



Transformation Manager Open Interfaces

Introduction

Transformation Manager (TM) provides in-built access to a wide variety of data formats. It supports XML files with model definitions either supplied as DTD or XSD, and is also able to derive models when only an XML instance is available. It can support many types of relational databases, including Oracle, Sybase, DB2, Microsoft SQL7 and also provides support for Java classes.

In-built support is provided in two parts:

1. in-built model loaders using introspection modules
2. in-built adapters for read and write of instance data

Although this in-built support provides access to a large variety of data, in order to cater for all forms of data, whether as a discrete unit such as a database or file or as integral part of a client architecture responding to transformation requests, ETL provide two forms of Open Interface:

1. The Loader Open Interface allows models, including those built against custom APIs to be loaded into the Transformation Manager (TM) repository at design time.
2. The Instance Data Open Interface allows access to any form of instance data at run time.

In addition TM Deployment Interface provides extensive control of transforms and permits Open Interfaces to be selected at run-time.

'TM Open Interfaces' were initially introduced, as a result of our work in the Geoscience sector, to allow Transformation Manager to be integrated into a number of client architectures. These interfaces are open in the sense that ETL publish a Java interface description. Thus any code, including user-written code that implements the appropriate description, can be called directly by Transformation Manager (at runtime or design time). As such, these interfaces support any data access method and any model. Models may be traditional, with objects, attributes and relationships, or hierarchical, i.e. also include child parent relationships. All of Transformation Manager's design, test and deployment tools use TM Open Interfaces.

TM Open Interfaces Concepts

Because Transformation Manager, especially at design time, works with a high level abstraction this approach is exceptionally powerful. For example all models, whether loaded by in-built Java, XML and RDBMS loaders or through the new TM Open Interfaces, Transformation Manager's common design paradigm remains the same:

- A Source and Target model may form the basis of one or many transformations
- Transformation Builder loads model information into Transformation Repository
- As the models are loaded, the Model Loader deduces information about the underlying model based on the model's metadata and stores models in one consistent form, consisting of elements, attributes and relationships such that:
 - an element has a name and a collection of attribute and relationship properties
 - an attribute has a name and is linked to a new or existing 'type'
 - a relationship has a name, properties defining its cardinality and rules which determine how to construct it and navigate it
- Once a transform has been written, transform code is generated.

- At run time, the generated code calls read and write adapters which access the source and target instance data

All that is required from the user is:

1. a description of the entities, attributes and relationships to be provided via TM Open Interfaces standard
2. implementations of simple data accessors according to the same standards

Once implemented models and instance data will be presented identically in the design and test tools.

TM Open Interfaces Use

There are three TM Open Interfaces provided by Transformation Manager:

1. TM Loader
2. TM Generic Adapter
3. TM Deployment

The use of these as required in the design and deployment stages is shown in Figure 1. The test stage (not shown) is very similar to the deployment stage, except that the transforms are executed directly by the Transformation Test Tool.

The TM Loader interface allows the user to define a model as a set of elements, attributes, and relationships. Transformation Manager's development environment, Transformation Builder, allows a user's model type and name to be entered so that it can be loaded into the Transformation Manager repository.

The TM Generic Adapter interface allows instance data, corresponding to models that have been defined using the TM Loader interface, to be read and written. Unlike the RDBMS, XML, and Java Adapters which are created implicitly, the TM Generic Adapter is created explicitly by the user. It is essentially user code that implements the relevant TM Generic Adapter interface read and write calls.

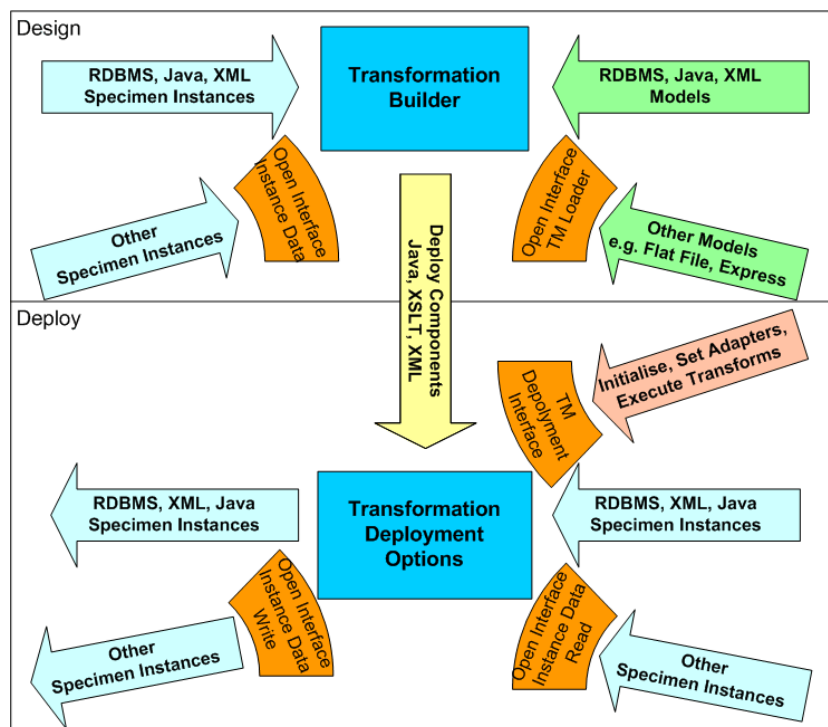


Figure 1 - Using TM Open Interfaces

The TM Deployment interface comprises both a short-lived interface and a long-lived interface. A short-lived interface completes a transform in one operation. In contrast the long-lived interface allows execution of a



transform on an instance by instance basis thus giving the user very fine control over sessions and connections. The latter approach is ideally suited for inclusion in client architectures. Both pull and push models are supported for selecting data. Using a pull model the selection criteria may be expressed in the context of the target model and is automatically converted to a selection query in the context of the source model, a feature described as 'Propagation of Selection'.

Example Uses

Support for the ISO 10303 Express model and associated Part 21 instance data, and Geoscience Log ASCII Standards (LAS) version 2 files are available as first examples of using TM Open Interfaces.

It is planned to apply TM Open Interfaces to extend Transformation Manager's support for an ever increasing range of interfaces. Initially work has begun to provide support for interfaces mainly in the Geoscience, Financial, and Engineering sectors, for example:

- General ASCII File support
- ISO Express models used in business sectors such as Aerospace, Defence, and Engineering
- Swift messages used in the Financial sector
- ISO 9735 Electronic data interchange for administration, commerce and transport (EDIFACT)
- DLIS, LIS, BIT, and other ASCII well log data file formats used in Geoscience
- Geoshare RP66 binary file formats also used in Geoscience

The list of possible interfaces is limitless and ETL are forming partnerships with companies in each domain with particular expertise in these areas. ETL intend to publicise TM Open Interfaces so that Transformation Manager users, expert in their domain, can extend support to their specific domain requirement.

User-defined adapters are subdivided into two general groups according to the model in use:

1. Those that give access to generic information which can be expressed as elements, attributes and relationships. This includes data modelled in Java or database and can also include data which is taken from APIs. We call these models generic.
2. Those that require information where elements are contained in a hierarchy or tree. We call these models hierarchical.

A number of typical examples are being constructed to cover the spectrum of models likely to be encountered:

JDBC/read	API/read
JDBC/write	API/write
Java/read	hierarchical/read/LAS
Java/write	hierarchical/read/Edifact
CSV/read	tree/target
CSV/write	Event/target
	flat/target
	Stored Procedure

Further description of the TM Open Interfaces are contained in the User Guides:

1. Instance Data Open Interface Guide
2. Loader Open Interface Guide

Further Information

Website : www.etlsolutions.com
Email : info@etlsolutions.com
Telephone : +44 (0) 1912 894040