

QuickStart Guide 4 - Merge

Document Version: v1.5

Product Version: v3.13

Date: 16th November 2018

This document provides an overview and Step-by-Step implementation instructions for the clearMDM Merge MDM operation.

The document Appendices also provide additional reference materials.

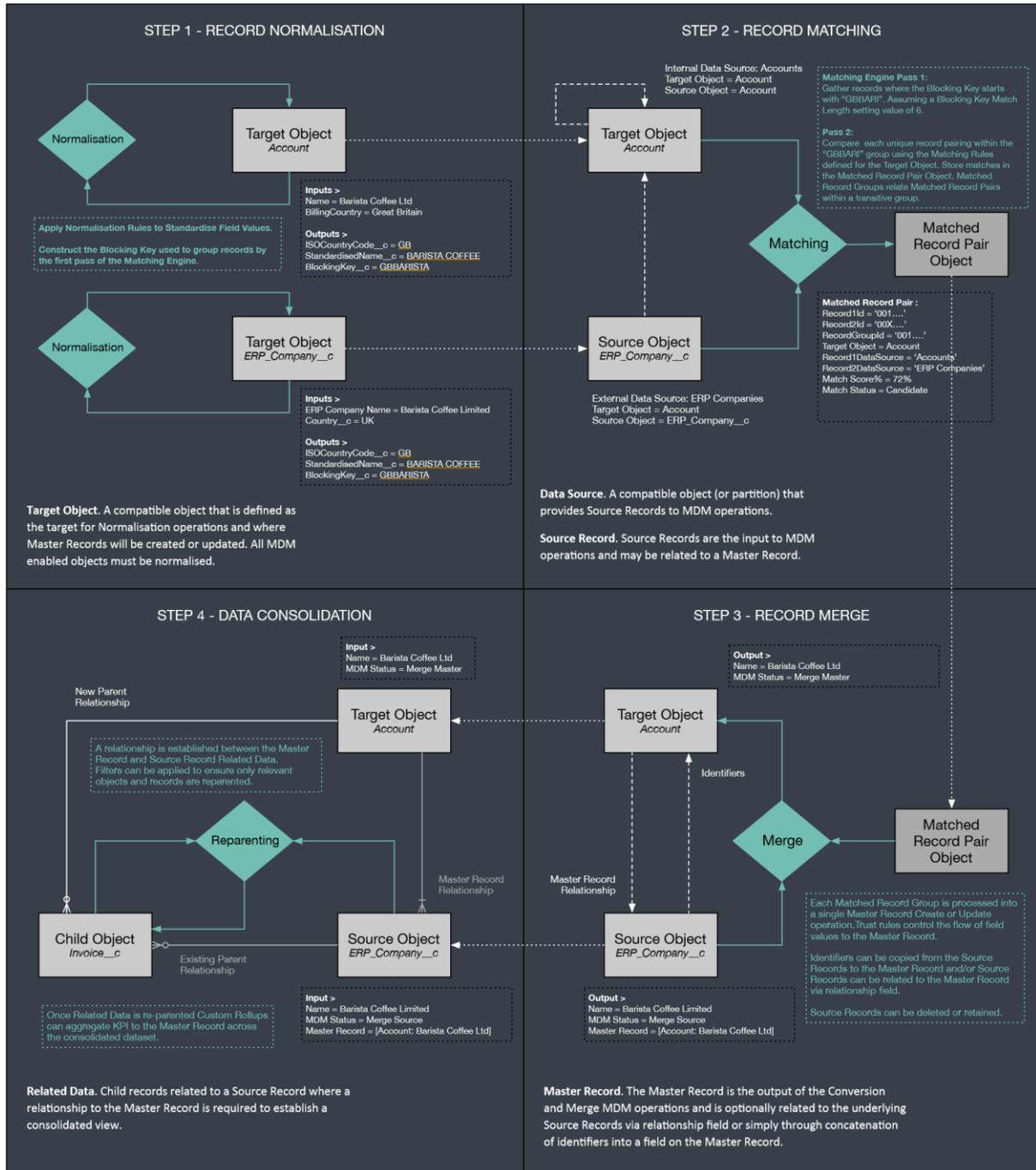
For practitioner guidance in respect to the implementation of clearMDM please refer to the Implementation Model documentation provided on the website, or upon request.

Table of Contents

QuickStart Guide 4 - Merge	1
MDM Process Overview	2
Merge Introduction	2
Definition.....	2
Key Concepts.....	3
Merge Methods	8
Step 1 Configure Application Settings	11
Step 2 Activate the Target Object for Merge	11
Step 3 Activate the Data Source for Merge	12
Step 4 Configure Merge Rules	13
Step 5 Run the Merge Job	14
Step 6 View Merge Results	15
Step 7 Manual Merge	16
Appendix A - Merge Settings Reference	19
Appendix B – Troubleshooting	21
Merge Log Types	21

MDM Process Overview

The diagram below provides an overview of the core MDM operations in sequence. This document relates to STEP3 – RECORD MERGE.



Merge Introduction

Definition

CLEARMDM

Record Merge is the process of transforming the Matched Record Groups produced by the Matching MDM operation into new Master Records or updates to existing Master Records. Where Source Records exist of the Target Object type (e.g. Account) then one of the Source Records will be designated as the Master Record for the group. Where no Target Object record exists then a new Master Record will be created. The Merge MDM operation provides three control points; master record determination, attribute group evaluation and field-level trust rules. In the former case, where multiple Target Object records exist in the group a means to apply custom business rules in determining the correct selection of the Master Record is required. In the latter cases, where multiple Source Records converge to a single Master Record merge rules are required that govern how field values from Source Records are composed to populate the Master Record.

For further information in relation to **Matched Record Groups** and related concepts please refer to the document *clearMDM – QuickStart Guide 3 – Matching*.

Key Concepts

Concept	Definition
Target Objects	<p>A compatible object that is defined as the target for Matching and Merge operations and where Master Records will be created or updated. Account, Contact, Lead Standard Objects are typically configured as Target Objects, where duplicate records may exist directly in the object or indirectly in a separate object.</p> <p>A broad range of Standard Objects (including Person Accounts) are supported as both Target Objects and Data Sources. Custom Objects are also supported.</p>
Data Sources	<p>A compatible object that provides data to MDM operations. Each Data Source has a Source Object and a Target Object setting.</p> <p>Internal Data Sources expose data held in the Target Object. Data Source (Account) > Target Object (Account)</p> <p>External Data Sources expose data held in a different object. Data Source (ERP Companies) > Target Object (Account)</p> <p>Partition Data Sources enable a single object to support multiple Data Sources with distinct settings. Partition Data Sources are typically used to group records relating to an external system (e.g. SAP, Sage X3) or to isolate records at different quality grades (e.g. High, Medium and Low).</p> <p>Master Record Data Sources allow Master Records to be exposed to MDM operations via a distinct Data Source with appropriate settings such as elevated merge field priorities. A Master Record Partition data source is implemented as a partition data source that references the record MDM status value.</p>

Master Records	<p>The Master Record is the output of MDM processing and is optionally related to the underlying Source Records (or duplicates) via relationship field or simply through the concatenation of record identifiers into a field on the Master Record.</p> <p>Most typically a Master Record is a de-duplicated Account, Person Account, Contact or Lead record enriched with data from its related Source Records. Where Source Records (or duplicates) are not removed, it is typical to use the Salesforce sharing model to present Salesforce end-users with access to Master Records only – thereby removing the visibility of duplicates.</p> <p>clearMDM implements a custom merge engine that works across object boundaries and can create as well as update Master Records.</p>
Source Records	<p>Source Records are the input to MDM operations and may be related to a Master Record. Source Records can be retained or deleted. Source Records can be considered as the underlying duplicates that are often hidden from view for Salesforce end-users or deleted entirely once processed.</p>
Merge Algorithm	<p>Step 1: Gather Matched Record Pairs by Matched Record Group</p> <p>Step 2: For each Matched Record Group</p> <ul style="list-style-type: none"> • Enumerate the Matched Record Pairs and gather the unique Source Record Ids. • If one Source Record exists of the Target Object type then designate this record as the Master Record. • If multiple Source Records exist of the Target Object type, then designate the record with the highest master record priority as the Master Record. • If no Source Record exists of the Target Object type, then create a new Master Record. <p>Step 3: For each Attribute Group</p> <ul style="list-style-type: none"> ○ If the designated Master Record is an existing Master Record then load all related Source Records. <ul style="list-style-type: none"> ○ Evaluate all Source Records to find the highest priority, valid Source Record for the Attribute Group. ○ Copy field values for all fields in the Attribute Group from the winning Source Record to the Master Record. <p>Step 4: For each Target Object Field</p> <ul style="list-style-type: none"> ○ If the field has been processed in Step 3 then skip field. <ul style="list-style-type: none"> ○ Find the highest priority, non-blank field value across the Source Records and set the Master Record field equal to this value. <ul style="list-style-type: none"> • Delete the Matched Record Pairs for the group. • On the Source Records set the MDM Status to “Merge Source” and set the [Is Active for Reparenting?] flag equal to True. • On the Master Record set the MDM Status to “Merge Master” and set the LastMergedDate field.

Merge Settings	Merge is configured per Target Object on the Target Object Settings page and also at the Data Source level on the Data Source settings page. A setting-by-setting definition is provided in Appendix A .
Merge Rules – Attribute Groups	<p>Attribute Groups are collections of fields on a Target Object that must be merged together from a single Source Record only. In determining how the Master Record fields (in the Attribute Group) are populated each related Source Record is evaluated for validity (typically completeness) and priority. Attribute Group priority can be defined on a Newest, Oldest or Dynamic Priority basis allowing custom business rules to direct which Source Record the field values are taken from.</p> <p>Attribute Group processing performs a full re-evaluation of all Source Records for a given Master Record whenever a new or existing Source Record modification is identified.</p> <p>Merge and Synchronisation MDM operations perform Attribute Group processing.</p> <p>New in v3.13: The ability to be able to set Multi-Select Picklist values to update to the Master Record. Where records have these types of fields, there may be different value(s) on each record. The Merge process will update these all to the Master Record.</p> <p>Check box fields can also be updated to the Master Record if checked. If the Master Record does not have the checkbox checked but the Source record does, then upon Merge, it will update the Master Record with a checked value.</p>

<p>Merge Rules – Attribute Group Templates</p>	<p>Attribute Group Templates support cross-field merge.</p> <p>A generic Attribute Group Template can be defined (e.g. Address, Phone or any other grouping) - with placeholder fields.</p> <p>Attribute Groups (e.g. Billing Address, Shipping Address) can then be mapped to the template (field-by-field) and the population order set; i.e. the order in which the Attribute Group fields will be populated on the master record.</p> <p>The Merge engine will gather all mapped values in priority order. The highest priority values are then added to the master record fields within the first Attribute Group (by population order) and so on.</p> <p>Basic Example.</p> <p>Attribute Group 1 - BillingAddress - Population Order 1 Attribute Group 2 - Shipping Address - Population Order 2 Attribute Group 3 - Custom Address - Population Order 3</p> <p>The prioritised list of actual addresses; record1:ShippingAddress record2:ShippingAddress record1:BillingAddress record2:BillingAddress</p> <p>The master record is populated as; Attribute Group 1 - BillingAddress = record1:ShippingAddress Attribute Group 2 - AG2 Shipping Address = record2:ShippingAddress Attribute Group 3 - AG3 Custom Address = record1:BillingAddress</p> <p>New in v3.13: Within the Attribute Group Template there is the ability to be able to create a cross-merge template that will update Master Record field values only if the value is unique. An example of this could be email address. If two matched records have the same email address and there is more than one email field on the master record with merge rules applied, the source record email will only update to the Master Record if that email does not exist.</p>
--	--

<p>Merge Rules – Field Priority</p>	<p>Each Data Source must specify a numeric Merge Priority value for each Target Object field. The priority is defined on a relative scale across Data Sources and provides control over which Data Sources are preferred for each field.</p> <p>For example, a Data Source that represents Source Records from an external Ecommerce system may be the most trusted source for physical address fields as the shipping address held in this system is likely to be accurate.</p> <p>Manual record changes applied to Master Records in the Salesforce UI are designated with a priority level and recorded in the Master Field Map to prevent overwrite by subsequent Source Record changes with a lower priority.</p> <p>Data Source field-level priorities that are higher than the manual change priority will overwrite, those that are lower will not.</p>
<p>Master Field Map</p>	<p>A JSON data structure held in a field on the Master Record for the purpose of tracking how the master record is currently populated per-field (from Source Record, Manual change etc.).</p> <p>The Master Field Map provides the stored knowledge upon which control is applied to flow of Source Record field changes to the Master Record.</p> <p>Where fields are populated via Attribute Group, the Master Field Map is used to track manual updates only (record changes applied in the Salesforce application). Where the Attribute Group priority type is Newest or Oldest then manual updates applied to any constituent field prevents any field in the group being overwritten. Where the Attribute Group priority type is Dynamic, fields will be overwritten if a valid Source Record is found where the calculated priority is higher than the default manual update priority.</p>
<p>Ignore Partial Groups</p>	<p>Matched Record Groups are considered to be partial where at least one “Candidate” status match exists. The Merge MDM operation by default will skip such groups until data stewarding has been applied to set each match to either “Accepted” or “Rejected” status.</p>
<p>Data Retention</p>	<p>The Merge MDM operation does not automatically delete Source Records instead linkages can be established between a Master Record and the Source Records that will be retained on-platform or identifiers can be copied to the Master Record where the Source Records will be deleted.</p> <p>There is no requirement to either retain or delete Source Records.</p>

CLEARMDM

Master Record Determination Algorithm	<p>By default the Merge MDM operation uses the logic below to determine which record in a given Matched Record Group becomes the Master Record.</p> <p>A) 1 Target Object record in the matched record group <= this record is the master</p> <p>B) No Target Object record <= create a new master</p> <p>C) Multiple Target Object records:</p>
	<p>.. if group contains a portal enabled record (Account and Contact only) that is an existing master record <= this is the master</p> <p>.. else if group contains a portal enabled record <= this is the master</p> <p>.. else if group contains an existing master record <= this is the master .. else strongest match record;</p> <p>(count of matches, then highest average match score, then most recent LastModifiedDate)</p> <p>The algorithm above can be overridden to implement a master record determination scheme that reflects custom business rules; please see Master Record Priority below for further information.</p>
Master Record Priority	<p>Where the Merge MDM operation processes a Matched Record Group containing multiple Target Object records (e.g. Accounts) it is imperative that the most appropriate record is designated as the Master Record. Often this selection logic is driven by custom business rules which recognise related data, data freshness, attribute population etc.</p> <p>A numeric formula field can be implemented that returns a record-level priority scale based on custom business rules, the Merge MDM operation will select the record with highest value. Alternatively, the best matched record within the group will be selected.</p> <p>The Data Quality MDM operation can be implemented to calculate the Master Record Priority based on various attributes of the record and the existence of related records (for example, active contracts).</p>

Merge Methods

The table below outlines the supported methods for invocation of the Merge MDM Operation. Setting references refer to the Target Object Merge settings.

Method	Definition
--------	------------

Batch Job	<p>The clearMDM Jobs page can be used to schedule the Merge MDM operation to run for a given Target Object immediately or on a scheduled basis.</p> <p>Where a daily batch processing model is implemented for MDM processing, the Merge MDM operation will typically be the fourth job and will invoke the next job in the sequence using the job chaining settings e.g. Reparenting Settings section, Is Invoked by Merge Job? flag.</p>
Action	<p>Records can be Merged by Process Builder as part of a custom process automation.</p> <p>To configure a Process Builder Action for this purpose add an Action with properties set as below.</p> <ol style="list-style-type: none"> 1. Action Type = Apex 2. Apex Class = "Merge Records Action" 3. Record ID parameter = Reference [Object Id field]. 4. Is Async? = True / False <p>The Action can also be implemented within Visual Workflow.</p>
API	<p>Records can be Merged by a custom action exposed via the standard Force.com REST API; endpoint below.</p> <p><code>/services/data/vXX.0/actions/custom/apex/clearmdm__RecordMergeAction</code></p> <p>The API operation takes a single recordId parameter. Further details can be found in the clearMDM API Guide.</p>
UI	<p>Record Merge can be applied manually via the Merge page accessible from the Matched Record Group page or from the MDM Reset and Matching Test features.</p> <p>Merge</p> <p>The Merge page supports the following functions:</p> <ol style="list-style-type: none"> 1. Selection of the Master Record for the group. 2. Selection of Master Record field values. 3. Automated invocation of the Reparenting MDM operation. <p>Note, the Custom Setting [Manual Merge Settings] controls whether the option to invoke the MDM Reparenting operation option is visible and its default state. This approach allows some user populations to be given the option whilst for</p>

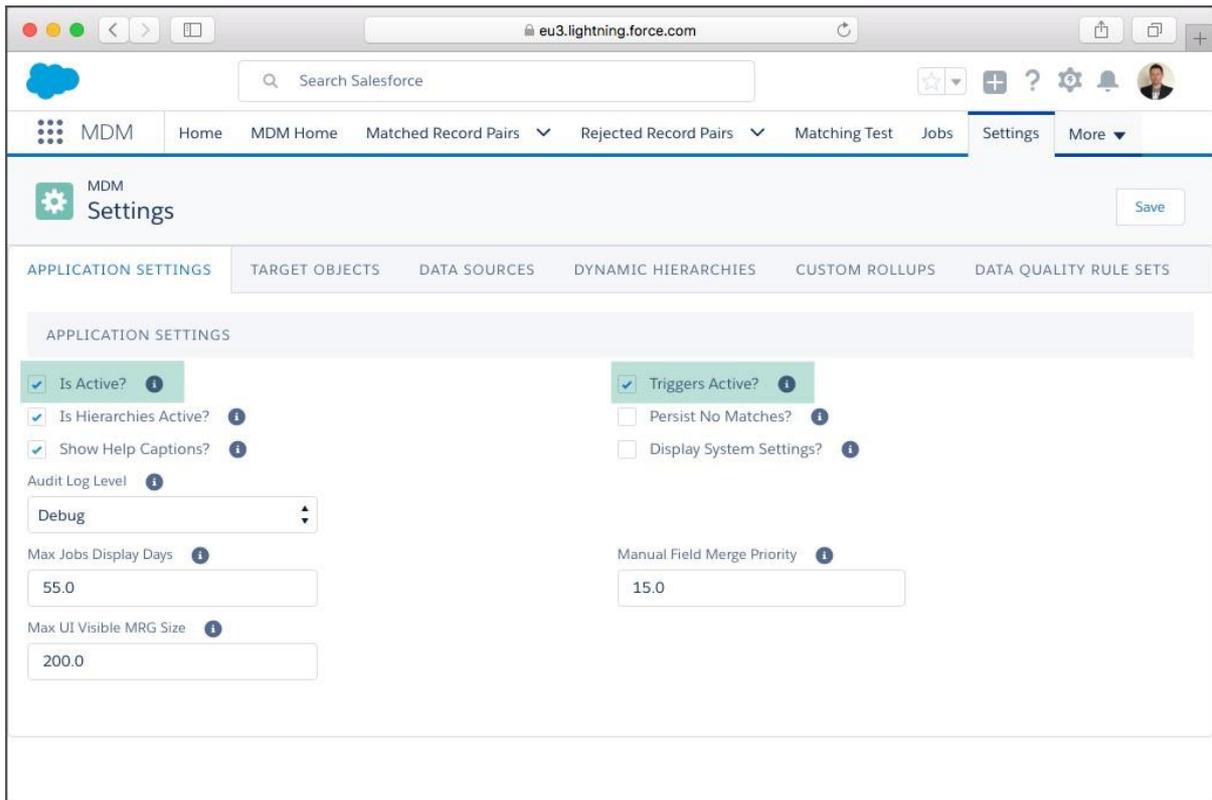
	<p>others the option is not displayed and reparenting runs automatically following every manual merge.</p> <p>Note, a Reparenting MDM operation invoked from the manual Merge page processes just the Matched Record Group that has been merged. The Batch Job type is set to "Reparenting Selective". Reparenting job chaining logic configured on the Target Object Settings page is not applicable to the selective variant; as such the Custom Rollup MDM operation will not be invoked as it may be when the Reparenting MDM operation runs in standard mode.</p>

Step 1 Configure Application Settings

Pre-requisite: clearMDM must be set to Active via the Application Settings page.

Pre-requisite: clearMDM must be set with Triggers Active via the Application Settings page.

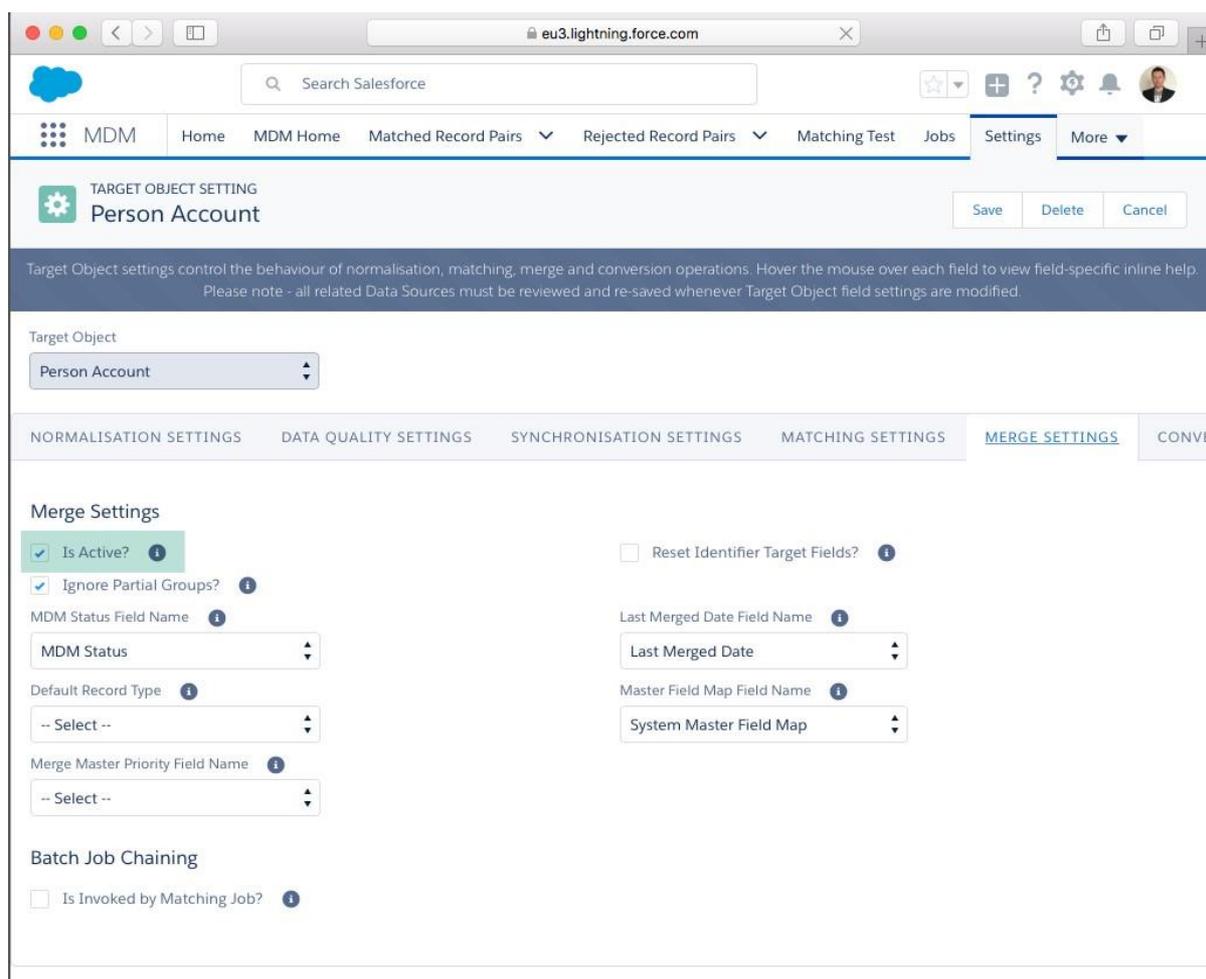
To complete this step, first navigate to the MDM App, open the Settings tab, tick the two fields highlighted below and click the Save button.



Step 2 Activate the Target Object for Merge

To complete this step, first navigate to the MDM App, open the Settings tab, select the Target Objects tab and click the Edit link next to the required Target Object. Note, for Lightning Experience the Edit menu is accessible via the Dropdown menu in the rightmost table column.

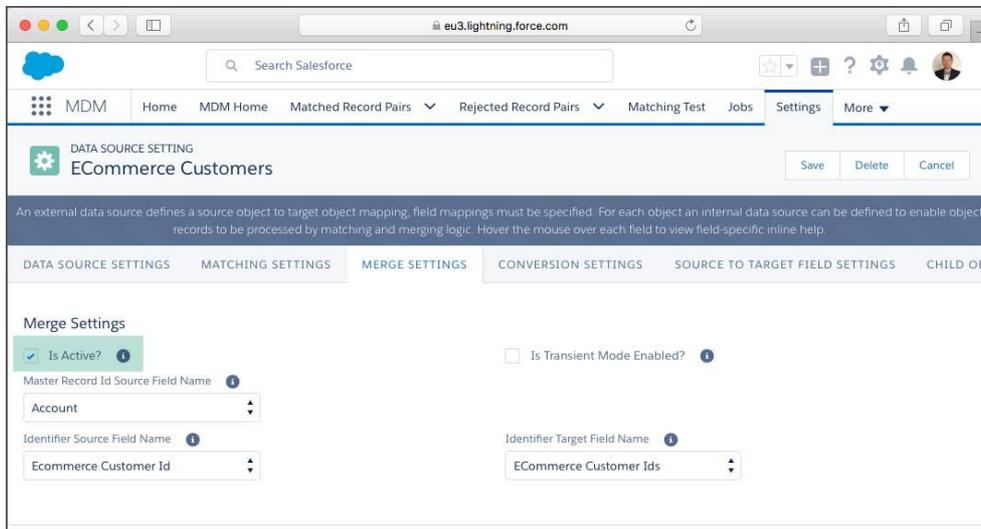
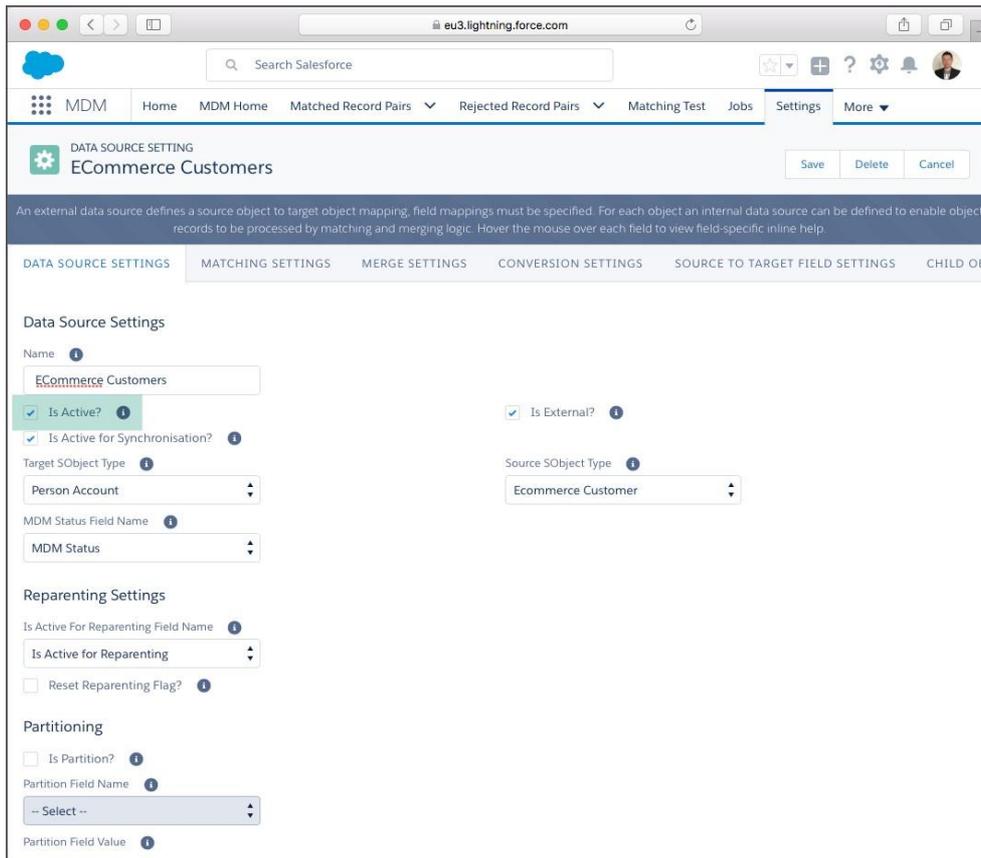
Next, set the [Is Active?] flag equal to true in the Merge Settings section (or Merge Settings tab in Lightning Experience) and ensure the relevant settings are configured correctly. [Appendix A](#) provides a Settings reference. Click Save to store the changes.



Step 3 Activate the Data Source for Merge

To complete this step, first navigate to the MDM App, open the Settings tab, select the Data Sources tab and click the Edit link next to the required Data Source. Note, for Lightning Experience the Edit menu is accessible via the Dropdown menu in the rightmost table column.

Next, set the [Is Active?] flags equal to true in the Data Source Settings and Merge Settings sections (or tabs in Lightning Experience) and ensure the relevant settings are configured correctly. [Appendix A](#) provides a Settings reference. Click Save to store the changes.



Step 4 Configure Merge Rules

To complete this step, first navigate to the MDM App, open the Settings tab, select the Data Sources tab and click the Edit link next to the required Data Source. Note, for Lightning Experience the Edit menu is accessible via the Dropdown menu in the rightmost table column. On the Data Source settings page scroll down to the “Source to Target Field Settings” section (or click the tab in the Lightning Experience).

Next, specify a Source Field mapping (External Data Sources only), cross-field matching settings and the merge priority for each field.

Finally click the Save button.

TARGET FIELD	DATA TYPE	MATCH TYPE	SOURCE FIELD	CROSS FIELD MATCHING	MANUAL FIELD MERGE PRIORITY	MERGE PRIORITY
Combination 1	Text	Key	Combination 1	-- Select -- Address Blocking Key Group Blocking Key	15	
Email	Text	Fuzzy	Email	-- Select -- Address Blocking Key Group Blocking Key	15	10
First Name	Text	Fuzzy	First Name	-- Select -- Address Blocking Key Group Blocking Key	15	10
Last Name	Text	Fuzzy	Last Name	-- Select -- Address Blocking Key Group Blocking Key	15	10
Mailing City	Text	Fuzzy	City	-- Select -- Address Blocking Key Group Blocking Key	15	10

Step 5 Run the Merge Job

To complete this step, first navigate to the MDM App, open the Jobs tab, select the Merge Job type and the required Target Object. Finally select the required Job Schedule and click the "Schedule Job" button. The job will sit in the scheduled jobs tab until complete.

Jobs
Monitor and Run

View active and completed Jobs. Run jobs immediately or schedule future job executions.

Schedule

Job: Merge | Object: PersonAccount

Data Sources: -- None -- | Process Count: 5

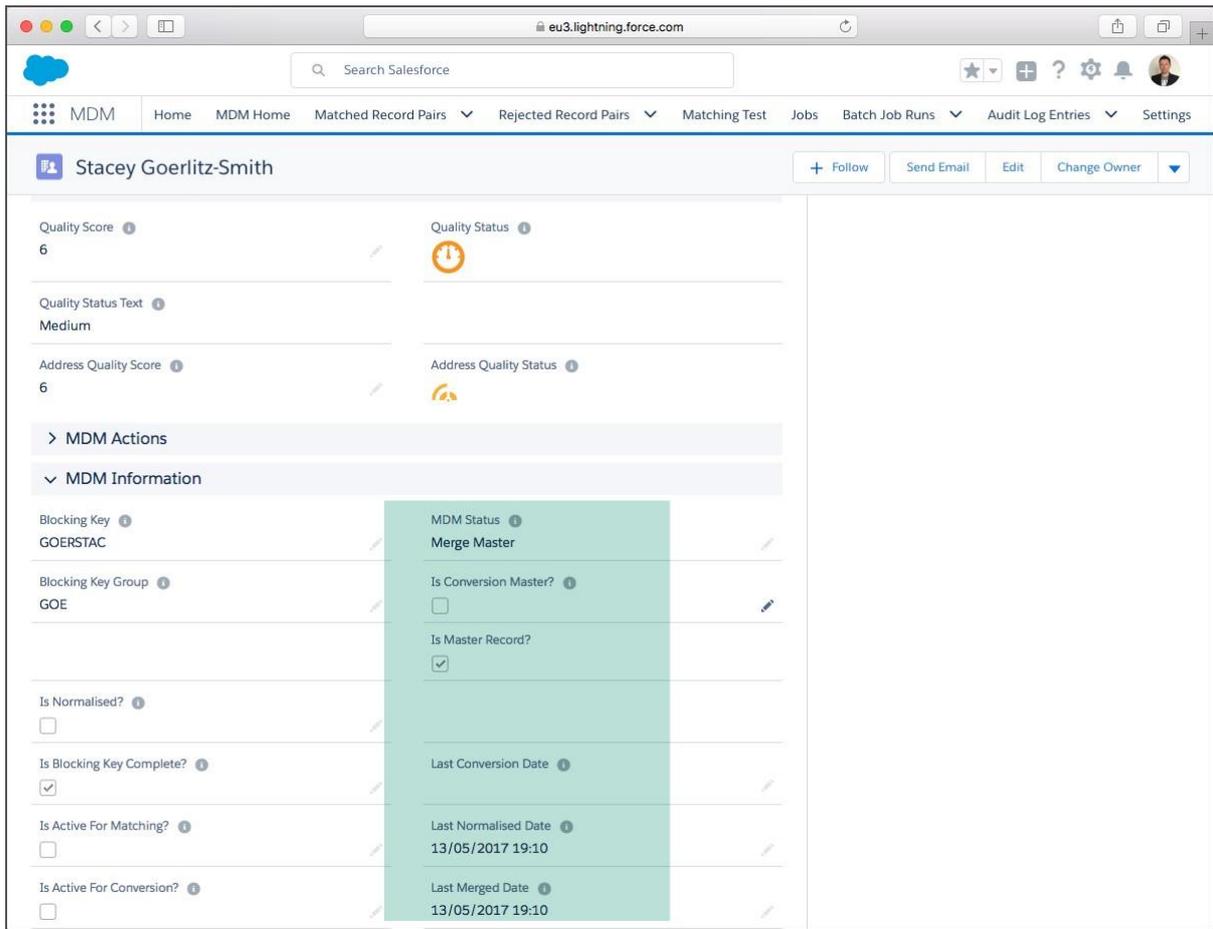
Schedule: Now

RUNNING JOBS | **COMPLETED JOBS - LAST 5 DAYS** | **SCHEDULED JOBS**

JOB TYPE	OBJECT	LAST RUN TIME	NEXT RUN TIME	RUN COUNT	ACTION
Normalisation	Ecommerce_Customer__c	13/05/2017 18:45	13/05/2017 19:45	411	Abort
Normalisation	Corporate_Customer__c	13/05/2017 18:00	13/05/2017 19:00	410	Abort

Step 6 View Merge Results

The screenshot below shows the Merge MDM fields displayed on an example record. Placement of MDM fields on layouts for administrators, data stewards and end-users is a key implementation decision.



Step 7 Manual Merge

The screenshots below show the steps of the manual “Merge” feature that allows a Matched Record Group to be manually merged.

Open a Matched Record Group and review the Match Status of each pairing. Next, click the Merge button to proceed to the Merge page.

eu3.lightning.force.com

Search Salesforce

MDM Home MDM Home Matched Record Pairs Rejected Record Pairs Matching Test Jobs Batch Job Runs More

Matching Matched Record Group [Back](#) [Save](#) [Delete](#) [Merge](#)

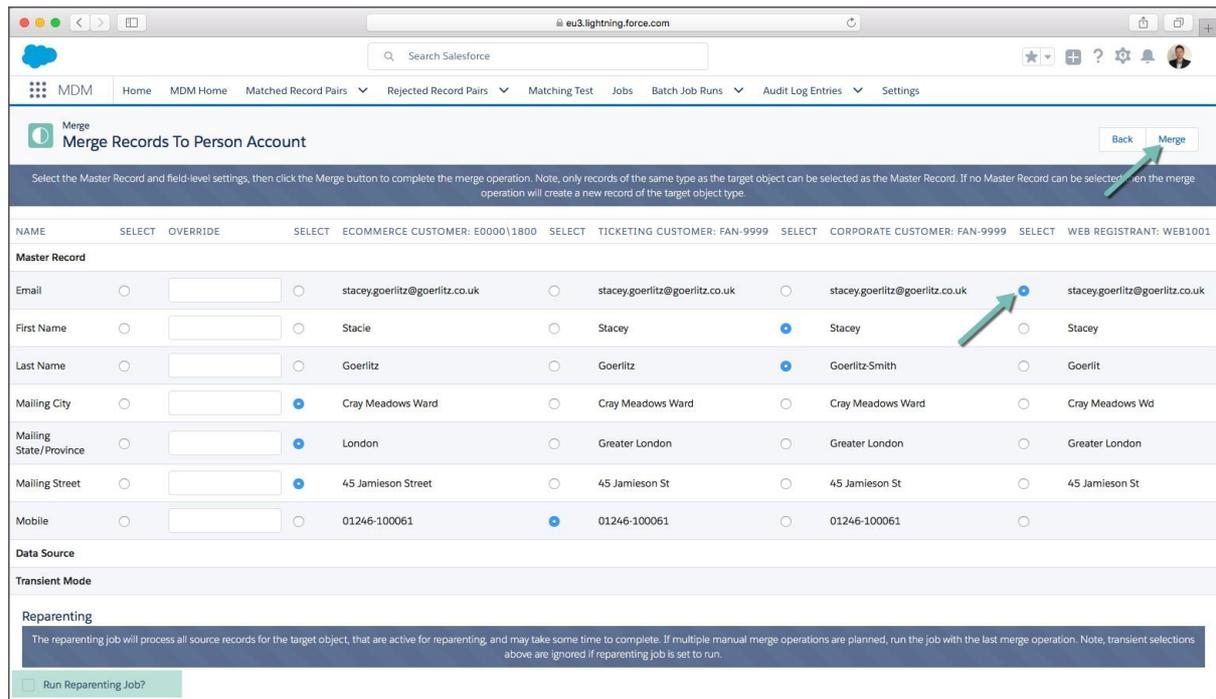
This page displays all Matched Record Pairs within a parent Matched Record Group. All record pairs share the same Blocking Key Match value and match to at least one other record in the group. Matched Record Groups contain transitive matches meaning if A=B and B=C then A=C. When matches are rejected that break the chain it is possible to Split the group into sub-groups using the "Split and Save" button.

Matched Record Group Information

Matched Record Group (Target Object)	Blocking Key Match Value
a00w0000009buCAAQ42339 (PersonAccount)	GOERS
Matching Date	Match Score Threshold %
2017-05-13 06:24	65.00 %

MATCHED RECORD PAIRS | MATCHED RECORDS

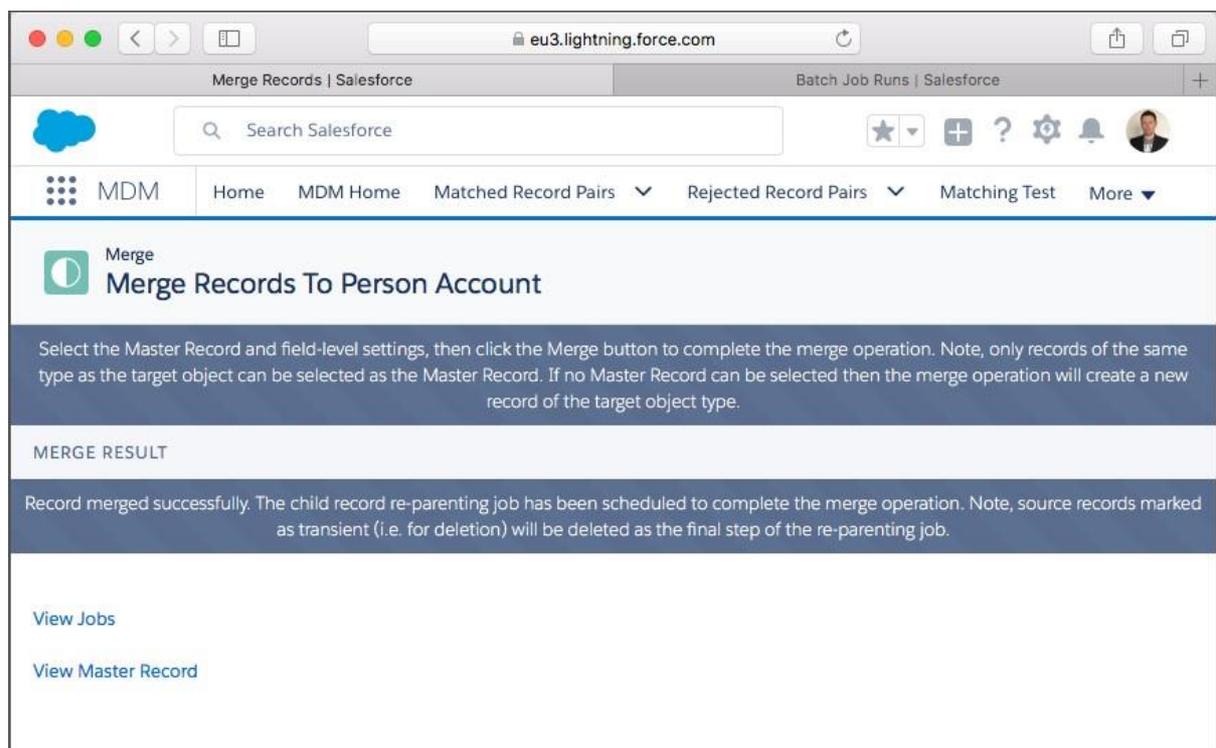
ACTION	MATCH TYPE	MATCH SCORE %	MATCH STATUS	RECORD 1 NAME	RECORD 1 DATA SOURCE	RECORD 2 NAME	RECORD 2 DATA SOURCE
Match Analysis	Fuzzy	90.00	Accepted	E0000\1800	ECommerce Customers	FAN-9999	Ticketing Customers
Match Analysis	Fuzzy	86.00	Accepted	E0000\1800	ECommerce Customers	FAN-9999	Corporate Customers
Match Analysis	Fuzzy	75.00	Accepted	WEB1001	Web Registrants	E0000\1800	ECommerce Customers



The Merge page allows individual field selections to be made across the Source Records; override values can be specified where all record values are blank or the required value is missing.

The option to run the Reparenting MDM operation (in selective mode) for the Matched Record Group post-merge is also offered.

Clicking on the Merge button (and providing confirmation to proceed), results in the page below being displayed with the option to open the Master Record.



Appendix A - Merge Settings Reference

Target Object Merge Settings are configured on the Target Object settings page.

Setting	Definition
Is Active?	If set to False the Target Object will be excluded from all Merge MDM operations.
Create Master for New Groups?	If set to True then a new Master Record will be created by the merge engine for Matched Record Groups that do not contain an existing Master Record.
MDM Status Field Name	A text type Custom Field on the Target Object that the Merge MDM operation will populate with the MDM Status value "Merge Source" or "Merge Master".
Last Merged Date Field Name	A datetime field on the Target Object that will be populated with the processing timestamp – Master Records only.
Ignore Partial Groups	If set to True then Matched Record Groups that contain at least one "Candidate" status Matched Record Pair will be skipped. If set to False then all groups that contain non-Rejected status pairings will be processed.
Default Record Type	The default Record Type for new Master Records.
Master Field Map Field Name	A text type Custom Field on the Target Object that the Merge MDM operation will populate with an internal field map which describes how the field population has been processed, i.e. for each Master Record field, which Source Record was used plus the Data Source priority for the Source Record. The field map is used to ensure that MDM field updates respect the data source priorities (or trust rules).
Merge Master Priority Field Name	A numeric type Custom Field on the Target Object that returns a relative priority scale (10 to 1 as an example). The Source Record within a given group with the highest value will be set as the Master Record. Typically Data Quality rules will be implemented to populate this field based on custom business rules. If this field is not populated the best matched record in the group will be selected (Match Score and Count).
Reset Identifier Target Fields	If set to True then the Merge MDM operation will set to blank all Data Source specified [Identifier Target Field Name] fields before population with the identifiers for the current operation.
Master Priority 2 nd Factor Field Name	Date or Datetime field used as the tiebreaker where the Master Record priority is equal – newest value wins.
Is Merge Recently Modified MRP?	If True , MRP are processed by the Merge Job only if the LastModifiedDate is within the date range set by the Application Setting (Recent Days Limit).
Is Invoked by Matching Job?	If set to True then Matching jobs will invoke a Merge job for the same Target Object upon completion.

Data Source Settings are configured on the Data Source settings page.

Setting	Definition
---------	------------

Data Source Merge Settings	
Is Active?	If set to False the Data Source will be excluded from all Merge MDM operations.
Master Record Id Source Field Name	<p>A Lookup field on the Source Object that references the Target Object SObjectType, e.g. Master Account.</p> <p>The Merge MDM operation will populate this field on Source Records with the Master Record Id to establish a relationship.</p>
Is Transient Mode Enabled?	If set to True the Data Source supports deletion of Source Records when the Merge MDM operation is invoked from the manual Merge page only.
Identifier Source Field Name	A text type Custom Field on the Source Object that provides value to be copied to the [Identifier Target Field Name] field on the Master Record.
Identifier Target Field Name	<p>A text type Custom Field on the Target Object that the Merge MDM operation will populate with the concatenation of values from the [Identifier Source Field Name] field on related Source Records.</p> <p>The pipe () character will be used to delimit the individual values.</p>

Appendix B – Troubleshooting

All clearMDM MDM operations log activity (Start and End times etc.) and errors to the Audit Log Entry object. The MDM application includes an **Audit Log Entry** tab to provide convenient access to this data. Each Audit Log Entry record is time-stamped and related to the parent Batch Job Run record: all MDM operations that run via the Job Method created a Batch Job Run record that records job statistics and status. Monitoring of the Audit Log should be a frequent activity performed by the Administrator or Data Steward responsible for the clearMDM implementation.

Best Practice: Salesforce Reporting Notifications provide a proactive means of reporting on errors generated by clearMDM operations. In this model a standard report is used to return data from the Audit Log Entry object where required conditions are met. Only when records meet the criteria is a report sent to the Administrator or Data Steward responsible.

Merge Log Types

Rule Type	Definition
Group Size Limit Reached - Source Records Will Be Ignored [XXXX] 7451:5000.0	<p>The number of Source Records for the Matched Record Group exceeds the setting value below.</p> <p>[System Settings] Max Records Per Iterable Cycle</p> <p>The above is necessary to limit the workload placed into each Batch Apex execute cycle such that CPU timeout errors are avoided.</p> <p>To mitigate this error, the System Settings below can be modified.</p> <p>Max Records Per Group : Controls the maximum Matched Record Group size. <i>Recommended Setting: 500</i></p> <p>Max Records Per Iterable Cycle : Controls the maximum number of records that can be placed into the workload for a single Batch Apex execute cycle. <i>Recommended Setting: 2000</i></p> <p>Max Records Per Iterable Cycle : Controls the maximum number of Matched Record Groups that can be placed into the workload for a single Batch Apex execute cycle. <i>Recommended Setting: 2 (this can be reduced to 1 where limit issues are record as below)</i></p>

CLEARMDM

Matched Record Group [XXXXX] Group Size Limit Reached - Source Records Will Be Ignored	The number of Matched Record Pairs for the Matched Record Group exceeds the setting value below. [System Settings] Max Records Per Group
Missing or Inactive Target Object Merge settings	This error can occur where the Target Object Merge settings are inactive, deleted or the Salesforce User does not have permissions to the object or fields referenced.