Organizational Identity Sources

About Organizational Identity Sources

Organizational Identity Sources allow for the creation of Organizational Identities linked to an external source or "system of record". These sources can include LDAP servers, REST APIs, SQL databases, flat files, and so on. Custom plugins can be written for arbitrary sources.

Organizational Identity Sources can only be defined on a per-CO basis. If org identities are pooled, Organizational Identity Sources are not supported. Once configured, Organizational Identities can be created from these sources in several ways:

- Manually, via *People >> Organizational Identities >> Add New Org Identity From Source* or by using the *Search* button from the list of Organizational Identity Sources.
- Using an Enrollment Flow, via Enrollment Sources.
- Via the Registry Job Shell, as described below.

Organizational Identity Sources can be linked to Registry Pipelines in order to automatically create CO Person records.

⚠️ When an Organizational Identity is created from a source, it is linked to that source and cannot be manually edited, not even by an administrator. However, it can be manually resynced to pull changes from the source.

⚠️ If the corresponding record is removed from the Organizational Identity Source, on the next sync the Org Identity will be set to status *Removed*, but the Org Identity itself will remain available – it is not deleted.

⚠️ If Attribute Enumerations are enabled for any attributes, permitted values for those attributes are constrained to the enumerated options. Source records containing a non-enumerated value will fail to process correctly.

Organizational Identity Sources are available in COmanage Registry v2.0.0 and later.

Terminology

The terminology used by Registry can be a little confusing when looking at person records related to Organizational Identity Sources.

- **View Organizational Identity**: Retrieves the current Org Identity operational record used by Registry in normal operations.
- **View Organizational Identity Source**: Performs a live query against the Org Identity Source backend and retrieves the current data as known to the backend. ie: This is the source's current data.
- **View Organizational Identity Source Record**: Retrieves the last data retrieved from the backend and used to create or update an Org Identity. ie: This is Registry's copy of the source data.
- **Add New Org Identity From Source**: Create a new Org Identity based on the Org Identity Source's data. In addition, this will create an Organizational Identity Source Record.
- **Resync Org Identity From Source**: Update the Org Identity and Organizational Identity Source Record using the latest (current) data available from the Org Identity Source.
- **Configuration >> Organizational Identity Sources**: Manage the plugins used to define and query one or more Org Identity Sources.

Sync Modes

Job Shell

Organizational Identity Sources can be configured in one of several sync modes to allow periodic automatic synchronization of records via the Registry Job Shell. Prior to Registry v4.0.0, this is handled with the `syncorgsources` and `forcesyncorgsources` tasks. As of Registry v4.0.0, this is handled via the Core Job Plugin Sync job.

- **Full**: Create new Org Identities for any record in the Organizational Identity Source that does not yet have one, and update (or delete, if appropriate) existing records.
- **Query**: Similar to Enrollment Sources Search mode, query the Organizational Identity Source for any records matching verified email addresses of all Org Identities, looking for new matching records to link. Also update (or delete, if appropriate) existing records.
  - ⚠️ Query mode should only be used for Organizational Identity Sources attached to a Registry Pipeline configured for email address-based matching. Otherwise, linking to existing CO People may not happen correctly.

    - In Query mode, if a Organizational Identity Source is queried for an email address and the Source returns a record with a different email address (eg: the person changed their email address in the other system), by default a new Org Identity (and probably CO Person) will be created. This is because Registry has not confirmed the alternate email address and cannot trust the Organizational Identity Source asserting a record linkage. This corresponds to the **Email Mismatch Mode** of **Create New Org Identity**. Alternately, Email Mismatch Mode can be set to **Ignore**, in which case no action is taken.
    - In Query mode, by default the Organizational Identity Source will be re-queried for all email addresses, even those already attached to an Org Identity associated with the Source. This is to allow for the checking of additional records associated with the same email address. However, this can also create a large number of extra queries, if the Source is known not to create such records (or if such records are not of interest). To disable this behavior, set (tick the box for) **Do Not Query for Known Email Addresses**.
- **Update**: Update and delete (if appropriate) records that are already synced to Org Identities.
    - The `syncorgsources` task will skip any unchanged records, while the `forcesyncorgsources` task will process records even when they are unchanged. `forcesyncorgsources` is available as of Registry v3.2.0. As of Registry v4.0.0, it is accessed via the `--force` flag of the CoreJob Sync job.
    - ⚠️ When the `forcesyncorgsources` task is run, the getChangeList() capability supported by some plugins cannot be used. This may result in significantly longer processing times.
- **Manual**: Do not automatically sync records. Currently, manual syncing is only available on an individual record basis. (CO-1309)

⚠️ Not all Organizational Identity Source plugins support all sync modes. Check the documentation for any limitations.

Syncing via Job Shell can be disabled on a per-CO basis via *CO Settings >> Disable Org Identity Source Sync*.

Login

As of Registry v3.1.0, Org Identities associated with an Organizational Identity Source can be resynced on user login to Registry. This is enabled on a per-Organizational Identity Source basis, by enabling the *Sync on Login* setting.

Because user login crosses COs, when a user logs in the identifier they used to login will be searched against all Organizational Identities (regardless of CO) for matching CO Person records. Then, any Org Identity associated with those CO Person records will be checked for an associated Organizational Identity Source for the *Sync on Login* setting. In other words, a login event will resync *any* suitably configured record, not just the one associated with the identifier used to login.

> ⚠️ If external databases are configured as Organizational Identity Sources to sync during login, users may experience login delays related to querying those databases.

> ⚠️ Sync on login is not supported when Organizational Identities are pooled. Unexpected results may occur.

Linking a Record to a CO Person

By default, creating an Org Identity (via *Add New Org Identity From Source* or any other mechanism) will not create a CO Person.

If the Org Identity Source is attached to a Pipeline, then that Pipeline will likely create a CO Person for the new Org Identity. If a Pipeline Match Strategy is configured, then the Pipeline may attach the new Org Identity to an existing CO Person if the match conditions are satisfied.

To manually link an Org Identity to an existing CO Person, there are two options:

1. If no Pipeline is attached to the Org Identity Source, simply link the record manually.
2. Define an Enrollment Flow. A typical configuration would be
    a. Authorization: CO Admin (or COU Admin)
    b. Identity Matching: Select
    c. Attach an appropriate Enrollment Source, in Select mode
    d. Do not define any Enrollment Attributes

Creating ePPNs

When syncing records from an Org Identity Source, Registry can automatically create an identifier of type ePPN to be injected into the Org Identity created from the Source. This can be useful for (eg) automatically calculating the ePPN of an IdP associated with the Source. There are two settings:

- **EPPN Identifier Type**: The Identifier of this type as found in the Org Identity created from the Source will be used as the left-hand side of the newly created ePPN.
- **EPPN Suffix**: The specified string will be used as the right-hand side of the newly created ePPN. Do not include the `@`.

An ePPN will not be generated if one is found in the Org Identity record created from the Source.

CO Group Mappings

Organizational Identity Sources can generate CO Group Memberships via *Group Mappings*, when the relevant OIS Plugin implements the appropriate interfaces. However, since group memberships attach to a CO Person and not an Organizational Identity, for this to be useful the OIS must typically be attached to a Pipeline, which will then create CO Group Memberships attached to the relevant CO Person record. For OIS Plugins that support this feature, the steps to enable it are:

1. Make sure the group(s) you want to add memberships to already exists. Automatic groups (such as *members* groups) cannot be used here.
2. View the OIS configuration (ie: use the *Edit* button, not the *Configure* button), and click *Configure Group Mapping*.
    a. ℹ️ In Registry versions prior to v3.1.0, the *Configure Group Mapping* button is available on the index page of Organizational Identity Sources.
    b. The OIS must be connected to a Pipeline for CO Group Memberships to be assigned.
3. Add one or more mappings.
    a. *Attribute*: The attribute found in the Organizational Identity Source.
    b. *Target Group*: The group for which a membership will be created, when the Org Identity has an *Attribute* matching the specified *Comparison* and *Pattern*.
4. At this point, if you search the Organizational Identity Source, any records matching the defined mappings will also show what CO Group memberships would be assigned if a CO Person record were created from or attached to this source record. However, an action triggering the Pipeline (as described above) must take place. ⚠️ This means group memberships for existing records will not be assigned until there is a manual sync or a change in the source record.

ℹ️ If a CO Person already has a CO Group Membership (either manually created or from another Organizational Identity Source), a new membership will not be created.

⚠️ Manually rerunning a Pipeline will not correctly recalculate group memberships, as the original source record is not available for processing. Completely resyncing the Org Identity from Source (which will in turn rerun the Pipeline) will work correctly.

Storing Cached Source Records

When a record has been synced to Registry from the Organizational Identity Source, a cached copy is stored in the Organizational Identity Source Record so that Registry may detect when the source record has been updated. By default, this is a full copy of the record as returned to Registry from the backend (in whatever format is returned from the source, eg JSON, XML, etc). This is also useful for tracing problems, as it is possible for an administrator to look at a cached copy of the data.

However, there may be privacy or data retention concerns that make storing a full copy less desirable for a given deployment. As an alternative, Registry can create a hash of the data to be stored instead. This can be enabled via the *Hash Source Records* configuration option.

When this configuration is changed, existing records are *not* affected. Furthermore, since the cached copy will no longer match the current source record, all records from the Source will be considered out of date the next time a sync is performed. It is best to determine the appropriate value for this setting prior to significant production usage.

An additional consideration when enabling *Hash Source Records* for privacy or data retention reasons is that older copies of the source records are maintained by Changelog Behavior. It is insufficient to enable this setting and perform a full sync to remove all old records from the database, rather manual intervention is required. The following SQL is for general guidance and should not be used directly without first testing against a test server:

```
SQL> update cm_org_identity_source_records set source_record=md5(source_record) where org_identity_source_id=?
and (deleted=true or org_identity_source_record_id is not null);
```

Filtering Inbound Records

As of Registry v4.1.0, Organizational Identity Sources support Data Filters. Data Filters apply to the structured Org Identity record created by the plugin, and not the cached source record. The cached source record will have the original values obtained from the source, unless *Hash Source Records* is enabled.