

Data Flow Transformations


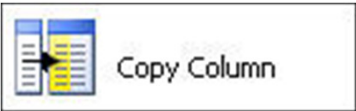

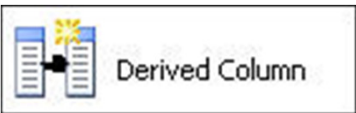

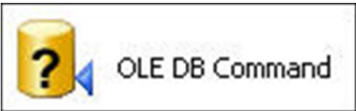
Modifying data structure and content



Overview







- Row Transformations
- Rowset Transformations
- Split and Join Transformations
- Business Intelligence Transformations
- Other Transformations

Row Transformations

| Transformation | ETL Example |
|--|--|
|  Character Map | Apply string conversions, such as uppercase, to character data in a column |
|  Copy Column | Create a new column containing data from an existing column |
|  Data Conversion | Convert data to a new data type |
|  Derived Column | Apply expressions to a data column, such as arithmetic or concatenation |
|  Script Component | Perform a custom transformation |
|  OLE DB Command | Perform a SQL Update statement |








Synchronous Transformations

Rowset Transformations






| Transformation | ETL Example |
|---|---|
|  Aggregate | Aggregate values by group |
|  Sort | Sort rows (and optionally eliminate duplicates) |
|  Unpivot | Convert columns to rows |
|  Pivot | Convert rows to columns |
|  Percentage Sampling | Randomly select percentage of rows |
|  Row Sampling | Randomly select specific number of rows |

Asynchronous Transformations

Split and Join Transformations



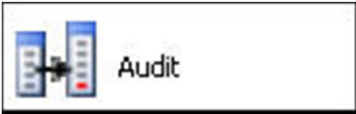

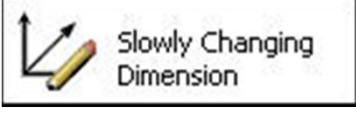
| Transformation | ETL Example |
|---|---|
|  Conditional Split | Apply rules each row to assign applicable output |
|  Multicast | Copy each row to each output |
|  Union All | Combine two inputs for a common output |
|  Merge | Combine two sorted inputs for a common output |
|  Merge Join | Join two sorted inputs for a common output |
|  Lookup | Decode a value in one output based on data from another source |
|  Cache Transform | Store data as a file or in memory for use in a lookup operation |

Business Intelligence Transformations

| Transformation | ETL Example |
|---|--|
|  Fuzzy Grouping | De-duplicate data |
|  Fuzzy Lookup | Perform lookup without requiring exact match |
|  Term Extraction | Generate set of English nouns from input text |
|  Term Lookup | Perform lookup from English nouns found in input text |
|  Data Mining Query | Applies prediction query using data mining model to input data |

Available in Enterprise Edition only

Other Transformations

| Transformation | ETL Example |
|---|---|
|  Export Column | Export column to a specified file named in a pair of columns |
|  Import Column | Imports data from file as a row column, such as a product image |
|  Audit | Add data about package execution to data flow (package name, execution start, etc.) |
|  Row Count | Count rows in the data flow and store as variable |
|  Slowly Changing Dimension | Automatically generate transformations for Type 1 and Type 2 SCDs |

Summary

- **Row Transformations**
 - Operate on data row by row
- **Rowset Transformations**
 - Restructure entire data set
- **Split and Join Transformations**
 - Separate or combine data
- **Business Intelligence Transformations**
 - Advanced transformations available only in Enterprise Edition
- **Other Transformations**
 - Blob data, auditing, row counts, and slowly changing dimension handling

References

- **Integration Services Transformations**
 - <http://msdn.microsoft.com/en-us/library/ms141713.aspx>
- **CodePlex – search for Integration Services Transformation**
 - <http://www.codeplex.com>
- **SQLIS.com**
 - <http://www.sqlis.com/sqlis/category/Component-Downloads.aspx>