SAP NetWeaver® Identity Management
White Paper

Implementing an Identity Management Solution

August 2007
Table of contents

Executive summary .................................................................................................... 4
Background ................................................................................................................ 5
The identity management architecture....................................................................... 6
The implementation process ...................................................................................... 7
  Step 1: Data services.......................................................................................................... 7
  Step 2: Identity services ..................................................................................................... 11
  Step 3: Identity applications .................................................................................................. 14

SAP NetWeaver Identity Management........................................................................... 16
Executive summary

Enterprises are under pressure to increase the speed of deploying new applications and systems across their global networks, both internally and in the context of e-Business with partners and customers. One of the challenges to achieving this is the difficulty in finding and bringing together information relating to identities and resources, which are distributed across multiple and often incompatible information sources. This is especially critical during corporate migrations, where a multitude of data repositories in different formats and with a variety of interfaces is in use.

The prime objective of identity management is to centrally manage and keep up to date all identity data within the enterprise. Identity data is often stored in many different applications throughout the enterprise, and maintained manually in different locations. This is cost-ineffective and poses a security risk. It can cause inconsistency and low data quality.

This white paper gives an overview of a proposed layered identity management architecture. It then goes on to describing the complex process of implementing an identity management solution; starting from a situation that is common in many organizations, where the identity data is not centralized, but stored in a multitude of applications. It also includes issues to consider when starting to join identities across the repositories before building functions and services like provisioning, authentication & authorization, workflow and single sign-on.

SAP NetWeaver Identity Management provides a robust product range that can be used for a wide range of functions and services within an identity management solution. The products have proven their quality over years in large organizations, where a high degree of security, scalability, availability, persistence and performance is a business-critical requirement.
Background

Identity Management is a challenge for most organizations today: The larger the organization, the greater the challenge. The user must present his or her identity to get access to the many ICT (Information and Communication Technology) applications within the organization. Examples include the various operating systems, the HR (Human Resources) system, CRM (Customer Relationship Management) systems, databases, directories, physical access control systems, e-mail systems and support systems.

The problem is that every one of the existing applications within the organization usually has its own identity management system: they all store information about the users accessing the application, such as authentication information (the user id and password needed for access to the application) and authorization information (the required access levels to the application data). Standardized ways of storing or managing this sort of identity information centrally, for different applications that are normally not connected, are still fairly rare.

As most applications are unaware of each other, the identity data must be maintained more or less manually for every application. Every time a new employee joins the organization, he or she must be manually added to every required application. The same is true when an employee changes their position (or role) within the organization. Access to new applications is required, as well as changing (or revoking) the access rights to existing applications.

The main problems with this are:

- **High maintenance cost**
  A system with many manual operations will have a high maintenance cost. For every new employee or employee moving within the organization, many manual interventions are usually involved. This demands resources that could be put to much better use. In addition, these manual steps are time-consuming, and the employee can do nothing but wait for the tasks to be completed.

- **Complexity**
  The complexity of the system is high. It is difficult, or even impossible, to get an overview of all employees, to find the correct information about them and what access rights they have to the various applications within the organization. This complexity will also lead to errors when entering the information, like duplicate entries for the same person.

- **Security risks**
  In addition to this, there are security risks – the major one being that employees leaving the organization may still have access to certain applications, if they have to be manually and individually revoked. There should be an Identity Management System preventing this situation. Even people moving within the organization may present a risk: they will normally be granted the access rights of the new position within the organization, but the previous access rights may not be revoked. Another security risk is that since there are manual procedures involved, human errors may cause security flaws.

New regulatory requirements, such as the Sarbanes-Oxley Act, place high demands on the organizations regarding reporting of authorization information. Failing to provide a complete set of documentation or uncovering the existence of accounts for ex-employees is considered a “significant deficiency”, and may result in heavy penalties.
The identity management architecture

Our understanding of identity management is based on the following definition\(^1\):

"A set of processes, and a supporting infrastructure, for the creation, maintenance, and use of digital identities".

The identity management area consists of many different services or components, like provisioning, directory services, authentication, authorization or data access. The SAP NetWeaver Identity Management provides products that cover parts of this area.

The components of identity management can also be illustrated by the following model\(^2\):

\(\text{Figure 1 Layers of the identity management architecture}\)

**Applications & repositories:**
All existing ICT\(^3\) infrastructure in an organization, including the data repositories and the applications/interfaces that are used to access them. These may be business applications of various kinds containing customer and product information; specific applications maintaining identity data, such as human resources applications; or applications used to maintain other types of information, such as document management systems.

**Data Services:**
The Data Services layer builds a uniform, normalized, integrated view of the Applications & Repositories layer. This is achieved through services/functions such as synchronizing, joining and publishing data and providing access to data. The Identity Store is a core component of the Data Services layer. It is used to gather information about all identities throughout all applications in the organization.

**Identity Services:**
This layer consists of the services that are offered on the bases of the Data Services layer. These include provisioning, authentication, authorization and virtualization.

**Identity Applications:**
This layer consists of all applications using the Identity Services. This may be the identity-management components of existing applications, or new functions such as workflow, single sign-on and self-services.

---

\(^1\) As defined in the Burton Group's presentation "Enterprise Management (IdM): The Strategic Infrastructure Imperative".

\(^2\) Based on the Burton Group's report "Beyond Application Integration: Concepts and Functions of Enterprise Data Networking".

\(^3\) Information and Communication Technology
The implementation process

In addition to representing layers in the architecture, the figure also illustrates the process of implementing an identity management solution. This process is described in the following sections.

Step 1: Data services

The first step is implementing the Data services layer. This will normally include extracting, normalizing, synchronizing, joining, transforming and publishing the identity data contained in the existing repositories and create a uniform view of the data (a identity store). The result of this process can then be used in the business applications, for instance for white pages and authentication.

Figure 2 The applications and repositories hold the disparate identity data

The initial situation is often that there are many applications, each maintaining their own user accounts or other identity data. There is no connection between these applications and they do not share any identity information. Also, there is a very high probability that the interface with the identity information is different for the various applications. This is illustrated by the different shapes of the "interface" between the two layers in the figure below.

Figure 3 The Data services layer

When the Data services are implemented, the identity data is collected from the various applications and joined to create unique entries for each employee (or other identity). The synchronized information can then be updated in the identity repository of each application, increasing the data quality of this information considerably. The same employee or customer can be identified across all applications.
The Identity store

One of the core components in the Data services layer is the Identity store. This identity store will hold identity information about all identities within the organization. The long-term goal is that the applications will not need to maintain their own identity information, but automatically synchronize this with the identity store.

Figure 4 The Identity store

This puts certain demands on the identity store:

- **Complete**
  All information must be present. The application or user connecting to the identity store must trust this to hold all information. If information is not found, it must be assumed that it does not exist.

- **Correct**
  The information within it must be accurate. Since the information is used for mission-critical operations, such as access control, it is important that the information in the identity store is correct.

- **Available**
  The information must be available whenever it is requested by the client. The applications will depend on the identity store, and its availability must therefore be high.

- **Relevant**
  Different clients will have a different view of the information, and each client must receive the information that is relevant for it. Different applications may want a different view of the identities. For example, the accounting department has a different view on the employees than the IT department has.

These requirements are fulfilled using different technologies. Completeness and correctness are achieved using data integration/synchronization and reconciliation. A high-availability environment gives the necessary availability. A virtual directory ensures relevance of the data to the various user groups and applications.
The process
The process of implementing an identity store includes a number of steps and decisions. Some of them are:

- Analyzing data quality.
- Joining the data.
- Providing data access.

Analyzing data quality
The initial task is to analyze the existing data sources to determine the quality of the attributes within the data sources. This analysis normally discovers that data is duplicated in many data repositories and also that the quality of data differs from one data source to another. The goal of the quality analysis is to decide which data source is authoritative on which attributes. As an example: Many repositories will have the employee telephone number, but in many of them this number was entered when the employee joined the organization, and may not have been updated since. The telephony system may be the only application which can be trusted to have the correct phone number.

Joining the data
One important requirement is the ability to combine attributes from different data sources into one object (or entry), known as the join operation. Rules must be defined for which attributes are included from which repositories and which key(s) to use to join the attributes. The identity store is the heart of this process and will (when the system is set into production) contain complete information for each entry.

The main problem with joining is that a common identifier across all systems does not exist. Some systems may use a unique identifier (for example the employee number) but in many cases there is no such identifier, which means that other means of joining must be defined. The employee name may be used, but this is often a challenge for several reasons: Different spellings, different punctuations, several people with the same name, different use of national characters, etc.

To solve this problem, a common identifier is generated within the identity store. Various mechanisms and algorithms are used to join data from the different repositories based on the available attributes.

Another important task is to determine ownership and future maintenance of entries and attributes. There are always security and political issues that need to be solved regarding these questions.

Providing data access
The result of the process must be made available to users and applications. This includes defining where to publish the data, how it can be accessed and who should be allowed access to the data elements.

The contents of the identity store can be published either in full or partially, for instance to a directory server to which applications and clients can connect and find the necessary information. It is also possible to update the identity repository of the business applications using synchronization to make the data available to them.
**The virtual directory**

In addition to creating an identity store, a virtual directory may provide a single point of access to a number of data sources and also take care of access control among other tasks. The virtual directory itself has no data storage, but contains knowledge about the location and format of data stored elsewhere. Sometimes the source data may be spread over different systems and platforms. In effect it transforms the incoming LDAP requests either to LDAP requests to other directory servers or to SQL statements sent to one or several relational databases.

The virtual directory also contains rules on how information should be displayed for a particular class of users. This can be used to give different users a different view of the same data.

An example is the use of the virtual directory as an LDAP proxy to give different users a different view of the information, based on credentials, or the location from which the user connects. So an external user may see a switchboard telephone number, rather than the real direct dial extension.

![Figure 5 The virtual directory](image)

© Copyright 2007 SAP AG. All rights reserved.
Step 2: Identity services

Identity services are offered based on the services on the Data services layer. These services can be of various kinds, including:

- Provisioning
- Authorization Authentication Services
- Audit
- Virtualization

Provisioning

As already mentioned, one of the main problems of identity management is the manual maintenance of the identity data of most applications. A provisioning system will automate this process by defining users in the applications based on a central repository (the identity store). Normally the authoritative source to determine whether a person should exist in the identity store is the HR system, as this is also responsible for authorizing salary payment for the employee. Often the HR system also contains information such as the e-mail address and phone number, but this information is in many cases not maintained after the first registration of the employee. The e-mail system must know the e-mail address of the employee to deliver the e-mail and is thus the authoritative source of e-mail addresses. If an electronic telephone switchboard is deployed, this may be queried for the telephone numbers. Customers may be provisioned from a completely different set of business applications.

The provisioning solution will, based on information from the authoritative systems, automatically create the user in the required repositories, and provide the necessary access rights. And even more important, when a user leaves the organization the accounts are disabled, and access rights are revoked. Similarly when a user moves between departments, access rights are granted and revoked accordingly.

It will now also be easier to have an overview of all the employees within the organization, and of the applications in which the employees are defined.

Provisioning can be used for handling much more than employees joining and leaving the organization. We already mentioned customer self-service. In a project-based organization, the provisioning system can be engineered to create the necessary accounts and project web space automatically when a new project is started. It can also make sure the project members have access to the project archive, and create e-mail lists of the people involved. It can be used to start a workflow that equips traveling salespersons with a mobile phone, a laptop computer, and remote access to the company network and web resources. Some of these provisioning tasks are in themselves workflows with ordered and dependent sub-tasks. If the workflow fails, it may be necessary to roll back tasks that had previously been completed successfully. This requires a high degree of sophistication in the chosen provisioning solution.
SAP NetWeaver Identity Management’s provisioning solution is:

- **Persistent**
  Created objects and variables continue to exist and retain their values between runs of the program. If a remote system fails to respond, the provisioning will perform a pre-defined number of re-tries over a defined period before reporting an error.

- **State driven**
  In provisioning, states are often modeled as being discrete (rather than continuous) and the transition from one state to another is considered to be instantaneous. A simple example is a user, which is either enabled or disabled. Multiple states can be defined, triggering different follow-up actions.

- **Auditable**
  The Identity Cente Provisioning system keeps a full record of which process changed what, who authorized the change and the ‘before’ and ‘after’ states. This is important to comply to regulatory requirements, such as the Sarbanes-Oxley Act and others.

![Figure 7 The provisioning task can be dependent on a user profile and have sub-tasks](image-url)
Authentication & authorization
At an early stage of identity management, the provisioning system will create accounts within the various applications within the organization. However, the goal is to centralize this functionality within the Identity Services layer. A central service, based on the Identity Store, will manage both authentication (determining the identity of the user or application) and authorization (verifying that the user or application is allowed to perform the requested task).

Currently, each application is responsible for implementing the mechanisms for verifying the user identity and based on this, which operations are allowed. Using identity services, the applications will trust the Authentication and Authorization Services (AAS) both to verify the identity and to handle the access rights. The application itself need not have any knowledge of the users at all, as it will depend on the AAS to grant or deny access to functionality within the applications.

This will also increase security for all applications, as well as simplify the authentication seen from the user's point of view, as there is one way of handling authentication, instead of depending on each application handling this with different rules for password syntax and expiry.

Audit
Using a centralized AAS will make it possible to set up a centralized audit application, which can track all operations performed by all users, and thus set up safeguard measures for detecting and fighting attacks on the organization. This will also give an overview of application usage.

Virtualization
When all identity data is centralized in the identity store, this information can be made available to the entire organization. For example, this may be the organization's white pages, containing the contact information for the employees. A different subset of this information may also be published externally. The identity store may also be the basis for showing the organizational structure, or the project organization. The data may be virtualized in a number of different ways, depending on the needs.
Step 3: Identity applications

When the identity services are in place, the enterprise can introduce new types of applications that were previously difficult or impossible to implement. The identity services handle all the identities within the organization in a consistent way, and this will simplify setting up Identity applications. In addition to providing the business applications with identity services, the Identity services enable other uses of the identity data. Identity applications add features like:

- Workflow
- Single sign-on
- Web portals

Figure 8 The Identity applications layer

Workflow

A workflow application allows businesses to model their processes in a graphical way. The workflow engine then automates the process based on the type of data that it is given.

The workflow application tends to comprise one-off event-driven actions with a transactional data association. An example: A user wants to book an airline ticket. The user would go to a web page and start the process. These types of solutions are usually part of a Business Process Management (BPM) or Enterprise Application Integration (EAI) solution.

A workflow application, like any other application, can provision a single user account based on a logical decision internally. However, when brought into the context of a provisioning system, it increases the potential of what can be done.

Step 5 of a workflow system might say: “Now do a bulk provision for all accounts in AD and the HR system”. The workflow application would transfer this task to the provisioning subsystem.

Single sign-on

Using an identity store and an identity management solution makes it possible to introduce single sign-on (SSO) for the applications. The application would have to be aware of the identity store, or a niche SSO solution could be used (e.g. WebSSO).

If the legacy applications cannot be made aware of the identity store, then a same sign-on or simplified sign-on solution could be produced using the Identity Management solution.
Web portals
It is more popular these days to use portals to access applications. Web based portals can then be incorporated with the WebSSO solution to provide a seamless Identity & Access Management (I&AM) solution. The portals can then be personalized using the content in the Identity store to provide a more individual user experience.

A self-service interface can also be incorporated to the identity store to provide updates to the authoritative data sources in disparate applications.
SAP NetWeaver Identity Management

SAP NetWeaver Identity Management consists of the following products.

**Data Synchronization Engine.** The Data Synchronization Engine is responsible for any low-level operation on the applications and repositories. It runs as part of the Identity Center.

**Identity Center.** The Identity Center is a high-end identity management solution, providing low latency and high availability. It uses a relational database for the configuration data and the logging and status information, as well as for the identity store and all provisioning and workflow states.

**Virtual Directory Server.** A virtual directory provides the organization with real-time access to the identity information, as well as to other critical information, by providing a single access point to all information. The Virtual Directory Server can also be used to control access to the identity data. It is able to present the same data in different ways to different groups of users. It can also be used to write-protect or hide certain attributes, for example when making information available externally.

More detailed information is found in the white papers describing each of the products.