IBM
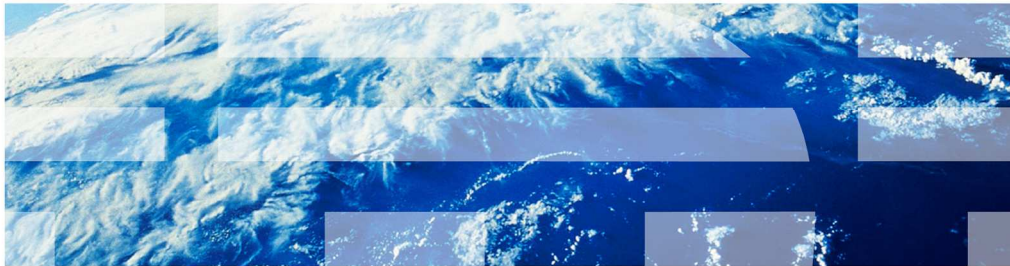
InfoSphere DataStage

Working with V-Pivot step in XML Assembly

This presentation provides a working example of how to use the Vertical-Pivot step, referred to as V-Pivot step, in the XML Assembly in the XML Pack version 3. The XML Assembly in the XML Pack version 3 is shipped with Information Server DataStage® version 8.5 fix pack 1 and is included in version 8.7 and version 9.1 as well.

IBM

Table of contents

2   © 2012 IBM Corporation

This presentation will provide an overview and working example of how to use the V-Pivot step in the XML Assembly. The example demonstrates how to map XML data that contains multiple list objects into new columns for a single output row. The presentation will review the source XML data and, using an existing XML Example job, show how to add and configure the V-Pivot step. The data mapping in the output step and review of the data produced will also be covered.

Overview

- V-Pivot step can transform records into fields of another record
- Useful to take multiple occurrence objects and add them to single row
- Presentation based on XML Pack v3 Example 1 job
  - 9.1
    http://publib.boulder.ibm.com/infocenter/iisinfsv/v9r1/topic/com.ibm.swg.im.iis.ds.stages.xml.core.usage.doc/topics/example1_parent_parsingxmldata.html
  - 8.7
    http://publib.boulder.ibm.com/infocenter/iisinfsv/v8r7/topic/com.ibm.swg.im.iis.ds.stages.xml.core.usage.doc/topics/example1_parent_parsingxmldata.html
  - 8.5
    http://publib.boulder.ibm.com/infocenter/iisinfsv/v8r5/index.jsp?topic=/com.ibm.swg.im.iis.ds.stages.xml.usage.doc/topics/example1_parent_parsingxmldata.html
- Data for XML example jobs
  - http://www-01.ibm.com/support/docview.wss?uid=swg27019894

3   © 2012 IBM Corporation

The V-Pivot step can help to manipulate XML data in a single stage so the required pivoting does not have to be done outside of the XML Stage. The V-Pivot step is capable of transforming multiple records into fields of another record. The V-Pivot Stage can be useful when XML data contains multiple child records that need to be added to a single row of output data.

This presentation is built from the XML Example 1 job that is shipped with Information Server and the XML pack. You should familiarize yourself with the example and have it working to get the most out of this presentation.

In the example displayed on this slide, the Input XML data contains multiple employee records and each employee can have multiple past and current job positions within different departments. In this example, the V-Pivot Stage is used to gather all current and past job titles from the child elements and puts them into a single output row for each employee. Displayed on this slide are arrows pointing at the two XML elements that are used in this example to pivot the data. The use of these two elements are explained in greater detail in later slides.

Job design (1 of 2)

- Edit XML Example 1 job
- Remove Address link and file
- Set null handling

© 2012 IBM Corporation
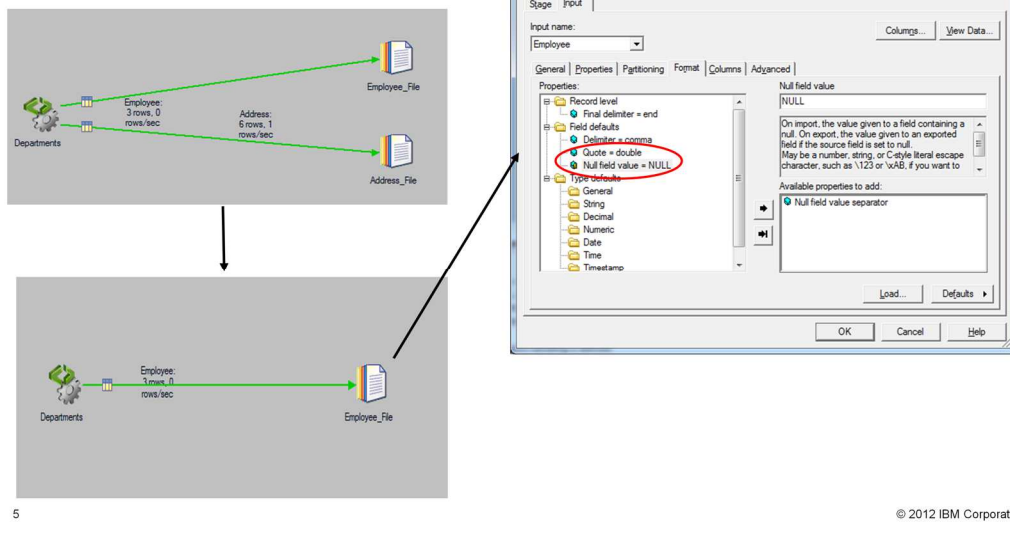
This V-Pivot example uses the XML Example 1 job as a starting point. The first step is to remove the Address link and Address_File stages from the canvas, since the focus is on the Employee_File output. The second modification is to add null handling in the Employee_File stage, since the XML Stage in this example will produce null values for some of the pivoted columns.

- Open XML Stage
- Click Edit assembly
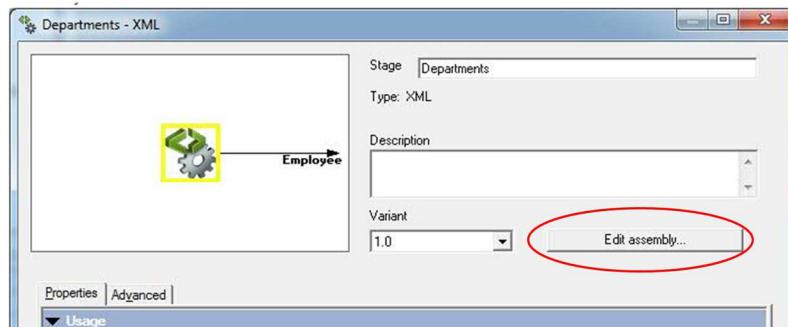
Working with V-Pivot step in XML Assembly

© 2012 IBM Corporation

Once the unnecessary link and stage have been removed, it is time to work on the XML assembly and add the V-Pivot step. First, open the XML Stage Properties and click the Edit assembly button to launch the Assembly Editor.

Adding V-Pivot step to XML Assembly

- Add V-Pivot between XML Parser and Output
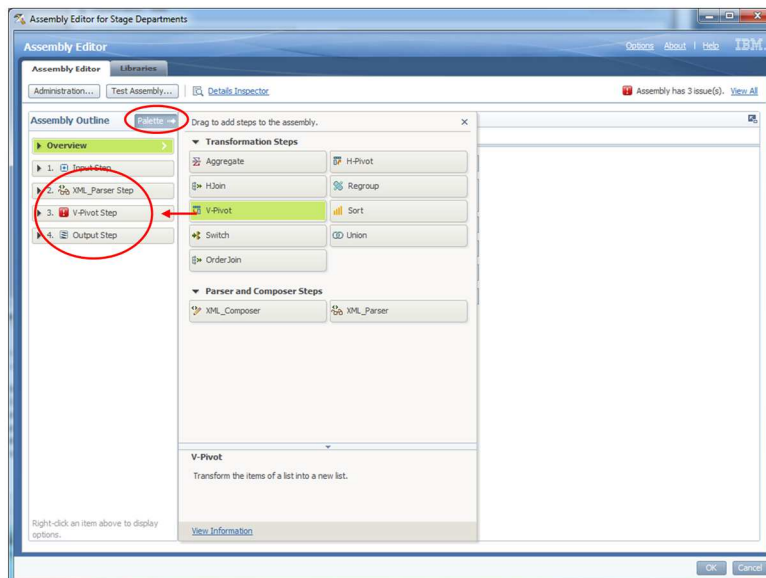
Once in the Assembly Editor, the V-Pivot step can be found in the Palette. If your Palette does not display like it does here on the slide, click the Palette button. Grab the V-Pivot on the Palette and drag it to the left. Drop it between the XML Parser and Output Step so it appears as it does in the screen capture displayed on this slide. Click the V-Pivot step to begin the required setup.
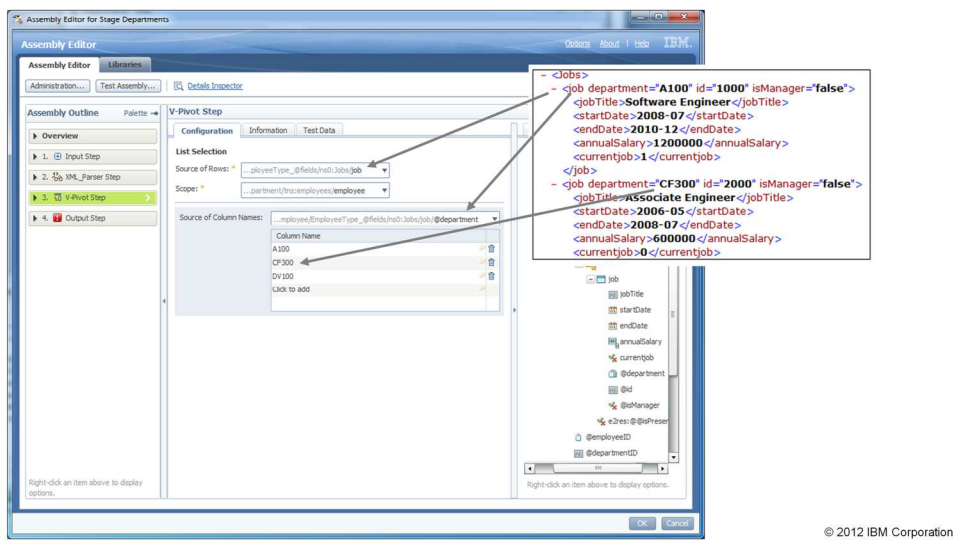
In the V-Pivot step there are several things that have to be setup correctly to obtain the desired output. The first step is to click the drop down for "Source of Rows". This needs to point to "job" since that record contains the jobTitle element that is going to be appended to the output records.

Next, click the drop down for Scope and set it to "employee", so the pivoting is limited to each employee's child records and not all job records within the XML data.

Next, the Source Of Column Names needs to be set. This needs to point to the @department element. Next, you need to create column names to represent each possible value for job department. You can call these columns anything you like but for simplicity, this example uses the same names as the job department. Basically, this sets the V-Pivot step to collect all the job records for an employee that has a matching @department value defined in the Column Name list. The matched records can then be mapped in the output step.

Output Step and Mapping data (1 of 2)

- Click Output Setup => Configuration => Output
- Add job title columns
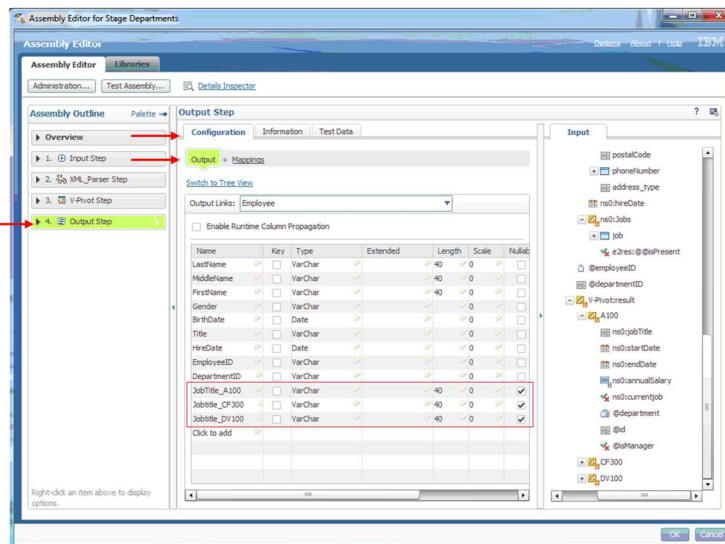
© 2012 IBM Corporation

The next step is to set up the output step. Click "Output Step" on the left side of the dialog box and then the Output link under Configurations. Before the data from the V-Pivot step can be mapped, three new columns need to be added to the output step as displayed in the red box on this slide. These three columns will hold the job titles that the employees hold or held at these departments. Again, these column names can be what ever you choose.

Output Step and Mapping data (2 of 2)

- Click Mappings link
- Match jobTile column to V-Pivot result set
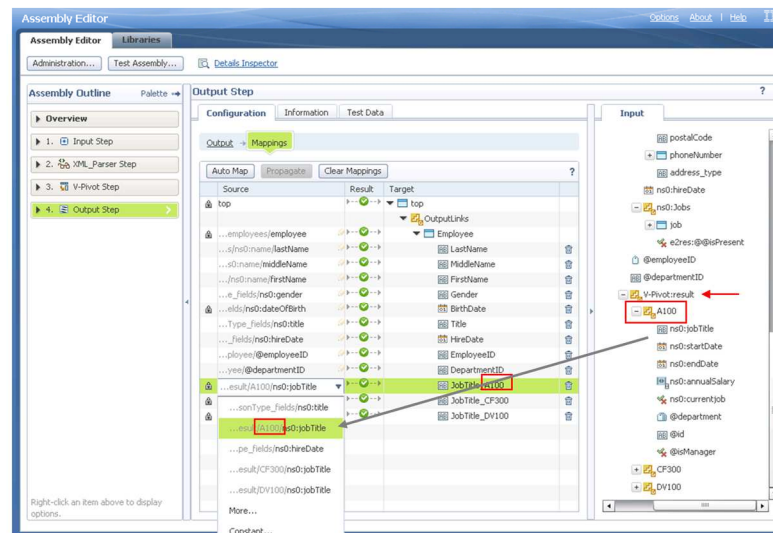
Next, click the Mappings link under the Configuration tab. Each jobTitle element from the matching V-Pivot results set, identified by the Column Names defined earlier in the V-Pivot step, must be mapped to target columns as seen in the screen capture displayed on this slide. Click the Source drop down for each of the new columns created on the previous slide and pick the matching jobTitle from the V-Pivot result set. Once this is completed, save, compile and run the job.

Output data produced

11

© 2012 IBM Corporation

Once the job is run, right click the sequential file stage and choose View Employee Data. The first output record clearly shows how the data from the multiple child records was pivoted to a single output row. Mr. Wright has not worked for department "DV100" thus, NULL is shown in the Jobtitle_DV100 column.

Feedback

Your feedback is valuable

You can help improve the quality of IBM Education Assistant content to better meet your needs by providing feedback.

- Did you find this module useful?

- Did it help you solve a problem or answer a question?

- Do you have suggestions for improvements?

Click to send email feedback:

mailto:iea@us.ibm.com?subject=Feedback_about_V_Pivot_XML.ppt

This module is also available in PDF format at: ../V_Pivot_XML.pdf

Working with V-Pivot step in XML Assembly

© 2012 IBM Corporation

You can help improve the quality of IBM Education Assistant content by providing feedback.

IBM

Trademarks, disclaimer, and copyright information