RSA Identity Governance and Lifecycle – Unification Guide

Version: 4.0
Issue Date: 20/July/2017

Revision History

Rev.	Date	Author(s)	Approver(s)	Description
1	2015	Waliszewski, Russell <Russell.Waliszewski@rsa.com>	N/A	Doc created
2	2016	Waliszewski, Russell <Russell.Waliszewski@rsa.com>	N/A	Various Updates
3	Nov 2016	Waliszewski, Russell <Russell.Waliszewski@rsa.com>	N/A	Various Updates
4	20th July 2017	Jamie Pryer	N/A	Removed section around duplicate users & Added Pre-req’s
5				

Table of Contents

1. OVERVIEW

The Unification process is a means by which we can collect the identity of a user from different applications using Identity Data Collectors and unify pieces of information from each application to form the user's corporate identity or master record. Consequently this master record will be used by the rest of ACM to associate all the user's entitlements within the company to provide a comprehensive view of the user. Since this master record is an accumulation of information it is imperative to know what information is available from the various applications and which ones have maintain the data while others may just consume the information.

1.1. Example1

Our HR system allows the user to maintain their First and Last name, while the Active Directory uses this information just for displaying the name. In this case we would want to use HR as the source for the user's name since this will be the most up to date information.

In order to join this information into a master record you must understand how a user is *uniquely* identified within an application as well as what types of information allow you to connect/join that user to other applications.

1.2. Example2

Using the previous example while we want the user's name from the HR system, but it may not be the best way to connect that user in another application. This fails when you have more common names like John Smith. During the unification process the single HR user would be matched with all the records found in Active Directory and each would produce a different master record. The HR system will have a way to uniquely identify the user within that application, and if this identifier is also stored within the Active Directory it should be considered as a possible way to join this information. One should verify that a search of Active Directory using that identifier would produce a single record, otherwise it would produce the same results as a join on the user's name.

You also need to identify what pieces of information you will use from each application to create the user. As above we decided that user's name would be taken from the HR system. If you want to include the user's email address you may find that is available in both the HR system and Active Directory. In that scenario you setup an attribute inheritance. When gathering the information for the master record you can take the email address from the Active Directory as long as the user has a record. If the user does not have a record then you could then specify to use the email address from the HR system.

2. DEFINITIONS

Unique Identifier – A piece of information that is used to identify a single user record within an application.

Master Record – This is the record that is used within ACM that will be an aggregation of information from various IDC's.

Primary IDC – This is an IDC which collects users and during the unification can have the ability to create master records, this option is set in the IDC Definition. These are the primary ways to identify a user record within the company.

Secondary IDC – Secondary IDC – This is an IDC which is used to collect additional information that will enrich the master record. It will have "Create Users" disabled; therefore if a record from this IDC does not match a record in the primary IDC during a join operation it will not create a separate master record.

Join – This is the process in the Unification to identify which IDC's information should be merged to a single record. The join is based on attribute(s) that are the collected values from the IDCs participating in the join. ACM will support joins between 2 Primary IDC's as well as a Primary and Secondary. Joining two Secondary IDCs is prohibited as the result of a join should be able to create a master record and by definition Secondary IDC's cannot.

Standalone IDC – This is an IDC which is not to be joined with any other, and does not have any overlap of users with any other IDC. If the collected users should appear in ACM, then the "Create Users" can be enabled. In this case the IDC will create a master record for each user by the IDC.
Ownership – As with the rest of ACM master records are "owned" by a collector within ACM. When a master user record is created it will be "owned" by an IDC.

3. PRE-REQ'S

Before starting on making ANY changes to collectors, or unification you should always do the following

1) Test properly on DEV, then UAT and finally into PROD, following your standard change management processes.
2) Take full DB and XML Config backups before you start, so you are able to roll-back if anything goes wrong.
3) Follow this guide carefully and reach out to RSA via the Link community if you have any questions/queries.

KEY NOTE: If you make changes to unification or collectors and this causes duplicate users, we CANNOT remove them as we have no way to delete users from the environment. You MUST ensure you can roll back if someone goes wrong, as simple mistakes with unification, can cause duplicates very easily.

Its is VITAL you take a full DB backup before you ever make changes, that will allow a roll-back

4. UNIFICATION CONFIGURATION

4.1. Attribute Processing

Attribute processing is a way to define the attributes from different Identity data collectors that will make up the definition of the Unified user. The processing order is Authoritative Source first and then the processing order for the Participating Collectors. For instance if the Unification configuration has the Authoritative source for the email address to be taken from an Active Directory server, but a user does not yet exist on the AD server then the email address will be taken from the Identity data collector that is first in the processing order. If the Unification configuration does NOT have an Authoritative source defined for the email address it will then default to the first identity collector in the processing order.

4.1.1. Attribute Sources

This is where the Authoritative source for an attribute is defined. This is a way to make sure the attribute is being used from the most up to date source.

4.1.2. Participating Collectors

This configuration will allow the order in which attributes are taken and applied to the Unified user. It is also used to determine the order the Identity data collectors are processed during a Unification run.

4.2. Joins

This section of the unification is to identify the common attributes between different Identity data collectors. These attributes are used to match one record for each collector in the join definition to be part of the Unified user. It is imperative that these attributes are unique to their respective Identity data collector (we do attempt to enforce this through the UI and rejecting collected records that contain duplicate data). If the data is not unique this can result in duplicate Unified users.

4.3. Reference Resolutions

This is the way to potentially identify multiple ways to resolve the Supervisor. Each Identity data collector which collects the Supervisor attribute will need to change that into a Unified user in the system. So the HR system may collect a Supervisor attribute which is an email address, but it might not have the email address. Therefore here the Supervisor can be resolved against the Identity data collector which is collecting the email address from Active Directory.

5. RECOMMEND PRACTICES

1) Any unique identifier within an application should also be collected as "User_Id" attribute by its collector.

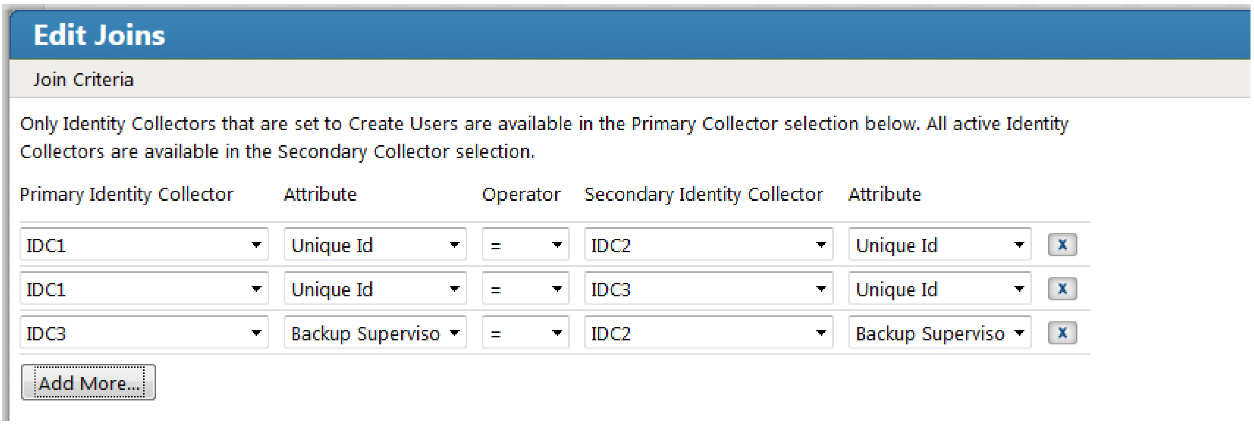
2) Use caution with Standalone IDC's. If they have "Create Users" enabled they will create a master record and potentially cause duplicate users.

3) Make sure the join between IDC's are based on a unique set of attributes. The IDC Collectors will reject user records when it detects multiple records with the same value for an attribute that is used in a Unification join.

4) Have your Primary(s) IDCs collect the largest # of users that will identify all the users in the company.

5) When joining primary idc's changing left and right IDC's will result in a change of ownership and will cause what appear to be duplicate users in the system.

6) When joining more than 2 primary idc's daisy chain them so there is no cycle.
Good:



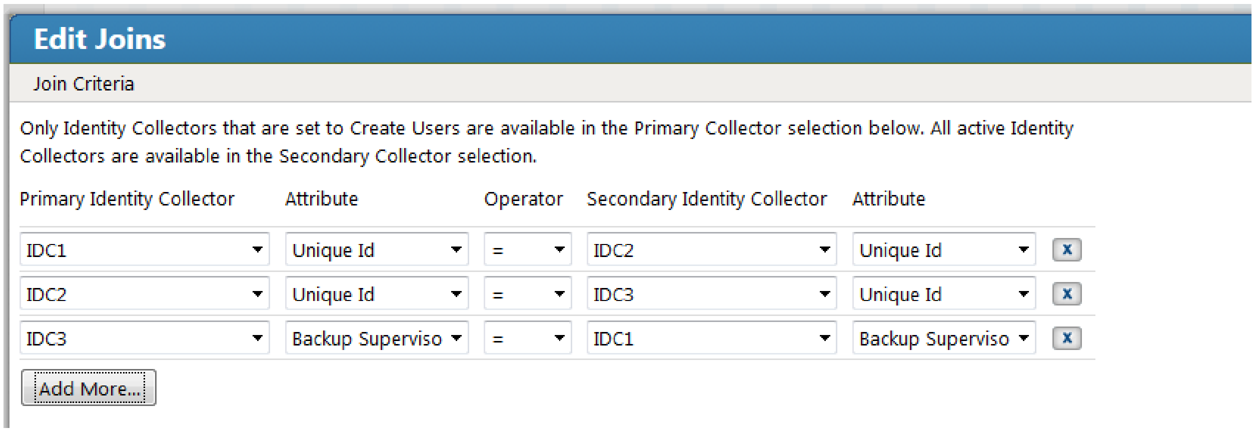
Edit Joins

Join Criteria

Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute	
IDC1	Unique Id	=	IDC2	Unique Id	X
IDC1	Unique Id	=	IDC3	Unique Id	X
IDC3	Backup Superviso	=	IDC2	Backup Superviso	X

Bad:



Edit Joins

Join Criteria

Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute	
IDC1	Unique Id	=	IDC2	Unique Id	X
IDC2	Unique Id	=	IDC3	Unique Id	X
IDC3	Backup Superviso	=	IDC1	Backup Superviso	X

7) Determine which IDC's are Primary (they can create master records). Are the primary IDC's stand alone or must they be combined to create all the master users. Is there overlap. If the primary IDC's are all standalone then they will each create all the users that they have collected and will

have ownership of them. If there are two or more Primary IDC's which have overlapping users then create all the joins between them first.

8) The joins between two IDC's **MUST** occur between a data set that uniquely defines another record. For instance you would want to avoid joining two IDC's on a Department Name, since one user's Department Name in an IDC could point to other users in another IDC that had the same Department Name. If each user had a unique identifier within that department you could use a join on Department Name & User Department Number.

9) If an IDC does not participate in the Unification (Standalone) and it has the "Create Users" option enabled it will be treated as a Primary and create users owned by itself. This should be used with caution.

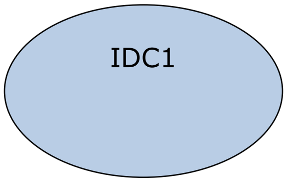
10) When joining two primary IDC's make sure they are maintaining up to date information on the attributes that are joined. If there is a delay in synchronizing information between the two IDC's on the joined attribute it will result in a failed join and potentially a duplicate master user records in ACM.

6. UNIFICATION PROCESSING

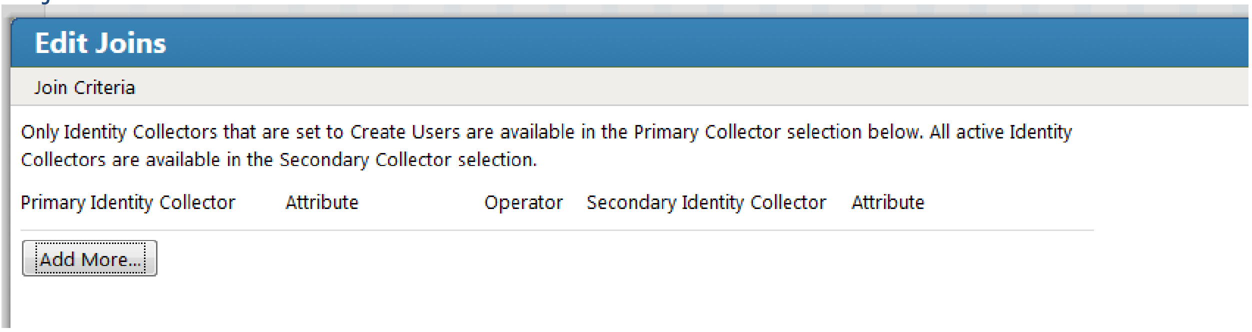
The unification starts by processing one IDC at a time. It finds the other IDCs that it joins to in the Unification configuration. For each IDC it will form a SQL statement that joining all the records in each IDC based on the join attribute. It is this reason that we require uniqueness on the join attribute. Since we are actually creating a SQL statement if the attribute is not unique the process ends up creating duplicate users. As it identifies user records in the joins it will then merge these records with other records found in other IDC joins. Once the users are merged they are then used to update the Master Enterprise Users with the updated information.

7. COMMON JOIN SCENARIOS

7.1. One Primary IDC



Example: IDC1 is a collector for employee data from the internal AD system.
When there are people in the corporate AD system we want to create and track new identities for them. It will also have the information that can be used to identify users within the corporate applications.
Resolution: IDC1 is set up as a primary IDC. Since it maintains all the information required there is no need to collect information from another IDC. ACM will contain all users that are collected from IDC1.
No joins.

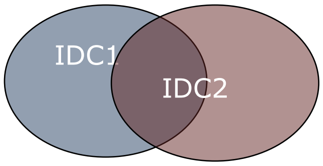

Edit Joins

Join Criteria

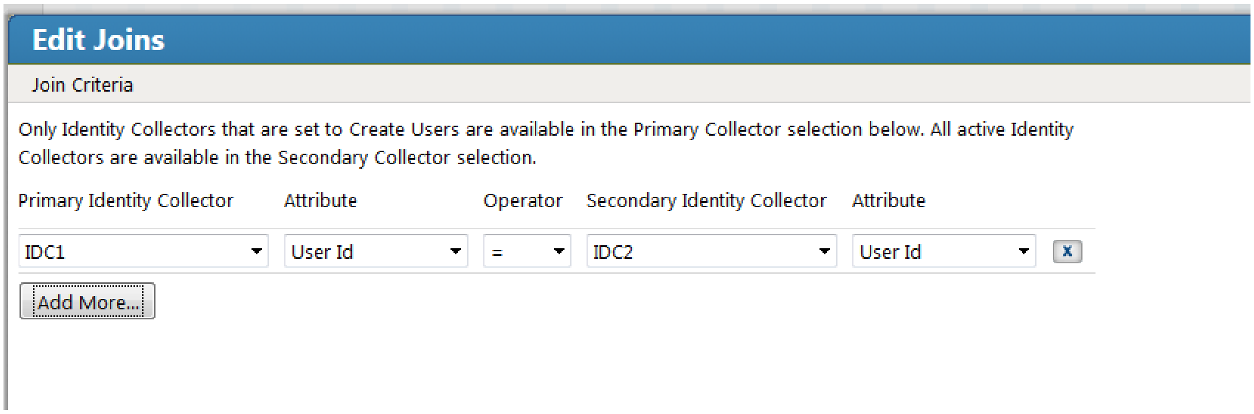
Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute

7.2. One Primary IDC & One Secondary IDC



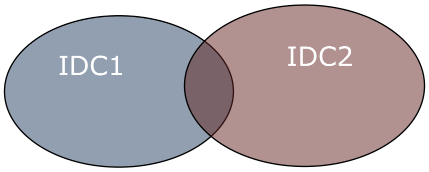
Example: IDC1 is a collector for employee data from the internal AD system. IDC2 collects HR data from the parent company's HR system, which will have additional data for many of the people in the internal AD system.
When there are people in the corporate AD system we want to create and track new identities for them. When there is information in the parent HR system on an employee, we want to include that information in ACM's representation of that. When there is information in the parent HR system on people who aren't employees (i.e., aren't in the AD system) we do not want new identities created.
Resolution: IDC1 is set up as a primary IDC. IDC2 is setup as a secondary IDC. A join is created between IDC1 and IDC2 on a common unique attribute (SSN?). ACM will contain all users collected from IDC1. The users in the area of overlap between IDC1 and IDC2 may have additional information in the master record.



Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute
IDC1	User Id	=	IDC2	User Id

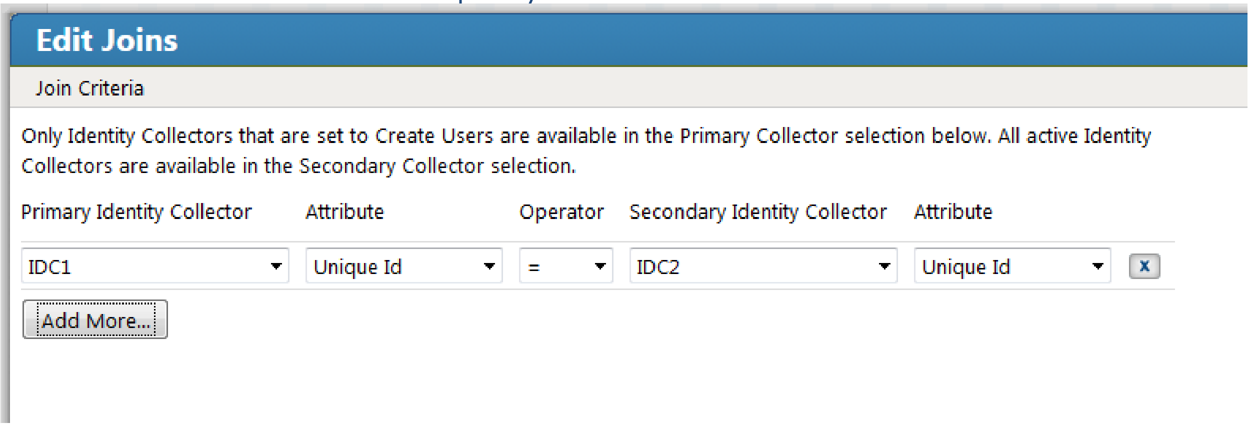
7.3. Two Primary IDCs



Example: IDC1 is a collector for customer data from the customer tracking system. IDC2 is a collector for employee data from the internal AD system. Some employees are also customers.

When there are customers who are not also employees we want to create and track new identities for them in ACM. When there are people in the corporate AD system who are not also customers we want to create and track new identities for them. When there are people who are both customers and employees, we want to create a single identity for each physical person in ACM. When someone who was only in the employee system also becomes a customer, we don't want to add a new identity, we want the system to know both pieces of data refer to the same person. (Same for the less likely case of customer who becomes an employee).

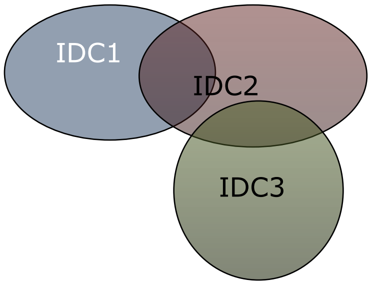
Resolution: IDC1 is set up as a primary IDC. IDC2 is setup as a primary IDC. A join is created between IDC1 and IDC2 on a common unique attribute (SSN?). All users collected from IDC1 and IDC2 will appear in ACM. The users are who overlap may contain additional information in the master record.



Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute
IDC1	Unique Id	=	IDC2	Unique Id

7.4. Two Primary IDCs & one Secondary IDC

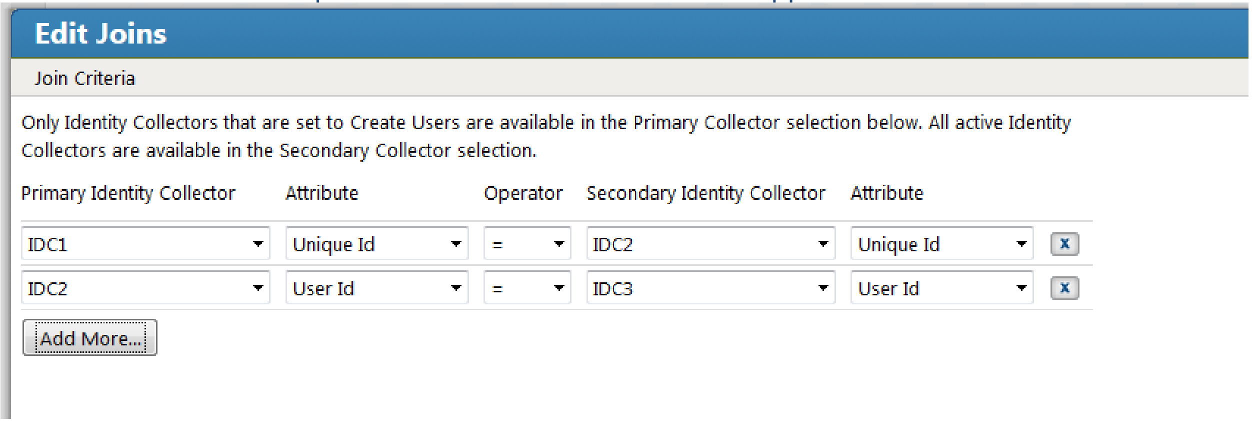


Example: IDC1 is a collector for customer data from the customer tracking system. IDC2 is a collector for employee data from the internal AD system. Some employees are also customers. IDC3 collects HR data from the parent company's HR system, which will have additional data for many of the people in the internal AD system, possibly including some that are also customers.

When there are customers who are not also employees we want to create and track new identities for them in ACM. When there are people in the corporate AD system who are not also customers we want to create and track new identities for them. When there are people who are both customers and employees, we want to create a single identity for each physical person in ACM. When someone who was only in the employee system also becomes a customer, we don't want to add a new identity, we want the system to know both pieces of data refer to the same person. (Same for the less likely case of customer who becomes an employee).

When there is information in the parent HR system on an employee, we want to include that information in ACM's representation of that person (including cases where the employee is also a customer). When there is information in the parent HR system on people who aren't employees (i.e., aren't in the AD system) we do not want new identities created.

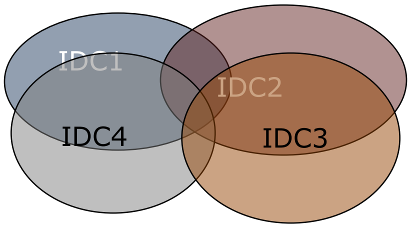
Resolution: IDC1 is set up as a primary IDC. IDC2 is setup as a primary IDC. A join is created between IDC1 and IDC2 on a common unique attribute (SSN?). IDC3 is set up as a secondary IDC and a join is created between it and IDC2 on a common unique attribute (corporate identity number). ACM will contain all users collected from IDC1 and IDC2. The users are in the overlap of IDC1 and IDC2 may contain additional information in the master record, as well the users who overlap with IDC3. Also the users from IDC3 that do not overlap with the IDC1 or IDC2 will not appear in ACM.



Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute	
IDC1	Unique Id	=	IDC2	Unique Id	x
IDC2	User Id	=	IDC3	User Id	x

7.5. Two Primary IDCs & two Secondary IDCs



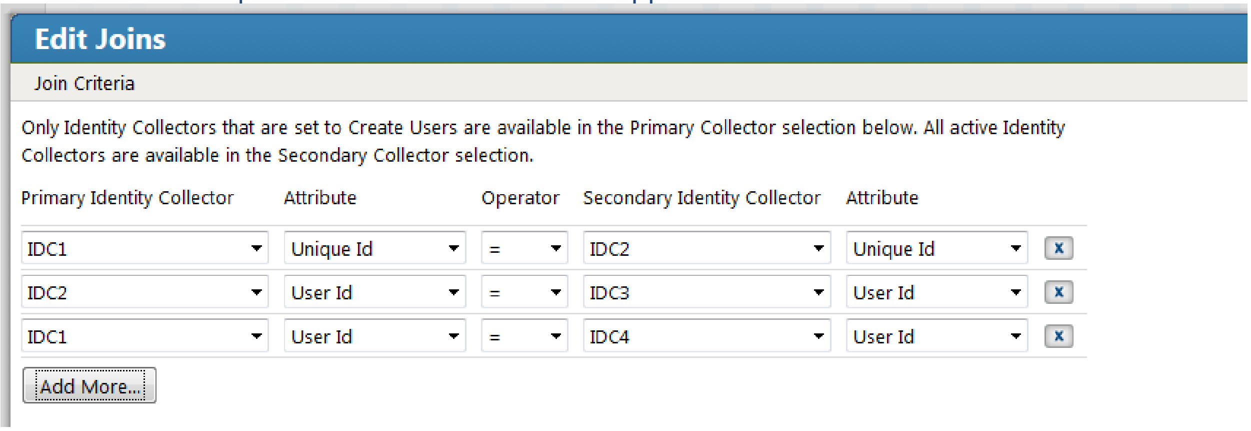
Example: IDC1 is a collector for customer data from the customer tracking system. IDC2 is a collector for employee data from the internal AD system. Some employees are also customers. IDC3 collects HR data from the parent company's HR system, which will have additional data for many of the people in the internal AD system, possibly including some that are also customers. IDC4 collects customer information from a 3rd party source.

When there are customers who are not also employees we want to create and track new identities for them in ACM. When there are people in the corporate AD system who are not also customers we want to create and track new identities for them. When there are people who are both customers and employees, we want to create a single identity for each physical person in ACM. When someone who was only in the employee system also becomes a customer, we don't want to add a new identity, we want the system to know both pieces of data refer to the same person. (Same for the less likely case of customer who becomes an employee).

When there is information in the parent HR system on an employee, we want to include that information in ACM's representation of that person (including cases where the employee is also a customer). When there is information in the parent HR system on people who aren't employees (i.e., aren't in the AD system) we do not want new identities created.

The information from the 3rd party source is used to identify customer prospects, but will also contain information for existing customers. Since this is a source of prospects, those that are not currently customers are not needed, but existing information is used.

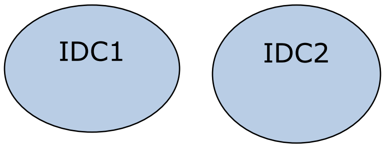
Resolution: IDC1 is set up as a primary IDC. IDC2 is setup as a primary IDC. A join is created between IDC1 and IDC2 on a common unique attribute (SSN?). IDC3 is set up as a secondary IDC and a join is created between it and IDC2 on a common unique attribute (corporate identity number). IDC4 is setup as a secondary IDC and a join is created between it and IDC1 on a common unique attribute. ACM will contain all users that are collected in IDC1 and IDC2. Any users in overlap areas may contain additional information from the secondary collectors IDC3 and IDC4. The users who are collected in IDC3 and IDC4 that do not overlap with IDC1 or IDC2 do not appear in ACM.



Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

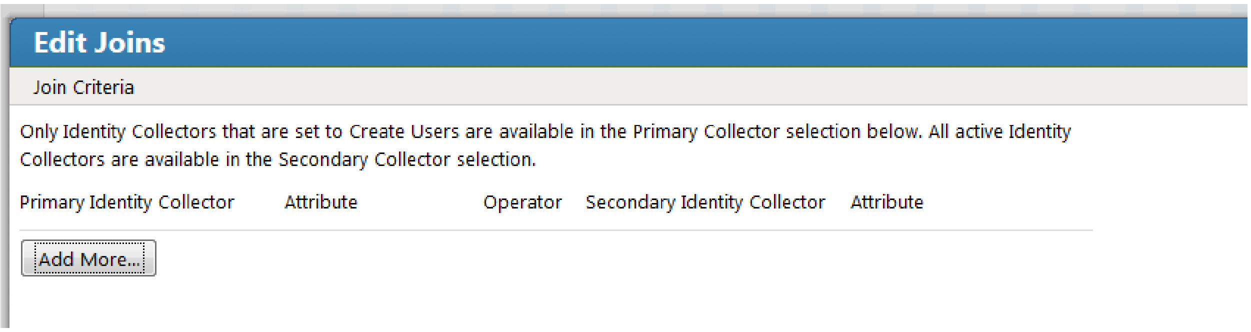
Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute	
IDC1	Unique Id	=	IDC2	Unique Id	X
IDC2	User Id	=	IDC3	User Id	X
IDC1	User Id	=	IDC4	User Id	X

7.6. Two Standalones



Example: IDC1.

Resolution: IDC1 is set up as a standalone IDC which can create users. IDC2 is setup as a standalone IDC which can create users. ACM will contain all users collected from both IDC1 and IDC2. There is no join between the IDC's.
No Joins



Edit Joins

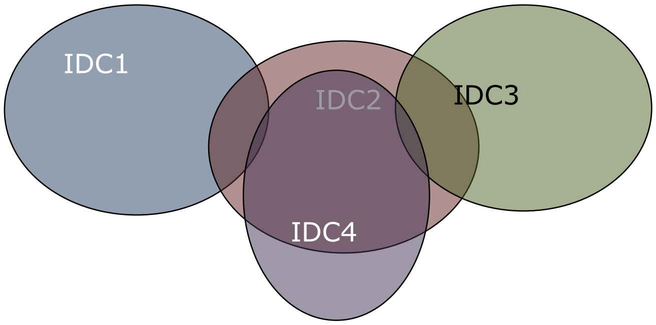
Join Criteria

Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute

NOTE:
Since both of these collectors are stand along and they have no join between them, if a user moves from one IDC to another IDC, the user will then have two records in the application. The record from the old IDC will be deleted / terminated and the record from the new IDC will be active.

7.7. Three Primary IDCs & one Secondary IDC

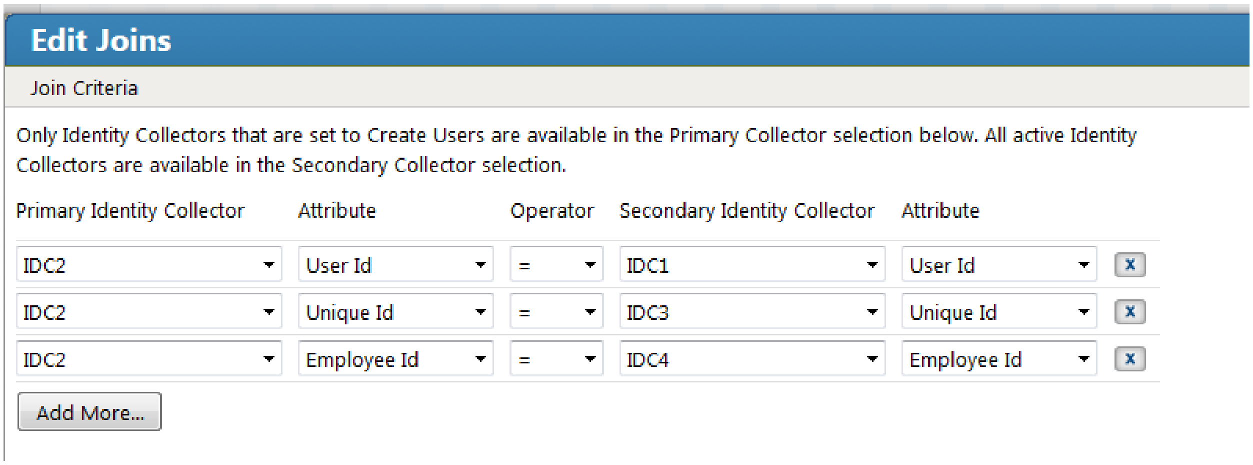


Example: IDC1 is a collector for customer data from the customer tracking system. IDC2 is a collector for employee data from the internal AD system. Some employees are also customers. IDC3 is a database that maintains information for Contractors for an internal system; some contractors may also have an AD entry. IDC4 collects HR data from the parent company’s HR system, which will ONLY have additional data for many of the people in the internal AD system.

When there are customers who are not also employees we want to create and track new identities for them in ACM. When there are people in the corporate AD system who are not also customers we want to create and track new identities for them. It follows that the database of contractors also needs to create records for tracking. When there are people who are both customers and employees, or employees and contractors, we want to create a single identity for each physical person in ACM. When someone moves from customer to employee or contractor, we don’t want to add a new identity, we want the system to know both pieces of data refer to the same person.

When there is information in the parent HR system on an employee, we want to include that information in ACM’s representation of that person (including cases where the employee is also a customer). When there is information in the parent HR system on people we do not want new identities created.

Resolution: IDC1, IDC2, and IDC3 are set up as primary IDCs. The joins are created between the IDCs on a common unique attribute (SSN?). IDC4 is setup as a secondary IDC and a join is created between it and IDC2 on a common unique attribute. ACM will contain all users that are collected in IDC1, IDC2, and IDC3. Any users in overlap areas between the primary may contain additional information. The users who are collected in IDC2 that overlap with IDC4 may also contain additional information, but users in IDC4 that do not overlap any users in IDC2 will not appear in ACM.



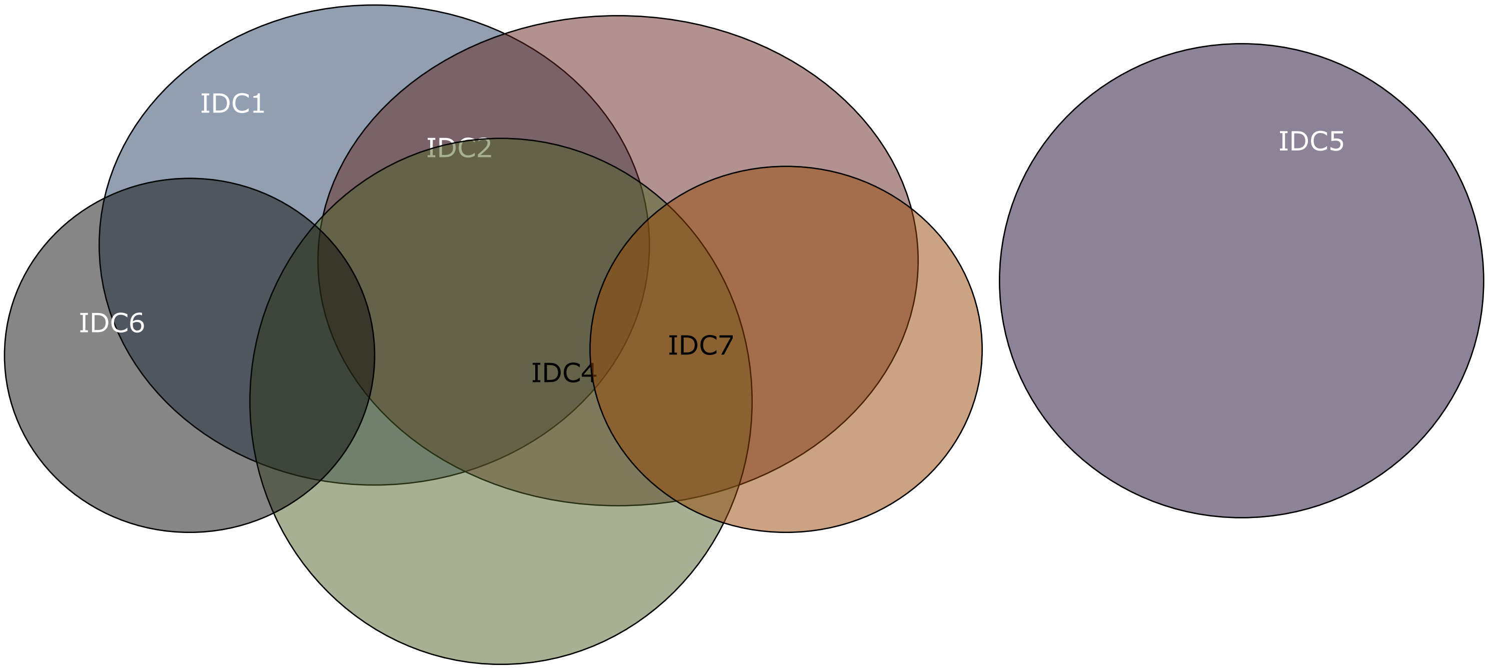
Edit Joins

Join Criteria

Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

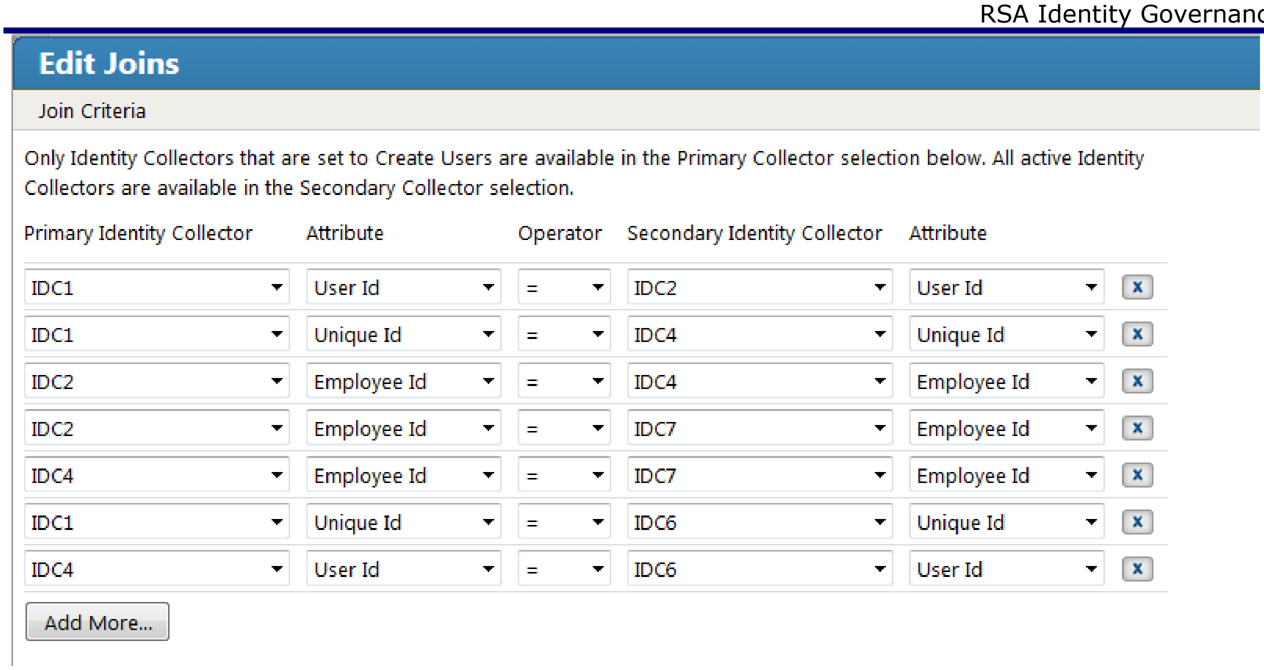
Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute	
IDC2	User Id	=	IDC1	User Id	x
IDC2	Unique Id	=	IDC3	Unique Id	x
IDC2	Employee Id	=	IDC4	Employee Id	x

7.8. Three primary IDCs & two Secondary IDCs & one Standalone



Example: IDC1 is a collector for customer data from the customer tracking system. IDC4 is a collector for employee data from the internal AD system. Some employees are also customers. IDC2 is an LDAP server that maintains information for contractors & customers for an internal system; some of the contractors or customers may also have an AD entry. IDC6 collects HR data from the parent company's HR system, which will have additional data for many of the people in the internal AD system, possibly including some that are also customers. IDC7 is database used by engineering maintaining additional user information, this information maybe for employees or contractors. IDC5 is a database of users of vendors.
Our goal is to track the identities of all Employees, Customers and Contractors within ACM. When there are people in the corporate AD system who are not also customers we want to create and track new identities for them. It follows that the database of contractors also needs to create records for tracking users who may or may not be in the Customer or Employee collection. When there are people who are both customers and employees, or employees and contractors, we want to create a single identity for each physical person in ACM. When someone moves from customer to employee or contractor, we don't want to add a new identity; we want the system to know both pieces of data refer to the same person.
When there is information in the parent HR system on an employee, we want to include that information in ACM's representation of that person (including cases where the employee is also a customer). When there is information in the parent HR system on people we do not want new identities created. Similarly with the engineering database we want to use that supply more information to the user, but we don't want any additional records created by this collection. The vendor information does not have any overlap with any other information collected by the other IDCs. The users collected from this system will not have any additional access for any systems within the company.

Resolution: IDC1, IDC2, and IDC3 are set up as primary IDCs. The joins are created between the IDCs on a common unique attribute (SSN?). Since Employees, Customers, and Contractors can overlap we want to make sure that we have two different joins between the three IDCs. IDC6 is setup as a secondary IDC and two joins are created between it and IDC1 and IDC4 on a common unique attribute. IDC7 is setup as a secondary IDC and two joins are created between it and IDC2 and IDC4 on a common unique attribute. IDC5 will not have any overlap with any other IDCs therefore there will be no joins. If the users collected by IDC5 are to be tracked within ACM then that IDC will need to have Create Users enabled, otherwise its users will not appear in ACM to be tracked.



Edit Joins

Join Criteria

Only Identity Collectors that are set to Create Users are available in the Primary Collector selection below. All active Identity Collectors are available in the Secondary Collector selection.

Primary Identity Collector	Attribute	Operator	Secondary Identity Collector	Attribute
IDC1	User Id	=	IDC2	User Id
IDC1	Unique Id	=	IDC4	Unique Id
IDC2	Employee Id	=	IDC4	Employee Id
IDC2	Employee Id	=	IDC7	Employee Id
IDC4	Employee Id	=	IDC7	Employee Id
IDC1	Unique Id	=	IDC6	Unique Id
IDC4	User Id	=	IDC6	User Id

8. DUPLICATE USERS

8.1. Definition

At times a user may appear as duplicated within ACM. There are different things that can cause this to happen; the first thing to understand is why it could happen. Users like other collected objects within ACM are “owned” by a collector. The owner will be a Primary IDC. If the collector owner changes, then the previous version of the user will be listed as a deleted user, and a new one is created for the different collector.

8.2. How/Why

There are different things that can cause duplicate users to appear:

1) If there are multiple Primary IDC’s the owner will be the defined by the ordering of the joins between the Primary IDC’s. Therefore once you identified the joins between the Primary IDC’s they should remain static.
2) Duplicate users may also appear if a join of user information between two Primary IDC’s fails and both IDC’s create a user record. You need to make sure that you have defined the joins between two Primary IDC’s on attributes which is consistent and complete in both IDC’s.
3) If a join between two Primary IDC’s is not based on data that identifies a unique record in the IDC’s duplicates can appear. If you have a user from one IDC and join to another IDC by Dept Name, unless the data is unique then it will match all user records that have the same Dept Name.
4) An IDC which does not participate in Unification (Standalone IDC), and the IDC is configured to Create Users, this can create duplicate users. In this scenario you would either want to disable the IDC’s ability to Create Users or have it participate in the Unification.

8.3. Resolution

In the event that duplicate users have occurred in the system, the first step is to determine how or why it occurred. If there was a unification configuration change that caused this then the correct configuration should be decided.

Once you have found the root cause, you will need to re-load the DB from a backup, update your unification configuration and then run again, to ensure you do not have any duplicates.

--- END OF DOCUMENT ---