INFRASTRUCTURE SYSTEM DESIGN METHODOLOGY FOR SMART ID CARDS DEPLOYMENT

Ramaswamy Chandramouli
National Institute of Standards & Technology, USA
mouli@nist.gov

ABSTRACT

With the increasing use of smart cards for identity verification of individuals, it has become imperative for organizations to properly design and engineer the expensive infrastructure system that supports smart card deployment. Apart from sound system design principles, this class of system (which we denote by the abbreviation IS-SCD – standing for Infrastructure System for Smart Card Deployment) requires rich body of domain knowledge in the areas of identity information, generation, storage , distribution and usage of different sets of credentials, the methods of capturing of credentials etc. The design methodology we describe for IS-SCD in this paper has as its basis a business process definition framework built using two credentialing specifications used for large scale smart card deployments. In addition, the methodology uses both top-down and bottom-up strategies for gathering all data requirements and employs a good blend of process and data analysis towards the goal of defining a final system architecture for IS-SCD.

KEYWORDS

Identity Management, Credentials, Smart Card

1. INTRODUCTION

Smart cards are now being increasingly employed as identity tokens and used to support authentication in several corporate applications. However, many of the infrastructure systems for smart card deployment (which we denote with the abbreviation IS-SCD) till date have been designed in an adhoc manner. Even with the presence of commercial product offerings such as Identity Management Systems (IDMS) for engineering such infrastructure systems, faulty and insufficient process analysis in the context of the organization's deployment scenario could result in a system architecture that is well over budget, fails to meet security goals and is difficult to scale. In this paper we present a design methodology for IS-SCD that is based on sound system design principles supported by a domain-specific business process definition framework. Our design methodology combines the best of top-down and bottom-up approaches with the goal of arriving at an overall architecture for IS-SCD. The design goal of the IS-SCD is to provide a high fidelity identity verification framework that meets the necessary trust requirements [Josang-2005] and will support the authentication needs of a number of systems in the enterprise (called as relying systems in this paper). These relying systems perform some key operations such as providing physical access to facilities and logical access to IT systems.

The overall organization of our paper is as follows: As with any system design methodology, we begin with the task of defining the various business processes that IS-SCD has to support. To achieve this we developed a business process definition framework based on domain knowledge provided by well established credentialing specifications. This is the focus of chapter 1. In chapter 2 we go on to identify the data flows stemming from instantiations of these business processes. Analyzing these data flows we identify the need for additional processes driven by the dynamics of these data flows. These additional processes are the focus of chapter 4. We then identify the system components that will enable these data flows in chapter 5. Detailed data content for each of the data flows is developed in chapter 6 and the collection of these data contents results in the development of the global data schema for IS-SCD. The requirements for this global schema are outlined in chapter 7. Last but not the least, the needs of interface components that will facilitate these data

flows are outlined in chapter 8 to complete the overall design methodology for IS-SCD. The summary and conclusions are given in chapter 9.

2. IDENTIFICATION AND DEFINITION OF BUSINESS PROCESSES (STEP 1)

The first step in the design of any large scale multi-component system such as IS-SCD is the identification and definition of business processes that the system has to support by adopting a business process definition framework. Our development of this framework consisted of abstracting the security principles (or control objectives) and functional description found in two recently published U.S government personal identity credentialing specifications [FIPS201-2006, TWIC-2007]. The reason for choosing these specifications is that they have spanned large smart card issuance programs that will eventually provide identity verification for over 10 million individuals spanning the civilian sectors of US government and maritime workers.

Our business process definition framework for IS-SCD consists of the following steps:

- Extraction of the core security principles from the two personal identity credentialing specifications referred above.
- By logically grouping the functions related to creation of personal identifiers, identity proofing, credential generation and distribution in these two credentialing specifications, a list of business processes can be developed and after incorporation of security principles can form the business process definitions in the context of our IS-SCD.

The core security principles we extracted from the two personal identity credentialing specifications are:

- Every Request for a smart card should be properly scrutinized by a responsible official of the organization who then should provide a complete set of initial identity information pertaining to the applicant along with information about himself/herself – the latter for auditing purpose (SP1)
- The identity of an individual is adequately vetted before he/she is issued credentials (SP2)
- Credentials are Tamper proof (SP3)
- Credentials are issued by the right authority (SP4)
- All credential and associated personal identity data must be stored and transmitted securely (SP5).
- The association between credentials and the card stock where it is carried is correctly tracked (SP6)
- The credential set stored in the card should be capable of supporting multiple authentication modes used by relying systems (SP7).

As stated above, the next step is the logical grouping of functions in the two credentialing specifications to generate the list of business processes. The input-output analysis of a business process along with applicable security principles results in a complete logical definition of the business process in the IS-SCD context.

- Sponsorship – This is the process that creates the initial identity information about the card applicant along with some basic information about the sponsor who is recommending the applicant for a smart card. The complete set of information generated by this process should consist of all the needed biographic information as well as the organizational affiliation information in order to satisfy the security principle (SP1).
- Identity Proofing – This process is a direct consequence of security principle SP2 and in practice is realized through verification of the authenticity of certain identity-vetting documents (e.g., birth certificate, passport etc) associated with an individual and then keeping the electronic images of those documents generated by scanning.
- Enrollment – This is the process that generates the credential. The credential can be a unique identifier generated by the enrollment system and/or can be one that is obtained from the card applicant (e.g., fingerprint image/minutiae). This process should be allowed to be executed only by authorized enrollers (to satisfy security principle SP4) and the generated/collected credentials should be encrypted, digitally signed and transmitted over a secure communication channel (to satisfy SP3). The end product of this process is the credential data set.
- Card Production – This is a process that prints on a smart card all the visual information consisting of a portion of biographic information (mostly just the name), organizational affiliation information

(employee, organizational unit, rank etc) and a transformed version of a credential (e.g., printed version of a digital photographic file). This is a process that does not generate any new data but merely uses a portion of data generated by sponsorship and enrollment processes. Hence it needs to incorporate just the generic security principle (SP5) which respect to storage and transmission (from the database to the card printer in this case) of credential and associated personal identity data.

- Card Issuance and Activation – This is a process that writes electronically credential information from many sources (enrollment output, PKI certificate from a Certificate Authority, Private key, PIN etc) to the card chip. Security principles such as SP5 (secure storage and transmission), SP3 (tamper proof) and SP 4 (credential from trusted authority) apply to this process.
- Identity Provisioning - This process transmits personal identity information and credentials to relying systems to perform the following: (a) Validation of the credentials read from the presented card and (b) Associate the credentials with identity information associated with the individual who is ultimately the target of authentication. The only security principle this process should incorporate is SP5 – secure transmission.

3. IDENTIFICATION OF DATA FLOWS (STEP 2)

Our analysis of the business processes defined in the previous section reveals that one set of processes generate data (i.e., sponsorship, identity proofing and enrollment processes) while the other set consumes the data (i.e., card production, card issuance/activation and identity provisioning processes). Let us call the collection of data that is either generated or consumed by a process as a data flow. Let us also categorize each data flow as an in-flow if it's source is a data-generating process and as an out-flow if it's target (or sink) is a data-consuming process. Based on this logic, our data flows and the associated category of each are:

- Sponsorship Data Flow (in-flow) (SPF)
- Identity Proofing Data Flow (in-flow) (IPF)
- Enrollment Data Flow (in-flow) (ENF)
- Card Production Data Flow (out-flow) (CPF)
- Card Issuance/Activation Data Flow (out-flow) (CIF)
- Relying System Data Flow(s) (out-flow) (RSF)

4. IDENTIFICATION OF STAGING PROCESS (STEP 3)

From the list of data flows shown above, we see that there are three in-flows and three or more out-flows (depending upon the number of relying systems to which the Identity provisioning process has to be established). It is also clear from the analysis of process definitions from section 2, that there is no 1:1 correspondence in data content between an in-flow and an out-flow. If that were the case, the system that generates an in-flow can directly feed into a system that consumes an out-flow. As an example, consider the card production data flow. This data flow consists of: (a) a subset of biographic information (b) a subset of organizational affiliation information and (c) possibly a single credential. While (a) and (b) come from Sponsorship Data Flow, (c) comes from Enrollment Data flow. Because of the lack of flow-level mappings between in-flows and out-flows, we need a staging process that acts as a sink for all in-flow data and as a source for all out-flow data.

4.1 Process Definition for Staging Process

In order to meet its design goal of acting as a buffer for capturing all data from in-flows to facilitate generation of various out-flows, the staging process should provide some artifacts (e.g., a global data schema). In addition, it has to provide some services relating to overall integrity of the multiple data flows and also enable complete credential lifecycle management. These artifacts and services that constitute the total process definition for the staging process given below:

- It should provide a global data schema for capturing all data from multiple in-flow data flows

- It should have a means of linking up information from different in-flow data flows. An example is the ability to link up an identity proofing information coming in for a person to the existing sponsorship information already stored in the system.
- It should support feedback data flows (and fields to capture the feedback information) from the targets of out-flow data flows for the purpose of credential lifecycle management. An example of a feedback data flow is the message from the card production process stating that cards have been printed for a specific batch of credential and personal identity data set that the staging process has sent out. To keep track of this feedback information, a flag field needs to be supported by the staging process. Additional examples of fields needed for credential lifecycle management (as well as security principle SP6) are the ones to keep track of the status of the credential (revoked, suspended, active), status of the issued card (lost, stolen, damaged), establish the association between the credential set for an individual and the physical device where the credential resides (e.g., card serial number written into a card chip) etc.
- It should contain the program logic to enforce precedence relationship between data flows and thus their associated processes. For example, the enrollment data flow (that contains the credentials) can only be uploaded into the staging platform only after the identity proofing data flow is complete.

5. IDENTIFICATION OF SYSTEM COMPONENTS (STEP 4)

Reviewing what we have done so far in our IS-SCD system design methodology, we find that we defined six primary business processes and derived six data flows based on the input-output analysis of those processes. We also derived an additional process called staging process while designing a conceptual architecture that will enable the various data flows generated by primary business processes. We thus have 6 primary business processes and an architecture-driven process. In spite of the different drivers, we will consider all the primary business processes on the one hand and the staging process on another to be semantically equivalent from the point of view of IS-SCD system design methodology.

Having analyzed the nature of the data flows and the having defined the seven processes, we will now consider them together to identify the system components of IS-SCD. Let us now take up the case of the staging process and its associated data flows. Reviewing its process definition from the last section, it is clear that a system component that supports this process should have features for defining a global schema that can capture all types of credentials and identity related data, support secure storage, access and validation of this data and secure transmission of subsets of data to different external systems. In addition it should provide ubiquitous access through web interfaces (accessible through web portals being built on a service oriented architecture) and should preferably support an access control framework that provides support for the concept of roles for grouping together all authorizations needed for a particular process. A class of systems that provides these features is the Identity Management System or IDMS.

Let us now look at other processes and associated data flows to identify system components that will meet those process requirements for the design of IS-SCD architecture.

- The sponsorship process being the generator of biographic information, organizational affiliation information and sponsor identifier information can only be supported using data from a Human Resource (HR) system of the enterprise.
- Identity Proofing process has to be supported by a customized application that can verify the authenticity of vetting documents such as driver's licenses, birth certificates and passports and generate scanned images of these documents.
- Enrollment (of credentials) process has to be supported by a customized application which makes use of various devices for collecting different types of credentials such as a digital camera for electronic facial images, fingerprint scanners and minutiae generators for capturing biometric information etc.
- Card Production process has to be supported by a specialized card production system that is capable of performing graphical printing on plastic cards with an integrated circuit chip (ICC).
- Card Issuance/Activation process has to be supported by a special system called Card Management system that is capable of establishing a secure communication session with the card and electronically populating the card chip.

The schematic diagram of all system components in IS-SCD, the data flows that are either generated or consumed by them are shown in Figure 1.

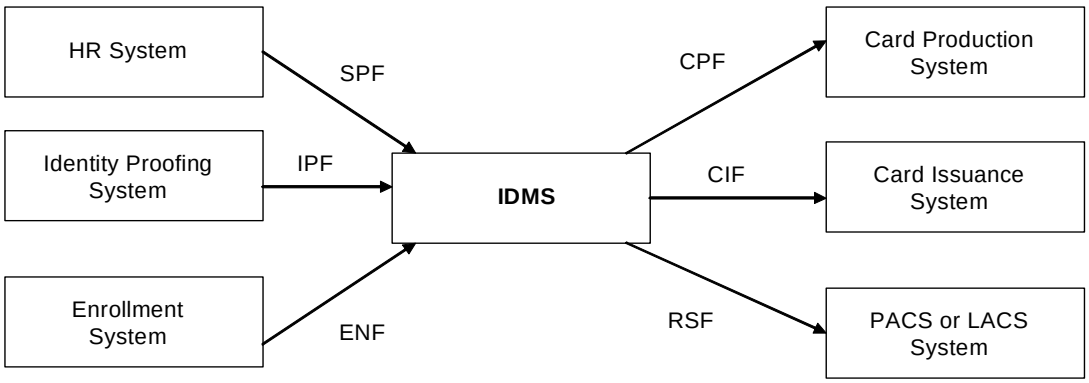

Figure 1. IS-SCD Components and Data Flows

6. IDENTIFICATION OF ALL DATA REQUIREMENTS (STEP 5)

Having identified the system components that can generate all in-flow data flows and consume all out-flow data flows including the one (i.e., IDMS) that acts as a repository or sink for all in-flow data flows and origin or source for all out-flow data flows, we are now ready to develop the data content for each of these data flows.

6.1 Data Content for Sponsorship Data Flow

It is important to remember that credentials are generated and provisioned based upon properly vetting the identity of the prospective card holder (or card applicant). Before this identity vetting or identity proofing, there should be enough information generated about the person (biographic information) and the nature of his/her affiliation with the organization. Further the eventual card holder (now the card applicant) should have been duly sponsored by an authorized official of the organization and hence some identification details regarding the sponsor should be an integral part of the sponsorship data flow. Hence the data content of this data flow will consists of:

- Biographic Information (Full Name, SSN, DOB, Address etc)
- Organizational Affiliation Information (Type of Employee – Regular or Contractor, Rank, Date of Joining, Business Email etc)
- Sponsor Identifier Information (The Unique ID and Name of the Sponsor, the organizational affiliation of the sponsor, date of sponsorship etc)

6.2 Data Content for Identity Proofing Data Flow

Identity proofing of an individual is usually done based on government issued documents that the individual possess since these documents carry a great degree of trust. Examples of such documents include birth certificate, driver's license and passport. The IS-SCD should carry the electronic form of these identity proofing documents as evidence that his/her identity has been vetted prior to issuing the card and hence the data content of this flow is made up of:

- Scanned Images of Identity Proofing documents

6.3 Data Content for Enrollment Data Flow

To determine the total set of credential data that needs to be stored in the smart card (and thus satisfy security principle SP7), we adopted the equivalent of backward chaining approach in logic. We looked at all potential corporate operations whose security can be improved using smart card-based authentication. The two broad

categories of this operation are secure access to physical facilities and secure access to corporate IT systems called respectively as:

- Physical Access Control Systems (PACS)
- Logical Access Control Systems (LACS)

Specifically, the authentication component of PACS and LACS applications make use of the identity credentials in the smart card to authenticate users who access these systems. We will call these systems that perform smart card-based authentication as *relying systems*. When we refer to these relying systems, we are only implicitly referring to the authentication component of these systems.

Relying systems, depending upon their criticality will make use of different set of credentials and consequently different methods for validating those set of credentials. These different methods are called smart card use cases (SC-U). For example, a PACS system (relying system) that provides access to a relatively non-critical facility may authorize entrants (open the door or turnstile) merely based on reading a unique identifier (that is associated with an authorized person such as an employee or contractor). On the other hand, a PACS system at the entrance of a secure facility such as a data center may in addition require the card holder to enter a PIN (what the person KNOWS) or present a finger print (what the person IS). We see that the first of the use case described above uses only one credential (i.e., the Unique Identifier) while the second use case uses two credentials (Unique Identifier + Fingerprint Biometric/PIN).

Hence to assess the total set of credentials that must be handled by IS-SCD, and therefore has to be present on the card, the organization has to decide on the set of relying systems and the authentication methods that will be used by those relying systems. The consolidation of authentication methods will in turn determine the total set of smart card use cases and confidence level for each use case. In general, a use case with a higher confidence level will use more credential data items than a use case with a lower confidence level and the credential validation process of the former is also more complex than the latter. The number of smart card use cases, the associated confidence levels etc will vary with the security profile of the relying systems in the organization. However in many deployment scenarios, the following are the common set of use cases and associated confidence levels.

1. Use case SC-U1: User Identified based on the Unique Identifier read from the card – LOW CONFIDENCE
2. Use case SC-U2: Use case SC-U1 augmented with successful verification of a digital signature associated with the Unique Identifier – SOME CONFIDENCE
3. Use case SC-U3: Use case SC-U2 augmented by verifying that the card possess a tamper-proof secret by performing a cryptographic challenge-response with the card which may require the card holder to provide a PIN (user showing proof of knowledge of a secret in order for the card to reveal another secret – in this case a private cryptographic key) – MEDIUM CONFIDENCE
4. Use case SC-U4: Use case SC-U3 augmented with verification of the binding between the card holder and the card by matching the live fingerprint provided by the card user with the digitally signed biometric data present on the card – HIGH CONFIDENCE

The rationale for increasing confidence levels assigned to use cases SC-U1 through SC-U4 is not only due to the fact that each successive use case subsumes the previous one by adding an incremental validation process but also due to the increasing trust requirement each successive one meets. In order to see this, we have to look at the trust requirements involved in smart card-based authentication processes. They are given below:

- The card contains the valid Unique Identifier created by the organization and associated with a legitimate affiliate of the organization (T1)
- The unique identifier or number present in the card has been put in by an authorized issuing authority and that it has not been altered since then (T2)
- The card that is presented is the card in which the credential was loaded by the authorized issuer and that the credential has not been duplicated in a cloned card (T3)
- The card itself has not been stolen from the legitimate holder and presented to the identity verification system by the impersonator or the card thief. In other words, the person presenting the card is the same person to whom the card was issued (T4)

From the description of the trust requirements it is clear that use case SC-U1 provides only trust requirement T1, SC-U2 provides T1 and T2, SC-U3 provides T1, T2 and T3 while the last use case SC-U4, the one with the highest confidence level provides all the 4 trust requirements (T1, T2, T3 and T4). Recall

that we stated that the total set of use cases compiled for the deployment scenario determines the total set of credentials needed on the card. For a card that needs to support use case SC-U4, therefore, needs the following credentials: (a) Card-based Unique Identifier associated with the person (b) Digital Identity token (PKI certificate) issued to the card holder by an authorized certificate issuing authority (CA) and the Card-based Unique Identifier digitally signed with a secret (called private) key associated with the token (c) Presence of the secret associated with the digital identity token (referred above) being stored in a tamper-proof way in the card (in other words the secret cannot be obtained without destroying the card itself) (d) Card Holder PIN and (e) Card Holder Biometric data.

6.4 Data Content for Card Production Data Flow

This is an out-flow data flow that feeds into card production systems and carries all printed information (as opposed to electronic information that is written to the card chip). Hence it contains:

- A rudimentary subset of biographic information (Full Name)
- A subset of organizational affiliation information (Rank, Organizational Unit)
- A transformed version of a credential (printed version of a digital facial image)

6.5 Data Content for Card Issuance/Activation Data Flow

This is again an out-flow data flow uploaded to card issuance/activation systems for populating the card chip with electronic credentials. Hence it is mainly made up of:

- Electronic Credentials such as digital facial image, biometric fingerprints, a unique identifier associated with the individual all from Enrollment data flow.
- The card issuance or card management system as it is called may acquire additional credentials such as PKI certificates directly from an authorized certificate issuance authority (CA) and the cryptographic programs in the card itself may generate data such as the private cryptographic key and digital signature of the unique identifier.

6.6 Data Content for Relying System Data Flows

This consists of one or more out-flow data flows to relying systems such as PACS or LACS. For PACS, the flow may consist of just the Full Name and the Unique Identifier from sponsorship data flow and staging process respectively. In some cases it may contain organizational affiliation information from sponsorship data flow. For LACS, the data flow may consist of Full Name from sponsorship data flow and a UPN number (generated by the staging process).

7. REQUIREMENTS FOR IS-SCD GLOBAL DATA SCHEMA (STEP 6)

The IS-SCD global data schema should meet the following requirements:

- It should be capable of capturing all data items coming into IDMS in all in-flow data flows
- It should contain additional items such as a system generated unique identifier for an individual to properly link up the information pertaining to that individual in various data flows (sponsorship, enrollment etc) so as to maintain the overall data integrity.
- It should contain additional fields (or at least have an extensible schema) that are needed for supporting certain classes of relying systems (e.g., a UPN field for supporting logical access application (LACS) through an enterprise directory)
- Should contain data items that can facilitate credential lifecycle management. These include fields that carry: (a) the status of the credential uploaded for card production & card issuance, (b) the status of the card that has been issued (lost, stolen, damaged etc), (c) the status of the individual who has been issued the card (fired, resigned, suspended etc) and (d) the status of the credential itself (expired, revoked etc).

8. DESIGN OF INTERFACE COMPONENTS TO FACILITATE DATA FLOWS

The system components (step 5), data content (step 6) , the global data schema (step 7) along with volume and granularity of data flows and integration capabilities of the IS-SCD system components are the drivers for the interface design.

- On looking at the in-flow data flows and the systems that generate them we find that sponsorship information flow comes from HR systems, while identity proofing data flow and enrollment data flows originate from identity proofing and enrollment applications respectively. Out of these 3 systems in question, the HR system in many enterprises will be a legacy system. Hence direct integration of an HR system and IDMS is not possible. The direct integration of identity proofing and enrollment applications is also ruled out since there is no standardized API for IDMS systems. Hence we find that for enabling all data flows into IDMS, we need to develop web portal or web service interface.
- To enter some data (e.g., UPN (for provisioning of credentials to LACS systems), missing biographic data etc) directly into IDMS, it should provide a web interface which also helps to view feedback data coming from targets of provisioning such as card production, card issuance, PACS and LACS systems.
- Web portals and web services interfaces are also needed to extract data from IS-SCD global data schema (stored as the database in IDMS) and uploaded to provisioning targets for relying systems.

9. SUMMARY & CONCLUSIONS

In summary, the IS-SCD design and engineering methodology consists of the following steps:

- Identification and definition of business processes that should be supported by IS-SCD based on two credentialing specifications used for large scale deployment.
- Identification of various data flows and nature of information handled from business process definitions.
- Identification of an additional process based on analyzing the dynamics of data flows
- Identification of system components that will generate or consume the data flows
- Identification of the detailed data content in each of the data flows
- Identifying the requirements for a global data schema that will capture all the information in all in-flow data flows, provide data integrity and support credential lifecycle management
- Design of interface components that will facilitate the secure transmission of data flows between the identified components of IS-SCD.

From the description of the steps it should be clear that the IS-SCD design methodology uses both top-down and bottom-up strategies for gathering all data requirements and employs a good blend of process and data analysis for defining the system components and their interface functions. This hybrid approach together with the initial step of defining business processes based on security principles and functional requirements in two government smart card specifications accepted for large scale deployments, results in a design methodology that has the potential to generate an infrastructure system architecture for smart card deployment that is both scalable and secure.

REFERENCES

FIPS 201-1, Personal Identity Verification of Federal Employees and Contractors,

http://csrc.nist.gov/publications/**fips/fips201-1/FIPS-201-1**-chng1.pdf, March 2006.

Josang, A., 2005, Trust Requirements in Identity Management, *Proceedings of the Australasian Information Security Workshop*,

Newcastle, Australia, pp 99-108.

TWIC, TWIC Reader Hardware and Card Application Specification,

www.tsa.gov/assets/pdf/**twic_reader**_card_app_spec__09-11_07.pdf, September 2007.