Asserting ePPN Across the Gateway

Asserting ePPN Across the Gateway

It is generally recognized that asserting scoped attributes across a gateway is problematic. Social gateways are particularly troublesome since few social IdPs assert an attribute that maps naturally to `eduPersonPrincipalName`, which is a scoped attribute known to be required by many RPs in the R&E space.

Executive Summary

1. The user's email address is a poor choice for `eduPersonPrincipalName` asserted by a gateway.
2. The OpenID Connect subject identifier (sub) more accurately maps to `eduPersonTargetedID` or `eduPersonUniqueId`, not `eduPersonPrincipalName`.
3. For a social gateway, the recommended value of `eduPersonPrincipalName` is

   ```
   user+domain1@social_idp.domain2
   ```

   where `user@domain1` is the email address of the user, `social_idp` is the name of the social provider, and `domain2` is a domain owned by the organization that owns and operates the gateway.

Introduction

It is well known that `eduPersonPrincipalName` (`ePPN`) is a globally unique, persistent identifier for the user. For level-setting purposes, we begin with the following facts about persistent identifiers and scoped attributes.

Persistent Identifiers

- **Definition**. A *persistent identifier* for the user is one that spans multiple SSO sessions.
- Although `ePPN` is a persistent identifier, it is not intended to be permanent. Relying parties certainly prefer that `ePPN` remain stable but users can and do change their `ePPN` for a variety of reasons.
- Although *non-reassignment* is a highly desirable property of **any** persistent identifier, `ePPN` deployments are not guaranteed to be non-reassigned (but often are since it is understood that a persistent, non-reassigned identifier is more valuable than one that is not).

Scoped Attributes

- `ePPN` is the primary example of a *scoped attribute*.
- `ePPN` is globally unique by virtue of its *scope*, which by convention is a DNS name.
- The scope part of a scoped attribute indicates the asserting authority. This is why a scope is a DNS name by convention.
- A trusted third party (such as a federation) ensures that the scopes listed in metadata are rooted in registered domains owned by the organization deploying the IdP.
- Normally an IdP asserts a scoped attribute with a scope part for which the IdP is authoritative. Likewise an SP filters scoped attributes for which the IdP is not authoritative, at least by default.

Email Address as ePPN?

If you were an enterprise architect designing an identity management system from scratch, it would be in your best interest to define `ePPN` such that it was a routable email address. There are many reasons for this, not the least of which is the fact that SaaS services invariably use email address as a user ID.

That said, when mapping attributes across a social gateway, resist the urge to map the user's email address to `ePPN`, even if the social IdP asserts email addresses known not to be re-assigned. Why? Because the right hand side of an email address asserted by a social IdP can be just about anything, which forces the RP to accept practically any scope from the corresponding gateway. That totally defeats the purpose of scoped attributes.

Consider Google, for example. Since a Google Apps subscriber provisions local email addresses in Google Apps (e.g., `user@university.edu`), the Google IdP will assert arbitrary email addresses (not just `@gmail.com` email addresses). Thus mapping email address to `ePPN` is quite possibly the worst thing you could do.

The OIDC Sub Claim as ePPN?

The OpenID Connect (OIDC) subject identifier (`sub`) is an opaque, non-reassigned identifier for the user, scoped to the issuer. Coupled with the OIDC issuer identifier (`iss`) the `sub` claim is a stable, globally unique identifier for the user. As such the `sub` claim closely aligns with the SAML2 Persistent NameID or the equivalent `eduPersonTargetedID` attribute. Unlike `eduPersonTargetedID`, however, the `sub` claim is not a targeted identifier.

The `sub` claim is also closely aligned with `eduPersonUniqueId`. The latter, however, is a scoped attribute, which leads to complications. The obvious choice of scope value is `@google.com` but this scope MUST NOT be asserted in gateway metadata. An RP would have to carefully configure the handling of scope `@google.com` in its SAML software. To make matters worse, `eduPersonUniqueId` is new and not widely deployed, so one should expect little support for it in existing SAML implementations.

Finally, mapping the `sub` claim to `ePPN` is least desirable for the following reasons:

- Both `eduPersonTargetedID` and `eduPersonUniqueId` are better suited to carry the `sub` claim.

- Like `eduPersonUniqueId`, `ePPN` is a scoped attribute, with all the same problems.
- The RP does not expect the left hand side of `ePPN` to be opaque.

All in all, the `sub` claim is perhaps best mapped to `eduPersonTargetedID`.

Best Practices for Gateway ePPNs

For a social gateway, the recommended value of `ePPN` is:

```
user+domain1@social_idp.domain2
```

where `user@domain1` is the email address of the user, `social_idp` is the name of the social provider, and `domain2` is a domain owned by the organization that owns and operates the gateway. For example, my `ePPN` asserted by the Internet2 Google Gateway is:

```
trscavo+gmail.com@google.incommon.org
```

Since Internet2 is a Google Apps for Education campus, I am also known as:

```
trscavo+internet2.edu@google.incommon.org
```

since Google will readily assert my Internet2 email address `trscavo@internet2.edu` if I happen to log in via the Internet2 IdP.

What happens when an RP that has been using a central gateway chooses to run its own local gateway? In that case, a migration will be necessary since the scope on the `ePPN` will no doubt change. Thus the best choice of scope in the first place is a stable value that won't change over time, regardless of who owns and operates the gateway.

For example, consider the Internet2 Google Gateway again. The scope `@google.incommon.org` was chosen because:

1. Internet2 owns the registered domain `incommon.org`, which is a required characteristic of all scopes in metadata.
2. The subdomain `google.incommon.org` makes it easy for Internet2 to support other social providers if and when the time comes. (For instance, a Facebook Gateway would have scope `@facebook.incommon.org`.)
3. If the Internet2 Google Gateway were promoted to a centrally-run gateway for other (non-Internet2) services, the scope would not have to change.

Now observe that the Internet2 service `wiki.shibboleth.net` does **not** use the Internet2 Google Gateway today…but it could. The best scope for this particular service would be `@google.shibboleth.net` since then the service could easily migrate to its own gateway in the future if desired. We could modify the existing gateway implementation to assert an `ePPN` with scope `@google.shibboleth.net` but we wouldn't be able to assert that scope in metadata since Internet2 does not own the registered domain `shibboleth.net`. In that case, the SAML software protecting `wiki.shibboleth.net` would have to be locally reconfigured to accept scoped attributes of the form `value@google.shibboleth.net` from the Internet2 Google Gateway.