Non-Intrusive Identity Management

Dr. Stefan Brands
McGill School of Computer Science & Credentica
brands@cs.mcgill.ca, brands@credentica.com
March 23, 2004

ABSTRACT: This paper presents a novel architecture for digital identity management. The proposed architecture is highly secure and scales seamlessly across organizational boundaries, while at the same time protecting the privacy interests of individuals and organizations. To achieve these properties, the architecture heavily relies on Digital Credentials, a cryptographic authentication technology specifically designed to allow data subjects and organizations to securely co-manage identity-related information. We also examine the use of the new architecture in the context of three emerging information-sharing applications: Electronic Health Record management, E-Government, and Digital Rights Management.

1. Introduction

Most people are registered in many hundreds if not thousands of databases scattered across disparate systems. In identity management jargon, individuals have multiple *network identities*: collections of information that relate to an individual, that are created and managed as single units in a network, and that are stored in electronic form. Advancements in networking technologies make it increasingly easy to collect and collate these network identities.

Of course, this cross-domain aggregation power by itself is not of much value to organizations, unless it is combined with the ability to determine which network identities correspond to the same individual. Traditionally, identifiers such as health insurance numbers and Social Security Numbers serve as keys to facilitate such cross-linking. The current efforts in the electronic world to enable cross-domain identity management and information sharing rely on their own unique cross-domain identifiers, such as biometric templates and digital certificates.

For businesses, an increase in cross-domain linking power ultimately translates into increased sales and cost reduction. For government organizations, the ability to share client information translates into more efficient interactions with citizens and an improved ability to detect and contain fraud. Individuals stand to benefit as well from these, at least in principle.

Privacy and security concerns

The increased introduction of (and reliance on) cross-domain identifiers also brings serious privacy risks. Target marketing can turn into spamming, service customization can turn into unfair price discrimination, hackers and insiders can cause systemic denial of access to targeted individuals, and so on. For these and other reasons, many people provide false identity information when accessing on-line services.

Indeed, the business goal of cross-domain digital identity management is directly at odds with the privacy interests of individuals. What businesses since a few years refer to as network identity is essentially what data protection legislation around the world already since the early eighties refers to as *personal information*: information about a "data subject" whose identity can reasonably be ascertained from the information. Data protection legislation requires organizations to protect personal information in accordance with several privacy principles, one of which is information security safeguards.

Intra-enterprise security needs, however, are much lower than cross-domain data protection requirements. Indeed, while password-only authentication is often adequate for internal access to organizational resources, in cross-organizational contexts it would give outside organizations unacceptable impersonation powers. In the context of cross-domain access management, traditional information security products (such as firewalls, anti-virus software, intrusion detection systems,

and vulnerability assessment tools) break down as well; with trust domains being logical rather than physical, security must be tied to the data itself rather than to the perimeter of its repository.

In short, organizations are starting to discover that the arsenal of security tools they use for intra-organizational data protection is not appropriate to protect information that is shared across organizational boundaries.

Federated identity management

The currently prevailing industry approach to address this situation is to centralize all the authentication power from different domains into a single trusted domain that acts on behalf of its constituent organizations. With *federated identity management* architectures, such as those pushed forward by Liberty Alliance, organizations do not authenticate access requestors themselves, but instead query a trusted Identity Provider that does the authentication for them. The Identity Provider simply returns an authentication assertion as to the validity of the identity claim of the access requestor, which the relying organization uses in its own authorization process. This approach in effect maps the cross-domain context back to the traditional single-domain context, which organizations know how to handle using traditional authentication techniques, be they password-only authentication, Kerberos, or perhaps PKI. (Indeed, PKI vendors generally consider federated identity management, and notably standardization efforts such as SAML, as what will rescue PKI from an untimely death, since a full-fledged certificate infrastructure is unnecessary.)

However, centralizing systems of an inherently decentralized nature brings its own administration, scalability, security, and privacy problems, which may be far worse than the original problem one was seeking to solve. In its original Passport architecture, for example, Microsoft relied on the centralization of all authorization data, and was forced to back down following complaints from consumer groups, EU officials, and organizations that were reluctant to entrust Microsoft with their customer data. The Liberty Alliance proposal improves over the original Passport scheme by leaving personal data at the organizations that collected it, but the authentication power (and therefore the ultimate access control power) remains centralized within each circle of trust.

At its core, federated identity management architectures such as the Liberty Alliance proposal and Microsoft's revised Passport scheme are centralized authentication architectures. Indeed, the Identity Provider's role in the Liberty Alliance architecture greatly resembles that of Visa or Mastercard among their respective "circles" of merchants: within its circle of trust, the Identity Provider can track, trace and link in real time all the interactions between users and organizations. (It may not know the transaction details itself, but that by no means is enough for information privacy.) The Identity Provider can even impersonate users and falsely deny them access everywhere. Furthermore, Identity Providers are highly appealing targets for fraudulent insiders and hackers. On top of that, relying organizations do not get the strength of the authentication mechanism used by the Identity Provider, but merely that of the session maintenance mechanism used when redirecting the user between the organization and the Identity Provider; impersonating a user or cloning access privileges depends merely on the difficulty of getting to a session cookie, rather than on the difficulty of getting to the user's secret key (which could be stored on a smartcard). For an in-depth analysis of the Liberty Alliance architecture, see [1].

More generally, the privacy, security, and scalability problems of centralized authentication architectures have been well-documented in the past two decades by the professional cryptography and security community. In the context of "unbalanced" B2B digital identity management (where organizations inherently place asymmetric trust in a central party), the shortcomings of industry's current federated identity management efforts may not be problematic. Beyond that, however, they may well turn out to be a showstopper. Collaborative enterprise efforts, where participating organizations are equals ("balanced" B2B), may already prove too much of a stretch, not to mention G2C, B2C, and C2C applications.

The need for new approaches

The growing mismatch between the security needs of cross-domain identity management and traditional security tools and practices is not all that surprising. The currently prevailing authentication techniques (password-only, biometrics, Kerberos, PKI) were all invented more than two decades ago, when open networks were hardly existent, let alone organizations seeking to securely share identity-related information over such networks. At that time, privacy legislation was virtually non-existent. The only privacy protection that the designers of the traditional security techniques had in mind was protection against unauthorized outsiders (e.g., wire-tapping). In the new frontier of cross-domain access and identity management, however, the biggest threats to privacy do not come from outsiders, but from insiders.

To better understand the shortcomings of PKI and other authentication mechanisms that were not designed with cross-domain identity management requirements in mind, it is important to understand the relation between (information) security and privacy. *Security* is generally defined as the extent to which information can be stored and transmitted in such a manner that data access is limited to authorized parties. *Privacy* is generally defined as "the claim of individuals to determine for themselves when, how, and to what extent information about them is communicated to others." In accordance with the Fair Information Principles of the OECD (which form the basis of most of today's data protection legislation around the world), "security safeguards" is only one of the eight principles necessary to achieve privacy. In contrast to security, which is aimed at preventing access by unauthorized outsiders, the other basic privacy principles are primarily aimed at unauthorized use by insiders. As such, security safeguards are necessary to achieve information privacy, but not sufficient. Ironically, traditional authentication technologies have a highly adverse impact on two of the most important privacy principles: collection and use limitation. They are, in fact, privacy-invasive technologies.

What this paper is about

Two decades of research in modern cryptography has shown that security and privacy are not trade-offs, but that they are mutually reinforcing when implemented properly. A fundamental premise of modern cryptography is that the need to rely for privacy on Trusted Third Parties (such as the Identity Provider in federated identity management) can be eliminated. This brings us to the goal of this paper: to present a non-intrusive identity management architecture that is highly secure and that scales seamlessly across organizational boundaries.

2. Non-Intrusive Identity Management

Before describing the proposed architecture, we give an overview of the state-of-the-art authentication primitive that is at the core of the new approach to cross-domain identity management: *Digital Credentials*. Our architecture will rely on this new primitive in four ways.

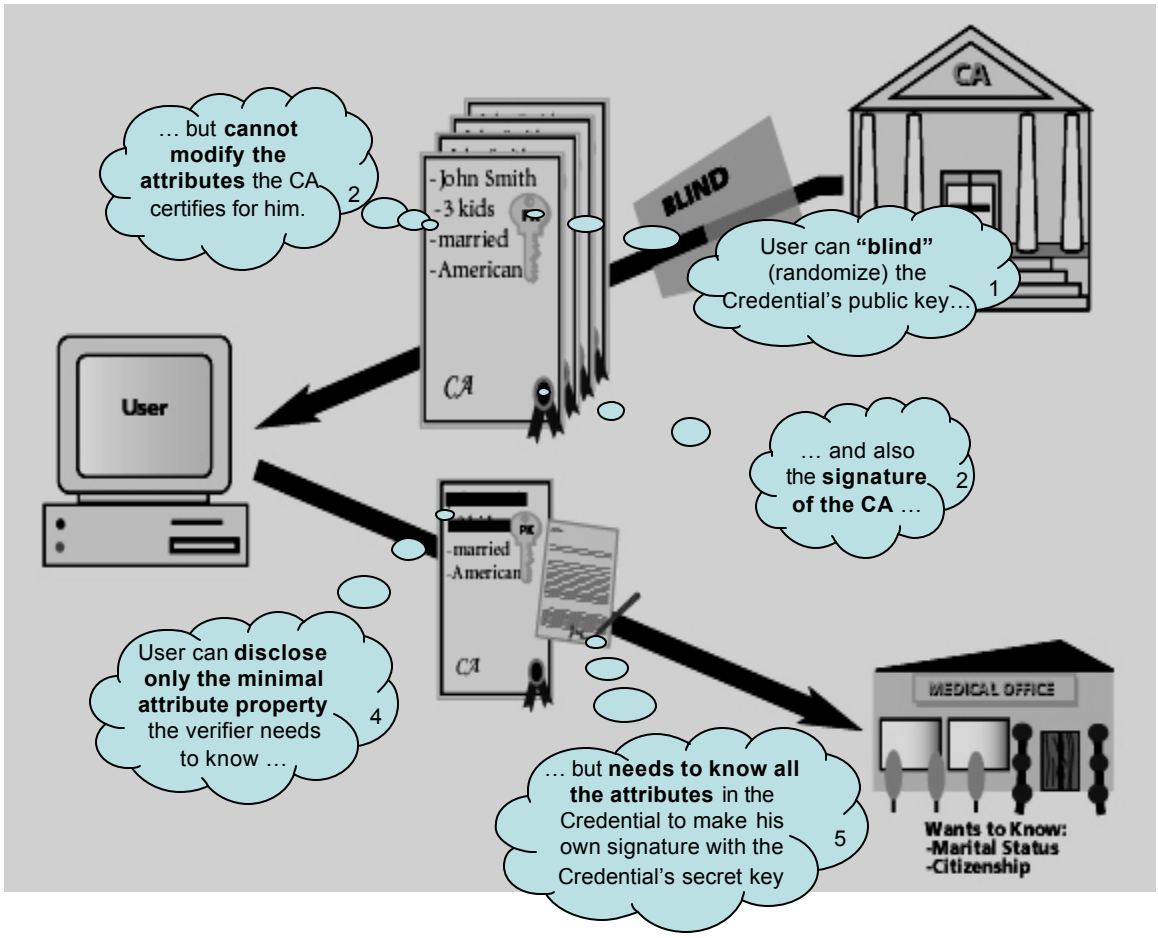
Digital Credentials

Digital Credentials are the culmination of two decades of scientific advances by dozens of professional cryptographers, starting in the early eighties. They are basic cryptographic constructs, much like digital signatures and equally efficient, but with much greater functionality. Specifically, Digital Credentials provide *fine-grained privacy control* at every step in the life-cycle of certified personal data that is being sent around. As well, they have security properties that go well beyond what can be achieved for X.509 identity and attribute certificates (in ways that may at times seem counter-intuitive), and can be implemented in low-cost smartcards without cryptographic coprocessors. They support all the traditional authentication strengths, from software-only protection to military-grade two-factor and three-factor security. Technically, Digital Credentials are issued and shown as follows:

- **(Issuing protocol)** Digital Credentials are issued to applicants by Credential Authorities. Each Credential Authority has its own key pair for 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 knows. (An attribute can be any information.) The entire package that Alice receives is called a Digital Credential. Although the sequences of zeros and ones that make up Alice's public key and the signature of the Credential Authority are unique for each Digital Credential, the Credential Authority cannot learn who obtains which sequences; they are *blinded* during the issuing process. At the same time, the blinding operations that Alice can perform are *restricted* in such a manner that Alice cannot modify the attributes that the Credential Authority encodes into her Digital Credential. What's more, some or all of the attributes in Alice's Digital Credential could initially be provided to the Credential Authority by Alice herself, by her smartcard, or by another organization, without the Credential Authority being able to learn them.

- **(Showing protocol)** To show her Digital Credential to Bob, Alice sends her Digital Credential public key and the signature of the Credential Authority. 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 Alice's information for his own benefit in another transaction, since in each showing protocol execution a new nonce must be signed; this requires knowledge of Alice's secret key, which never leaves Alice's device. At the same time, Alice can *selectively disclose* to Bob a Boolean property of the attributes in her Digital Credential (this goes well beyond what can be done with a paper-based certificate and a marker), while hiding any other information about them. Importantly, however, it is infeasible for Alice to demonstrate any property without her actually knowing all the attributes encoded into her Digital Credential (including those that she does not disclose). To convince Bob that the claimed property is true, Alice's signature on Bob's nonce doubles up as a proof of correctness.



A detailed description of how these basic properties are achieved in a highly practical manner is outside the scope of this paper. A technical overview of Digital Credentials can be found in [2], and the full details appear in [3]. As explained in these references, by carefully exploiting these basic properties of Digital Credentials, one can efficiently realize all of the following features:

- **(Privacy of Credential holders)** Digital Credentials accommodate fully adaptable levels of privacy ranging from user-driven anonymity to government/enterprise-mandated identification. They support automated negotiation of credential information, ensuring the disclosure of only the minimum credential information needed to meet the authorization requirements of an access provider; this minimizes the risk of identity theft, and preserves privacy. The selective disclosure technique can be applied not only to attributes encoded into a single Digital Credential, but also to attributes in different Digital Credentials, possibly certified by different Credential Authorities.[1] There is no need to trust third parties to protect one's privacy: even if all the parties that rely on Digital Credentials actively conspire with all Digital Credential issuers and have unlimited computing resources, they cannot learn more than what can be inferred from the assertions that Digital Credential holders willingly and knowingly disclose.
- **(Privacy of Credential verifiers)** In many situations, verifiers may want or need to pass on Digital Credential evidence to central parties (e.g., for online revocation status checking, to enable fraud detection on behalf of multiple access providers, to allow statistical data gathering, or to serve as transaction receipts). A Digital Credential verifier can *selectively hide* any or all of the information that a Digital Credential holder selectively disclosed to it, before forwarding that Digital Credential. In other words, the verifier can forward non-repudiable transaction evidence that proves to third parties no more than exactly what it wants the evidence to prove; this may be much less than what the Digital Credential holder selectively disclosed to the verifier. By way of example, consider a patient-physician interaction or a consumer-merchant transaction; while the customer may have no problem identifying himself to his doctor or to the merchant, the latter parties may not want to disclose their customer's identity to third parties.
- **(Strong accountability)** Digital Credentials offer audit capability for non-repudiation and to assess compliance with regulatory requirements, through digital audit trails and receipts that facilitate automated dispute resolution. Malicious parties, including Credential Authorities, cannot frame a Digital Credential holder by making it look as if he or she participated in a transaction, even if they would collude and would have unlimited computing power or special knowledge of trapdoor information. Audit trails can be kept in the form of role-based digital signatures; in case of a dispute, the transaction originator cannot disavow the origin.
- **(Pooling protection)** Different people can be prevented from pooling together multiple Digital Credentials in order to enjoy access privileges that they would not enjoy on their own. Hereto the access provider requires the access requestor to demonstrate that any Digital Credentials that he or she shows contain the same built-in identifier. Owing to the selective disclosure property, an honest Digital Credential holder can demonstrate this without disclosing the built-in identifier.
- **(Discarding protection)** Digital Credentials can be used to prevent the discarding of authenticated information that an access requestor would rather not show. A mark for drunk driving, for instance, can be tied into a driver's license Digital Credential that specifies that the holder is authorized to drive. Once again owing to the selective disclosure property, the owner can hide the mark whenever it need not be disclosed.
- **(Lending protection)** Lending of credential information can be discouraged by wrapping the information into a Digital Credential and encoding confidential data of the legitimate owner into it. The legitimate owner can hide this data (again owing to the selective disclosure property), but

[1] Rather than encoding many attributes into a single Digital Credential, it may be preferable to distribute them across multiple Digital Credentials. This helps avoid the aggregation of an individual's attributes by a single Credential Authority, improves efficiency when many attributes need to be encoded independently, and removes the need to update certificates more frequently than otherwise needed.

the Digital Credential cannot be used without actually knowing the confidential data. (Note that this measure does not rely on credential holders using tamper-resistant devices.)

- **(Dossier-resistance)** A Digital Credential can be presented to an organization in such a manner that the organization is left with no evidence at all of the transaction (much like showing a passport without letting the other party make a photocopy) or such that the verifier is left with self-authenticating evidence of only a part of the disclosed property. Furthermore, the self-authenticating evidence can be limited to designated parties. In case of a dispute, the disclosed property can always be revealed in full.
- **(Limited-show credentials)** A limited-use Digital Credential can contain a built-in identifier, value token, or self-signed fraud confession, that will be exposed if (and only if) the Digital Credential is shown more than a pre-authorized number of times.[2] These *limited-show* Digital Credentials (which can be used to design the digital equivalent of stamps, coins, tickets, and so on) have no obvious paper-based analogue. The limited-show property holds even when Digital Credential holders are free at each occasion to choose the attribute properties that they demonstrate, and even if they conspire with verifiers (who, as mentioned, are able to hide any information disclosed to them before forwarding transaction evidence). Limited-show Digital Credentials are highly practical: to be able to compute a built-in identifier in case of fraud, a footprint of a mere 60 bytes must be stored for each Digital Credential shown, regardless of the complexity of the property disclosed and regardless of the number of encoded attributes.
- **(Negative authentication)** This property allows the holder of a Digital Credential to demonstrate that he or she is *not* someone listed on a blacklist, without enabling identification. More generally, the holder of a Digital Credential can demonstrate that the data in the Digital Credential does *not* meet certain conditions, without revealing more.
- **(Recertification and updating)** In many cases the right to access a service comes from a pre-existing relationship in which identity has already been established. An individual can present a certified public key for recertification or for updating to a Credential Authority, without enabling it to learn the current values of the attributes in the Digital Credential. Of course, the Credential Authority could require the individual to demonstrate an attribute property before certifying the Digital Credential or its updated version.
- **(Information can reside anywhere)** Digital Credentials can be held both locally (on a device of the user) or remotely, and can be managed using roaming. In the extreme, organizations can do away entirely with central databases containing sensitive personal information, by securely distributing each database entry to the individual to whom it pertains; the unique security properties of Digital Credentials ensure that unauthorized users cannot modify, discard, pool, or lend their own credential information, nor can they prevent it from being updated (without locking themselves out of the entire system).
- **(Smartcard Implementation)** Digital Credentials can be issued to, or embedded in, smartcards and other tamper-resistant devices; this provides a second layer of protection (on top of the abovementioned cryptographic protections) against loss, theft, lending, pooling, copying, and discarding of Digital Credentials. As well, the Digital Credential holder's smartcard can prevent other kinds of unauthorized behavior by its owner, and can protect him against "virtual" extortion attempts. The storage and computational burden for the tamper-resistant device can be off-loaded almost entirely to another user device that need not be tamper-resistant (such as a handheld device, a laptop, or another chip on the same smartcard that need not be trusted by the system provider), while preserving all of the smartcard's security benefits; literally billions of Digital Credentials can be securely managed in this manner using a single 8-bit smartcard chip without a cryptographic co-processor.

---

[2] Alternatively, copying and reuse can be prevented by resorting to online Digital Credential validation by a central party, but this may pose a serious performance bottleneck.

- **(Secure multi-application smartcards)** Smartcards 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 smartcard to derive the security benefits of that smartcard. The certificates will have uncorrelated secret keys which cannot be determined by anyone including the smartcard supplier, and all Digital Credentials can be revoked separately. The application software on the user's trusted computer ensures that smartcards attacks and data leakages are impossible. Moreover, different applications relying on the same smartcard can be fire-walled through the application software running on the patient's trusted computer, rather than the application providers and the card holder having to trust the smartcard issuer.
- **(Managed security services)** With an increasing number of incompatible authentication mechanisms in use, organizations that need to make authorization decisions will increasingly ask trusted authorities to issue and/or verify the credential information presented by their clients. With Digital Credentials, Credential Authorities 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 clients disclose their embedded identities to the organizations they transact with). In this manner, organizations can outsource core tasks related to digital authentication and authorization, without having to provide their managed security services provider with competitive data or customer information for which they could incur legal liabilities. Even the role of the tamper-resistant smartcard can be outsourced, removing the logistical problem of securely distributing tamper-resistant devices.[3]

With this set of features in mind, we are now sufficiently prepared to discuss our approach to cross-domain digital identity management.

Identity management based on Digital Credentials

We will refer to our proposed architecture as the Credential Management Platform (CMP). CMP is characterized by three central notions: records, participants, and protocols.

A *record* is a logical collection of information. Records may be held in a central database, may be distributed across multiple databases, or may be held locally on a user device. In the first two cases the record is called a Remote record; in the latter case it is called a Local record. In general, Local records offer greater security and privacy to access requestors, but may be less convenient. Implementations of CMP could facilitate the automated sharing and synchronization of Local and Remote records in accordance with application-specific administrative data, to allow multiple records to be managed electronically as one logical entity.

A record contains two kinds of information:

- **Attributes**: An attribute is any personal data, corporate intelligence data, or otherwise sensitive information to which access must be guarded. Attributes may be encrypted by a key known only to a participant; this is useful for instance when attributes that are normally held in a Local record are temporarily stored on a public network to support roaming access by other devices.
- Related **administrative data**: The administrative data describes rules that specify by whom each attribute in the record may be read, written, modified, or otherwise accessed. Administrative data can include audit trails (possibly digitally signed) for access events.

CMP distinguishes between three kinds of attributes in a record:

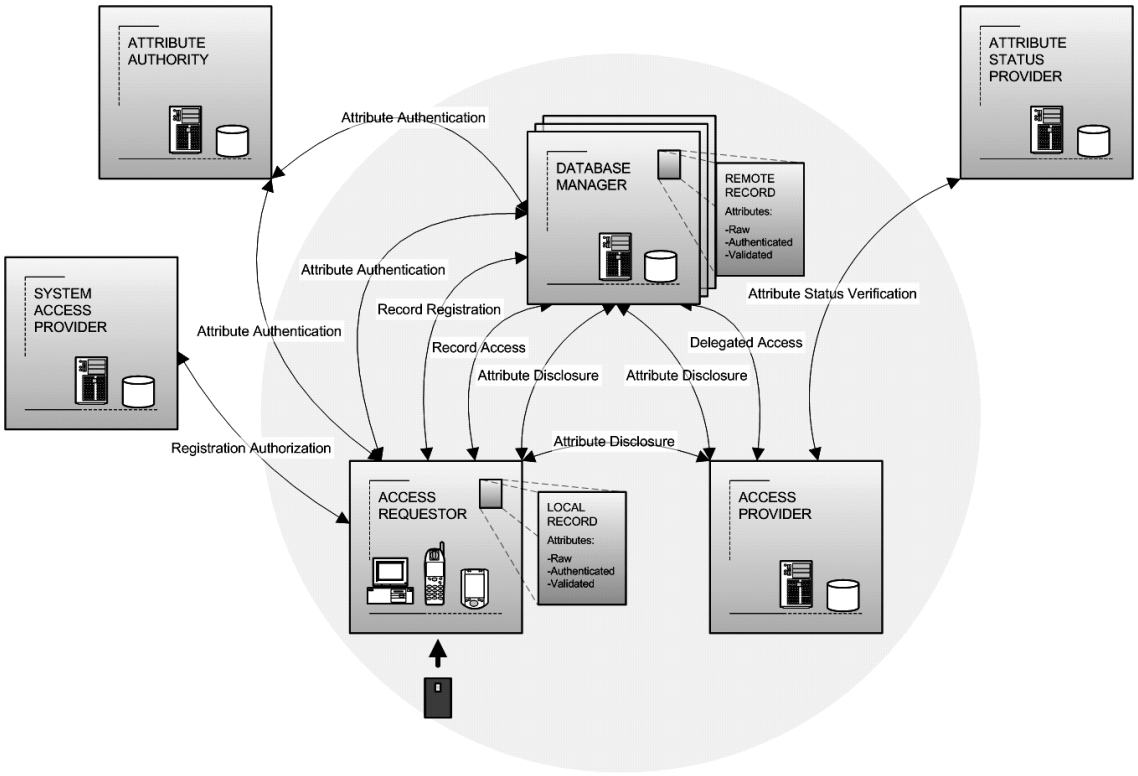
- **Raw attributes**: These are attributes specified by any party without any guarantee as to their validity. Personalized display or content preferences for a Web site are an example. Modification

[3] Although every transaction of a Digital Credential holder will now require the real-time involvement of a third party that guarantees protection of the user's secret key, that third party cannot learn any details that could lead to a privacy compromise (other than knowing the transaction times of pseudonymous users).

or discarding of raw attributes by unauthorized participants might cause inconvenience to the party to whom the data pertains, but would not adversely affect the security of any other party.

- **Authenticated attributes**: Attributes that are digitally authenticated by a participant by means of a digital signature, but without prior verification of their validity. This prevents other participants from modifying the attribute. In an on-line chat group discussing gender-related issues, for example, a person might wrongly specify his own gender but would be stuck with it in future sessions.
- **Validated attributes**: Attributes that are digitally authenticated by an "Attribute Authority" only after the validity of the attribute has been verified by that Attribute Authority.[4]

Authentication of authenticated and validated attributes takes place by wrapping one or more attributes into a Digital Credential, to offer unique security, privacy, and usability benefits. Different attributes may be packaged either into separate Digital Credentials or into the same Digital Credential. Furthermore, different Attribute Authorities may authenticate the same attribute by packaging it in different manners. By way of example, consider an electronic patient record: multiple doctors may digitally "sign off" on the same entries in a patient record. More generally, multiple Attribute Authorities may package the same or different overlapping subsets of attributes in a record in different ways into Digital Credentials. In this manner, access providers can be assured that the data entries on which they rely have been entered by authorized parties, and different parties can effectively maintain partial ownership of information in a record. Not even the party (or combination of parties) controlling the storage of a record can modify, delete, or add information, unless they are properly authorized.



A *participant* is a device or application (or a collection of devices or applications) that acts either autonomously or on behalf of an individual, a group, or an organization. For simplicity we will interchangeably refer to participants as both devices or applications and the parties they represent. CMP distinguishes between six types of participant:

[4] Attribute information may be supplied to Attribute Authorities by "Registration Authorities" who are responsible for validation; we do not explicitly show Registration Authorities in our architecture, however.

- **Database Manager**: A party that controls the physical storage of records.
- **Attribute Authority**: A party that issues authenticated or validated attributes. These attributes may be valid only a limited number of times or only for a limited-time period.
- **System Access Provider**: A special Attribute Authority responsible for granting participants the right to "initialize" Remote records (and possibly to subsequently manage it in a co-owner role). The System Access Provider issues Registration tokens, either one per participant until expiry of the token or a new one at regular time intervals or when requested.
- **Access Requestor**: A party interested in accessing a service that requires an authorization decision. The Access Requestor may be represented by a PC, a handheld device, a mobile phone, a smartcard, or any other device capable of computing and communicating.
- **Access Provider**: A party that relies on some or all of the attribute information in a record in order to make an authorization decision pertaining to an Access Requestor. Attributes in the record (more generally, properties about attributes in one or more records) are presented to the Access Provider either by the Access Requestor or by the Database Manager. In the latter case, either the Access Requestor's active involvement or prior explicit consent (in the form of a Delegation token) is needed. The Access Provider may resort to an Attribute Status Provider to complete the verification of authenticated and validated attributes.
- **Attribute Status Provider**: A party that verifies the status of one or more attribute-related requests presented by an Access Provider. Its primary role is to verify the revocation status of validated Attributes, to manage and issue updates of revocation lists, and to keep track of the number of times a limited-show attribute has been used.

In a real-world application, there will normally be many instantiations of most types of participant. For instance, in an electronic health record management system, each doctor authorized to update patient records would be an Attribute Authority. Of course, the roles of multiple participants from the same or from different systems may in practice all be performed by the same party.

Participants interact with each other by means of *protocols*. CMP distinguishes between seven basic protocols:

- **Registration Authorization**: A protocol between a System Access Provider and an Access Requestor whereby the Access Requestor obtains a Registration token allowing him to subsequently initiate a Remote record. The Registration token may be issued to a tamper-resistant device (e.g., a smartcard) of the Access Requestor for greater security.
- **Record Registration**: A protocol between an Access Requestor and a Database Manager whereby the Access Requestor presents a Registration token to initialize a record. As part of the protocol, the Access Requestor and the Database Manager specify administrative data.
- **Record Access**: A protocol between an Access Requestor and a Database Manager whereby the Access Requestor accesses a record stored by the Database Manager in order to read, write, or modify attribute information. The Access Requestor must show an Authorization credential (which may be the Registration token) to demonstrate proper access rights.
- **Attribute Authentication**: A protocol whereby an Attribute Authority issues an authenticated or validated attribute for entry into a Local or Remote record. The Attribute Authority issues the authenticated attribute either upon receiving an Authorization credential or upon receiving authenticated attribute information issued by another Attribute Authority.
- **Attribute Disclosure**: A protocol whereby an Access Requestor discloses attribute information to an Access Provider. The protocol can be conducted either with or without the assistance of the Database Manager. For Local records, there is no need to involve a third party in order to disclose attribute information to the Access Provider. For Remote Records the Access Requestor can disclose the attribute information to the Access Provider either by directly retrieving it online

and forwarding it to the Access Provider, or by routing its own access request through the Access Provider to the Database Manager.

- **Delegated Access**: CMP allows the Access Requestor to provide the Access Provider with a digitally authenticated Delegation token specifying the latter's access rights, so that the latter can later on access a record (perhaps for a limited period of time or a limited number of times) without further involvement from the Access Requestor's side.
- **Attribute Status Verification**: A protocol between an Access Provider and an Attribute Status Provider whereby the Access Provider requests and obtains information on the status of an attribute beyond what it can infer from the attribute itself. Attribute Status Verification may take place either on-line (in conjunction with an Attribute Disclosure protocol) or off-line. For short-lived authenticated and validated attributes, the Attribute Status Verification protocol may not be needed.

All tokens, access requests, and other forms of authentication in CMP are implemented using Digital Credentials. Specifically, CMP relies on Digital Credentials in four basic manners:

- To implement access privileges, entitlements, delegations, and any other attributes that access requestors show to access providers to allow them to make local authorization decisions;
- To implement privacy-enhanced digital identity certificates (usable as digital pseudonyms where identification is not required) that allow the separation of different spheres of activity;
- To authenticate the entries of electronic records stored in central or distributed databases; and
- To implement digital audit trails and digital receipts that witness details of access requests.

Unique benefits of CMP

As a direct consequence of using Digital Credentials throughout the CMP architecture, a number of unique benefits arise, including the following:

- The Registration token can be presented in a manner that does not enable identification of the Access Requestor. (Digital Credentials encompass identity certificates as a special case: an identifier is just one of infinitely many attributes that can be encoded into a Digital Credential, and the Digital Credential holder can disclose it whenever desired.)
- For Remote records, the Access Requestor can choose to be identified or to remain pseudonymous. The ability to pseudonymously hold a Remote record reduces the risk of identity fraud, and minimizes the damage that can be done by malicious insiders and outside attackers. In the case of a dispute a pseudonymous Access Requestor will not be able to deny having accessed the record; only pseudonymous Access Requestors who did not access the record can prove they did not do so.
- In the case of Local records, CMP allows the Access Requestor to be fully anonymous. The authenticators of attributes in the record can strongly discourage the Access Requestor from cloning or lending his attributes. Furthermore, the Access Requestor can present the Attribute Authority with a previously issued authenticated or validated attribute in order to have it re-authenticated or updated, without enabling the Attribute Authority to learn more than it strictly needs to. In the case of a limited-show attribute, a built-in identifier, value token, or self-signed fraud confession will be exposed if the attribute is used more times than allowed.
- The Access Requestor can disclose only the minimum attribute information (such as a particular property of multiple attributes) needed to meet the authorization requirements of the Access Provider. (In case the attribute is stored in a Remote record, this requires the Access Requestor to have some trust in the Database Manager.)
- Access Providers that know an Access Requestor under different unlinkable pseudonyms can enable the Access Requestor to transfer authentication attribute information from one pseudonym to another without creating pseudonym linkage, while at the same time preventing

the Access Requestor from showing attributes that belong to another Access Requestor (even if Access Requestors collude).

- In case Digital Credentials are issued to smartcards, all computationally expensive operations for the smartcard can be off-loaded to a more powerful device; virtually no smartcard storage space is required in that case, so that plenty of room is left for a software solution to protect against sophisticated attacks such as differential power analysis. Also, CMP can offer protection against fake-terminal attacks and smartcard data leakage by routing communications from and to the smartcard through a device trusted by the card holder.
- Attribute Authorities can digitally authenticate information on behalf of others without being able to learn attribute data that they have no need to know. Likewise, Attribute Status Providers can validate certificates without being able to learn the identities of access requestors and access providers. In this manner, Access Providers and Database Managers can outsource core tasks related to digital authentication and authorization to security specialists, without having to provide them with sensitive information.
- Attribute information can be presented to the Access Provider in such a manner that the Access Provider is left with self-authenticating evidence that proves only a part of the Attribute property disclosed by the Access Requestor; this enables the Access Provider to pass on the evidence to third parties (such as the Attribute Status Provider), while protecting its own privacy, complying with privacy legislation, and avoiding leakage of competitive intelligence.
- For Access Providers, Record Access can be identified, pseudonymous, or anonymous. The latter two cases prevent the Database Manager or the Attribute Status Provider from gaining competitive intelligence on Access Providers or from improperly rejecting valid requests for access on the basis of the identity of the Access Provider. At the same time, the Access Provider can disclose exactly that which is required to enable the Database Manager to make its own authorization decision: CMP provides for role-based access. The Database Manager and other parties can strongly discourage the Access Requestor from reusing, lending, pooling, discarding, or cloning his access rights, even for pseudonymous access.

3. Example Applications

We now discuss the benefits of using CMP in the context of several emerging applications that fundamentally rely on cross-domain identity management.

Electronic health record management

An Electronic Health Record (EHR) is defined as the health record of an individual that is accessible online from many separate, interoperable automated systems within an electronic network. EHRs can contain a variety of data and can be used for different purposes by different parties involved in health care. The grand vision of EHR infrastructures is the interconnection and reusability of all recorded health information, regardless of where it is stored, so that all relevant health information can electronically flow to wherever it is needed.

Nothing will become of this vision, however, unless critical privacy and security problems are overcome. Studies reveal that most patients do not trust the administrators of national health services and other insiders in the health care system with the control over their personal health information. Often, their trust does not extend beyond their own care providers, and indeed the opportunities for privacy invasions due to secondary use of health record information are enormous. Organizations with a justified need (according to current widespread regulations) to access health information include government and private health plans, insurance companies, administrators, hospitals, doctors, pharmacies, employers, schools, researchers, data clearinghouses, accreditation and standard-setting organizations, laboratories, pharmaceutical companies, practice management system vendors, and billing agents.

Privacy is also sought by medical practitioners. Many doctors do not like the idea of central parties (such as health insurance organizations) being able to monitor all their actions, since they feel this negatively impacts their autonomy; in many situations, they would prefer to be able to access information on the basis of their role rather than their identity, and they certainly do not want identifiable digital evidence of all their interactions with patients to automatically flow to central parties. Role-based access is also preferred by medical researchers, for accessing online disease registers and other medical databases.

With CMP, an EHR is simply a Local or Remote record, or the logical combination of several such records. Attribute Authorities are health care professionals and possible other entities (including the patient himself) who add digitally authenticated statements to EHRs. EHRs can be securely managed by both the data subject and his health care professionals, in a manner that simultaneously protects the data subject's privacy interests, the professional's liability interests, and the legitimate interests of researchers and other third parties:

- Each patient can co-manage his health information together with selected physicians. A record can be managed electronically as one logical entity, even though different parts may reside in different physical locations. Each party with access rights can be assured that the data entries on which it relies have been entered by authorized parties, through either role-based or identity-based digital signatures. In this manner, health care service providers can effectively maintain partial ownership of a data subject's health information.
- By providing patients with tamper-resistant smartcards, health care providers can maintain even greater control over their own contributions to EHRs, since the cards can further limit the ability of patients and others to manipulate entries. Literally billions of authenticated EHR entries (possibly originating from different health professionals) can be securely managed using a single 8-bit smartcard. Cards can be issued to patients by a central entity that cannot compromise the legitimate privacy and security interests of patients and health care providers that ride along on the added security provided by the card.
- At the same time, patients as well as health professionals are able to selectively disclose authenticated health data in anonymous or pseudonymous form (with or without certifications). Patients can also delegate the right to do so to their doctors (e.g., to over-ride protections in emergency situations) or to third parties (e.g., for research purposes).

CMP in effect creates a continuum between health records maintained by health professionals and health records maintained by data subjects, seamlessly unifying the two approaches and covering the entire spectrum of possible rights management settings. In the CMP approach, the issue of where the health data resides hardly matters anymore; it is all about who has electronic access to which parts of a record.

E-government

E-government refers to the electronic delivery of government services to citizens. The primary objective is to simplify the interaction with citizens and institutions. In the past three years, many municipal, provincial, and federal governments around the world have established an on-line presence. Among the leading countries to bring government online are the United Kingdom, Canada, and the United States. Market analysts distinguish between five phases of e-government: (1) providing information via Web sites; (2) electronic service delivery; (3) improving operations through Web interfaces and electronic data exchanges; (4) moving toward more personalized electronic service delivery ("e-CRM"); and, (5) introducing Web-based collaborative technologies. Implementing CRM initiatives is widely considered a key priority to provide personalized citizen self-service.

In most cases, government organizations will need to be able to securely make authentication and authorization decisions about citizens who request electronic access to their services. Liberty Alliance is already being viewed with increasing interest by e-government architects. Indeed, the considerations of government for managing identity-related information in part match those of

industry. Governments however have a stronger interest in protecting the security and privacy of individuals and private sector organizations, and for good reasons. For instance, an August 2000 survey by Hart-Teeter about U.S. citizens' view of e-government services found that 53% of respondents were extremely concerned with the potential loss of privacy, and in a Gartner survey in 2001, nearly 70% of consumers cited privacy concerns as one reason that could make them stop using e-government services. As well, it would be most awkward for government not to live up to the spirit of its own data protection legislation, and ultimately the stability of democracy may be put on the line if a privacy-threatening infrastructure would be implemented.

On the security side, progress is being made in the right direction. Indeed, according to the Giga Information Group in June 2002, "in some technologies, like smartcards, biometrics and electronic records management, the government is ahead of business." Many governments are keen on access management systems based on smartcards, not only for citizens but also for its own employees and to replace driver's licenses, airport security documents, passports, and so on. On the privacy side, however, governments are struggling. Consumer outcry, trade group complaints, potential violation of privacy laws, and complaints by data protection commissioners have already lead to the suspension of several national PKI e-government initiatives.

Using CMP, it is easy to see how security, scalability, privacy, and general performance requirements can be reconciled. Consider the case of personalized access to on-line government services. A user would retrieve digital pseudonyms in batch from a central certificate issuer. The user would register a different pseudonym with each on-line service provider, which the service provider would link to its own program identifier for that user (following its own one-time authentication of the user's program-specific identity, or following an "introduction" by another organization). Due to the unlinkability of pseudonyms, government service providers do not gain cross-domain profiling powers that were not present in the legacy system. The certificate issuer can serve as a managed security services provider to government organizations, by providing some or all of the security features described previously. At the same time, the certificate issuer can be prevented from gaining any tracking and tracing powers, and can even be prevented from learning the identities of the certificate requestors as they retrieve pseudonyms. By embedding a unique "identifier" (e.g., a random number) into all of a user's pseudonyms, the certificate issuer can ensure that users can transfer certified personal information from one government organization to another without users being able to lend or pool personal information; in this manner, government organizations can reliably share user information without obtaining cross-domain profiling powers. As well, the certification of identity information by each government organization can be delegated to the certificate issuer without the latter being able to learn the information itself.

Digital rights management

Digital Rights Management (DRM) is generally defined as the collection of tools and technologies for protecting copyrights and other rights on digital media. DRM is an umbrella term: no single tool or technology suffices to guarantee access and content usage controls throughout a digital content distribution infrastructure.

DRM deals with authorization decisions about access to resources, and as such it is an application of access management. However, DRM places stronger requirements on fraud prevention than general access management. Namely, access management in general does not deal with long-lived access, while DRM also seeks to control usage by authorized users after they gain access to a resource.

All modern DRM systems have at their core the notion of a digital license, and most deal with content and licenses in a separate manner, along the following lines:

- Licenses are issued when access is requested, while content is made freely available in encrypted form to prevent access by unauthorized parties. To access protected content, the client must obtain a digital license that specifies how the content may be used.

- To consume protected content, the client connects to a clearing house and requests a digital license for the content. The request requires the client to send a unique identifier that identifies him and/or a specific client device that will play the content. The request is typically initiated by the client's software application or hardware device upon the client's first attempted access.
- Assuming the clearing house makes a favorable authorization decision for the client, it sends the requested digital license to the client. The client's device or application, which is presumed to be secure against tampering, then decrypts the license and displays or otherwise makes available the content to its user in accordance with the usage rules.

By separating content from licenses, content providers can issue one license for multiple sources of content, can issue different licenses for the same content, and can support business models that cleanly separate the interests of copyright holders, content distributors, service provider networks, and others. Of course, this basic DRM architecture inherits all the security, privacy, and performance problems of general cross-domain access management. Several authors have already noted the unique security and privacy benefits that Digital Credentials could bring to DRM; see, for instance, http://www.w3.org/2000/12/drm-ws/pp/hp-poorvi2.html.

By building DRM on top of CMP, consumers can control and limit the correlations that content distributors can establish about their consuming habits and identity, and content distributors can protect their intellectual properties more securely without infringing fair use rights. Let us walk through a simple CMP-based DRM scenario:

- Bob visits MusicPortal, an Internet portal where the latest album of his favorite band is available via download. MusicPortal groups content distributors together so that customers can buy their music from a single point-of-sale. The portal offers various subscription packages, such as monthly fees for unlimited downloads, prepaid number of downloads, and so on. It also offers Web site personalization and can make recommendations to Bob by keeping track of his musical preferences.
- Bob chooses to purchase a subscription which entitles him to limited number of download every month for a fixed monthly fee. He pays for the subscription with his credit card in a special section of the portal. To protect his privacy, his subscription is delivered in the form of a Digital Credential. This ensures that the subscription cannot be forged, while at the same time the portal will not be able to trace which credit card was used to buy the subscription.
- After making his music selection and the usage rights he wishes to acquire, Bob goes to the checkout section of the portal. To acquire the rights on a specific album Bob presents his subscription to the portal. The portal processes the payment through the clearing house and in exchange emits a digital license describing the rights and privileges associated with the music file. The music file is encrypted specifically for Bob using an encryption key that can be found in the digital license. The digital license is packaged into a Digital Credential as well, to provide lending protection and possibly other protections.
- To play the music, Bob needs a player that understands and enforces the license. Bob's player can permit him to copy the file from one player to another, so that he can play the file from many places. A lending disincentive placed in the licenses (as the credit card information that Bob used to purchase the subscription) would strongly discourage Bob from copying the music to his friends even if he could bypass the hardware protections of his player.
- For extra security, all subscriptions and digital licenses could be managed using a simple 8-bit smartcard, by off-loading all expensive computations and storage to the user's PC, a laptop, or PDA, while preserving all the smartcard's security benefits. Multiple license issuers could all ride along on the security of the same card, without needing to trust each other.

4. Closing Remarks

Currently, the visible battle over user identities is between organizations. The interests of individuals are not seriously taken into consideration by businesses; individuals can only rely on

government legislation and on themselves to reduce the power of organizations to profile them across domains. With ever-increasing advances in data storage, communication, processing, and analysis, both of these are rapidly losing their effectiveness. While some individuals may continue to provide organizations with even more polluted information, others may avoid them altogether. This power struggle between organizations and individuals is in nobody's best interests.

The key to getting everyone on the same side of the table is to adopt identity and access management technologies that give the owners of identity information direct control over how their profile information can be used by others. The notion of ownership of identity information is not always easy to define and capture, however; while often the data subject must be considered to be the legitimate owner of personal information, there are numerous instances where ownership legitimately resides in the hands of one or more organizations, possibly jointly with the data subject. The CMP architecture proposed in this paper has been designed to take this into account; it allows identity-related information to be securely co-managed by both data subjects and organizations, in a manner that simultaneously protects the privacy interests of data subjects and the business and liability interests of organizations.

5. References

[1] "An In-Depth Analysis of the Liberty Alliance Architecture," Credentica whitepaper, April 2004. http://www.credentica.com/technology/LA.pdf

[2] "A Technical Overview of Digital Credentials," by S. Brands, International Journal on Information Security, 2004 (to appear). http://www.credentica.com/technology/overview.pdf

[3] "Rethinking Public Key Infrastructures and Digital Certificates; Building in Privacy," by S. Brands, 315 pages, August 2000, MIT Press, ISBN 0-262-02491-8. With a foreword by prof. Ronald. L. Rivest. http://www.credentica.com/technology/book.html