

1. Splitting Source Data

Splitting feature enables you to process large files. You must use splitting if your source file is huge. If you do not use splitting in this case, then Data Mapper may not be able to map it. If the data in the source file is huge, you can split the source data into different chunks. You can specify the size of the chunks in terms of the number of records or data streams. In the output, however, only one file is generated at the location specified in the target activity.

Steps to split the source data

1. Click **[+] Advanced Properties** in the *Edit Mapping* screen to expand the advanced properties of the new Mapping activity (see Figure 1.1).

Design > Services > Data Transform > Data Mapping

[-] Standard properties

Name *

Description *

Data Mapper

[-] Advanced properties

Splitting Data

Split Number of Records

Parallel processing

Maximum Concurrent Processors

Owner*

Permissions*

	Read	Write	Execute
Owner	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Group	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Other	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Mapping XSL*

Mapping XML*

* Mandatory fields.

Save **Cancel**

Figure 1.1: Advanced Properties

2. Select the *Splitting Data* checkbox to enable the splitting of data.



If you select this checkbox, then you have to split the data. You can do this by [setting the Splitter XPath](#) in the Data Mapper applet. Alternately, you can split the data by **Enabling Quick Splitting** feature.

3. Enter the number of records into which you want to split the data, in the *Split Number of Records* field. For example, if you have a source file of 100,000 records and you want split this file into records of 10,000 each, then you need to enter 10,000 in this field.
4. Select the *Parallel Processing* checkbox to parallel process these records. By default, records are processed sequentially.
5. Enter the number of records you want to process concurrently, in the *Maximum Concurrent Processes* field. By default, value of 1 is entered.



It is advised to change this value, if *Parallel Processing* is selected. Else, the records will be processed at one time only and hence sequentially.

6. Select the owner from the *Owner* drop-down list. By default, Administrator is selected.
7. Select the *Read*, *Write* and *Execute* checkboxes to set permissions for the Owner, Group or Other entities. By default, permissions are selected for the Owner and Group entities.
8. If mapping has been saved in the Data Mapper applet, then the XSL code is displayed in *Mapping XSL* field. This is a very useful feature. You can enter new XSL code or edit existing code to test or debug the mapping. If you save these changes, then they are reflected in the applet. For example, if you need to change the value of a constant, you can change it from this screen itself, instead of the applet. Additionally, any modifications made in the Mapping XSL code in the Data Mapper applet, will overwrite the XSL code in this field.



If the Advanced Properties and the Data Mapper screens are open simultaneously, then the changes made in the Mapping XSL field in the Advanced Properties screen will not be reflected in the Data Mapper screen.

9. The mapping XML code is displayed in *Mapping XML* field. You can enter new code or edit existing code. If you save these changes, they will be reflected in the applet.
10. Click **Save** to save the splitting details. A screen is displayed where you need to enter comments related to the splitting done.

Set Splitter XPath

To split data you need to set the Splitter XPath on the source tree. The data is split and sent to the transformer which executes the split data. You can set the Splitter XPath in case of a single source schema only. You can also set multiple Splitter XPaths in a source schema.

The following example explains the use of multiple Xpath splitting:

```
Root
|
|- R1
|   |-R1C1
|   |-R1C2
|- R2
|   |-R2C1
```

| |-R2C2

In this scenario we have **Root** element as the Root for schema. It has two child elements **R1** and **R2** at the first level. **R1** element has two children **R1C1** and **R1C2**. **R2** element has two children **R2C1** and **R2C2**.

If in the source xml, there are multiple occurrences for **R1** and **R2** elements, then we need to set split XPath on both **R1** and **R2** elements. If we set split XPath on **R1** element only then the Data Mapper will assume that **R2** has only one occurrence and vice versa. Similarly, if child elements **R1C1**, **R1C2**, **R2C1** and **R2C2** have multiple occurrences, then also **R1** and **R2** should be set as split XPath. Thus the thumb rule is that the top most node which is repeating, should be set as splitter XPath.

Once splitter XPath is defined, during execution, mapping will split the source tree using each splitter XPath and each chunk will have records specified by *Split Number of Records*.

In this scenario, when we set **R1** and **R2** as split XPath then *Split Number of Records* property is applied for both **R1** and **R2**. For example if the value of *Split Number of Records* property is 500 then 500 records of **R1** and 500 records of **R2** are selected.

Splitting data and setting Splitter XPath is very effective when the source data is very large. You can set the Splitter XPath in case of a single source and target schema. You can also set multiple Splitter XPaths in a source and target schema.



It is advised to always split the file and set the Splitter XPath if the file size is greater than 50 MB. This enhances the performance and increases data reliability, as in some cases such execution of large data can fail and generate errors.

Steps to set Splitter XPath

1. Ensure that all the [source and target schemas are loaded](#) and all their elements are listed under their respective nodes.
2. Right-click the source schema element that you want to split and select the **Set as Splitter XPath** option. A screen showing the Splitter XPath details is displayed (see Figure 1.2).

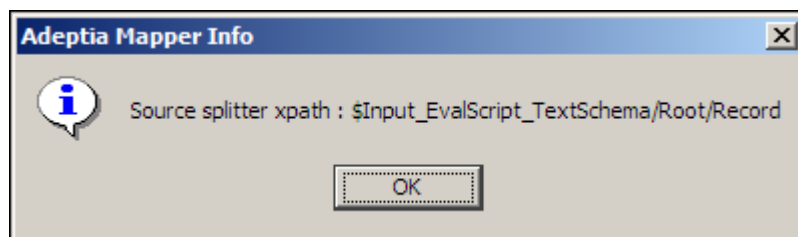



Figure 1.2: Splitter XPath Details



Splitting can be set only on root or record levels. You cannot split on an element level.

3. Click **OK**. This splits and sets the Splitter XPath on the selected source element.

	<p>Once a Splitter XPath is set for a source element, the letter (S) is displayed next to that source element. This signifies that the source element has been split and the Splitter XPath has been set. Refer to Table of Suffixes for details on suffixes displayed next to an element.</p> <p>Once you set the Splitter XPath, you have to set the Merger XPath. To know how to Set Merger XPath, refer to section</p>
---	---

Once you have set a Splitter XPath, you can view it in the Output XML.

Steps to view Splitter XPath

1. Click **View** menu and select **View Mapping XML**. The Mapping XML window is displayed. The Splitter XPath details are displayed with a \$Input tag prefixed to Source XPath code (see Figure 1.3).

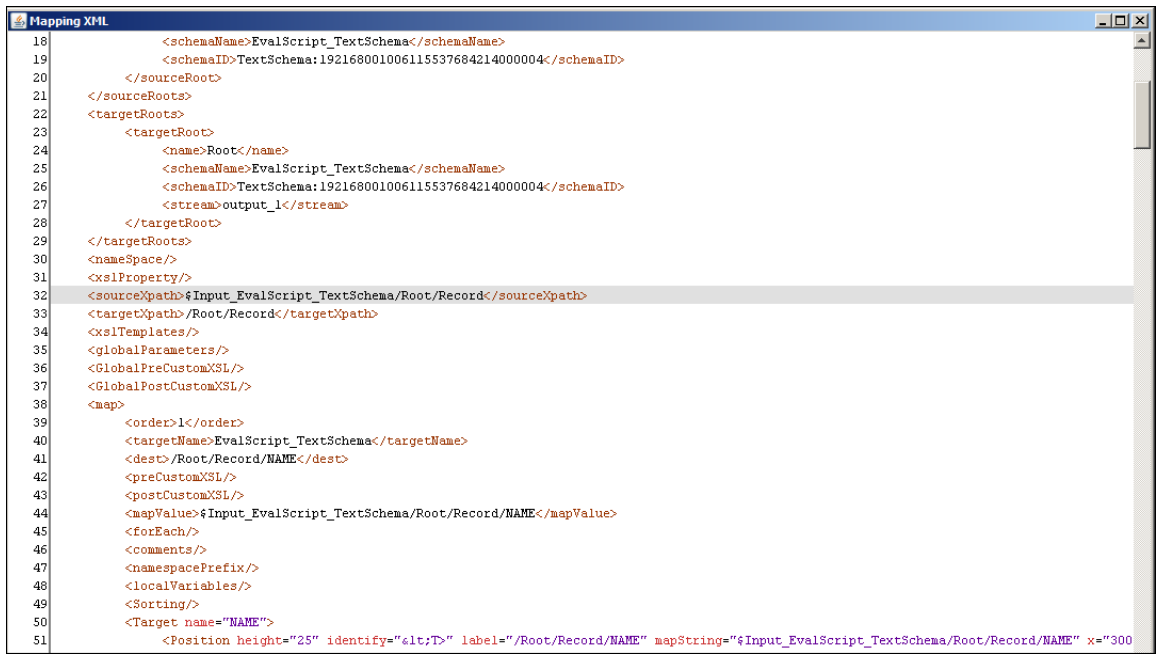


Figure 1.3: Splitter XPath in Mapping XML

You can remove a Splitter XPath that has been set for a source element.

Steps to remove Splitter XPath

1. Right-click the source schema element whose Splitter XPath you want to remove and select the **Remove Splitter XPath** option. A screen showing the Splitter XPath details is displayed (see Figure 1.4).

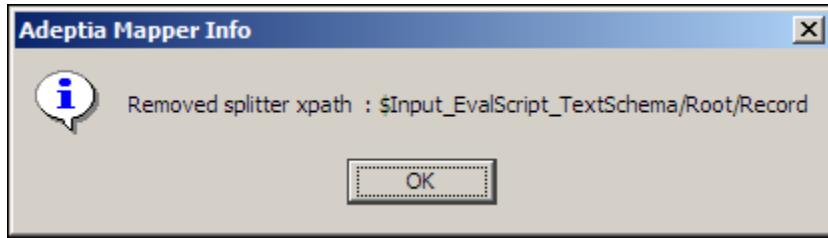


Figure 1.4: Remove Splitter XPath Details

2. Click **OK**. This removes the Splitter XPath set on the selected source element.

Set Merger XPath

Once you have split the source data by setting the Splitter XPath on a source element, you need to merge the split data at the target end. For this, you need to set the Merger XPath on a target element.

Setting the Merger XPath is largely dependent upon the Splitter XPath. You need to ensure that the number of Splitter XPaths and the Merger XPaths at the source and target schemas respectively are the same.

Steps to set Merger XPath

1. Ensure that all the [source and target schemas are loaded](#) and all their elements are listed under their respective nodes.
2. Ensure that a Splitter XPath has been set for a source schema element.
3. Right-click the target schema element on which you want to merge the split data and select the **Set as Merger XPath** option. A screen showing the Merger XPath details is displayed (see Figure 1.5).

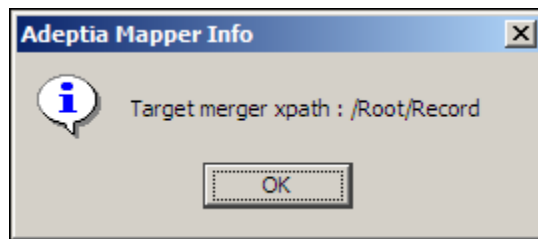





Figure 1.5: Merger XPath Details


	Merging can be set only on root or record levels. You cannot merge on an element level.
---	---

4. Click **OK**. This merges the split data and sets the Merger XPath on the selected target element.

	Once a Merger XPath is set for a target element, the letter (S) is displayed next to that target element. This signifies that the split data has been merged on the target element and the Merger XPath has been set. Refer to Table of Suffixes for details on suffixes displayed next to an
---	--

	element.
--	----------

	You can view Merger XPath details in the output XML. A \$ input tag is prefixed to the line containing the Splitter XPath details.
---	--

	When using Enable Quick Splitting feature, the Merger Xpath is also automatically set when this feature is selected.
---	--

Once you have set a Merger XPath, you can view it in the Output XML.

Steps to view Merger XPath

1. Click **View** menu and select **View Mapping XML**. The Mapping XML window is displayed. The Merger XPath details are displayed with *Target XPath* code (refer to Figure 1.3).

You can remove a Merger XPath that has been set for a target element.

Steps to remove Merger XPath

1. Right-click the target schema element whose Merger XPath you want to remove and select the **Remove Merger XPath** option. A screen showing the Merger XPath details is displayed (see Figure 1.6).

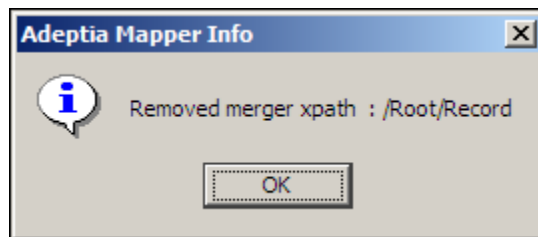


Figure 1.6: Remove Merger XPath Details

2. Click **OK**. This removes the Merger XPath set on the selected target element.

Enabling Quick Splitting

Another way of splitting source data is to enable the quick splitting feature. This splits the source and sets the Splitter XPath.

Steps to split data by enabling quick splitting

1. Ensure that all the **source and target schemas are loaded** and all their elements are listed under their respective nodes.
2. Select **Actions** menu and select *Enable Quick Splitting* option. Alternately, you can press **<Ctrl> + <Q>**. This splits the data and displays the splitter XPath details (see Figure 1.7).

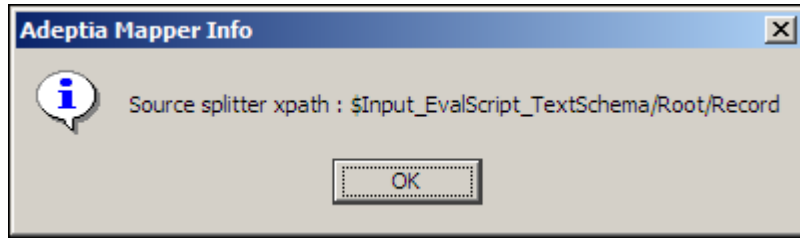


Figure 1.7: Splitter Xpath Details(Enable Quick Splitting)

3. Click **OK**. This splits and sets the Splitter XPath on the selected source element.

However, quick splitting splits the records in default numbers. To split into a specific number, you need to set the Splitter XPath explicitly.

	Once a Splitter XPath is set for a source element, the letter (S) is displayed next to that source element. This signifies that the source element has been split and the Splitter XPath has been set. Refer to Table of Suffixes for details on suffixes displayed next to an element.
--	--