2 Introduction

This section is informative.

Federation is a process that allows for the conveyance of authentication attributes and subscriber attributes across networked systems. In a federation scenario, the verifier or CSP is referred to as an identity provider, or IdP. The RP is the party that receives and uses the information provided by the IdP.

Federated identity systems use assertions to accomplish this task. Assertions are statements from an IdP to an RP that contain information about a subscriber. Federation technology is generally used when the RP and the IdP are not a single entity or are not under common administration. The RP uses the information in the assertion to identify the subscriber and make authorization decisions about their access to resources controlled by the RP. An assertion typically includes an identifier for the subscriber, allowing association of the subscriber with their previous interactions with the RP. Assertions may additionally include attribute values or attribute references that further characterize the subscriber and support the authorization decision at the RP. Additional attributes may also be available outside of the assertion as part of the larger federation protocol. These attribute values and attribute references are often used in determining access privileges for Attribute Based Access Control (ABAC) or facilitating a transaction (e.g., shipping address).

In a federated identity scenario, the subscriber does not authenticate directly to the RP. Instead, the federation protocol defines a mechanism for an IdP to generate an assertion for the identifier associated with a subscriber, usually in response to a request from the RP. The IdP is responsible for authenticating the subscriber (though it may use session management as described in SP 800-63B, Section 7). This process allows the subscriber to obtain services from multiple RPs without the need to hold or maintain separate credentials at each. This process can also be used to support single sign on, where subscribers authenticate once to an IdP and subsequently obtain services from multiple RPs.

Federation requires relatively complex multiparty protocols that have subtle security and privacy requirements and require careful consideration. When evaluating a particular federation structure, it may be instructive to break it down into its component interactions. Generally speaking, authentication between the subscriber and the IdP will be based on the authentication mechanisms presented in SP 800-63B, while interactions between the IdP and RP will convey attributes established using procedures in SP 800-63A and other self-asserted attributes. Many of the requirements presented in this document, therefore, have some relationship with corresponding requirements in those two documents.

The following table states which sections of the document are normative and which are informative:

Table 2-1 Normative and Informative Sections of 800-63C

Section Name	Normative/Informative
1. Purpose	Informative
2. Introduction	Informative
3. Definitions and Abbreviations	Informative
4. Federation Assurance Level (FAL)	Normative
5. Federation	Normative
6. Assertion	Normative
7. Assertion Presentation	Normative
8. Security	Informative
9. Privacy Considerations	Informative
10. Usability Considerations	Informative
11. Examples	Informative
12. References	Informative

3 Definitions and Abbreviations

See SP 800-63, Appendix A for a complete set of definitions and abbreviations.

4 Federation Assurance Levels

This section is normative.

This section defines allowable Federation Assurance Levels, or FAL. The FAL describes requirements for how assertions are constructed and secured for a given transaction. These levels can be requested by an RP or required by the configuration of both the RP and the IdP for a given transaction.

All assertions SHALL be used with a federation protocol as described in Section 4. All assertions SHALL comply with the detailed requirements in Section 6. All assertions SHALL be presented using one of the methods described in Section 7. While many different federation implementation options are possible, the FAL is intended to provide clear implementation recommendations representing increasingly secure deployment options. Combinations of aspects not found in the FAL table are possible but outside the scope of this volume. See SP 800-63 Section 6.3 for details on how to choose the most appropriate FAL.

This table presents different requirements for each FAL. Each successive level subsumes and fulfills all requirements of lower levels. Federations presented through a proxy SHALL be represented by the lowest level used during the proxied transaction.

Table 4-1 Federation Assertion Levels

FAL	Requirement
1	Bearer assertion, signed by IdP.
2	Bearer assertion, signed by IdP and encrypted to RP.
3	Holder of key assertion, signed by IdP and encrypted to RP.

For example, FAL1 maps to the OpenID Connect Basic Client profile or Security Assertion Markup Language (SAML) Web SSO Artifact Binding profile with no additional features. FAL2 additionally requires that the assertion (e.g., the OpenID Connect ID Token or SAML Assertion) be encrypted to a public key representing the RP in question. FAL3 requires the subscriber to cryptographically prove possession of a key bound to the assertion (e.g., the use of a cryptographic authenticator) along with all requirements of FAL2. The additional key presented at FAL3 need not be the same key used by the subscriber to authenticate to the IdP.

Regardless of what the RP requests or what the protocol requires, the RP can easily detect the FAL in use by observing the nature of the assertion as it is presented as part of the federation protocol. Therefore, the RP is responsible for determining which FALs it is willing to accept for a given authentication transaction and ensuring that the transaction meets that FAL's requirements.

If the RP is using a front-channel presentation mechanism, as defined in Section 7.2 (e.g., the OpenID Connect Implicit Client profile or the SAML Web SSO profile), it SHALL require FAL2 or greater in order to protect the information in the assertion from disclosure to the browser or other parties in the transaction other than the intended RP.

Additionally, the IdP SHALL employ appropriately-tailored security controls (to include control enhancements) from the moderate or high baseline of security controls defined in SP 800-53 or equivalent federal (e.g., FEDRAMP) or industry standard.

4.1 Key Management

At any FAL, the IdP SHALL ensure that an RP is unable to impersonate the IdP at another RP by protecting the assertion with a signature and key using approved cryptography. If the assertion is protected by a digital signature using an asymmetric key, the IdP MAY use the same public and private key pair to sign assertions to multiple RPs. The IdP MAY publish its public key in a verifiable fashion, such as at an HTTPS-protected URL at a well-known location. If the assertion is protected by a MAC using a shared key, the IdP SHALL use a different shared key for each RP.

Government-operated IdPs asserting authentication at AAL2 and all IdPs asserting authentication at AAL3 SHALL protect keys used for signing or encrypting those assertions with mechanisms validated at FIPS 140 Level 1 or higher.

4.2 Runtime Decisions

The fact that parties have federated SHALL NOT be interpreted as permission to pass information. The decision of whether an authentication can occur or attributes may be passed can be determined by the use of a whitelist, a blacklist, or a runtime decision by an authorized party.

IdPs MAY establish whitelists of RPs authorized to receive authentication and attributes from the IdP without a runtime decision from the subscriber. All RPs in an IdP's whitelist SHALL abide by the provisions and requirements in the SP 800-63 suite. IdPs SHALL make whitelists available to subscribers as described in Section 9.2. IdPs MAY also establish blacklists of RPs not authorized to receive authentication or attributes from the IdP, even when requested by the subscriber. Both whitelists and blacklists identify RPs by their domain or other sufficiently unique identifier, depending on the federation protocol in use. Every RP not on a whitelist or a blacklist SHALL be placed by default in a gray area where runtime authorization decisions will be made by an authorized party, usually the subscriber. The IdP MAY remember a subscriber's decision to authorize a given RP, provided that the IdP SHALL allow the subscriber to revoke such remembered access at a future time.

RPs MAY establish whitelists of IdPs that the RP will accept authentication and attributes from without a runtime decision from the subscriber. All IdPs in an RP's whitelist SHALL abide by the provisions and requirements in the 800-63 suite. RPs MAY also establish blacklists of IdPs that the RP will not accept authentication or attributes from, even when requested by the subscriber. Both whitelists and blacklists identify IdPs by their domain or other sufficiently unique identifier, depending on the federation protocol in use. Every IdP that is not on a whitelist or a blacklist SHALL be placed by default in a gray area where runtime authorization decisions

will be made by an authorized party, usually the subscriber. The RP MAY remember a subscriber's decision to authorize a given IdP, provided that the RP SHALL allow the subscriber to revoke such remembered access at a future time.

A subscriber's information SHALL be transmitted between IdP and RP only for identity federation transactions or support functions such as identification of compromised accounts as discussed in Section 5.2. A subscriber's information SHALL NOT be transmitted for any other purposes, even when those parties are whitelisted.

To mitigate the risk of unauthorized exposure of sensitive information (e.g., shoulder surfing), the IdP SHALL, by default, mask sensitive information displayed to the subscriber. The IdP SHALL provide mechanisms for the subscriber to temporarily unmask such information in order for the subscriber to view full values. The IdP SHALL provide effective mechanisms for redress of applicant complaints or problems (e.g., subscriber identifies an inaccurate attribute value). For more details on masking and redress, please see Section 10 on usability considerations.

When the subscriber is involved in a runtime decision, the subscriber SHALL receive explicit notice and be able to provide positive confirmation before any attributes about the subscriber are transmitted to any RP. At a minimum, the notice SHOULD be provided by the party in the position to provide the most effective notice and obtain confirmation, consistent with Section 9.2. If the protocol in use allows for optional attributes, the subscriber SHALL be given the option to decide whether to transmit those attributes to the RP. An IdP MAY employ mechanisms to remember and re-transmit the exact attribute bundle to the same RP.

5 Federation

This section is normative.

In a federation protocol, a three-party relationship is formed between the subscriber, the IdP, and the RP, as shown in Figure 5-1. Depending on the specifics of the protocol, different information passes between the participants at different times. The subscriber communicates with both the IdP and the RP, usually through a browser. The RP and the IdP communicate with each other in two ways:

- The *front channel*, through redirects involving the subscriber; or
- The *back channel*, through a direct connection between the RP and IdP, not involving the subscriber.

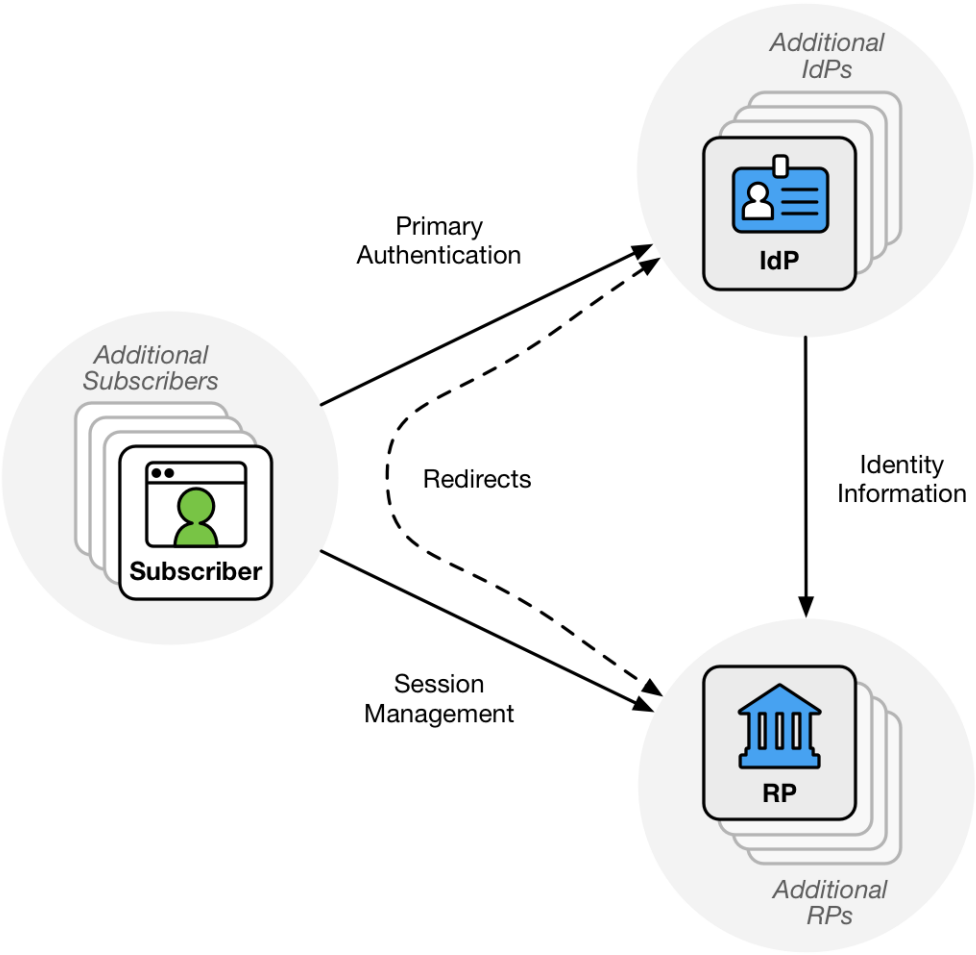

Figure 5-1 Federation

The subscriber authenticates to the IdP and the result of that authentication event is asserted to the RP across the network. In this transaction, the IdP acts as the verifier for the credential, as described in SP 800-63B. The IdP can also make attribute statements about the subscriber as part of this process. These attributes and authentication event information are carried to the RP

through the use of an assertion, described in Section 6. Additional attributes MAY be made available through a secondary protocol protected by an authorized credential.

5.1 Federation Models

IdPs that provide authentication services and RPs that consume those services are known as members of a federation. From an IdP's perspective, the federation consists of the RPs that it serves. From an RP's perspective, the federation consists of the IdPs that it uses. This section provides an overview of and requirements for common identity federation models currently in use. In each model, relationships are established between members of the federation.

5.1.1 Manual Registration

In the manual registration model, the IdP and RP manually provision configuration information about parties with which they expect to interoperate. IdPs MAY configure RPs using an explicit whitelist, allowing these RPs to receive authentication and attribute information as part of the authentication transaction. In cases where an RP is not whitelisted, the IdP SHALL require runtime decisions (see Section 4.2) to be made by an authorized party (such as the subscriber) before releasing user information.

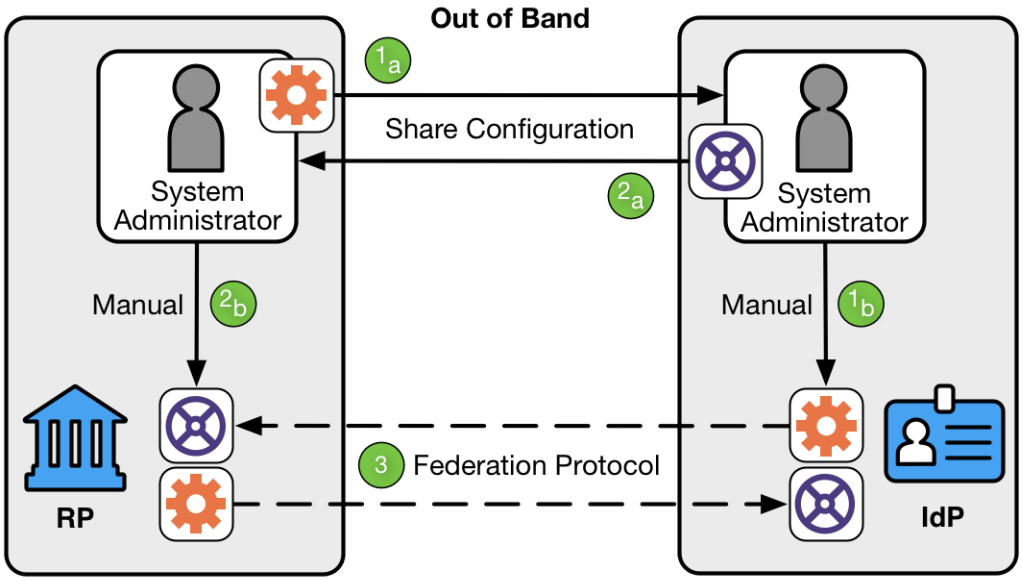

Figure 5-2 Manual Registration

As shown in Figure 5-2, manual registration involves three steps:

1. The RP's system administrator shares the RP's attributes with the IdP's system administrator, who associates those attributes with the RP.
2. The IdP's system administrator shares the IdP's attributes with the RP's system administrator, who associates those attributes with the IdP.
3. The IdP and RP then communicate using a standard federation protocol.

IdPs and RPs MAY act as their own authorities on who to federate with or MAY externalize those authority decisions to an external party as in Section 5.1.3.

Protocols requiring the transfer of keying information SHALL use a secure method during the registration process to exchange keying information needed to operate the federated relationship, including any shared secrets or public keys. Any symmetric keys used in this relationship SHALL be unique to a pair of federation participants.

Federation relationships SHALL establish parameters regarding expected and acceptable IALs and AALs in connection with the federated relationship.

5.1.2 Dynamic Registration

In the dynamic registration model of federation, it is possible for relationships between members of the federation to be negotiated at the time of a transaction. This process allows IdPs and RPs to be connected together without manually establishing a connection between them using manual registration (see Section 5.1.1). IdPs that support dynamic registration SHALL make their configuration information (such as dynamic registration endpoints) available in such a way as to minimize system administrator involvement.

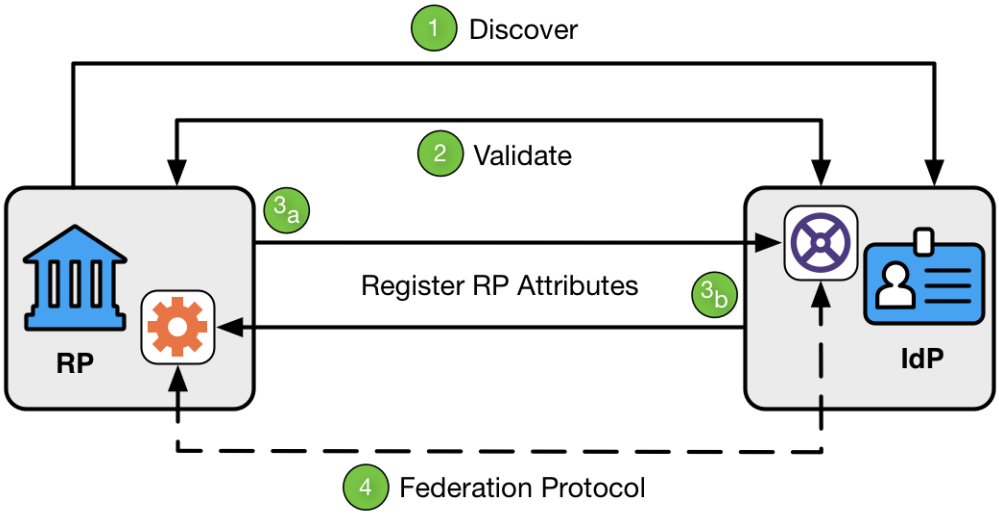

Figure 5-3 Dynamic Registration

As shown in Figure 5-3, dynamic registration involves four steps:

1. Discover. The RP goes to a well-known location at the IdP to find the IdP's metadata.
2. Validate. The RP and IdP determine each other's validity. This can be accomplished through keying information, metadata, software statements, or other means.
3. Register RP attributes. The RP sends its attributes to the IdP, and the IdP associates those attributes with the RP.
4. Federation Protocol. The IdP and RP then communicate using a standard federation protocol.

Protocols requiring the transfer of keying information SHALL use a secure method during the registration process to establish such keying information needed to operate the federated relationship, including any shared secrets or public keys. Any symmetric keys used in this relationship SHALL be unique to a pair of federation participants.

IdPs SHALL require runtime decisions (see Section 4.2) to be made by an authorized party (such as the subscriber) before releasing user information. An IdP accepting dynamically registered RPs MAY limit the types of attributes and other information made available to such RPs. An RP capable of dynamically registering MAY limit which IdPs it is willing to accept identity information from.

Parties in a dynamic registration model frequently do not know each other ahead of time. Where possible, this SHOULD be augmented by software statements, which allow federated parties to cryptographically verify some attributes of an RP being dynamically registered. Software statements are lists of attributes describing the RP software, cryptographically signed by an authority (either the IdP itself, a federation authority as in Section 5.1.3, or another trusted party). This cryptographically-verifiable statement allows the connection to be established or elevated between the federating parties without relying solely on self-asserted attributes. (See RFC 7591 Section 2.3 for more information on one protocol's implementation of software statements.)

5.1.3 Federation Authorities

Some federated parties defer to an authority, known as a *federation authority*, to assist in making federation decisions and to establish the working relationship between parties. In this model, the federation authority generally conducts some level of vetting on each party in the federation to verify compliance with predetermined security and integrity standards. The level of vetting — if it occurs at all — is unique to the use cases and models employed within the federation. This vetting is depicted in the left side of Figure 5-4.

Federation authorities approve IdPs to operate at certain IALs, AALs, and FALs. This information is used by relying parties, as shown in the right side of Figure 5-4, to determine which identity providers meet their requirements.

Federation authorities SHALL establish parameters regarding expected and acceptable IALs, AALs, and FALs in connection with the federated relationships they enable. Federation authorities SHALL individually vet each participant in the federation to determine whether they adhere to their expected security, identity, and privacy standards.

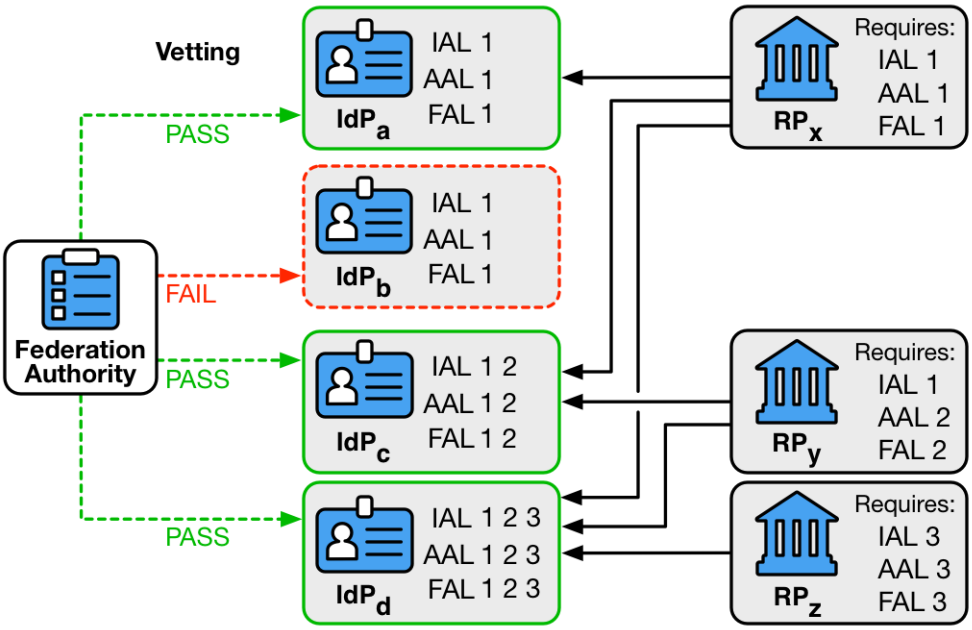

Figure 5-4 Federation Authority

Vetting of IdPs and RPs SHALL establish, as a minimum, that:

- Assertions generated by IdPs adhere to the requirements in Section 6.
- RPs adhere to IdP requirements for handling subscriber attribute data, such as retention, aggregation, and disclosure to third parties.
- RP and IdP systems use approved profiles of federation protocols.

Federation authorities MAY assist the technical connection and configuration process between members, such as by publishing configuration data for IdPs or by issuing software statements for RPs.

Most federations managed through authorities have a simple membership model: parties are either in the federation or they are not. More sophisticated federations MAY have multiple membership tiers that federated parties can use to tell whether other parties in the federation have been more thoroughly vetted. IdPs MAY decide that certain subscriber information is only releasable to RPs in higher tiers and RPs MAY decide to accept certain information only from IdPs in higher tiers.

5.1.4 Proxied Federation

In a proxied federation, communication between the IdP and the RP is intermediated in a way that prevents direct communication between the two parties. There are multiple methods to achieve this effect. Common configurations include:

- A third party that acts as a federation proxy (or *broker*)
- A network of nodes that distributes the communications

Where proxies are used, they function as an IdP on one side and an RP on the other. Therefore, all normative requirements that apply to IdPs and RPs SHALL apply to proxies in their respective roles.

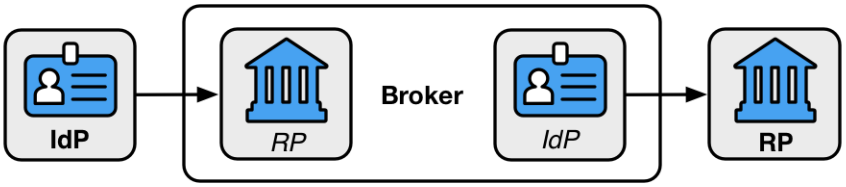

Figure 5-5 Federation Proxy

A proxied federation model can provide several benefits. Federation proxies can simplify technical integration between the RP and IdP by providing a common interface for integration. Additionally, to the extent a proxy effectively blinds the RP and IdP from each other, it can provide some business confidentiality for organizations that want to guard their subscriber lists from each other. Proxies can also mitigate some of the privacy risks described in Section 5.2 below.

See Section 9.5 for further information on blinding techniques, their uses, and limitations.

5.2 Privacy Requirements

Federation involves the transfer of personal attributes from a third party that is not otherwise involved in a transaction — the IdP. Federation also potentially gives the IdP broad visibility into subscriber activities. Accordingly, there are specific privacy requirements associated with federation.

Communication between the RP and the IdP could reveal to the IdP where the subscriber is conducting a transaction. Communication with multiple RPs allows the IdP to build a profile of subscriber transactions that would not have existed without federation. This aggregation could enable new opportunities for subscriber tracking and use of profile information that do not always align with subscribers' privacy interests.

If an IdP discloses information on subscriber activities at an RP to any party, or processes the subscriber's information for any purpose other than identity proofing, authentication, or attribute assertions (collectively "identity service"), related fraud mitigation, to comply with law or legal process, or in the case of a specific user request, to transmit the information, the IdP SHALL implement measures to maintain predictability and manageability commensurate with the privacy risk arising from the additional processing. Measures MAY include providing clear notice, obtaining subscriber consent, or enabling selective use or disclosure of attributes. When an IdP uses consent measures, the IdP SHALL NOT make consent for the additional processing a condition of the identity service. The IdP SHOULD employ technical measures, such as the use of pairwise pseudonymous identifiers described in Section 6.3 or privacy-enhancing cryptographic protocols, to provide disassociability and discourage subscriber activity tracking and profiling.

An IdP MAY disclose information on subscriber activities to other RPs within the federation for security purposes, such as communication of compromised subscriber accounts.

The following requirements apply specifically to federal agencies:

1. The agency SHALL consult with their Senior Agency Official for Privacy (SAOP) to conduct an analysis determining whether the requirements of the Privacy Act are triggered by the agency that is acting as an IdP, by the agency that is acting as an RP, or both (see Section 9.4).
2. The agency SHALL publish or identify coverage by a System of Records Notice (SORN) as applicable.
3. The agency SHALL consult with their SAOP to conduct an analysis determining whether the requirements of the E-Government Act are triggered by the agency that is acting as an IdP, the agency that is acting as an RP, or both.
4. The agency SHALL publish or identify coverage by a Privacy Impact Assessment (PIA) as applicable.

5.3 Reauthentication and Session Requirements in Federated Environments

In a federated environment, the RP manages its sessions separately from any sessions at the IdP. The session at the RP starts when the RP processes the federation protocol from the IdP. At the time of a federated login, the subscriber MAY have an existing session at the IdP which MAY be used as part of the authentication process to the RP. The IdP SHALL communicate any information it has regarding the time of the latest authentication event at the IdP, and the RP MAY use this information in determining its access policies. Depending on the capabilities of the federation protocol in use, the IdP SHOULD allow the RP to request that the subscriber re-authenticate at the IdP as part of a federation request.

Due to the distributed nature of a federated system, the subscriber is capable of terminating sessions with the IdP and RP independently of one another. The RP SHALL NOT assume that the subscriber has an active session at the IdP past the establishment of the federated log in. The IdP SHALL NOT assume that termination of the subscriber's session at the IdP will propagate to any sessions that subscriber would have at downstream RPs.

See SP 800-63B Section 7 for more information about session management requirements.

6 Assertions

This section is normative.

An assertion used for authentication is a packaged set of attribute values or attribute references about or associated with an authenticated subscriber that is passed from the IdP to the RP in a federated identity system. Assertions contain a variety of information, including: assertion metadata, attribute values and attribute references about the subscriber, information about the subscriber's authentication at the IdP, and other information that the RP can leverage (such as restrictions and expiration time). While the assertion's primary function is to authenticate the user to an RP, the information conveyed in the assertion can be used by the RP for a number of use cases — for example, authorization or personalization of a website. These guidelines do not restrict RP use cases nor the type of protocol or data payload used to federate an identity, provided the chosen solution meets all mandatory requirements contained herein.

Assertions MAY represent only an authentication event, or MAY also represent attribute values and attribute references regarding the subscriber.

All assertions SHALL include the following assertion metadata:

1. Subject: An identifier for the party that the assertion is about (i.e., the subscriber).
2. Issuer: An identifier for the IdP that issued the assertion.
3. Audience: An identifier for the party intended to consume the assertion (i.e., the RP).
4. Issuance: A timestamp indicating when the IdP issued the assertion.
5. Expiration: A timestamp indicating when the assertion expires and SHALL no longer be accepted as valid by the RP (i.e., the expiration of the assertion and not the expiration of the session at the RP).
6. Identifier: A value uniquely identifying this assertion, used to prevent attackers from replaying prior assertions.
7. Signature: Digital signature or message authentication code (MAC), including key identifier or public key associated with the IdP, for the entire assertion.
8. Authentication Time: A timestamp indicating when the IdP last verified the presence of the subscriber at the IdP through a primary authentication event (if available).

Assertions MAY also include the following information:

1. Key binding: Public key or key identifier of subscriber-held key to demonstrate their binding with the assertion described in Section 6.1.2.
2. Attribute values and attribute references: Information about the subscriber.
3. Attribute metadata: Additional information about one or more subscriber attributes, such as that described in NIST Internal Report 8112 [NISTIR 8112].

Assertions SHOULD specify the AAL when an authentication event is being asserted and IAL when identity proofed attributes (or references based thereon) are being asserted. If not specified, the RP SHALL NOT assign any specific IAL or AAL to the assertion.

An RP SHALL treat subject identifiers as not inherently globally unique. Instead, the value of the assertion's subject identifier is usually in a namespace under the assertion issuer's control. This allows an RP to talk to multiple IdPs without incorrectly conflating subjects from different IdPs.

Assertions MAY include additional attributes. Section 7 contains privacy requirements for presenting attributes in assertions. The RP MAY fetch additional identity attributes from the IdP in one or more separate transactions using an authorization component issued alongside the original assertion. The ability to successfully fetch such additional attributes SHALL NOT be treated as equivalent to processing the assertion.

Although details vary based on the exact federation protocol in use, an assertion SHOULD be used only to represent a single login event at the RP. After the RP consumes the assertion, session management by the RP comes into play (see SP 800-63B Section 7); an assertion SHALL NOT be used past the expiration time contained therein. However, the expiration of the session at the RP MAY occur prior to the assertion's expiration. See Section 5.3 for more information.

The assertion's lifetime is the time between its issuance and its expiration. This lifetime needs to be long enough to allow the RP to process the assertion and create a local application session for the subscriber, but should not be longer than necessary for such establishment. Long-lived assertions have a greater risk of being stolen or replayed; a short assertion lifetime mitigates this risk. Assertion lifetimes SHALL NOT be used to limit the session at the RP. See Section 5.3 for more information.

6.1 Assertion Binding

Assertion binding can be classified based on whether presentation by a claimant of an assertion, or an assertion reference, is sufficient for binding to the subscriber, or if the RP requires additional proof that the assertion is bound to the subscriber.

6.1.1 Bearer Assertions

A bearer assertion can be presented by any party as proof of the bearer's identity. If an attacker can capture or manufacture a valid assertion or assertion reference representing a subscriber and can successfully present that assertion or reference to the RP, then the attacker could be able to impersonate the subscriber at that RP.

Note that mere possession of a bearer assertion or reference is not always enough to impersonate a subscriber. For example, if an assertion is presented in the back-channel federation model (described in Section 7.1), additional controls MAY be placed on the transaction (such as identification of the RP and assertion injection protections) that help further protect the RP from fraudulent activity.

6.1.2 Holder-of-Key Assertions

A holder-of-key assertion contains a reference to a key possessed by and representing the subscriber. The key referenced in a holder-of-key represents the subscriber, not any other party

in the system including the browser, IdP, or RP. Note that the reference to the key is asserted (and signed) by the issuer of the assertion.

When the RP receives the holder-of-key assertion, the subscriber proves possession of the key referenced in the assertion directly to the RP. While the subscriber could also have used a key-based means of authenticating to the IdP, the primary authentication at the IdP and the federated authentication at the RP are considered separately and are not assumed to use the same keys or related sessions.

In proving possession of the subscriber's key to the RP, the claimant also proves with a certain degree of assurance that they are the rightful subject of the assertion. It is more difficult for an attacker to use a stolen holder-of-key assertion issued to a subscriber, since the attacker would need to steal the referenced key material as well.

The following requirements apply to all holder-of-key assertions:

1. The subscriber SHALL prove possession of that key to the RP, in addition to presentation of the assertion itself.
2. An assertion containing a reference to a key held by the subscriber for which key possession has not been proven SHALL be considered a bearer assertion by the RP.
3. Reference to a given key SHALL be trusted at the same level as all other information within the assertion.
4. The assertion SHALL NOT include an unencrypted private or symmetric key to be used with holder-of-key presentation.
5. The key MAY be distinct from any key used by the subscriber to authenticate to the IdP.
6. The key MAY be a symmetric key or a public key that corresponds to a private key.
7. The RP MAY verify the claimant's possession of the key in conjunction with the IdP, for example, by requesting that the IdP verify a signature or MAC calculated by the claimant in response to a cryptographic challenge.

6.2 Assertion Protection

Independent of the binding mechanism (discussed in Section 6.1) or the federation model used to obtain them (described in Section 5.1), assertions SHALL include a set of protections to prevent attackers from manufacturing valid assertions or reusing captured assertions at disparate RPs. The protections required are dependent on the details of the use case being considered, and recommended protections are listed here.

6.2.1 Assertion Identifier

Assertions SHALL be sufficiently unique to permit unique identification by the target RP. Assertions MAY accomplish this by use of an embedded nonce, issuance timestamp, assertion identifier, or a combination of these or other techniques.

6.2.2 Signed Assertion

Assertions SHALL be cryptographically signed by the issuer (IdP). The RP SHALL validate the digital signature or MAC of each such assertion based on the issuer's key. This signature SHALL cover the entire assertion, including its identifier, issuer, audience, subject, and expiration.

The assertion signature SHALL either be a digital signature using asymmetric keys or a MAC using a symmetric key shared between the RP and issuer. Shared symmetric keys used for this purpose by the IdP SHALL be independent for each RP to which they send assertions, and are normally established during registration of the RP. Public keys for verifying digital signatures MAY be fetched by the RP in a secure fashion at runtime, such as through an HTTPS URL hosted by the IdP. Approved cryptography SHALL be used.

6.2.3 Encrypted Assertion

When encrypting assertions, the IdP SHALL encrypt the contents of the assertion using either the RP's public key or a shared symmetric key. Shared symmetric keys used for this purpose by the IdP SHALL be independent for each RP to which they send assertions, and are normally established during registration of the RP. Public keys for encryption MAY be fetched by the IdP in a secure fashion at runtime, such as through an HTTPS URL hosted by the RP.

All encryption of assertions SHALL use approved cryptography.

When assertions are passed through third parties, such as a browser, the actual assertion SHALL be encrypted. For example, a SAML assertion can be encrypted using XML-Encryption, or an OpenID Connect ID Token can be encrypted using JSON Web Encryption (JWE). For assertions that are passed directly between IdP and RP, the actual assertion MAY be encrypted. If it is not, the assertion SHALL be sent over an authenticated protected channel.

> Note: Assertion encryption is required at FAL2 and FAL3.

6.2.4 Audience Restriction

Assertions SHALL use audience restriction techniques to allow an RP to recognize whether or not it is the intended target of an issued assertion. All RPs SHALL check that the audience of an assertion contains an identifier for their RP to prevent the injection and replay of an assertion generated for one RP at another RP.

6.3 Pairwise Pseudonymous Identifiers

In some circumstances, it is desirable to prevent the subscriber's account at the IdP from being easily linked at multiple RPs through use of a common identifier.

6.3.1 General Requirements

When using pairwise pseudonymous subject identifiers within the assertions generated by the IdP for the RP, the IdP SHALL generate a different identifier for each RP as described in Section 6.3.2 below.

When pairwise pseudonymous identifiers are used with RPs alongside attributes, it may still be possible for multiple colluding RPs to re-identify a subscriber by correlation across systems using these identity attributes. For example, if two independent RPs each see the same subscriber identified with different pairwise pseudonymous identifiers, they could still determine that the subscriber is the same person by comparing the name, email address, physical address, or other identifying attributes carried alongside the pairwise pseudonymous identifier in the respective assertions. Privacy policies SHOULD prohibit such correlation, and pairwise pseudonymous identifiers can increase effectiveness of these policies by increasing the administrative effort in managing the attribute correlation.

Note that in a proxied federation model, the initial IdP may be unable to generate a pairwise pseudonymous identifier for the ultimate RP, since the proxy could blind the IdP from knowing which RP is being accessed by the subscriber. In such situations, the pairwise pseudonymous identifier is generally established between the IdP and the federation proxy itself. The proxy, acting as an IdP, can itself provide pairwise pseudonymous identifiers to downstream RPs. Depending on the protocol, the federation proxy may need to map the pairwise pseudonymous identifiers back to the associated identifiers from upstream IdPs in order to allow the identity protocol to function. In such cases, the proxy will be able to track and determine which pairwise pseudonymous identifiers represent the same subscriber at different RPs. The proxy SHALL NOT disclose the mapping between the pairwise pseudonymous identifier and any other identifiers to a third party or use the information for any purpose other than federated authentication, related fraud mitigation, to comply with law or legal process, or in the case of a specific user request for the information.

6.3.2 Pairwise Pseudonymous Identifier Generation

Pairwise pseudonymous identifiers SHALL contain no identifying information about the subscriber. They SHALL also be unguessable by a party having access to some information identifying the subscriber. Pairwise pseudonymous identifiers MAY be generated randomly and assigned to subscribers by the IdP or MAY be derived from other subscriber information if the derivation is done in an irreversible, unguessable manner (e.g., using a keyed hash function with a secret key). Normally, the identifiers SHALL only be known by and used by one pair of endpoints (e.g., IdP-RP). However, an IdP MAY generate the same identifier for a subscriber at multiple RPs at the request of those RPs, provided:

- Those RPs have a demonstrable relationship that justifies an operational need for the correlation, such as a shared security domain or shared legal ownership; and
- All RPs sharing an identifier consent to being correlated in such a manner.

The RPs SHALL conduct a privacy risk assessment to consider the privacy risks associated with requesting a common identifier. See Section 9.2 for further privacy considerations.

The IdP SHALL ensure that only intended RPs are correlated; otherwise, a rogue RP could learn of the pseudonymous identifier for a set of correlated RPs by fraudulently posing as part of that set.

7 Assertion Presentation

This section is normative.

Assertions MAY be presented in either a *back-channel* or *front-channel* manner from the IdP to the RP. There are tradeoffs with each model, but each requires the proper validation of the assertion. Assertions MAY also be proxied to facilitate federation between IdPs and RPs under specific circumstances, as discussed in Section 5.1.4.

The IdP SHALL transmit only those attributes that were explicitly requested by the RP. RPs SHALL conduct a privacy risk assessment when determining which attributes to request.

7.1 Back-Channel Presentation

In the back-channel model, the subscriber is given an assertion reference to present to the RP, generally through the front channel. The assertion reference itself contains no information about the subscriber and SHALL be resistant to tampering and fabrication by an attacker. The RP presents the assertion reference to the IdP, usually along with authentication of the RP itself, to fetch the assertion.

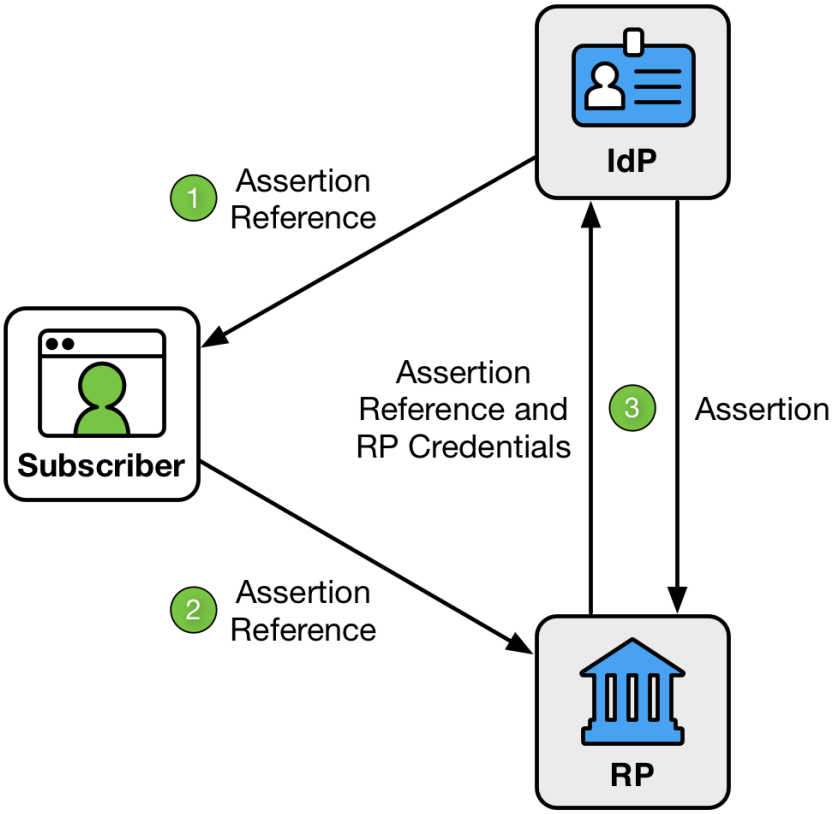

Figure 7-1 Back Channel Presentation

As shown in Figure 7-1, the back-channel presentation model consists of three steps:

1. The IdP sends an assertion reference to the subscriber through the front channel.

2. The subscriber sends the assertion reference to the RP through the front channel.
3. The RP presents the assertion reference and its RP credentials to the IdP through the back channel. The IdP validates the credentials and returns the assertion.

The assertion reference:

1. SHALL be limited to use by a single RP.
2. SHALL be single-use.
3. SHOULD be time limited with a short lifetime of seconds or minutes.
4. SHOULD be presented along with authentication of the RP.

In this model, the RP directly requests the assertion from the IdP, minimizing chances of interception and manipulation by a third party (including the subscriber themselves).

This method also allows the RP to query the IdP for additional attributes about the subscriber not included in the assertion itself, since back-channel communication can continue to occur after the initial authentication transaction has been completed without sending the user back to the IdP. This query occurs using an authorization component issued alongside the assertion, as described in Section 6.

More network transactions are required in the back-channel method, but the information is limited to only those parties that need it. Since an RP is expecting to get an assertion only from the IdP directly, the attack surface is reduced. Consequently, it is more difficult to inject assertions directly into the RP.

The RP SHALL protect itself against injection of manufactured or captured assertion references by use of cross-site scripting protection or other accepted techniques.

Elements within the assertion SHALL be validated by the RP, including:

- Issuer verification: ensuring the assertion was issued by the IdP the RP expects it to be from.
- Signature validation: ensuring the signature of the assertion corresponds to the key related to the IdP sending the assertion.
- Time validation: ensuring the expiration and issue times are within acceptable limits of the current timestamp.
- Audience restriction: ensuring this RP is the intended recipient of the assertion.

Conveyance of the assertion reference from the IdP to the subscriber, as well as from the subscriber to the RP, SHALL be made over an authenticated protected channel. Conveyance of the assertion reference from the RP to the IdP, as well as the assertion from the IdP to the RP, SHALL be made over an authenticated protected channel.

When assertion references are presented, the IdP SHALL verify that the party presenting the assertion reference is the same party that requested the authentication. The IdP can do this by requiring the RP to authenticate itself when presenting the assertion reference to the IdP or through other similar means (see RFC 7636 for one protocol's method of RP identification).

Note that in a federation proxy described in Section 5.1.4, the IdP audience restricts the assertion reference and assertion to the proxy, and the proxy restricts any newly-created assertion references or assertions to the downstream RP.

7.2 Front-Channel Presentation

In the front-channel model, the IdP creates an assertion and sends it to the subscriber after successful authentication. The assertion is used by the subscriber to authenticate to the RP, often through mechanisms within the subscriber's browser.

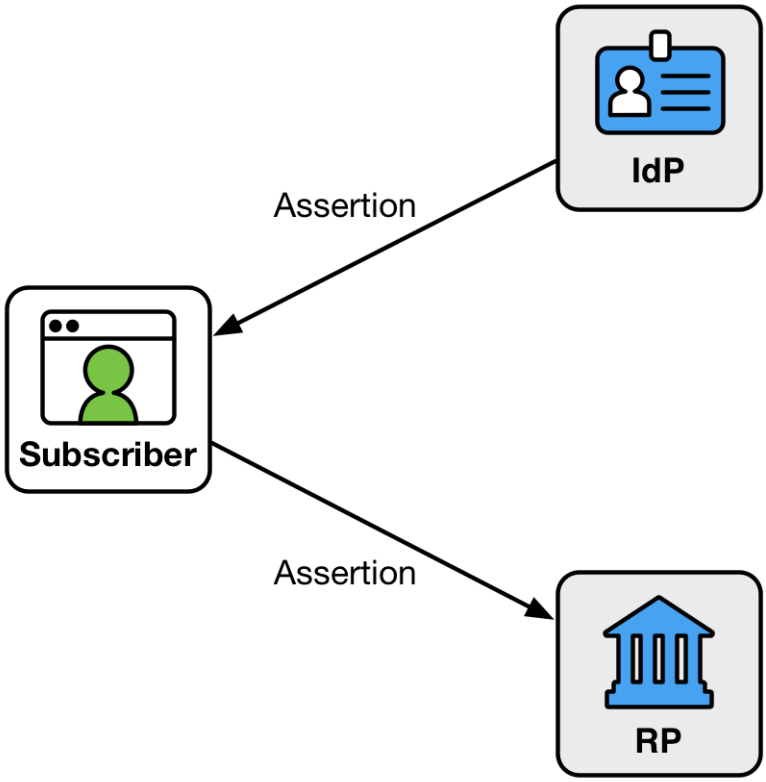

Figure 7-2 Front Channel Presentation

An assertion is visible to the subscriber in the front-channel method, which could potentially cause leakage of system information included in the assertion. Further, it is more difficult in this model for the RP to query the IdP for additional attributes after the presentation of the assertion.

Since the assertion is under the subscriber's control, the front-channel presentation method also allows the subscriber to submit a single assertion to unintended parties, perhaps by a browser replaying an assertion at multiple RPs. Even if the assertion is audience-restricted and rejected by unintended RPs, its presentation at unintended RPs could lead to leaking information about the subscriber and their online activities. Though it is possible to intentionally create an assertion designed to be presented to multiple RPs, this method can lead to lax audience restriction of the assertion itself, which in turn could lead to privacy and security breaches for the subscriber

across these RPs. Such multi-RP use is not recommended. Instead, RPs are encouraged to fetch their own individual assertions.

The RP SHALL protect itself against injection of manufactured or captured assertions by use of cross-site scripting protection or other accepted techniques.

Elements within the assertion SHALL be validated by the RP including:

- *Issuer verification*: ensuring the assertion was issued by the expected IdP.
- *Signature validation*: ensuring the signature of the assertion corresponds to the key related to the IdP making the assertion.
- *Time validation*: ensuring the expiration and issue times are within acceptable limits of the current timestamp.
- *Audience restriction*: ensuring this RP is the intended recipient of the assertion.

Conveyance of the assertion from the IdP to the subscriber, as well as from the subscriber to the RP, SHALL be made over an authenticated protected channel.

Note that in a federation proxy described in Section 5.1.4, the IdP audience restricts the assertion to the proxy, and the proxy restricts any newly-created assertions to the downstream RP.

7.3 Protecting Information

Communications between the IdP and the RP SHALL be protected in transit using an authenticated protected channel. Communications between the subscriber and either the IdP or the RP (usually through a browser) SHALL be made using an authenticated protected channel.

Note that the IdP may have access to information that may be useful to the RP in enforcing security policies, such as device identity, location, system health checks, and configuration management. If so, it may be a good idea to pass this information along to the RP within the bounds of the subscriber's privacy preferences described in Section 9.2.

Additional attributes about the user MAY be included outside of the assertion itself as part of a separate authorized request from the RP to the IdP. The authorization for access to these attributes MAY be issued alongside the assertion itself. Splitting user information in this manner can aid in protecting user privacy and allow for limited disclosure of identifying attributes on top of the essential information in the authentication assertion itself.

The RP SHALL, where feasible, request attribute references rather than full attribute values as described in Section 9.3. The IdP SHALL support attribute references.

8 Security

This section is informative.

Since the federated authentication process involves coordination between multiple components, including the CSP which now acts as an IdP, there are additional opportunities for attackers to compromise federated identity transactions. This section summarizes many of the attacks and mitigations applicable to federation.

8.1 Federation Threats

As in non-federated authentication, attackers' motivations are typically to gain access (or a greater level of access) to a resource or service provided by an RP. Attackers may also attempt to impersonate a subscriber. Rogue or compromised IdPs, RPs, user agents (e.g., browsers), and parties outside of a typical federation transaction are potential attackers. To accomplish their attack, they might intercept or modify assertions and assertion references. Further, two or more entities may attempt to subvert federation protocols by directly compromising the integrity or confidentiality of the assertion data. For the purpose of these types of threats, any authorized parties who attempt to exceed their privileges are considered attackers.

In some cases, the subscriber is issued some secret information so they can be recognized by the RP. Knowledge of this information distinguishes the subscriber from attackers who wish to impersonate them. In the case of holder-of-key assertions, this secret could have been established with the IdP prior to the initiation of the federation protocol.

Table 8-1 Federation Threats

Federation Threat/Attack	Description	Example
Assertion Manufacture or Modification	The attacker generates a false assertion	Compromised IdP asserts identity of a claimant who has not properly authenticated
	The attacker modifies an existing assertion	Compromised proxy that changes AAL of an authentication assertion
Assertion Disclosure	Assertion visible to third party	Network monitoring reveals subscriber address of record to an outside party
Assertion Repudiation by the IdP	IdP later claims not to have signed transaction	User engages in fraudulent credit card transaction at RP, IdP claims not to have logged them in

Federation Threat/Attack	Description	Example
Assertion Repudiation by the Subscriber	Subscriber claims not to have performed transaction	User agreement (e.g., contract) cannot be enforced
Assertion Redirect	Assertion can be used in unintended context	Compromised user agent passes assertion to attacker who uses it elsewhere
Assertion Reuse	Assertion can be used more than once with same RP	Intercepted assertion used by attacker to authenticate their own session
Assertion Substitution	Attacker uses an assertion intended for a different subscriber	Session hijacking attack between IdP and RP

8.2 Federation Threat Mitigation Strategies

Mechanisms that assist in mitigating the above threats are identified in Table 8-2.

Table 8-2 Mitigating Federation Threats

Federation Threat/Attack	Description	Example
Assertion Manufacture or Modification	Cryptographically sign the assertion at IdP and verify at RP	4.1, 6
	Send assertion over an authenticated protected channel authenticating the IdP	7.1, 7.2
	Include a non-guessable random identifier in the assertion	6.2.1
Assertion Disclosure	Send assertion over an authenticated protected channel authenticating the RP	7.1, 7.2
	Encrypt assertion for a specific RP (may be accomplished by use of a	6.2.3

Federation Threat/Attack	Description	Example
	mutually authenticated protected channel)	
Assertion Repudiation by the IdP	Cryptographically sign the assertion at the IdP with a key that supports non-repudiation; verify signature at RP	6.2.2
Assertion Repudiation by the Subscriber	Issue holder-of-key assertions; proof of possession of presented key verifies subscriber's participation	6.1.2
Assertion Redirect	Include identity of the RP ("audience") for which the assertion is issued in its signed content; RP verifies that they are intended recipient	6, 7.1, 7.2
Assertion Reuse	Include an issuance timestamp with short validity period in the signed content of the assertion; RP verifies validity	6, 7.1, 7.2

9 Privacy Considerations

This section is informative.

9.1 Minimizing Tracking and Profiling

Federation offers numerous benefits to RPs and subscribers, but requires subscribers to have trust in the federation participants. Sections 5, 5.1.4, and 6.3 cover a number of technical requirements, the objective of which is to minimize privacy risks arising from increased capabilities to track and profile subscribers.

For example, a subscriber using the same IdP to authenticate to multiple RPs allows the IdP to build a profile of subscriber transactions that would not have existed absent federation. The availability of such data makes it vulnerable to uses that may not be anticipated or desired by the subscriber and may inhibit subscriber adoption of federated services.

Section 5.2 requires CSPs to use measures to maintain the objectives of predictability (enabling reliable assumptions by individuals, owners, and operators about PII and its processing by an information system) and manageability (providing the capability for granular administration of PII, including alteration, deletion, and selective disclosure) commensurate with privacy risks that can arise from the processing of attributes for purposes other than identity proofing, authentication, authorization, or attribute assertion, related fraud mitigation, or to comply with law or legal process [NISTIR8062].

CSPs may have various business purposes for processing attributes, including providing non-identity services to subscribers. However, processing attributes for purposes other than the identity service can create privacy risks when individuals are not expecting or comfortable with the additional processing. CSPs can determine appropriate measures commensurate with the privacy risk arising from the additional processing. For example, absent applicable law, regulation or policy, it may not be necessary to get explicit consent when processing attributes to provide non-identity services requested by subscribers, although notices may help subscribers maintain reliable assumptions about the processing (predictability). Other processing of attributes may carry different privacy risks that call for obtaining explicit consent or allowing subscribers more control over the use or disclosure of specific attributes (manageability). Subscriber consent needs to be meaningful; therefore, when CSPs do use consent measures, they cannot make acceptance by the subscriber of additional uses a condition of providing the identity service.

Consult the SAOP if there are questions about whether the proposed processing falls outside the scope of the permitted processing or the appropriate privacy risk mitigation measures.

Section 5.2 also encourages the use of technical measures to provide disassociability (enabling the processing of PII or events without association to individuals or devices beyond the operational requirements of the system) and prevent subscriber activity tracking and profiling [NISTIR8062]. Technical measures, such as those outlined in Section 5.1.4 for proxied federation and Section 6.3 for pairwise pseudonymous identifiers, can increase the effectiveness

of policies by making it more difficult to track or profile subscribers beyond operational requirements.

9.2 Notice and Consent

To build subscriber trust in federation, subscribers need to be able to develop reliable assumptions about how their information is being processed. For instance, it can be helpful for subscribers to understand what information will be transmitted, which attributes for the transaction are required versus optional, and to have the ability to decide whether to transmit optional attributes to the RP. Accordingly, Section 7 requires that positive confirmation be obtained from the subscriber before any attributes about the subscriber are transmitted to any RP. In determining when a set of RPs should share a common pairwise pseudonymous identifier as in Section 6.3.2, the IdP considers the subscriber's understanding of such a grouping of RPs and the role of notice in assisting such understanding. An effective notice will take into account user experience design standards and research, as well as an assessment of privacy risks that may arise from the information processing. There are various factors to be considered, including the reliability of the assumptions subscribers may have about the processing and the role of different entities involved in federation. However, a link to a complex, legalistic privacy policy or general terms and conditions that a substantial number of subscribers do not read or understand is never an effective notice.

Section 7 does not specify which party should provide the notice. In some cases, a party in a federation may not have a direct connection to the subscriber in order to provide notice and obtain consent. Although multiple parties may elect to provide notice, it is permissible for parties to determine in advance, either contractually or through trust framework policies, which party will provide the notice and obtain confirmation, as long as the determination is being based upon factors that center on enabling the subscriber to pay attention to the notice and make an informed choice.

If an IdP is using a whitelist of RPs as described in Section 4.2, any RPs on that list are not presented to the subscriber during an authentication transaction. Since the IdP does not provide notice to the subscriber at runtime, the IdP makes its list of whitelisted RPs available to the subscriber so that the subscriber can see which RPs on the whitelist have access to which of the subscriber's attributes in an authentication transaction. Since IdPs cannot share a subscriber's authentication information or attributes with a whitelisted RP outside of an authentication transaction involving the subscriber (see Section 5.2), the existence of an RP on a list of IdPs does not indicate that the subscriber's information will be shared. However, if the subscriber logs into any of the whitelisted RPs using the IdP, the attributes indicated will be shared as part of the authentication transaction.

If a subscriber's runtime decisions were stored by the IdP to facilitate future transactions, the IdP also needs to allow the subscriber to view and revoke any RPs that were previously approved during a runtime decision. This list includes information on which attributes were approved.

9.3 Data Minimization

Federation enables the data exposed to an RP to be minimized — resultantly, the subscriber's privacy is enhanced. Although an IdP may collect additional attributes beyond what the RP

requires for its use case, only those attributes that were explicitly requested by the RP are to be transmitted by the IdP. In some instances, an RP does not require a full value of an attribute. For example, an RP may need to know whether the subscriber is over 13 years old, but has no need for the full date of birth. To minimize collection of potentially sensitive PII, the RP may request an attribute reference (e.g., Question: Is the subscriber over 13 years old? Response: Y/N or Pass/Fail). This minimizes the RP's collection of potentially sensitive and unnecessary PII. Accordingly, Section 7.3 requires the RP to, where feasible, request attribute references rather than full attribute values. To support this RP requirement IdPs are, in turn, required to support attribute references.

9.4 Agency-Specific Privacy Compliance

Section 5.2 identifies agency requirements to consult their SAOP to determine privacy compliance requirements. It is critical to involve the agency's SAOP in the earliest stages of digital authentication system development to assess and mitigate privacy risks and advise the agency on compliance obligations such as whether the federation triggers the Privacy Act of 1974 or the E-Government Act of 2002 requirement to conduct a PIA. For example, if the Agency is serving as an IdP in a federation, it is likely that the Privacy Act requirements will be triggered and require coverage by either a new or existing Privacy Act system of records since credentials would be maintained at the IdP on behalf of any RP it federates with. If, however, the agency is an RP and using a third-party IdP, digital authentication may not trigger the requirements of the Privacy Act, depending on what data passed from the RP is maintained by the agency as the RP (in such instances the agency may have a broader programmatic SORN that covers such data).

The SAOP can similarly assist the agency in determining whether a PIA is required. These considerations should not be read as a requirement to develop a Privacy Act SORN or PIA for use of a federated credential alone. In many cases it will make the most sense to draft a PIA and SORN that encompasses the entire digital authentication process or includes the digital authentication process as part of a larger programmatic PIA that discusses the program or benefit the agency is establishing online access.

Due to the many components of digital authentication, it is important for the SAOP to have an awareness and understanding of each individual component. For example, other privacy artifacts may be applicable to an agency offering or using federated IdP or RP services, such as Data Use Agreements, Computer Matching Agreements, etc. The SAOP can assist the agency in determining what additional requirements apply. Moreover, a thorough understanding of the individual components of digital authentication will enable the SAOP to thoroughly assess and mitigate privacy risks either through compliance processes or by other means.

9.5 Blinding in Proxied Federation

While some proxy structures — typically those that exist primarily to simplify integration — may not offer additional subscriber privacy protection, others offer varying levels of privacy to the subscriber through a range of blinding technologies. Privacy policies may dictate appropriate use of the subscriber attributes and authentication transaction data (e.g., identities of the ultimate IdP and RP) by the IdP, RP, and the federation proxy. Technical means such as blinding can

increase effectiveness of these policies by making the data more difficult to obtain. As the level of blinding increases, the technical and operational implementation complexity may increase. Proxies need to map transactions to the appropriate parties on either side as well as manage the keys for all parties in the transaction.

Even with the use of blinding technologies, a blinded party may still infer protected subscriber information through released attribute data or metadata, such as by analysis of timestamps, attribute bundle sizes, or attribute signer information. The IdP could consider additional privacy-enhancing approaches to reduce the risk of revealing identifying information of the entities participating in the federation.

The following table illustrates a spectrum of blinding implementations used in proxied federation. This table is intended to be illustrative, and is neither comprehensive nor technology-specific.

Table 9-1 Federation Proxies

Proxy Type	RP Knows IdP	IdP Knows RP	Proxy can Track Subscriptions between RP and IdP	Proxy Can See Attributes of Subscriber
Non-Blinding Proxy with Attributes	Yes	Yes	Yes	Yes
Non-Blinding Proxy without Attributes	Yes	Yes	Yes	N/A
Double Blind Proxy with Attributes	No	No	Yes	Yes
Double Blind Proxy without Attributes	No	No	Yes	N/A
Triple Blind Proxy with or without Attributes	No	No	No	No

10 Usability Considerations

This section is informative.

ISO/IEC 9241-11 defines usability as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use." This definition focuses on users, goals, and context of use as key elements necessary for achieving effectiveness, efficiency and satisfaction. A holistic approach considering these key elements is necessary to achieve usability.

From the usability perspective, one of the major potential benefits of federated identity systems is to address the problem of user fatigue associated with managing multiple authenticators. While this has historically been a problem with usernames and passwords, the increasing need for users to manage many authenticators — whether physical or digital — presents a usability challenge.

While many other approaches to authentication have been researched extensively and have well-established usability guidelines, federated identity is more nascent and, therefore, lacks the depth and conclusiveness of research findings. As ongoing usability research matures, usability guidelines for federated identity systems will have stronger supporting data. For example, additional data is needed to support guidance on the translation of technical attribute names and values into user-friendly language.

As stated in the usability sections in 800-63A and 800-63B, overall user experience is critical to the success of any authentication method. This is especially true for federated identity systems as federation is a less familiar user interaction paradigm for many users. Users' prior authentication experiences may influence their expectations.

The overall user experience with federated identity systems should be as smooth and easy as possible. This can be accomplished by following usability standards (such as the ISO 25060 series of standards) and established best practices for user interaction design.

ASSUMPTIONS

In this section, the term "users" means "claimants" or "subscribers." The terms "entity" and "entities" refer to the parties of federated systems.

Guidelines and considerations are described from the users' perspective.

Accessibility differs from usability and is out of scope for this volume. Section 508 was enacted to eliminate barriers in information technology and requires federal agencies to make their electronic and information technology public content accessible to people with disabilities. Refer to Section 508 law and standards for accessibility guidance.

10.1 General Usability Considerations

Federated identity systems should:

- Minimize user burden (e.g., frustration, learning curve)

    - Minimize the number of user actions required.
    - Allow users to quickly and easily select among multiple accounts with a single IdP. For example, approaches such as Account Chooser allow users to select from a list of accounts they have accessed in the recent past, rather than start the federation process by selecting their IdP from a list of potential IdPs.
    - Balance minimizing user burden with the need to provide sufficient information to enable users to make informed decisions.
- Minimize the use of unfamiliar technical jargon and details (e.g., users do not need to know the terms IdP and RP if the basic concepts are clearly explained).
- Strive for a consistent and integrated user experience across the IdP and RP.
- Help users establish an understanding of identity by providing resources to users such as graphics, illustrations, FAQs, tutorials and examples. Resources should explain how users' information is treated and how transacting parties (e.g., RPs, IdPs, and brokers) relate to each other.
- Provide clear, honest, and meaningful communications to users (i.e., communications should be explicit and easy to understand).
- Provide users online services independent of location and device.
- Make trust relationships explicit to users to facilitate informed trust decisions. Trust relationships are often dynamic and context dependent. Users may be more likely to trust some IdPs and RPs with certain attributes or transactions more than others. For example, users may be more hesitant to use federated identity systems on websites that contain valuable personal information (such as financial or health). Depending on the perceived sensitivity of users' personal data, users may be less comfortable with social network providers as IdPs since people are often concerned with the broadcasting nature of social networking implementations.
- Follow the usability considerations specified in SP 800-63A, Section 9 for any user-facing information.
- Clearly communicate how and where to acquire technical assistance. For example, provide users with information such as a link to an online self-service feature, chat sessions or a phone number for help desk support. Avoid redirecting users back and forth among transacting parties (e.g., RPs, IdPs, and brokers) to receive technical assistance.
- Perform integrative and continuous usability evaluations with representative users and realistic tasks in an appropriate context to ensure success of federated identity systems from the users' perspectives.

10.2 Specific Usability Considerations

This section addresses the specific usability considerations that have been identified with federated identity systems. This section does not attempt to present exhaustive coverage of all usability factors related to federated identity systems. Rather, it is focused on the larger, more pervasive themes in the usability literature, primarily users' perspectives on identity, user adoption, trust, and perceptions of federated identity space. In some cases, implementation examples are provided. However, specific solutions are not prescribed. The implementations mentioned are examples to encourage innovative technological approaches to address specific usability needs. See standards for system design and coding, specifications, APIs, and current

best practices (such as OpenID and OAuth) for additional examples. Implementations are sensitive to many factors that prevent a one-size-fits-all solution.

10.2.1 User Perspectives on Online Identity

Even when users are familiar with federated identity systems, there are different approaches to federated identity (especially in terms of privacy and the sharing of information) that make it necessary to establish reliable expectations for how users' data are treated. Users and implementers have different concepts of identity. Users think of identity as logging in and gaining access to their own private space. Implementers think of identity in terms of authenticators and assertions, assurance levels, and the necessary set of identity attributes to provide a service. Given this disconnect between users' and implementers' concepts of identity, it is essential to help users form an accurate concept of identity as it applies to federated identity systems. A good model of identity provides users a foundation for understanding the benefits and risks of federated systems and encourage user adoption and trust of these systems.

Many properties of identity have implications for how users manage identities, both within and among federations. Just as users manage multiple identities based on context outside of cyberspace, users must learn to manage their identity in a federated environment. Therefore, it must be clear to users how identity and context are used. The following factors should be considered:

- Provide users the requisite context and scope in order to distinguish among different user roles. For example, whether the user is acting on their own behalf or on behalf of another, such as their employer.
- Provide users unique, meaningful, and descriptive identifiers to distinguish among entities.
- Provide users with information on data ownership and those authorized to make changes. Identities, and the data associated with them, can sometimes be updated and changed by multiple actors. For example, some healthcare data is updated and owned by the patient, while some data is only updated by a hospital or doctor's practice.
- Provide users with the ability to easily verify, view, and update attributes. Identities and user roles are dynamic and not static; they change over time (e.g., age, health, and financial data). The ability to update attributes or make attribute release decisions may or may not be offered at the same time. Ensure the process for how users can change attributes is well known, documented, and easy to perform.
- Provide users means for updating data, even if the associated entity no longer exists.
- Provide users means to delete their identities completely, removing all information about themselves, including transaction history. Consider applicable audit, legal, or policy constraints that may preclude such action. In certain cases, full deactivation is more appropriate than deletion.
- Provide users with clear, easy-to-find, site/application data retention policy information.
- Provide users with appropriate anonymity and pseudonymity options, and the ability to switch among such identity options as desired, in accordance with an organization's data access policies.
- Provide means for users to manage each IdP to RP connection, including complete separation as well as the removal of RP access to one or more attributes.

10.2.2 User Perspectives of Trust and Benefits

Many factors can influence user adoption of federated identity systems. As with any technology, users may value some factors more than others. Users often weigh perceived benefits versus risks before making technology adoption decisions. It is critical that IdPs and RPs provide users with sufficient information to enable them to make informed decisions. The concepts of trust and tiers of trust — fundamental principles in federated identity systems — can drive user adoption. Finally, a positive user experience may also result in increased user demand for federation, triggering increased adoption by RPs.

This sub-section is focused primarily on user trust and user perceptions of benefits versus risks.

To encourage user adoption, IdPs and RPs need to establish and build trust with users and provide them with an understanding of the benefits and risks of adoption. The following factors should be considered:

- Allow users to control their information disclosure and provide explicit consent through the appropriate use of notifications (see SP 800-63C, Section 9.2, Notice and Consent). Balancing the content, size, and frequency of notifications is necessary to avoid thoughtless user click-through.
- For attribute sharing, consider the following:
    - Provide a means for users to verify those attributes and attribute values that will be shared. Follow good security practices (see Section 7).
    - Enable users to consent to a partial list of attributes, rather than an all-or-nothing approach. Allow users some degree of online access, even if the user does not consent to share all information.
    - Allow users to update their consent to their list of shared attributes.
    - Minimize unnecessary information presented to users. For example, do not display system generated attributes (such as pairwise pseudonymous identifiers) even if they are shared with the RP as part of the authentication response.
    - Minimize user steps and navigation. For example, build attribute consent into the protocols so they're not a feature external to the federated transaction. Examples can be found in standards such as OAuth or OpenID Connect.
    - Provide effective and efficient redress methods such that a user can recover from invalid attribute information claimed by the IdP (see Section 7).
    - Minimize the number of times a user is required to consent to attribute sharing. Limiting the frequency of consent requests avoids user frustration from multiple requests to share the same attribute.
- Collect information for constrained usage only, and minimize information disclosure (see Section 9.3). User trust is eroded by unnecessary and superfluous information collection and disclosure or user tracking without explicit user consent. For example, only request attributes from the user that are relevant to the current transaction, not for all possible transactions a user may or may not access at the RP.
- Clearly and honestly communicate potential benefits and risks of using federated identity to users. Benefits that users value include time savings, ease of use, reduced number of passwords to manage, and increased convenience.

User concern over risk can negatively influence willingness to adopt federated identity systems. Users may have trust concerns, privacy concerns, security concerns, and single-point-of-failure concerns. For example, users may be fearful of losing access to multiple accounts if a single IdP is unavailable, either temporarily or permanently. Additionally, users may be concerned or confused about learning a new authentication process. In order to foster the adoption of federated identity systems, the perceived benefits must outweigh the perceived risks.

10.2.3 User Models and Beliefs

Users' beliefs and perceptions predispose them to expect certain results and to behave in certain ways. Such beliefs, perceptions, and predispositions are referred to in the social sciences as mental models. For example, people have a mental model of dining out which guides their behavior and expectations at each establishment, such as fast food restaurants, cafeterias, and more formal restaurants. Thus, it is not necessary to be familiar with every establishment to understand how to interact appropriately at each one.

Assisting users in establishing good and complete mental models of federation allows users to generalize beyond a single specific implementation. If federated identity systems are not designed from users' perspectives, users may form incorrect or incomplete mental models that impact their willingness to adopt these systems. The following factors should be considered:

- Clearly explain the working relationship and information flow among the transacting parties (e.g., RPs, IdPs, and brokers) to avoid user misconceptions. Use the actual names of the entities in the explanation rather than using the generic terms IdPs and RPs.
    - Provide prominent visual cues and information so that users understand why seemingly unrelated entities have a working relationship. For example, users may be concerned with mixing online personal activities with government services due to a lack of understanding of the information flow in federated identity systems.
    - Provide prominent visual cues and information to users about redirection when an RP needs to redirect control from their site to an IdP. For example, display RP branding within the IdP user interface to inform users when they are logging in with their IdP for access to the destination RP.
- Provide users with clear and usable ways (e.g., visual assurance) to determine the authenticity of the transacting parties (e.g., RPs, IdPs, and brokers). This will also help to alleviate user concern over leaving one domain for another, especially if the root domain changes (e.g., .gov to .com). For example, display the URL of the IdP so that the user can verify that they are not being phished by a malicious site.
- Provide users with clear information, including visual cues, regarding implicit logins and explicit logouts. Depending on the implementation, logging into an RP with an IdP account may authenticate users to both the IdP and RP. Users may not realize that ending their session with the RP will not necessarily end their session with the IdP; users will need to explicitly "log out" of the IdP. Users require clear information to remind them if explicit logouts are required to end their IdP sessions.

11 Examples

This section is informative.

Three types of assertion technologies are discussed below: SAML assertions, Kerberos tickets, and OpenID Connect tokens. This list is not inclusive of all possible assertion technologies, but does represent those commonly used in federated identity systems.

11.1 Specific Assertion Markup Language (SAML)

SAML is an XML-based framework for creating and exchanging authentication and attribute information between trusted entities over the internet. As of this writing, the latest specification for SAML is SAML v2.0, issued 15 March 2005.

The building blocks of SAML include:

- The Assertions XML schema, which defines the structure of the assertion.
- The SAML Protocols, which are used to request assertions and artifacts (the assertion references used in the indirect model described in Section 7.1).
- The Bindings, which define the underlying communication protocols (such as HTTP or SOAP), and can be used to transport the SAML assertions.

The three components above define a SAML profile that corresponds to a particular use case such as “Web Browser SSO”.

SAML Assertions are encoded in an XML schema and can carry up to three types of statements:

- *Authentication statements* include information about the assertion issuer, the authenticated subscriber, validity period, and other authentication information. For example, an Authentication Assertion would state the subscriber “John” was authenticated using a password at 10:32pm on 06-06-2004.
- *Attribute statements* contain specific additional characteristics related to the subscriber. For example, subject “John” is associated with attribute “Role” with value “Manager”.
- *Authorization statements* identify the resources the subscriber has permission to access. These resources may include specific devices, files, and information on specific web servers. For example, subject “John” for action “Read” on “Webserver1002” given evidence “Role”.

Authorization statements are beyond the scope of this document and will not be discussed.

11.2 Kerberos Tickets

The Kerberos Network Authentication Service [RFC 4120] was designed to provide strong authentication for client/server applications using symmetric-key cryptography on a local, shared network. Extensions to Kerberos can support the use of public key cryptography for selected steps of the protocol. Kerberos also supports confidentiality and integrity protection of session data between the subscriber and the RP. Even though Kerberos uses assertions, it was designed for use on shared networks and, therefore, is not truly a federation protocol.

Kerberos supports authentication of a subscriber over an untrusted, shared local network using one or more IdPs. The subscriber implicitly authenticates to the IdP by demonstrating the ability to decrypt a random session key encrypted for the subscriber by the IdP. (Some Kerberos variants also require the subscriber to explicitly authenticate to the IdP, but this is not universal.) In addition to the encrypted session key, the IdP also generates another encrypted object called a Kerberos ticket. The ticket contains the same session key, the identity of the subscriber to whom the session key was issued, and an expiration time after which the session key is no longer valid. The ticket is confidentiality and integrity protected by a pre-established key that is shared between the IdP and the RP during an explicit setup phase.

To authenticate using the session key, the subscriber sends the ticket to the RP along with encrypted data that proves that the subscriber possesses the session key embedded within the Kerberos ticket. Session keys are either used to generate new tickets or to encrypt and authenticate communications between the subscriber and the RP.

To begin the process, the subscriber sends an authentication request to the Authentication Server (AS). The AS encrypts a session key for the subscriber using the subscriber's long-term credential. The long-term credential may either be a secret key shared between the AS and the subscriber, or in the PKINIT variant of Kerberos, a public key certificate. Most variants of Kerberos based on a shared secret key between the subscriber and IdP derive this key from a user-generated password. As such, they are vulnerable to offline dictionary attacks by passive eavesdroppers, unless Flexible Authentication Secure Tunneling (FAST) [RFC 6113] or some other tunneling and armoring mechanism is used.

In addition to delivering the session key to the subscriber, the AS also issues a ticket using a key it shares with the Ticket Granting Server (TGS). This ticket is referred to as a Ticket Granting Ticket (TGT), since the verifier uses the session key in the TGT to issue tickets rather than to explicitly authenticate the verifier. The TGS uses the session key in the TGT to encrypt a new session key for the subscriber and uses a key it shares with the RP to generate a ticket corresponding to the new session key. The subscriber decrypts the session key and uses the ticket and the new session key together to authenticate to the RP.

When Kerberos authentication is based on passwords, the protocol is known to be vulnerable to offline dictionary attacks by eavesdroppers who capture the initial user-to-KDC exchange. Longer password length and complexity provide some mitigation to this vulnerability, although sufficiently long passwords tend to be cumbersome for users. However, when Kerberos password-based authentication is used in a FAST (or similar) tunnel, a successful Man-in-the-Middle attack is additionally required in order to perform the dictionary attack.

11.3 OpenID Connect

OpenID Connect [OIDC] is an internet-scale federated identity and authentication protocol built on top of the OAuth 2.0 authorization framework and the JSON Object Signing and Encryption (JOSE) cryptographic system.

OpenID Connect builds on top of the OAuth 2.0 authorization protocol to enable the subscriber to authorize the RP to access the subscriber's identity and authentication information. The RP in both OpenID Connect and OAuth 2.0 is known as the client.

In a successful OpenID Connect transaction, the IdP issues an ID Token, which is a signed assertion in JSON Web Token (JWT) format. The client parses the ID Token to learn about the subscriber and primary authentication event at the IdP. This token contains at minimum the following information about the subscriber and authentication event:

- iss - An HTTPS URL identifying the IdP that issued the assertion.
- sub - An IdP-specific subject identifier representing the subscriber.
- aud - An IdP-specific audience identifier, equal to the OAuth 2.0 client identifier of the client at the IdP.
- exp - The timestamp at which the ID Token expires and after which SHALL NOT be accepted the client.
- iat - The timestamp at which the ID Token was issued and before which SHALL NOT be accepted by the client.

In addition to the ID Token, the IdP also issues the client an OAuth 2.0 access token which can be used to access the UserInfo Endpoint at the IdP. This endpoint returns a JSON object representing a set of attributes about the subscriber, including but not limited to their name, email address, physical address, phone number, and other profile information. While the information inside the ID Token is reflective of the authentication event, the information in the UserInfo Endpoint is generally more stable and could be more general purpose. Access to different attributes from the UserInfo Endpoint is governed by the use of a specially-defined set of OAuth scopes, openid, profile, email, phone, and address. An additional scope, offline_access, is used to govern the issuance of refresh tokens, which allow the RP to access the UserInfo Endpoint when the subscriber is not present. Access to the UserInfo Endpoint is structured as an API and may be available when the subscriber is not present. Therefore, access to the UserInfo Endpoint is not sufficient for proving a subscriber's presence and establishing an authenticated session at the RP.