

## Implications of Federated Identity Formats

**Paul Madsen**  
Security Consultant

April, 2004



Entrust is a registered trademark of Entrust, Inc. in the United States and certain other countries. Entrust is a registered trademark of Entrust Limited in Canada. All other Entrust product names and service names are trademarks or registered trademarks of Entrust, Inc. or Entrust Limited. All other company and product names are trademarks or registered trademarks of their respective owners.

The material provided in this document is for information *purposes* only. It is not intended to be advice. You shall be solely responsible for acting or abstaining from acting based upon the information in this document. ENTRUST DOES NOT WARRANT THE QUALITY, ACCURACY OR COMPLETENESS OF THE INFORMATION CONTAINED IN THIS DOCUMENT. THIS INFORMATION IS PROVIDED "AS IS" WITHOUT ANY REPRESENTATIONS WARRANTIES AND/OR CONDITIONS OF ANY KIND, WHETHER EXPRESS, IMPLIED, STATUTORY, BY USAGE OF TRADE, OR OTHERWISE, AND ENTRUST SPECIFICALLY DISCLAIMS ANY AND ALL REPRESENTATIONS, WARRANTIES AND/OR CONDITIONS OF MERCHANTABILITY, SATISFACTORY QUALITY, NON-INFRINGEMENT, TITLE AND FITNESS FOR A SPECIFIC *PURPOSE*.



## Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>2</b>
<b>2</b>	<b>Federated Identity Systems.....</b>	<b>3</b>
<b>3</b>	<b>Federated Identifiers.....</b>	<b>4</b>
3.1	Linkage Mechanism Options .....	4
3.1.1	<i>Global ID.....</i>	<i>4</i>
3.1.2	<i>Opaque Global ID.....</i>	<i>5</i>
3.1.3	<i>Opaque Pairwise ID.....</i>	<i>6</i>
3.1.4	<i>Opaque Directional Pairwise ID.....</i>	<i>7</i>
3.1.5	<i>Roles-based .....</i>	<i>7</i>
3.1.6	<i>Summary.....</i>	<i>8</i>
3.2	Linkage Establishment options .....	9
3.2.1	<i>Batch</i>	<i>9</i>
3.2.2	<i>Interactive Principal.....</i>	<i>9</i>
3.2.3	<i>Roles</i>	<i>10</i>
3.3	Applications .....	10
3.3.1	<i>Single Sign On.....</i>	<i>10</i>
3.3.2	<i>Single Sign Out.....</i>	<i>11</i>
3.3.3	<i>Linkage Management.....</i>	<i>11</i>
3.3.4	<i>Encrypted Identifiers .....</i>	<i>11</i>
3.4	Conclusion.....	13
<b>4</b>	<b>About Entrust.....</b>	<b>13</b>

# 1 Introduction

Organizations today need to extend access to sensitive corporate resources to an ever-growing number of employees, partners, suppliers and customers. Effectively managing the increasing number of users is a significant challenge in itself, adding to this challenge is the complexity of delivering access to enterprise resources in multiple ways such as through client-server, Web and Web Services applications.

As is the case for its local identities, an enterprise must be able to manage the federated identities used in transactions with employees, partners, and customers. Fundamentally, the enterprise must be able to create (and destroy) these federated identities, and specify what internal resources these identities can access. The demands of federated identity management are consequently the same as for managing local identities within the enterprise boundary, complicated by the issues of interoperability and the privacy concerns inherent in the flow of identity data.

## 2 Federated Identity Systems

As originally conceived, the Internet was anonymous. However, while still relevant, anonymity is not supportive of using the Internet for commercial activity - where there is often the need to uniquely identify participants to ensure that suppliers and customers can exchange goods and services and that the appropriate parties can be billed. To address this requirement, Internet identity systems have emerged, ranging from those maintained by merchants to support their B2C transactions to those established within enterprises for employees.

These identity systems are generally not interoperable, identity information held in one system is not generally consumable by another. This lack of interoperability inhibits many emerging business scenarios, including businesses joining together to provide affiliated services to consumers and collaborations and B2B transactions among business partners.

Federated identity addresses this interoperability issue and enables organizations to share trusted identities across the boundaries that separate them. The details and complexity of the identity systems of each are hidden from the other through standards for XML messaging. The OASIS SAML TC, the Liberty Alliance, and WS-Federation are current proposals for these standards.

If identity information is to flow across boundaries, it must carry information about the principal with which it is associated. Consequently, the different organizations need to be able to agree on how to identify their principals to the other. This paper explores issues associated with different mechanisms/formats for this identifier. We will see that the different options differ significantly in their privacy implications.

## 3 Federated Identifiers

Federation refers here to the concept of extending the scope of identity information beyond that for which it may have been originally intended. Single Sign On (SSO) is a good example; after authenticating to one Web site a principal is able to access their account at another without logging in there – they are able to leverage the identity they have at the first site (and their authentication event) at the second site.

Federation requires that both sites be able to agree on how to refer to the Principal. For instance, if the first site is to assert 'X authenticated to me on July 10/2003 at 14.30' to the second site, then for many cases the second site can only do something meaningful if it knows to whom 'X' refers to, (e.g. if the second site is to grant access to the Principal's account, it has to know who that Principal is.)

Note: in most cases the Principal will be oblivious to the existence or nature of this federated identifier, it will not be an account name which they present through an HTML form (although the effect of presenting it will have the same effect).

Consequently, some form of connection or linkage must be established between the two sites.

### 3.1 Linkage Mechanism Options

In the following discussion, an identity provider (IDP) will issue assertions (authentication, attribute, authorization) to be consumed by two service providers SP1 and SP2 (we show two providers because different models for the shared identifier are distinguished by their support for preventing collusion between service providers).

Each provider will typically have their own internal identifier for a Principal (e.g. account name or email address) - for each provider we show the different identifiers for a single Principal named James Smith, known to the IDP as 'Jsmith', to SP1 as 'jj', and to SP2 as 'jimmys'. We distinguish between these 'private' identifiers and the 'public' identifier (that which is used for inter-provider communication).

We assume that federation does not change the local private identifiers, it is the nature of the public identifiers that characterizes the different models shown below.

#### 3.1.1 Global ID

In this model, all providers agree to use the same Public identifier for the Principal, this global identifier being 'jsmith'. There are two main problems with this model. Firstly, a hacker would be able to build up a comprehensive record of the Principal's surfing activities by recording the assertions issued by the IDP, this an unacceptable compromise of the Principal's privacy. Secondly, this global identifier makes trivial collusion between the providers.

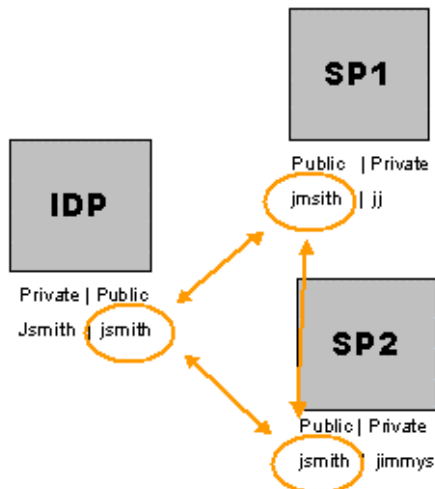
Although the Principal willingly maintains identity information at both service providers, its likely that this Principal would not approve of the two providers comparing notes (i.e. sharing with the other the information and habits) on that Principal. The outcome of this conversation would likely be advantageous to the two service providers but not to the Principal. For instance, if SP1 sells Western boots and SP2 sells Cowboy hats, then the

fact that a particular Principal had an account at both sites would presumably be of some value to a travel agent specializing in Dude Ranches.

Note: Federation requires that the IDP and individual SPs be able to communicate about the Principal. However, the assumption is that the Principal trusts the IDP and is given suitable control of the SPs with which the IDP federates their identity. Collusion here refers to undesired and inappropriate 'federation'.

Note: If the two service providers wish to collude without the approval of the Principal, its likely, independent of federated identity, that they already have sufficient shared information on which to do so, e.g. an email address. The issue is whether or not the mechanisms put in place to support federated identity contribute to the problem by simplifying or enabling inappropriate sharing.

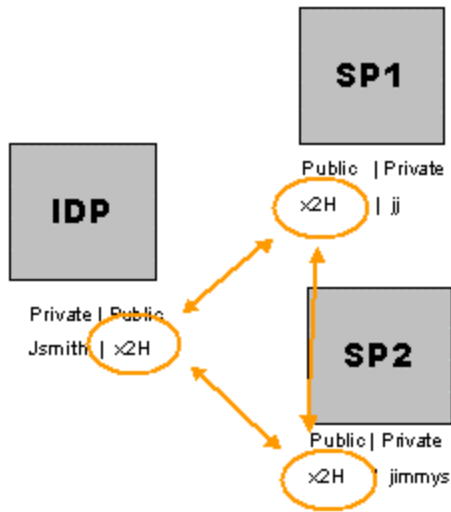
The stylized directory entries for the three providers in the diagram below illustrates how all three providers use the same external identifier for the Principal, namely 'jsmith'. Each provider is able to map from this external identifier to their local, internal, and likely unique identifier for the Principal.



### 3.1.2 Opaque Global ID

In this model, all providers agree to use the same Public identifier for the Principal, but rather than using a meaningful identifier, use an opaque (i.e. meaningless) identifier for the Principal. In the diagram below, all providers use the string 'x2H' as a handle for the Principal. Although the identifier is not meaningful and so does not allow the Principal's Web activities to be so easily tracked, the problem of SP1 and SP2 colluding together inappropriately, using this global (albeit opaque) identifier as a key for a Principal, remains.

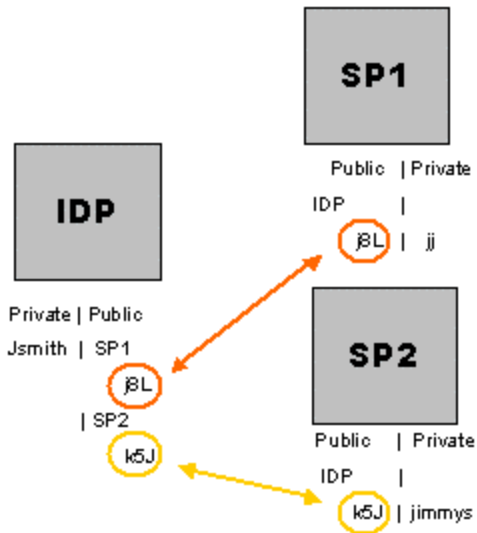
The stylized directory entries for the three providers in the diagram below illustrates that, even though the global identifier is no longer meaningful, e.g. it's the meaningless string 'x2H', the two service providers can still easily collude together because of this common linkage. For instance, one service provider could say to the other 'I know these things about the Principal with identifier 'x2H', why don't we compare notes'.



### 3.1.3 Opaque Pairwise ID

Collusion between SP1 and SP2 can be prevented (with the previously identified caveat) by removing from them the trivial mechanism for linking together their two sets of identity information for the Principal. Rather than the IDP using the same opaque identifier for federating the Principal's identity to both SP1 and SP2, it uses different identifiers (still opaque) for the different providers. In the diagram below, the IDP and SP1 agree to use the string 'j8L' for the Principal, the IDP and SP2 use the different string 'k5J'.

As before, the diagram below shows the stylized directory entries for the Principal at the three providers. The entry at the IDP shows how the IDP would need to index the opaque identifier for the Principal by the service providers to which they correspond (e.g. 'k5J' would be meaningless to SP1 as it is specific to SP2). If the service providers were to federate the Principal's identity with another identity provider, then they would face the same requirement and their directory entries would need to be indexed accordingly.

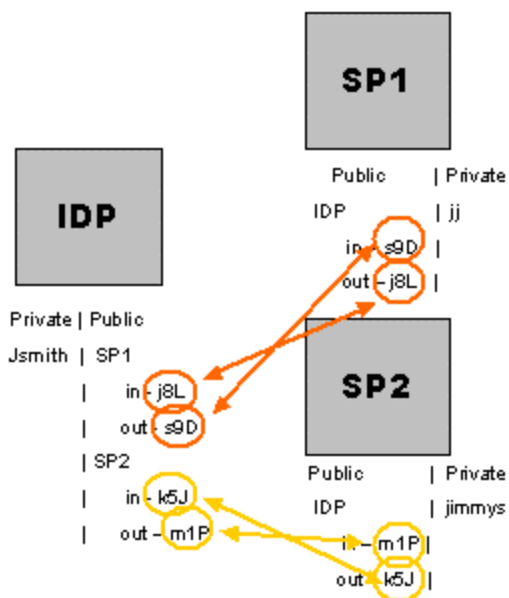


The different colours represent the fact that each pair of identity provider and service provider maintains a unique identifier for a given Principal. Because SP1 and SP2 no longer share a common handle for the Principal, they are unable to use this handle for collusion. No longer can one service provider, based on an external identifier, discuss a Principal with another service provider without the Principal's consent.

### 3.1.4 Opaque Directional Pairwise ID

A variant of the Opaque Pairwise model discussed in the previous section is for the identity provider and each service provider to use different opaque identifiers for each direction of communication. For instance, when the IDP refers to the Principal in a message to a service provider, they would use a different opaque identifier than the service provider would use in a message to the identity provider.

This is shown in the diagram below. There are now 4 different opaque identifiers for the Principal, two between IDP and SP1, distinguished by whether they are used on inbound or outbound messages; two more between IDP and SP2. As before, the IDP must remember with which service provider and when the different opaque identifiers are to be used.



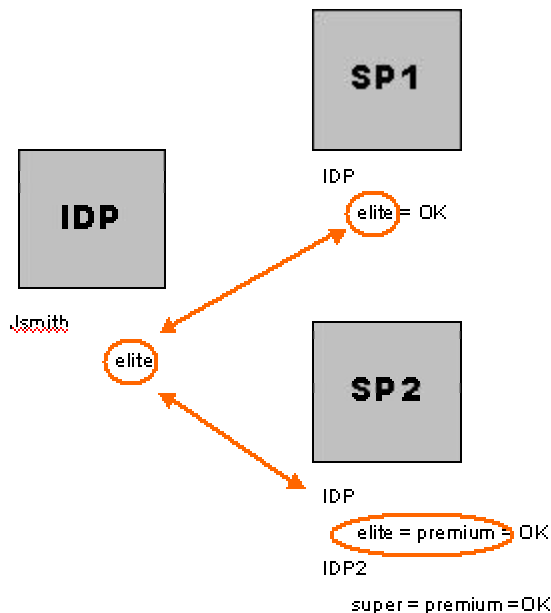
While not significantly impacting the privacy characteristics of the system, such a model allows both identity providers and service providers to have greater control over the form of the identifiers.

### 3.1.5 Roles-based

In the options listed so far, both identity and service provider have had to, a priori, establish an identifier(s) for the principal to be used for subsequent interactions – this identifier unique to the principal (but varying on its uniqueness relative to different service providers). This is relevant if the service provider needs to know precisely who the principal is.

There are however, scenarios in which the service provider doesn't actually care about the precise identity of the principal and so, there is no need for an identifier unique to different principals. For these scenarios, more relevant may be the roles that a principal has or the groups of which they are members. For instance, in a B2B scenario where an employee with the role of 'purchasing officer' needs to be able to access the appropriate Web resources of a supplier - ultimately the supplier cares only that the principal has the appropriate role (implying certain authorizations) and not the specific name of that purchasing officer.

Roles may also be relevant in a more B2C oriented scenarios For instance, assuming that Jsmith is 'elite' according to the identity provider IDP (perhaps reflecting the fact that they fly a lot). The service provider SP1 has specified that principals for whom this identity provider asserts to be 'elite' are accorded certain entitlements. . The diagram below now shows that, rather than the identity provider and service providers sharing some identifier unique to a principal, they must share an understanding of the possible roles that the identity provider's principals might play.



The implication of the above is that this service provider has defined its access policies in terms of a role named and defined by the identity provider. If the service provider interacted with other identity providers in a similar manner, it would need to ensure that its access control policy to account for the likely differently named but logically equivalent roles that these other providers might use. For instance, another identity provider might use the role 'super' to describe a group of its principals; if, to the service provider, 'super' and 'elite' are equivalent, then the service provider can simplify its access policy management by first mapping both these external roles to an internal role called 'premium' and then defining entitlements against this local role. This scenario is shown in the bottom part of the diagram above.

### 3.1.6 Summary

We summarize here the relevance of the different options for the linkage identifier format for two common federated identity scenarios, B2C and B2B.

	<b>B2C</b>	<b>B2B</b>
<b>Global</b>	no	maybe
<b>Opaque Global</b>	no	yes
<b>Opaque Pairwise</b>	yes	yes
<b>Roles-based</b>	maybe	yes

## 3.2 Linkage Establishment options

Federation requires that the Identity and service provider be able to come to an agreement on what strings to use as identifiers for the Principal. There are, at a high level, two options for this establishment mechanism, differing in whether or not the Principal play an active role in the identifier establishment.

### 3.2.1 Batch

For the Batch model of Identifier establishment, the Identity and service providers will engage in a conversation of the following sort (using the opaque identifiers of the previous diagram):

**IDP:** I will use 's9D' to refer to 'James.Smith@email.com' when I communicate to you

**SP1:** I will use 'j8L' to refer to 'James.Smith@email.com' when I communicate to you

The Principal's email address is the necessary piece of shared information from which the two providers will bootstrap a linkage for the Principal, the function could be played by any other common identifier (e.g. employee number, Social Security Number, etc).

There is nothing (technologically) preventing an identity provider and service provider from having this conversation without the Principal's involvement or consent, the Principal needn't even be online at the time. If they can determine that they do indeed share some key identifier for Principals, they can federate the Principal's accounts. Of course, without the Principal's consent, either explicit ('I, James Smith, consent to IDP.com and SP1.com federating my identity') or implicit ('By signing this employment contract, I consent to my employer maintaining my identity information'), this is collusion by our previous definition.

### 3.2.2 Interactive Principal

If the identity provider and service provider don't share a common key for bootstrapping the identifier establishment, or are prevented from doing so through practice or regulatory environment, establishment of shared identifiers requires that the Principal be involved.

The Principal's participation provides two features:

- Their consent to the federation operation can be obtained real-time
- Their simultaneous authentication at both providers provides the necessary shared context between the two providers

The most likely scenario is that the Principal, after authenticating (normally) at the service provider, is given the option of clicking on a 'Federate my identity with identity provider' link. If they do, they are redirected to the identity provider site where they authenticate (normally) again. The fact that they have authenticated at both sites, and that the two sites are connected through the redirect action, allows both providers to determine who

the Principal is without actually sharing any Personally Identifiable Information of that Principal.

For the Interactive Principal model, the conversation between IDP and SP1 appears more like:

**IDP:** I will use 's9D' to refer to the Principal (who you just redirected to me) when I communicate to you.

**SP1:** I will use 'j8L' to refer to the Principal (who I just redirected to you) when I communicate to you.

### 3.2.3 Roles

If the identity and service providers are to communicate roles to each other rather than individual identifiers for principals, then they must agree on both the names for those roles and the semantics of these roles. Unless there is some accepted standard for role definitions against which they can operate, for the two providers to be able to agree on role semantics will not be possible in an automated fashion. The scenario when there is such a standard is presented below

**IDP (to SP):** I will use 'elite' to refer to 'globalstandard:specialuser' when I communicate to you

**SP (to itself):** Well I know that 'globalstandard:specialuser' is equivalent to the role that I call 'premium' so therefore it must be the case that 'IDP:elite' is equivalent to my 'premium'.

More likely is that the two providers would need to determine the mapping through some out of band process.

## 3.3 Applications

### 3.3.1 Single Sign On

Federated SSO is accomplished through the identity provider creating a statement asserting that the Principal authenticated to them (at a certain point in time, with certain technology, etc). This assertion will have been requested by a service provider after a Principal tries to access a protected (or customizable) resource, The conversation between Identity and service providers to enable SSO is logically:

**SP1 (to IDP):** A Principal is trying to access a protected resource, can you tell me if they've authenticated and who they are?

**IDP (to itself):** The Principal in question did recently authenticate as 'Jsmith'. Because it is SP1 asking I will use the identifier for this Principal that SP1 and I previously agreed upon.

**IDP (to SP1):** Indeed, they authenticated 5 minutes ago with an 8 character password and they are the Principal that you and I previously agreed to refer to as 's9D'.

**SP1 (to itself):** Ah, I know that the Principal for which I established the identifier 's9D' with IDP is actually that which I know as 'jj'. Since I'm confident that IDP would only make such a statement if it were true and I can verify that the message wasn't tampered with, I'll log this Principal in as 'jj'.

**SP1 (to Principal):** Welcome jj. How can I help you?

### 3.3.2 Single Sign Out

When a Principal has sessions established at multiple service providers (through the actions of an identity provider), when they log out at one provider, it may be the case that the others should be notified. The conversation between the different providers would appear as:

**SP1** (to IDP): Principal 'j8L' has logged out.

**IDP** (to SP2): Principal 'm1P' has logged out.

The IDP is responsible for understanding that, when received from SP1, 'j8l' refers to 'Jsmith' and that, for this Principal, when communicating with SP2, the appropriate identifier to use is 'm1P'.

If the log out were instead initiated by SP2, the sequence would be:

**SP2** (to IDP): Principal 'k5J' has logged out.

**IDP** (to SP1): Principal 's9D' has logged out.

### 3.3.3 Linkage Management

Once established between an identity and service provider, the linkage must be managed. For instance, it may be relevant for the identifier (when neither global nor meaningful) to be refreshed on some schedule, this operation initiated by either provider. Likewise, it will occasionally be necessary to terminate the linkage, either at the behest of the principal or by either of the providers.

### 3.3.4 Encrypted Identifiers

We identified the importance of preventing service providers from inappropriately communicating with each other regarding Principals. Such collusion is prevented (or at least not made trivial) by ensuring that the service providers do not share the same opaque identifiers for Principals.

There will however be circumstances where two service providers are authorized to communicate in order to perform some service to the Principal. As is the case for communication between identity provider and service provider, such communication between service providers 'about' a Principal requires some common identifier(s) for that Principal.

One possibility would be to federate the Principal's identities at the two service providers through either of the two Identifier Establishment options described above. This would result in another opaque identifier for the principal, this one unique to the pair of service providers. While this is definitely possible, the implication is that the Principal would need to consent to this federation and would represent another trust decision for that principal.

The alternative to having the two service providers federate the Principal's identities is to leverage the existing federations that exist between the two service providers and the identity provider. Looking at the diagram in Section Opaque Directional Pairwise ID we see that, although the two service providers SP1 and SP2 don't share any common identifier for the Principal, the identity provider IDP has it all. The identity provider can act as a bridge between the two providers. One mechanism would then be for the IDP to simply provide to a service provider the appropriate identifier to use in discussions with another service provider. For instance,

**SP1** (to IDP): I wish to talk to SP2 about the Principal with identifier 'j8L'

**IDP** (to SP1): Use 'm1P' in your discussions with SP2.

**SP1** (to SP2): Let's talk about 'm1P'.

The two service providers can now communicate on behalf of the principal. The price however is that the principal is now known as 'm1p' at all three providers. As the IDP facilitates such discussions between other providers, the situation quickly degrades to that of a global identifier for the Principal, with the attendant privacy concerns.

What is required is a mechanism that allows the identity provider to facilitate the discussion between the SPs in a privacy respecting manner – the IDP can map between the Principal's identifiers as necessary to enable SP1 and SP2 to communicate on behalf of the Principal.

The conversation might go as follows (refer to the diagram)

**SP1** (to IDP): I wish to talk to SP2 about the Principal with identifier 'j8L'

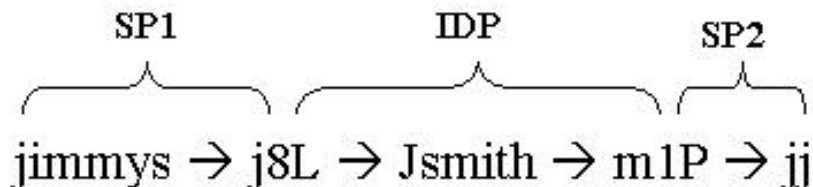
**IDP** (to SP1): Use ESP2('m1P') in your discussions with SP2.

**SP1** (to SP2): Let's talk about ESP2('m1P')

Where Ex('id') represents that the identifier 'id' is encrypted for the provider 'x'.

IDP provides to SP1 an identifier for the principal that will be meaningful to SP2, but protects this from SP1 by encrypting for SP2. Because IDP doesn't want SP1 to know the opaque identifier 'm1P' that IDP and SP2 previously agreed to use for the Principal (this would enable SP1 to subsequently collude), IDP encrypts this identifier for SP2 (as shown by the notation ESP2('m1P'))

The IDP maps between the identifiers for the principal it maintains at SP1 and SP2. There are a number of mappings occurring, these shown below



Before posing the initial question to IDP, SP1 must map from its local identifier for the principal 'jimmys' to the federated identifier 'j8L'. IDP uses its local identifier 'Jsmith' to map to the second federated identifier 'm1P', this understandable by SP2. When SP2 receives this identifier (after decrypting) it maps to its local identifier for the Principal, 'jj', and is then able to access the appropriate directory entry for the Principal.

To avoid the encrypted identifier serving as a shared identifier, the form of encryption must ensure that successive encryptions of a persistent identifier will yield distinct results that cannot be meaningfully correlated to one another.

### **3.4 Conclusion**

Organizations are feeling increased pressure to lower the costs of doing business with their partners, suppliers, and customers. As they leverage the Internet to streamline business processes, organizations must effectively extend the boundaries of their enterprise to ensure that these external clients are given appropriate access to internal resources in a timely and secure manner. If an enterprise is to control access to its sensitive internal business resources, it must be able to compare the identity of the entity requesting access to those resources against those identities that are authorized to access the resources. Consequently, the external clients must be assigned identities against which appropriate authorizations have been defined – this is federated identity management.

This paper discussed different options for the forms of this identifier (as well mechanisms for its establishment) and presented the advantages and disadvantages of each, with consequent implications for their relevance to different deployment scenarios. Ultimately, a comprehensive federated identity management solution should support the full range of options.

## **4 About Entrust**

Entrust, Inc. [Nasdaq: ENTU] is a world leader in securing digital identities and information, enabling businesses and governments to transform the way they conduct online transactions and manage relationships with customers, partners and employees. Entrust's solutions promote a proactive approach to security that provides accountability and privacy to online transactions and information. Over 1,200 enterprises and government agencies in more than 50 countries use Entrust's portfolio of security software solutions that integrate into a broad range of applications organizations use today to leverage the Internet and enterprise networks. For more information, please visit <http://www.entrust.com>