beyond pre-established trust domains. The actual authorization decisions are left to the access manager, which must look up elsewhere the relevant information needed to make an access control decision. With increasing numbers of individuals and organizations seeking to share information and computing resources, it becomes an administrative nightmare to keep track of each potential person's access rights, even when switching to low-grained role-based access control. Furthermore, when information is shared amongst different trust domains, the parties that have to make authorization decisions may not be able to gain access to all the data they need for making their decision: the missing data may be contained in databases outside of their control, and as a consequence may be incomplete, erroneous, or simply not available.

4.3 Central point of attack

It is difficult for organizations to protect their online central data repositories against misuse by hackers and insiders. A summer 2002 survey by Evans Data Corporation revealed that of 700 database specialists surveyed, 20% have experienced a direct breach in their database security. According to a 2000 CSI/FBI computer crime survey, 71% of unauthorized break-ins are by corporate insiders. Furthermore, data records may be outdated, and may be the result of misattributions due to identity theft. With X.509-style PKI, however, organizations cannot securely give individuals control over their own information, since data subjects typically should not be able to copy, lend, modify, discard, or prevent updating of information pertaining to them.

4.4 Poor performance on low-cost devices

In principle, both physical and logical security could be integrated by implementing PKI in smart cards or other portable devices (including Web-enabled cell phones and handheld computers). This would allow companies to migrate multiple disparate security systems into one integrated PKI system. However, in the words of the Aberdeen Group in their January 2001 white paper, CPU drain prevents X.509-style PKI from being "a solo building block". Indeed, the computational requirements of processing an X.509-style certificate are well beyond today's popular smart cards and devices. Addressing the problem by adding advanced circuitry (such as cryptographic co-processors) seriously increases the price of these devices. According to the PKI Forum in an April 2002 report, "Price competitiveness is the overriding driving factor for the card industry and will continue to commoditize the cards and components." Worse, the addition of sophisticated circuitry can easily lead to new weaknesses in the internal defence mechanisms, and adversely affects reliability. These problems get worse in the case of multi-

application smart cards, which are desirable to prevent system providers from needing separate card platforms for each individual application and to decrease the number of lost cards.

4.5 Identity theft

Systems that systemically rely on user identification give rise to a fraud known as identity theft, whereby fraudsters assume the identities of their victims. According to the US Federal Trade Commission (FTC), identity theft is the fastest growing crime in America, affecting approximately 900 000 new victims in 2001. The FTC expects that its cost will reach $8 billion by the year 2005. A recent study by the UK Department of Trade and Industry warns that in the not–too–distant future criminals will be as interested in stealing victims' identities as they are in stealing possessions. Notwithstanding the fact that X.509-style PKI provides for message encryption, it seriously increases the risk of identity theft, since its fundamental premise is that of inescapable system-wide identification. Criminals who manage to steal digital identity certificates or to assume the identities of unwitting people will be able to misuse certificates on a global scale, while their victims take the blame. Also, Certificate Authorities will have to establish identities on the basis of legacy paper-based systems, and will thus inherit their insecurity. This exposes organizations to potentially unlimited legal liability. (Legislation such as the US E-Sign Law, passed in 2000, and the EU Digital Signature Law, passed in 2001, recognize digital signatures as legally binding.)

4.6 Privacy violations

According to an April 2002 report by Gartner, individuals distrust online authentication systems, their skepticism resting in great part on privacy concerns. In many contexts, to gain access to resources the requestor would prefer to present just enough credentials to be granted access. With X.509-style PKI, however, the real name of the requestor is exposed. Numerous studies show that individuals are increasingly concerned about who has access to their personal information and how it might be used. Failure to protect privacy can damage an organization's reputation, brand image, and valuation. It can also lead to litigation, fines, criminal sanctions, and civil liability. Moreover, it has been shown to be the leading cause of losses in sales opportunities.

The fundamental problem with applying X.509-style PKI to the problem of access management is that public keys are globally unique identification numbers. They are even more discernable than social security numbers, credit card numbers, and health registration numbers. These identifiers travel along with every action taken by system participants, and can be readily and automatically linked to the identity of their owners by a myriad of parties, even if the owner's name is not explicitly stated in the certificate. PKI vendors misleadingly define privacy as "communications are safe from eavesdropping". This works in the context of sending over the wire a message to an intended recipient, but in the context of access management, encryption has very little to do with privacy. Indeed, according to a January 2001 executive white paper by the Aberdeen Group, "as currently sold and implemented, PKI is incompatible with the coming privacy era [and] eliminates even the pretense of ensuring user privacy."

4.7 Violation of data protection laws

In response to the growing security and privacy concerns, many countries have started to enact data protection legislation that place

stringent requirements on use, retention and disclosure of information. Companies that fail to comply may run into serious problems with government, ranging from fines to operational suspension. Data protection laws pertain to 'personally identifiable' information, which is any information that can be linked (directly or indirectly) to an individual. Deleting an individual's name from his record does not imply that the record is no longer personally identifiable, since identification may take place indirectly on the basis of social security numbers, health insurance numbers, and so on. Most European countries have adopted national legislation implementing the 1995 European Data Protection Directive. The United States has adopted sectoral regulations to protect the privacy of personal financial information, including the Gramm-Leach-Bliley Act, the Healthcare Insurance Portability and Accountability Act, the Family Educational Rights and Privacy Act, the Children's Online Privacy Protection Act, the pending On-line Privacy Protection Act, and an abundance of privacy bills filed at the state level. Many other countries have also adopted or are in the process of adopting stringent privacy legislation, based on the "Fair Information Practice" principles of the Organization for Economic and Cooperative Development (OECD) in 1980.

According to Wedbush Morgan Securities in their February 2002 industry report, "The need to comply with these government regulations and preferences will be a steady and driving force in the adoption of access management technologies." X.509-style PKI, however, is a poor match with the fair information practice principles, since it makes everything fully identifiable. In fact, it is quite possible that the uncontrolled use of PKI will be found unconstitutional when challenged in court. Some precedents: the Hungarian Constitutional Court in 1991 decided that multi-use personal identification numbers violate the constitutional right of privacy, the Portuguese Constitution states that "Citizens shall not be given an all-purpose national identity number", and SSN legislation in many countries prohibits the use of SSNs beyond very specific purposes.

4.8 Untrustworthy smart cards

If the secret key of a digital identity certificate is generated and stored on a personal computer or the like, it is virtually impossible to prevent its compromise, loss, disclosure, modification, and unauthorized use. However, when using X.509-style PKI in combination with smart cards, it is virtually impossible to verify that the cards do not leak their secret keys, card identifiers, access control codes, data from other applications running on the same device, and so on. Moreover, a variety of fake-terminal attacks become possible, and it cannot be guaranteed that the smart card supplier cannot simply reconstruct all the secret keys. As a result, application providers must have unconditional trust in the honesty of their smart card suppliers. National defence networks and other critical information infrastructures cannot reasonably place such trust in outsiders.

To summarize, X.509-style PKI cannot address the growing needs of access management systems.

5. A solution based on digital credentials

Credentica is currently building its Credential Management Platform (CMP), a set of server and client components that provide all the authentication and authorization services essential to today's information infrastructure applications. Applications can directly access these services to provide security, privacy, efficiency, and user-control for all system

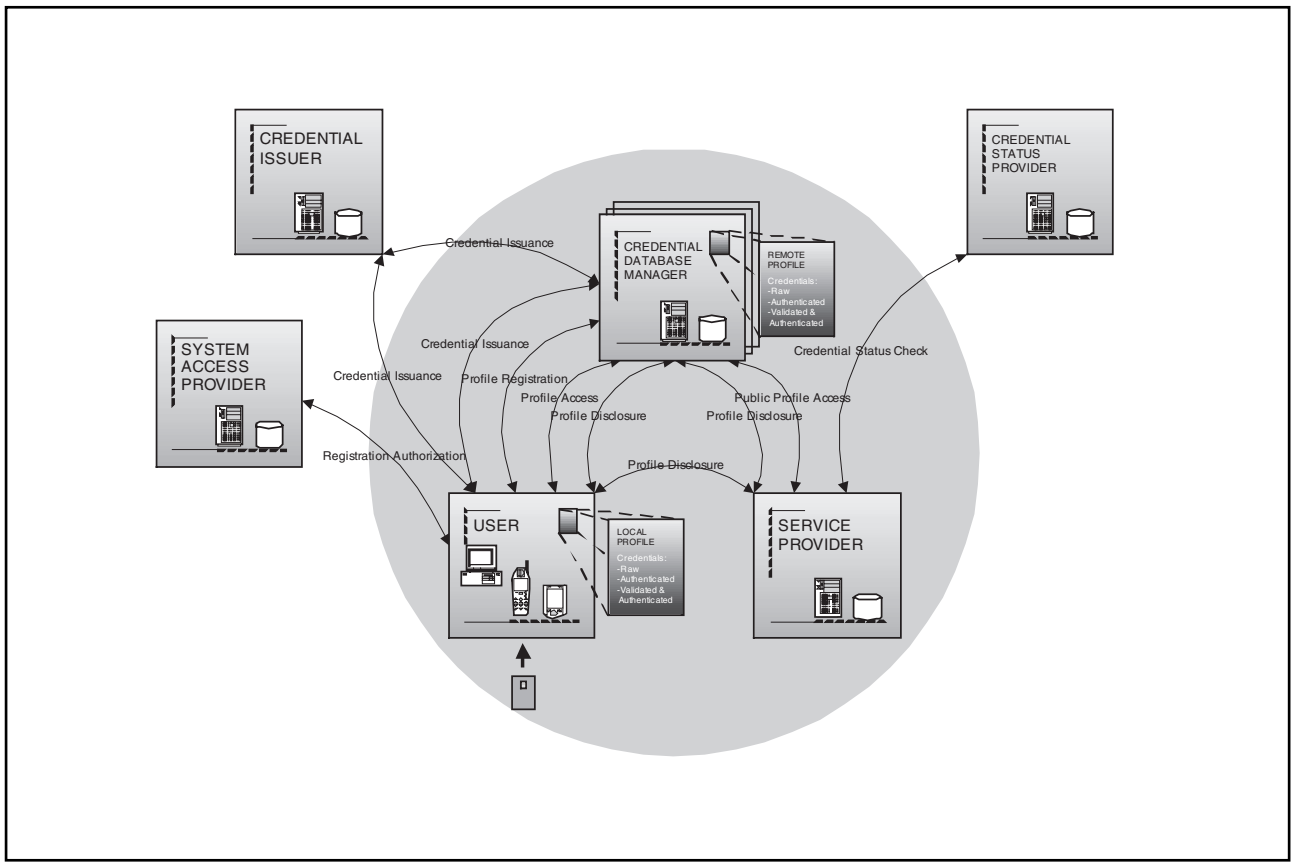

Figure 1: Credential Management Platform.

participants. The platform holistically overcomes the shortcomings of access management systems based on X.509-style PKI.

Secure single sign-on: The functionality of single sign-on is achieved in a manner that is vastly more secure and scalable. A user gains access to services by presenting a digital certificate that becomes usable by him upon unlocking a (password-encrypted) secret key that is stored on his computer. Users may be represented by a PC, a hand-held, a mobile phone, a smart card, or any other device (or combination of devices) capable of computing and communicating.

Strong security: The platform provides data integrity, data confidentiality (through encryption), and secure time-stamping. Audit capability is provided for non-repudiation and to assess compliance with regulatory requirements, through secure audit trails and digital receipts. It supports authentication strengths ranging from weak to military-grade two-factor and three-factor security. CMP provides for the management of digitally authenticated credential information; different Credential Issuers can vouch for the authenticity of identity-related information by digitally certifying that information. Organizations can strongly discourage credentials holders from lending or cloning

their access rights (even for pseudonymous access) by embedding disincentives that will be disclosed if and only if the legitimate holder commits a fraud.

Negotiable privacy: Adaptable privacy levels range from user-driven anonymity to enterprise-mandated identification. In particular, the platform allows for pseudonymous as well as role-based access (both server-driven for scalability and user-driven for privacy). There are multiple protocols for gaining access to credential information, with varying levels of involvement from the credential database manager and the user. Automated trust negotiation permits the exchange of credential information, ensuring that only the minimum credential information needed to meet the authorization requirements of the service provider is disclosed. In particular, identity-related information can be selectively disclosed in a manner that does not enable identification.

Limited-use access rights and credentials: Issuance and management of limited-use access rights is possible, and Credential Issuers can issue credentials that are valid a limited number of times. A built-in identifier, value token, or self-signed fraud confession will be exposed if and only if the credential is shown more times than allowed.

Information can reside anywhere: Credential information can be held both locally and remotely. Federation of remotely stored credential information is possible: credential information pertaining to the same entity can be accessed and managed as one logical entity even if it is distributed across different storage locations. CMP facilitates automated sharing and synchronization of credentials between local and remote credential information in accordance with application-specific administrative data. Roaming is also feasible.

Efficient smartcard implementations: The storage and computational burden for the smart card can be off-loaded almost entirely to a user-controlled device (such as a PC, a laptop, or a PDA), while preserving all of the smart card's security benefits. Billions of digital certificates, which may come from disparate systems that do not trust one another, can be securely managed using a single 8-bit smart card.

Secure multi-application smart cards: Smart cards can be used as multi-application devices, without introducing any of the privacy and security problems caused by other technologies. Specifically, different application providers can all share the same secret key stored in a user's smart card to derive the security benefits of that smart card. The certificates will have uncorrelated secret keys which cannot be determined by anyone including the smart card supplier, and all the certificates can be revoked separately. The application software on the user's trusted computer ensures that smart cards attacks are impossible, and that different applications using the same smart card remain fire-walled.

Managed Services: Credential Issuers can certify sensitive information on behalf or organizations without being able to learn that data, and Revocation Authorities can validate certificates (using OCSP or other standards) without being able to learn the identities of the clients of organizations (even when these expressly identify themselves to the organizations they transact with through the certificates themselves). In this manner, organizations can outsource core tasks related to digital authentication and authorization, without having to provide their managed services providers with competitive data or

customer information for which they could incur legal liabilities.

Peer-to-peer support: Organizations can securely give individuals control over some or all of their own credential information by allowing them to store and manage the information locally on their own computer. This information is cryptographically protected to ensure that users cannot modify, discard, lend, or prevent updates of information for which they have no right to do so. In the extreme, an organization can do away with its central databases, by securely distributing each database entry to the individual to whom it pertains. By basing authorization decisions directly on authenticated attributes shown by the requestor himself, trust can be established offline on first contact, with no prior knowledge of the requestor.

CMP is based on so-called Digital Credentials. Digital Credentials are basic cryptographic constructs, much like digital signatures but with much greater functionality. They are issued to applicants by trusted parties, referred to as Credential Authorities. Each Credential Authority has its own key pair for digitally signing messages. When issuing a Digital Credential to Alice, the issuing Credential Authority (through its own digital signature) binds one or more attributes to a Digital Credential public key, the secret key of which only Alice should know. The whole package that Alice receives is called a Digital Credential. When Alice shows her Digital Credential to Bob, she not only sends her Digital Credential public key and the signature of the Credential Authority, but she also digitally signs a nonce using her secret key. A nonce is a random number, the concatenation of Bob's name and a counter, or any other fresh data provided by Bob. Bob cannot replay the data, since in each showing protocol execution a new nonce must be signed, which requires knowledge of Alice's secret key. At the same time, Alice selectively discloses to Bob a property of the attributes in her Digital Credential, while hiding any other information about them. To convince Bob that the claimed property is true, Alice's signature on Bob's nonce doubles up as a proof of correctness. In the issuing protocol, Alice can ensure that the Credential Authority can learn neither the Digital Credential public key it certifies nor its own digital signature, while the Credential Authority can prevent Alice from modifying the attributes that end up in the Digital Credential. The Credential Authority can issue and update Digital Credentials without ever needing to see the attributes within it. The Credential Authority can also encode disincentive attributes into Digital Credentials that will be disclosed if and only if Alice copies or lends the Digital Credentials or (in the case of limited-use Digital Credentials) shows them more times than allowed. Furthermore, the Credential Authority can ensure that Alice cannot show a Digital Credential without needing the assistance of a smart card that has been issued to her; at the same time, the smart card can be prevented from leaking sensitive information, and need not do any resource-intensive cryptographic operations.

A detailed description of how these and other properties are achieved is outside the scope of the present paper. For full details the interested reader is referred to a book published by MIT Press [1], and for non-technical and technical overviews to [2] and [3], respectively. For the present purpose, it suffices to think of a Digital Credential as a hybrid form of a digital identity certificate and a digital attribute certificate, encompassing them both as degenerate cases and offering a bundle of security, privacy, and performance benefits. CMP makes use of Digital Credentials in three basic ways: to authenticate the data entries of local and remote records, to serve as authenticated

pointers to records, and to provide digitally signed audit trails.

6. Outlook

As information sharing applications are increasingly spanning multiple trust domains, the fundamental shortcomings of X.509-style PKI are becoming painfully clear. With the increased sharing of information across applications, the boundaries between the enterprise, government, and consumer space will continue to blur. In the next five to 10 years, all three spaces will develop highly similar needs for access control solutions: strong security, adaptable levels of privacy and user control, multi-application smart cards, and cost-effective implementations in low-cost portable devices.

At present, the leading PKI vendors are working to integrate PKI into access management solutions in a user-friendly manner. At the same time, they are struggling to match their PKI technologies with the growing needs of access management, and have yet to gain momentum in the new market. With the market in an embryonic stage, organizations that require secure and scalable access management solutions would do well to consider superior alternatives that enable the separation of the concepts of authentication and authorization in a secure, efficient, and privacy-friendly manner.

References

[1]Brands, S., 2000. "Rethinking Public Key Infrastructures and Digital Certificates; Building in Privacy," MIT Press, 2000. With a foreword by Prof. Ronald L. Rivest. This 350-page book describes the mathematics of Digital Credentials and analyzes their security. See http://www.credentica.com/technology/book.html for excerpts and endorsements.

[2]Brands, S., 2002. "Towards Digital Credentials," April 2002, ERCIM News #49. Available for download from http://www.ercim.org/publication/Ercim_News/enw49/brands.html

[3]Brands, S., 2002. "A Semi-Technical Overview of Digital Credentials," scheduled to appear in the November 2002 issue of the International Journal on Information Security, available for download from http://www.credentica.com/technology/overview.pdf