

Rational Reporting

Module 3: IBM Rational Insight and IBM Cognos Data Manager

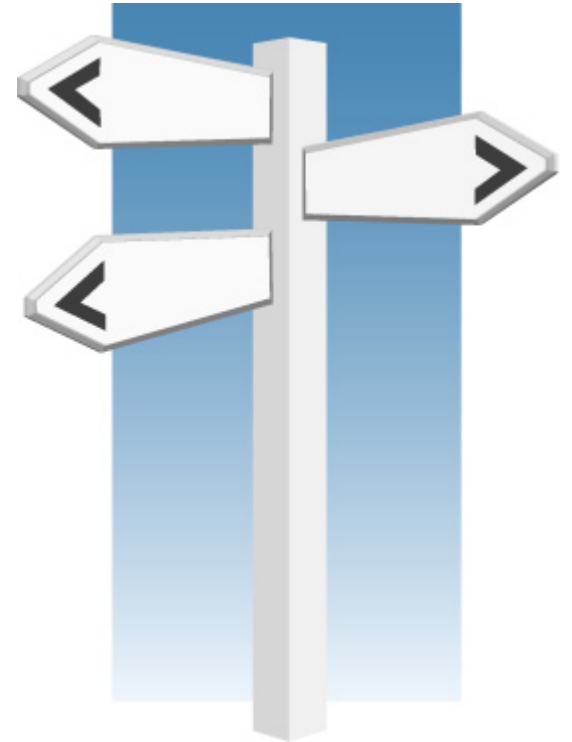


Rational. software



What's next?

- Module 1: RREDI and IBM Rational Insight Introduction
- Module 2: IBM Rational Insight Data Warehouse
- **Module 3: IBM Rational Insight and IBM Cognos Data Manager**
- Module 4: IBM Rational Insight and IBM Cognos Framework Manager
- Module 5: Third-party integrations



Target audience and objectives

Target audience:

- IBM® Rational® Insight users who want an introduction to ETL (Extraction-Transformation-Loading) implemented by IBM Cognos® Data Manager

Objectives: After reading this document, you will be able to:

- Understand the Rational Insight Extraction-Transformation-Loading implementation based on Cognos Data Manager

Prerequisite:

- Review modules 1 and 2

Your primary sources of information are the information centers for Rational Insight and the Rational solution for Collaborative Lifecycle Management (CLM).

These presentations are intended as conceptual introductions.

Introduction

This presentation provides an introduction to IBM Rational Insight Extraction-Transformation-Loading implemented by IBM Cognos Data Manager.

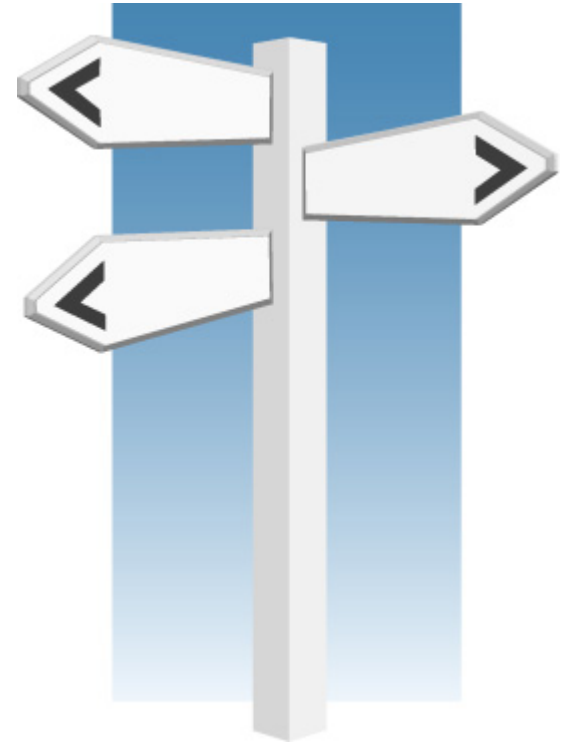
Note: This presentation *does not* provide you with technical knowledge on how to perform tasks. You should take formal Cognos Data Manager and Framework Manager courses before working with Cognos Data Manager and Framework Manager.

For recommended training courses, see the IBM Rational Insight Data Modeler training path:

<https://www-304.ibm.com/jct03001c/services/learning/ites.wss/us/en?pageType=page&c=L806673M20298F27>

What's next?

- Module 3: IBM Rational Insight and IBM Cognos Data Manager
 - Understanding ETL and Data Manager concepts
 - Review data warehouse structure
 - Rational Insight ETL catalog structure
 - Fact builds
 - Job streams



Data Manager (ETL) overview

The primary purpose of Cognos Data Manager is to deliver data into data warehouses and calculate data for reporting, analysis, and performance management.

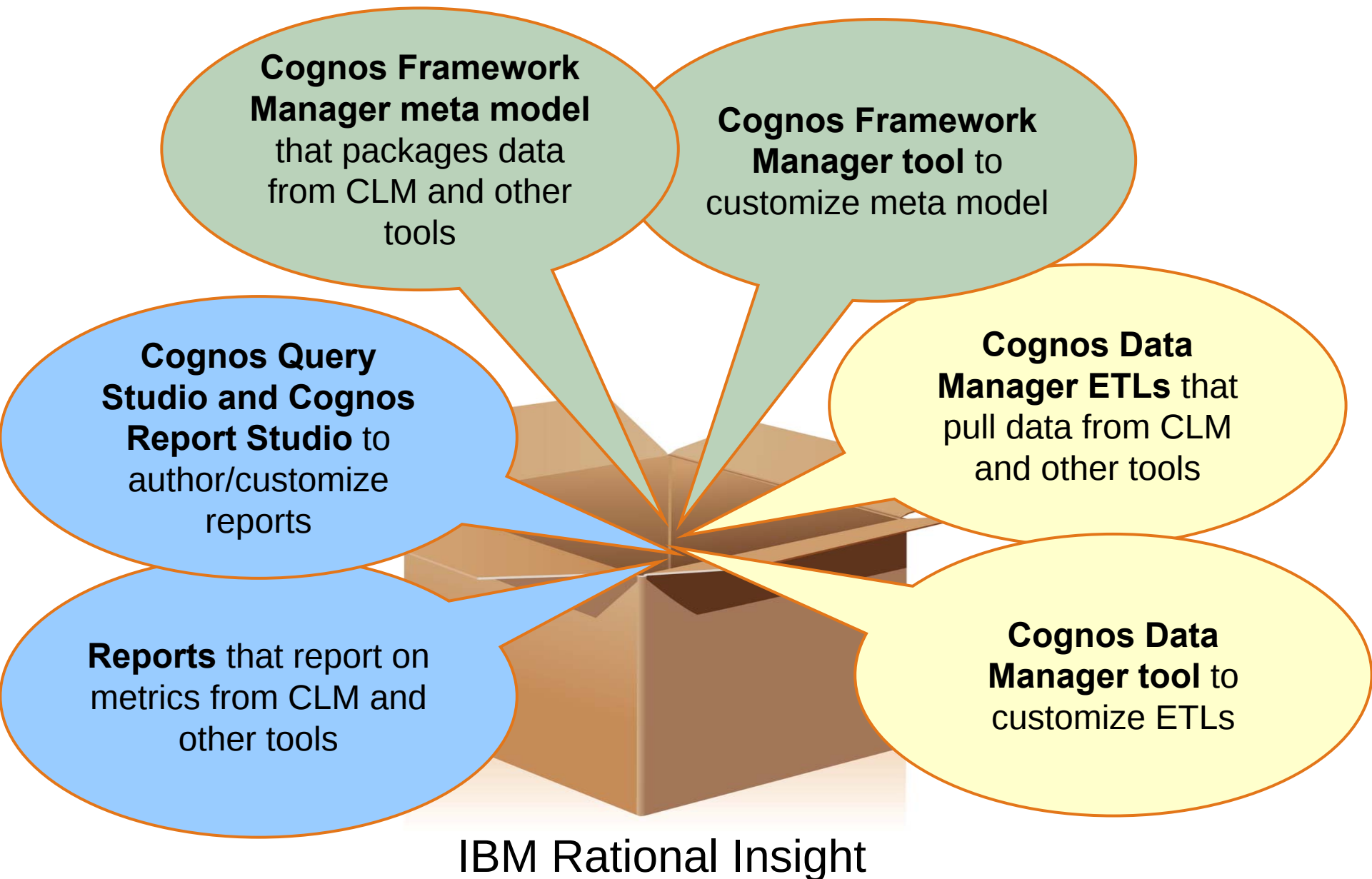
Data Manager does the above through:

- **Extracting** operational data from multiple sources
- **Transforming** and merging the data to facilitate enterprisewide reporting and analysis
- **Loading** the transformed data into coordinated metrics tables

ETL can deliver the following types of data:

- **Operational data store:** Pick up necessary data from various data sources, such as test cases from quality management tools and defects from change management tools.
- **Fact data:** Contains values that represent facts, such as the number of defects filed against a project in a week.
- **Dimension data:** Defines the core business components, such as the range of products that the company produces.

Value of IBM Rational Insight



Rational Insight provides prebuilt ETLs

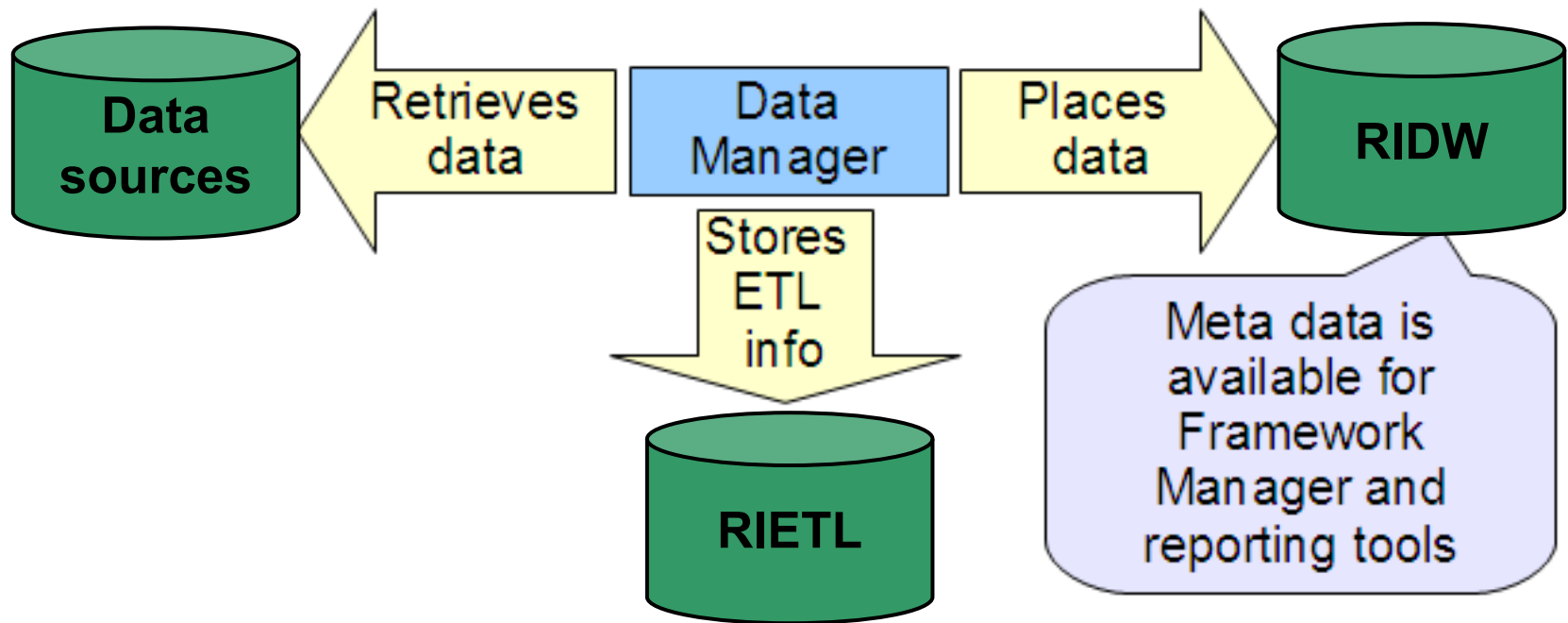
	RRDI includes	Insight includes
IBM Cognos Connection	x	x
IBM Cognos Query Studio	x	x
IBM Cognos Report Studio	x	x
IBM Cognos Active Report*		x
IBM Cognos Business Insight*		x
IBM Cognos Framework Manager		x
IBM Cognos Data Manager		x

Rational Insight includes:

- Prebuilt Data Manager ETL logic that allows Rational Insight to report on software development metrics from CLM and other tools.
- The Data Manager tool. This tool allows you to customize the prebuilt ETLs.

Data Manager (ETL) overview (continued)

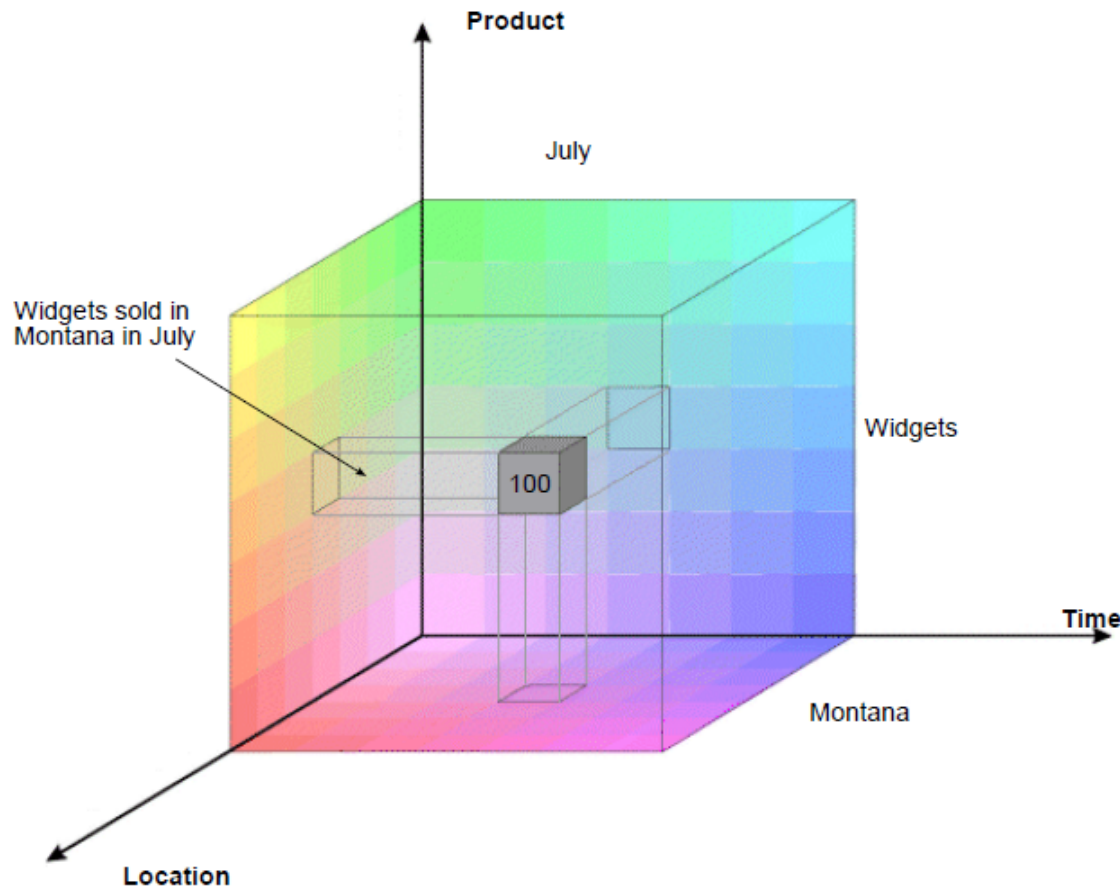
- Data Manager integrates with Rational Insight components through metadata that is defined in the data warehouse.
- This metadata allows target data warehouse and data repositories to be modeled and used in Rational Insight reports.



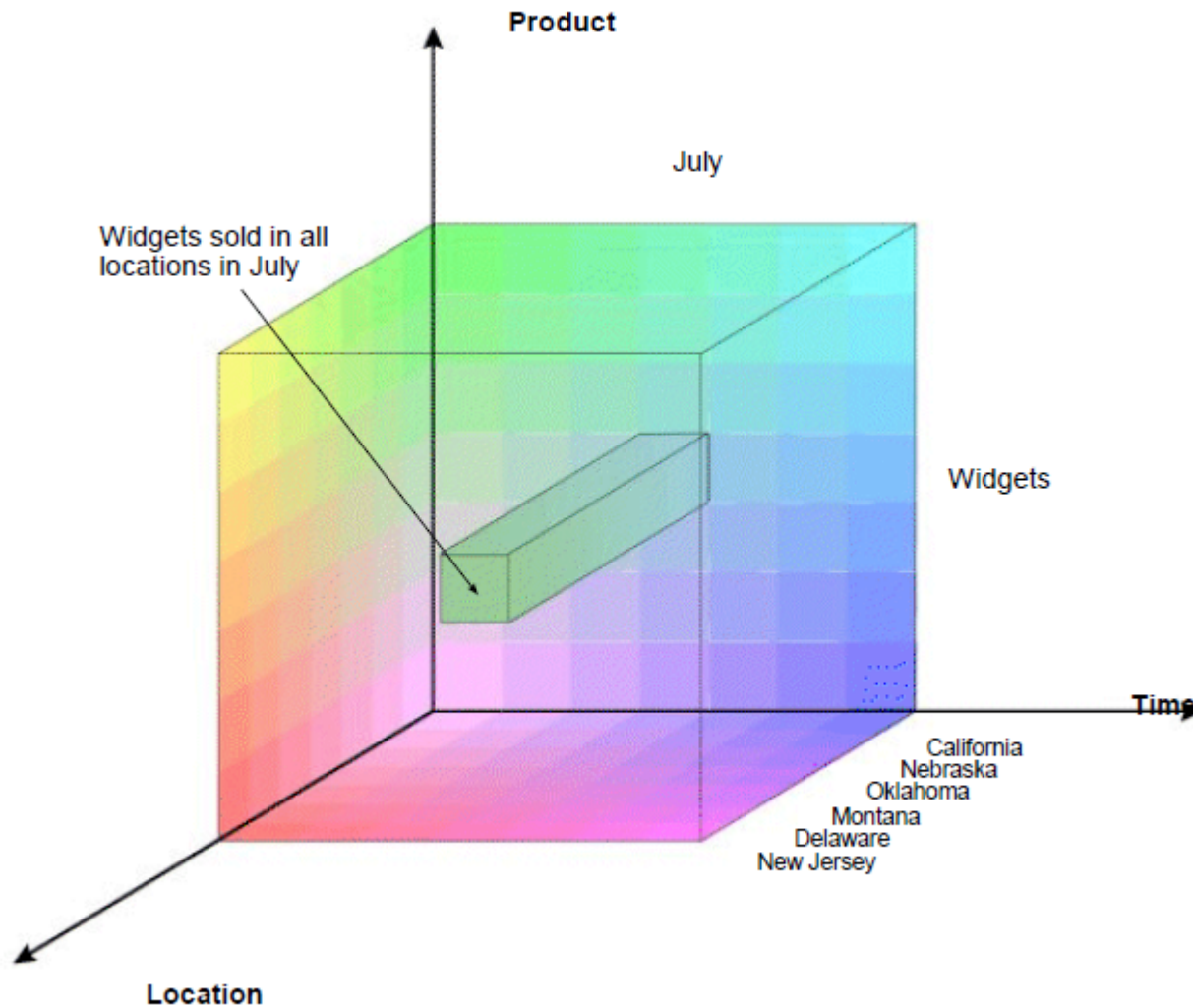
Data Manager terminology – business dimensions

Business dimensions are:

- Core components or categories of a business – anything that you want to analyze in reports
- Fundamental to data organization, and provide context for numeric data items or measures



Data Manager terminology – business dimensions (continued)



Dimension mapping category

- A set of similar dimension maps
- Example: A category can be created for days, weeks, and hours, with different levels of granularity

Dimension table

- A table that contains attributes that are used to constrain and group data when performing data warehousing queries

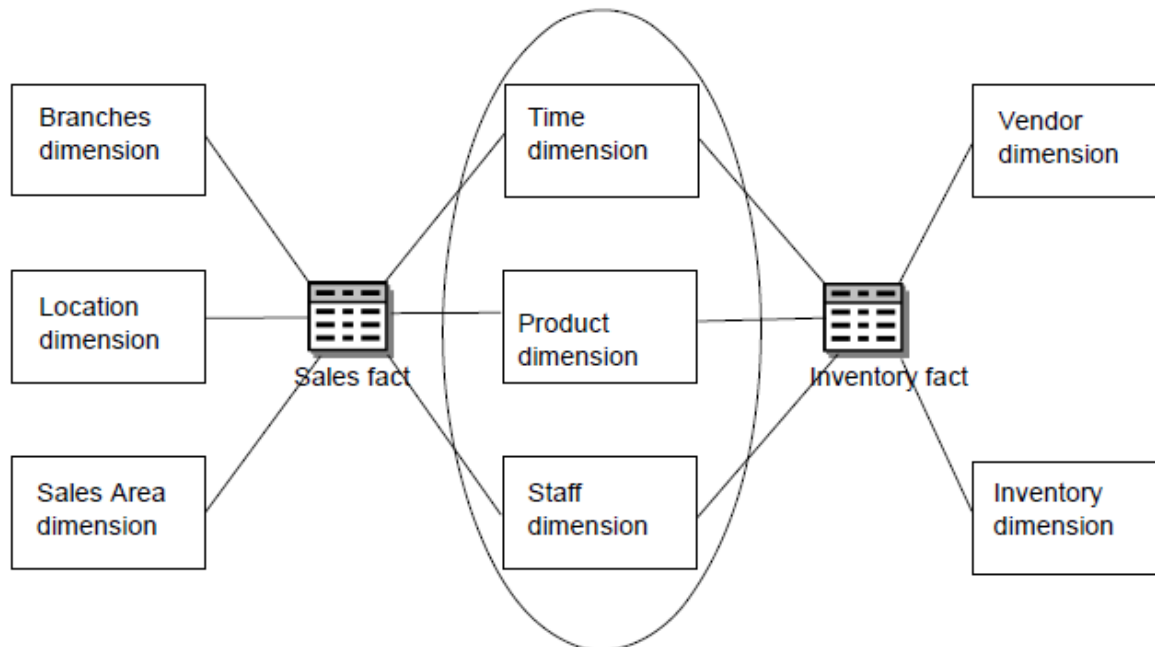
Dimension build

- A build that delivers data to describe a single business dimension, such as a product or a customer
- Includes source dimension table, a hierarchy for the data, a target database, and a target table

Data Manager terminology – fact table

A *fact table* is table with two types of columns: a measure and a dimension.

1. **Measure columns:** Represent an aggregation, such as a sum or count
2. **Dimension columns:** Foreign keys for dimension tables that are used to describe the measure



Data Manager terminology (continued)

Look-up build

- A simple, single-level reference structure in an ETL catalog
- Track the relationships between fact data and the operational data store that contributes to the fact

Measure

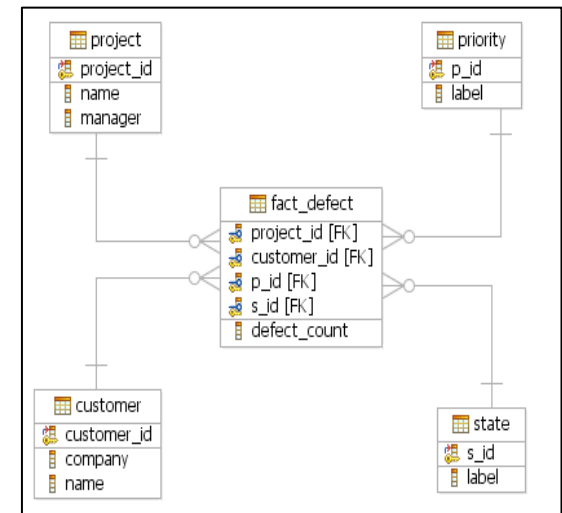
- A value that can be analyzed, such as the number of defects

Metric

- A measure to assess performance in a key area of a business

Star schema

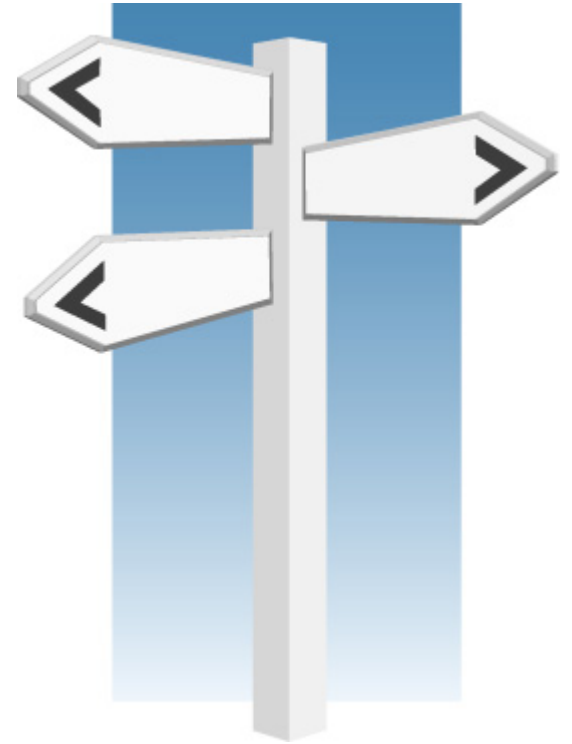
- A type of relational database schema that is composed of a set of tables making up a single, central fact table that is surrounded by dimension tables



star schema

What's next?

- Module 3: IBM Rational Insight and IBM Cognos Data Manager
 - Understanding ETL and Data Manager concepts
 - Review data warehouse structure
 - Rational Insight ETL catalog structure
 - Fact builds
 - Job streams



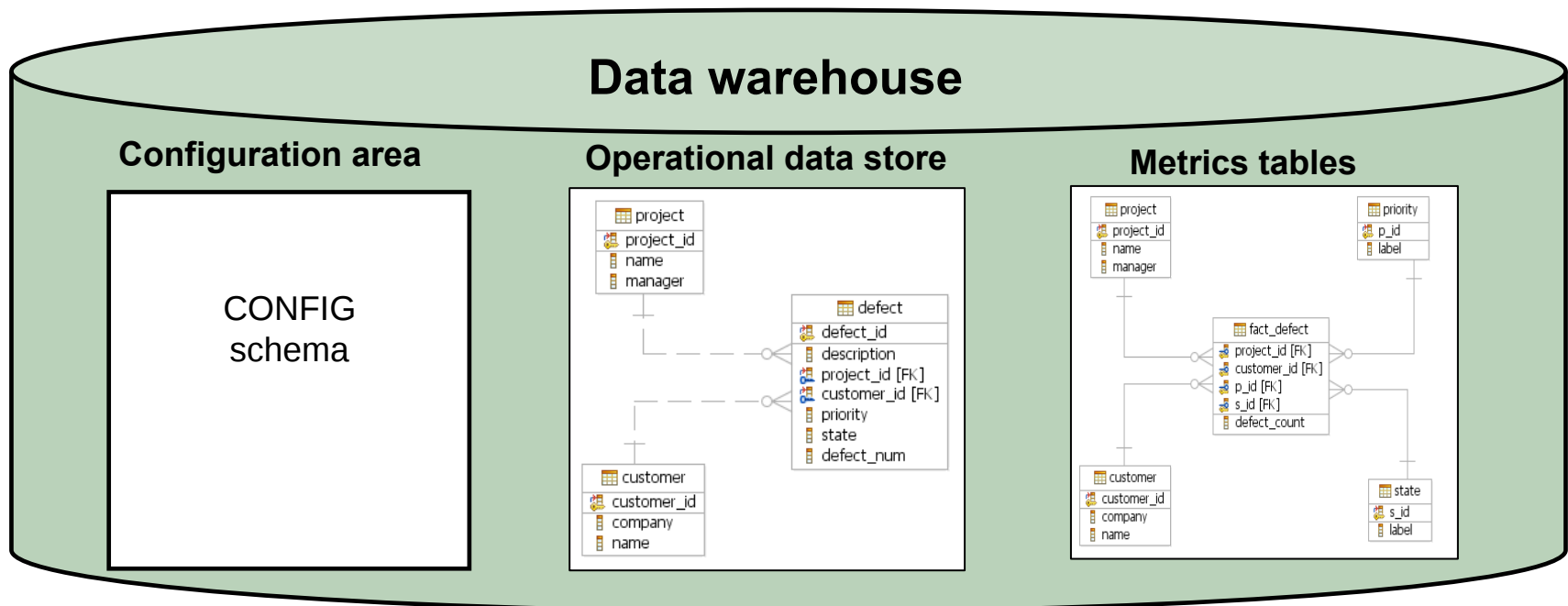
Data warehouse structure

IBM Rational Insight data warehouse (RIDW) is the component that:

- Stores data from various operational systems in a format that facilitates analytical reporting
- Provides the engine that is required for query optimization, multidimensional analysis, and historical trending

RIDW contains three physical areas, all defined and residing within one database:

- **Configuration area**
- **Operational data store (ODS)**
- **Metrics tables** (implemented as star and snowflake schemas)



The configuration area of the Rational Insight data warehouse:

- Identified in the physical model by the CONFIG schema
- Contains and maintains warehouse configuration information
- Serves as a reference point for auditing and tracking the ETL (extract, transform, and load) activity

The operational data store (ODS) area of the Rational Insight data warehouse:

- Identified in the physical model by the RIODS schema
- Integrates data from various operational sources
- Location used by ETL process to populate and store data
- Data is manipulated and cleansed to meet quality levels
- Maintains pertinent information that is needed in analytics (it is not intended to be a copy of the operational source)
- Implemented as a relational model

RIDW – Metrics tables area (data mart)

The metrics tables area of the Rational Insight data warehouse:

- Identified in the physical model by the RIDW schema
- Data is derived from the operational data store
- Consists of fact and dimension tables that are arranged in a star schema form (fact tables that are surrounded by dimension tables)

Note that RIDW has an open schema.

Query optimization:

- A common data model for all data, regardless of data source
- Facilitates easy reporting of data from multiple data sources

Cross-product reporting:

- Extracts data from various development domain areas
- Relates them together in a meaningful way, enabling cross-product correlation reporting

Historical trending:

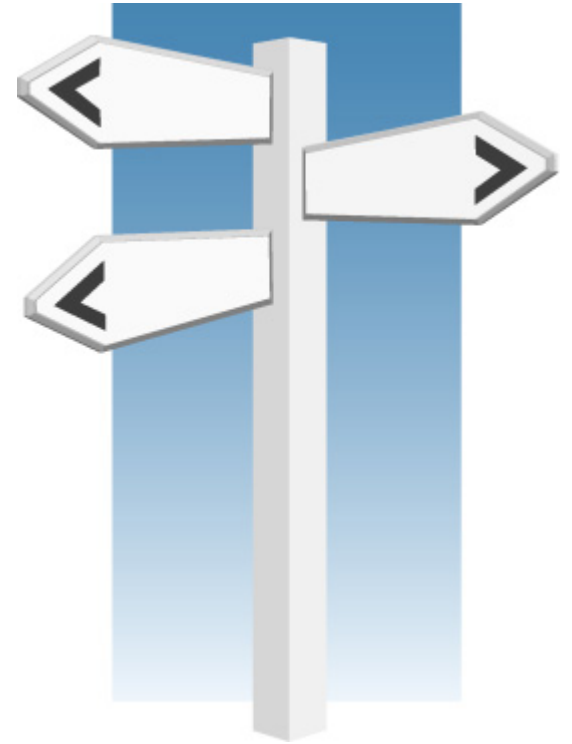
- The RIDW stores historical data and facilitates decision support system applications such as trend reports

Aggregation of information:

- Useful for strategic planning and reporting

What's next?

- Module 3: IBM Rational Insight and IBM Cognos Data Manager
 - Understanding ETL and Data Manager concepts
 - Review data warehouse structure
 - Rational Insight ETL catalog structure
 - Fact builds
 - Job streams



Extract, transform, and load (ETL) is a data warehousing process that includes:

- *Extracting* data from external sources
- *Transforming* it to fit business needs
- Ultimately, *loading* it into the data warehouse

The definition of ETL is stored in the ETL catalog.

The ETL catalog is a collection of tables, which are located in the Rational Insight ETL (RIETL) schema.

Rational Insight comes with a prebuilt ETL catalog.

Step 1: Extraction

- The ETL framework extracts data through both direct database access and ODBC drivers, including the Rational Insight XML ODBC driver.

Step 2: Transformation

- Transform the data based on business rules.
- Normalize the data and modify based on business rules. These business rules, when run, perform business logic such as calculating or deriving a column that is based on other columns.
- Store data in physical tables that implement the star schema convention.
- Add dimensions and facts – the characteristic elements of a data warehouse.

Step 3: Load

- Load transformed data into the metrics tables.

ETL catalog structure (continued)

In Cognos Data Manager, each piece of the ETL process is a fact or dimension build.

The builds can be organized into job streams for different sets of data or tables. The job streams can be:

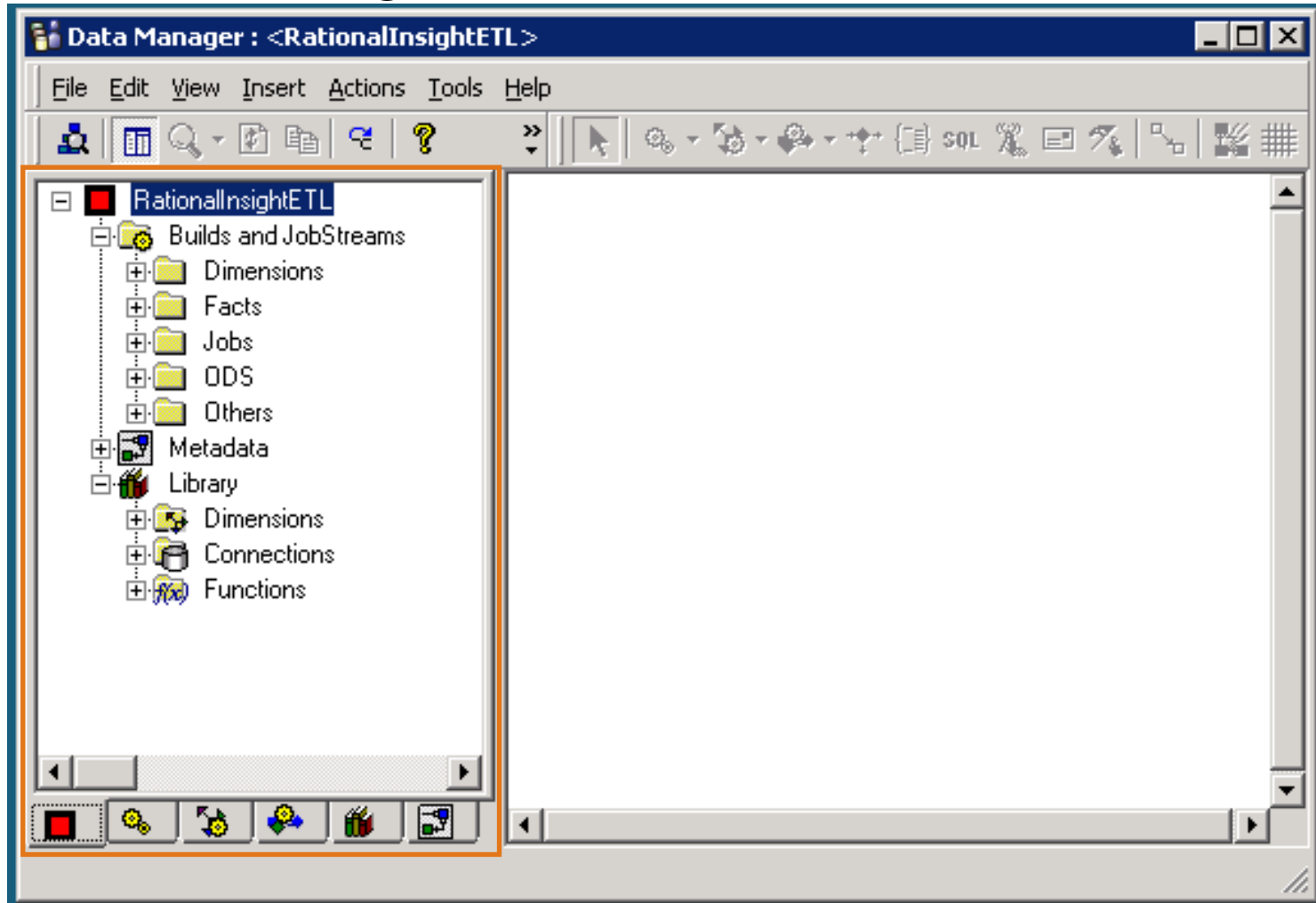
- Directly run in Data Manager.
- Published as a data move task and then scheduled for later execution in the IBM Rational Insight Report Server.

Insight implements and supports sample ETL transactions for products such as:

- Rational DOORS
- Rational Team Concert
- Rational Quality Manager
- Rational Test Manager
- Microsoft Project

ETL catalog structure (continued)

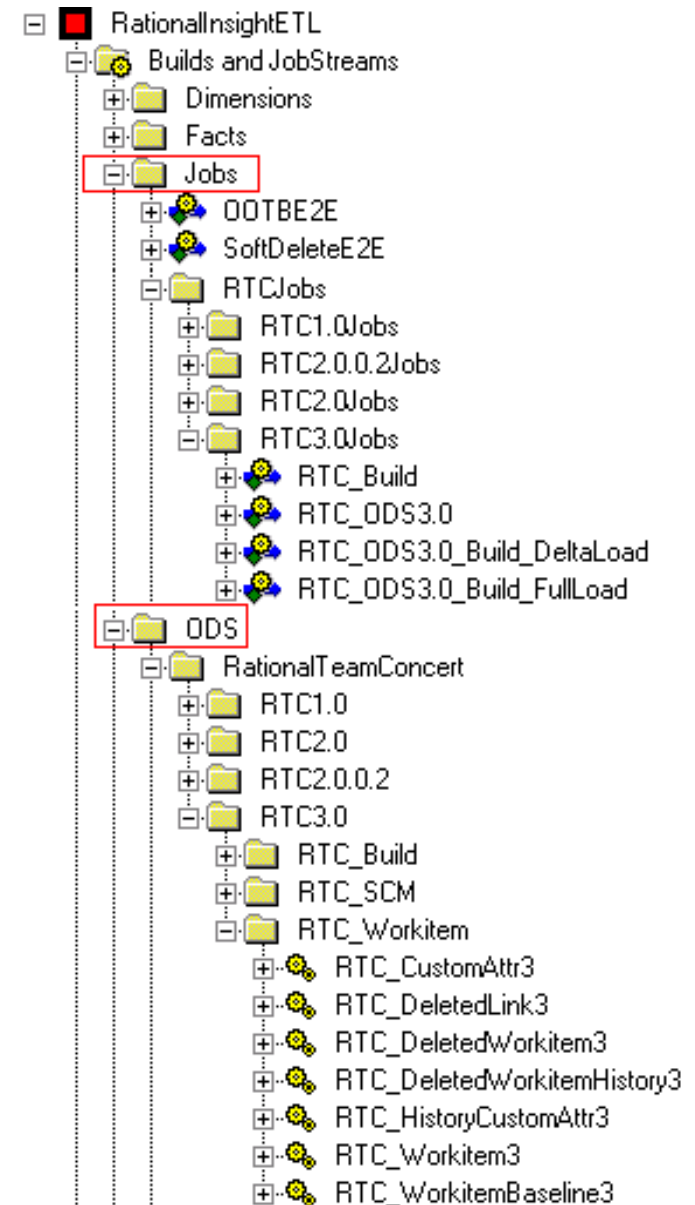
Cognos Data Manager interface showing Rational Insight ETL catalog:



ETL catalog structure – Builds and JobStreams folder

Jobs folder: Contains jobs that consist of a sequence of fact builds. Jobs define the order in which builds are run.

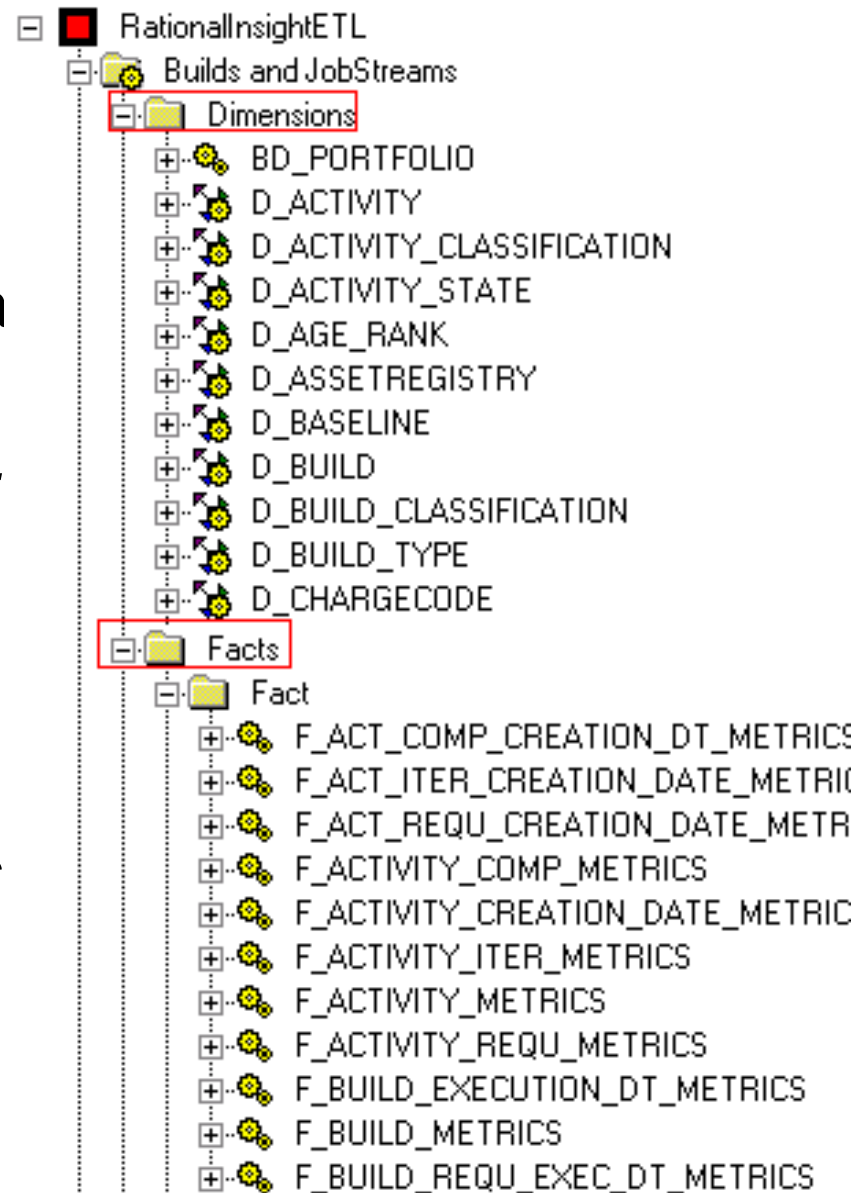
ODS folder: Contains fact build definitions for transferring the data from data sources to the Rational Insight data warehouse.



ETL catalog structure – Builds and JobStreams folder

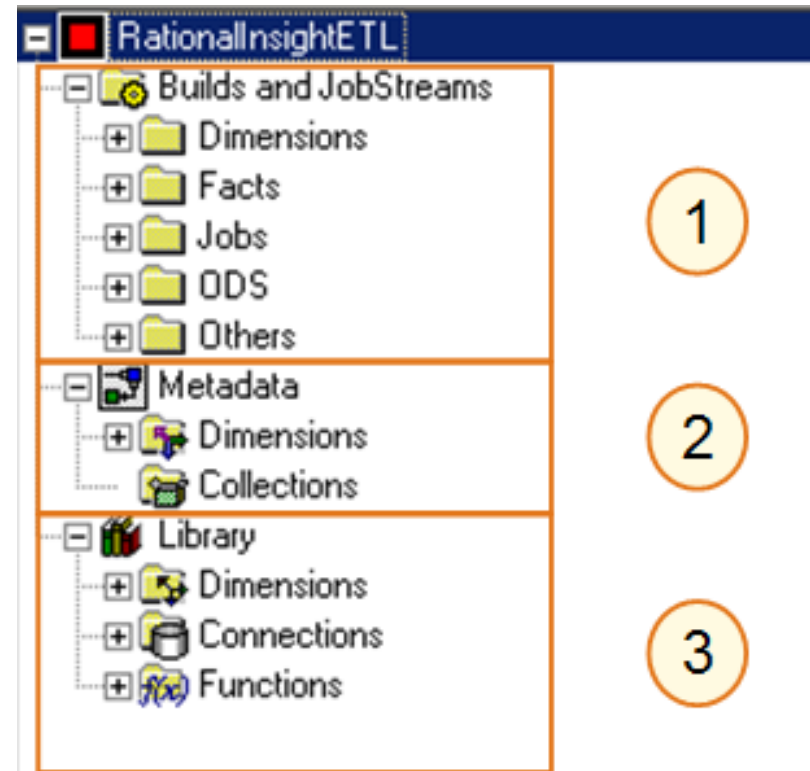
Dimensions folder: Contains dimension builds that extract data from a data source, transform it, and load it into a *dimension table*.

Facts folder: Contains facts and historical builds that extract data from a data source, transform it, aggregate it, and load it into a *fact table*.



ETL catalog structure – Metadata folder

In Rational Insight, the metadata folder is *not* used.

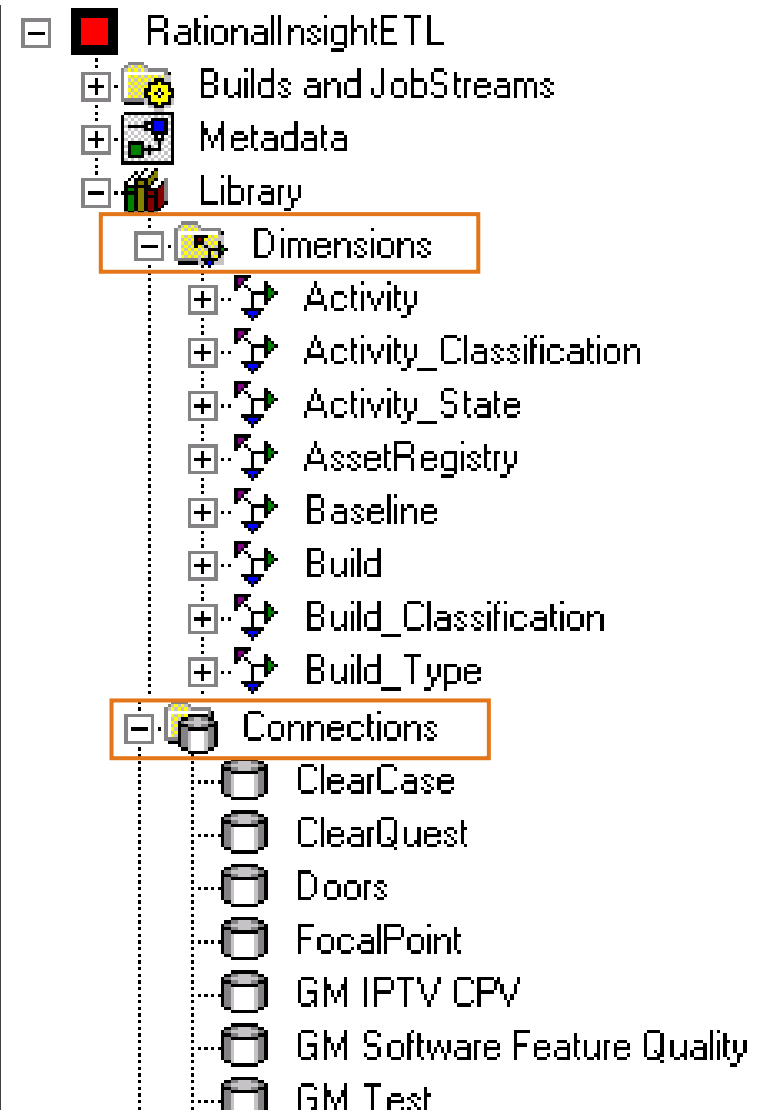


Dimensions folder:

Contains reference dimensions.

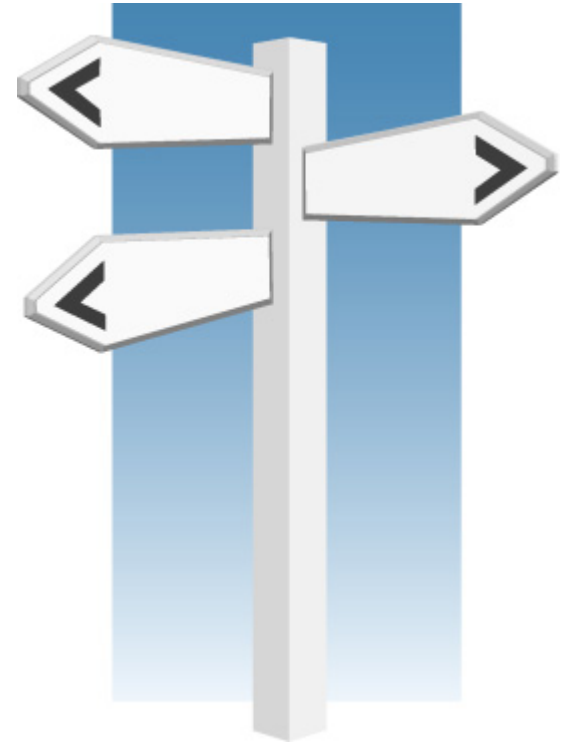
Connections folder:

Contains connection definitions.



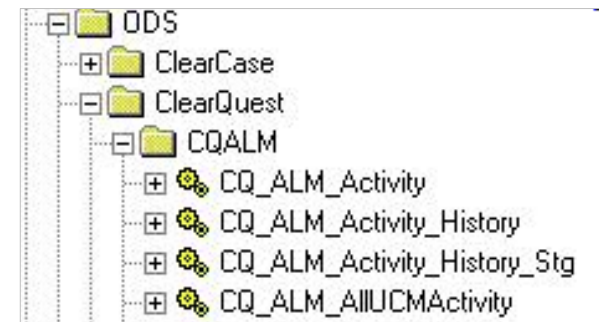
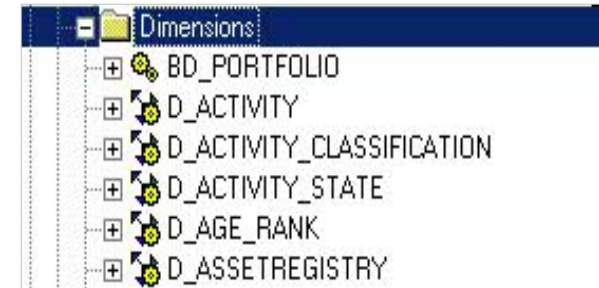
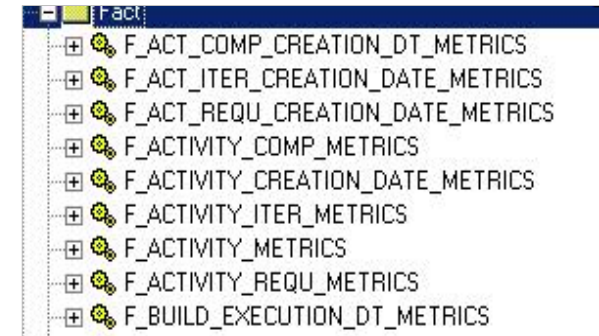
What's next?

- Module 3: IBM Rational Insight and IBM Cognos Data Manager
 - Understanding ETL and Data Manager concepts
 - Review data warehouse structure
 - Rational Insight ETL catalog structure
 - Fact builds
 - Job streams



Data Manager concepts: Fact and dimension builds

- Data manager has only two types of tables:
 - Fact
 - Dimension
- *Fact builds* are used to extract data from a data source, transform it, aggregate it and load it into a fact table.
- *Dimension builds* are used to extract data from a data source, transform it, and load it into a dimension table.
- Question: How do you deliver data to a relational table such as the tables in the ODS?
 - Answer: *Use the fact build*. In this case, there is no data aggregation; the fact build is only used as a data transfer medium.





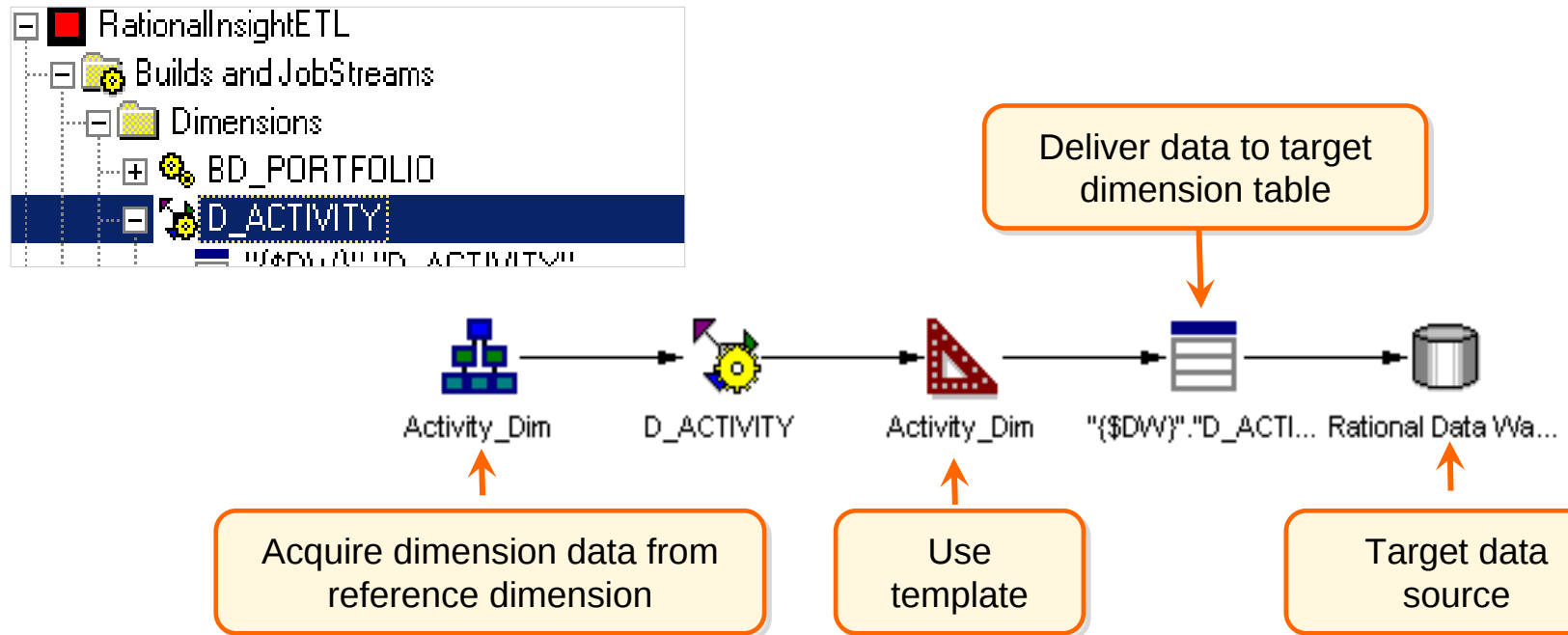
Fact and dimension builds

- A fact build acquires transactional data from the data source, transforms the data with reference to the dimensional framework, and delivers the data to target metrics tables.
- A dimension build delivers data that describes a single business dimension, such as Product or Customer.
- Each dimension build:
 - Acquires dimension data from the reference structure you specify.
 - Delivers the dimension data to the target metrics tables. The dimension build uses a template to define the behavior of each column.

Rational Insight example dimension build

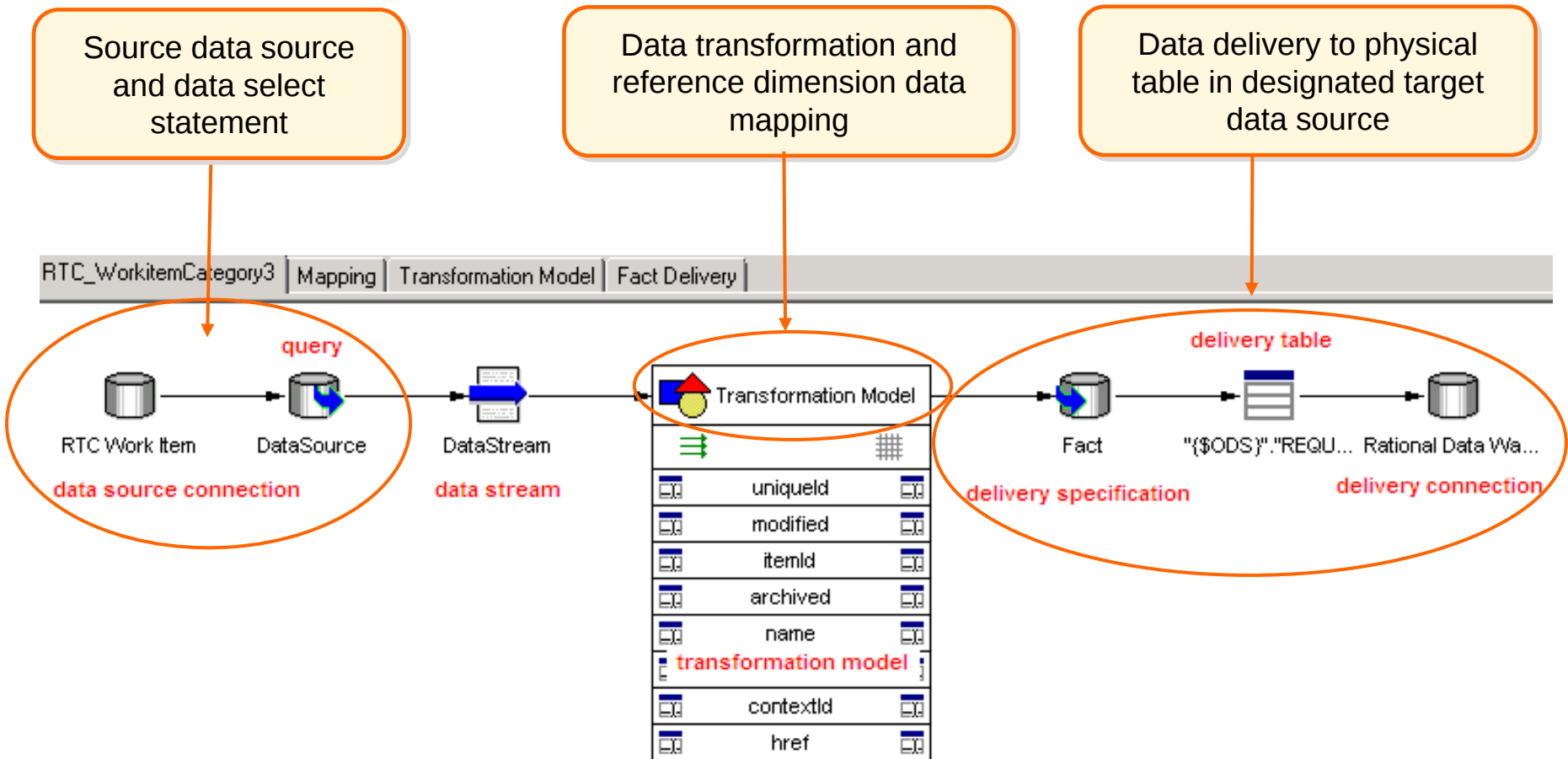
Dimension builds

- Used to deliver data into a physical table that represents a dimension in a data mart
- Uses a reference dimension in the process to extract the data from a data source
- Delivers data into the target dimension table using a template that defines the behavior of the columns



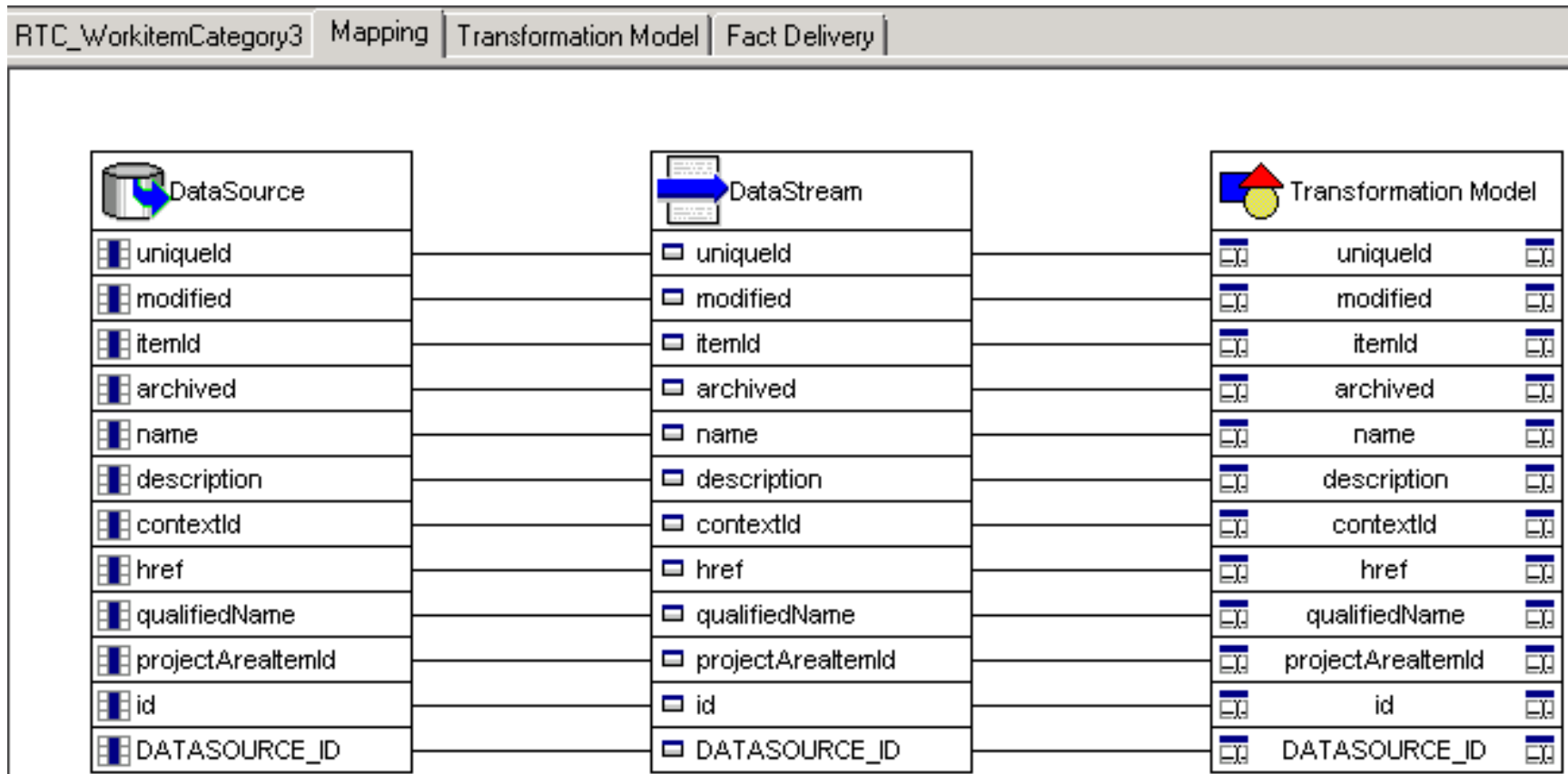
Fact build – Simple

Fact build: Data Manager mechanism for delivering data from a data source into a table.



Fact build – Simple (continued)

- The Mapping tab shows the mapping between DataSource columns, DataStream items, and the Transformation Model.



Fact build – Simple (continued)

Transformation Model tab

RTC_WorkitemCategory3 | Mapping | Transformation Model | Fact

Transformation Model

uniqueId
modified
itemId
archived
name
description
contextId
href
qualifiedName
projectAreaItemId
id
DATASOURCE_ID

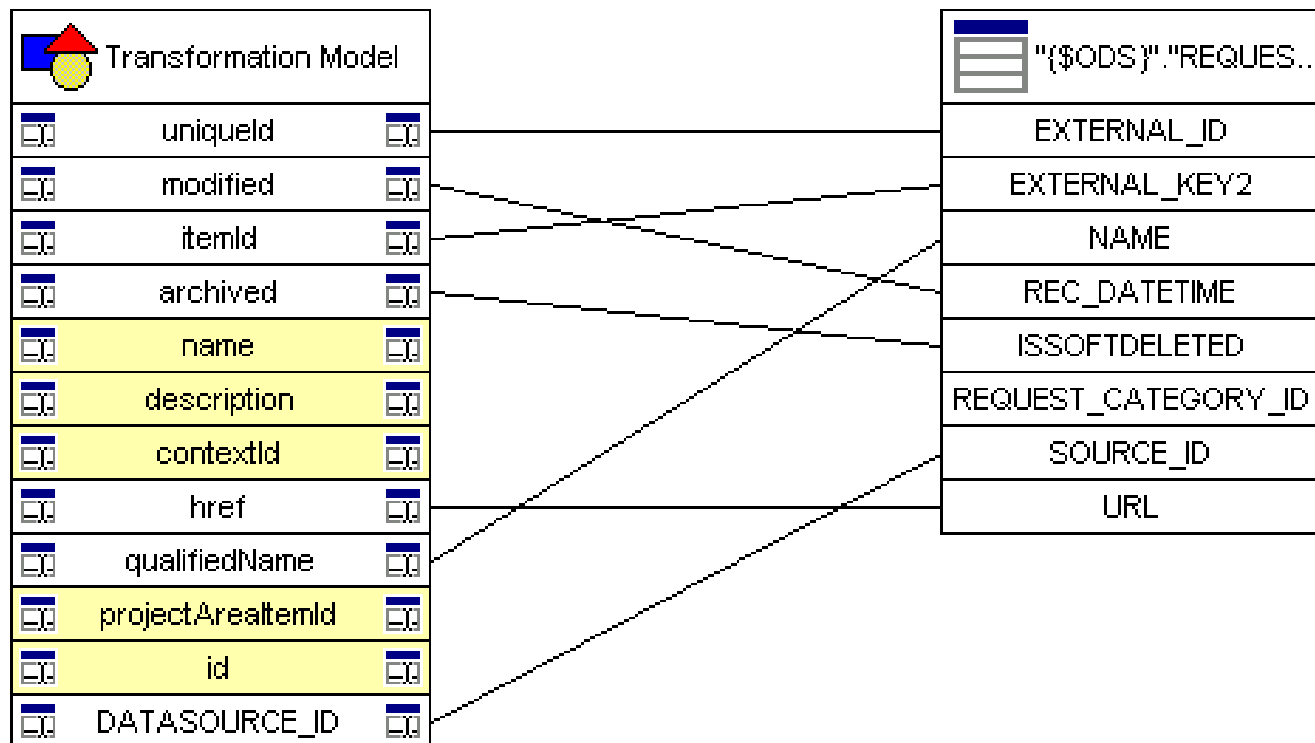
RTC_WorkitemCategory3 - Transformation Model Mapping

DataStream Item	Maps To	Transformation Model
uniqueId	uniqueId	uniqueId
modified	modified	modified
itemId	itemId	itemId
archived	archived	archived
name	name	name
description	description	description
contextId	contextId	contextId
href	href	href
qualifiedName	qualifiedName	qualifiedName
projectAreaItemId	projectAreaItemId	projectAreaItemId
id	id	id
DATASOURCE_ID	DATASOURCE_ID	DATASOURCE_ID

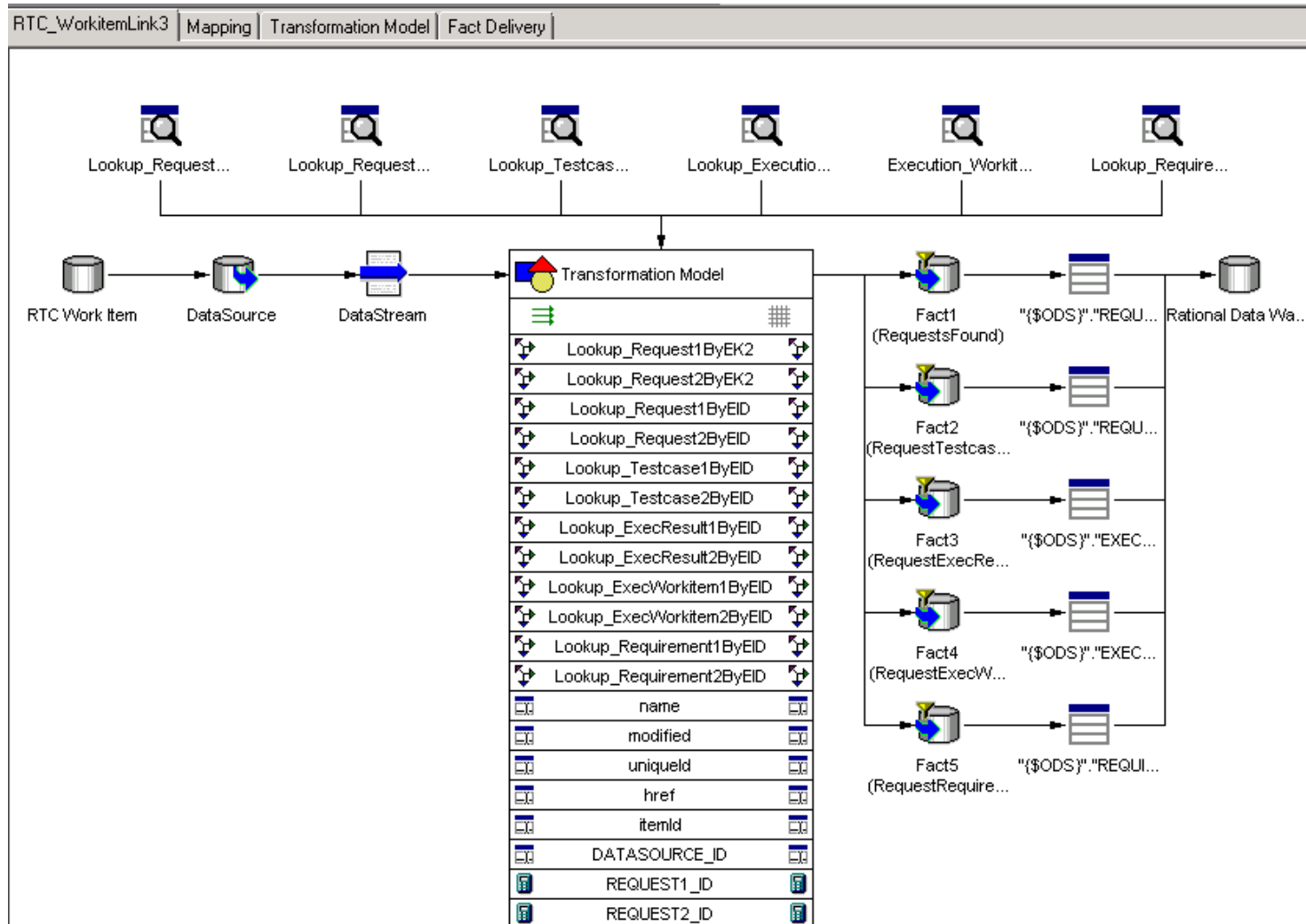
Fact build – simple (continued)

Fact Delivery tab

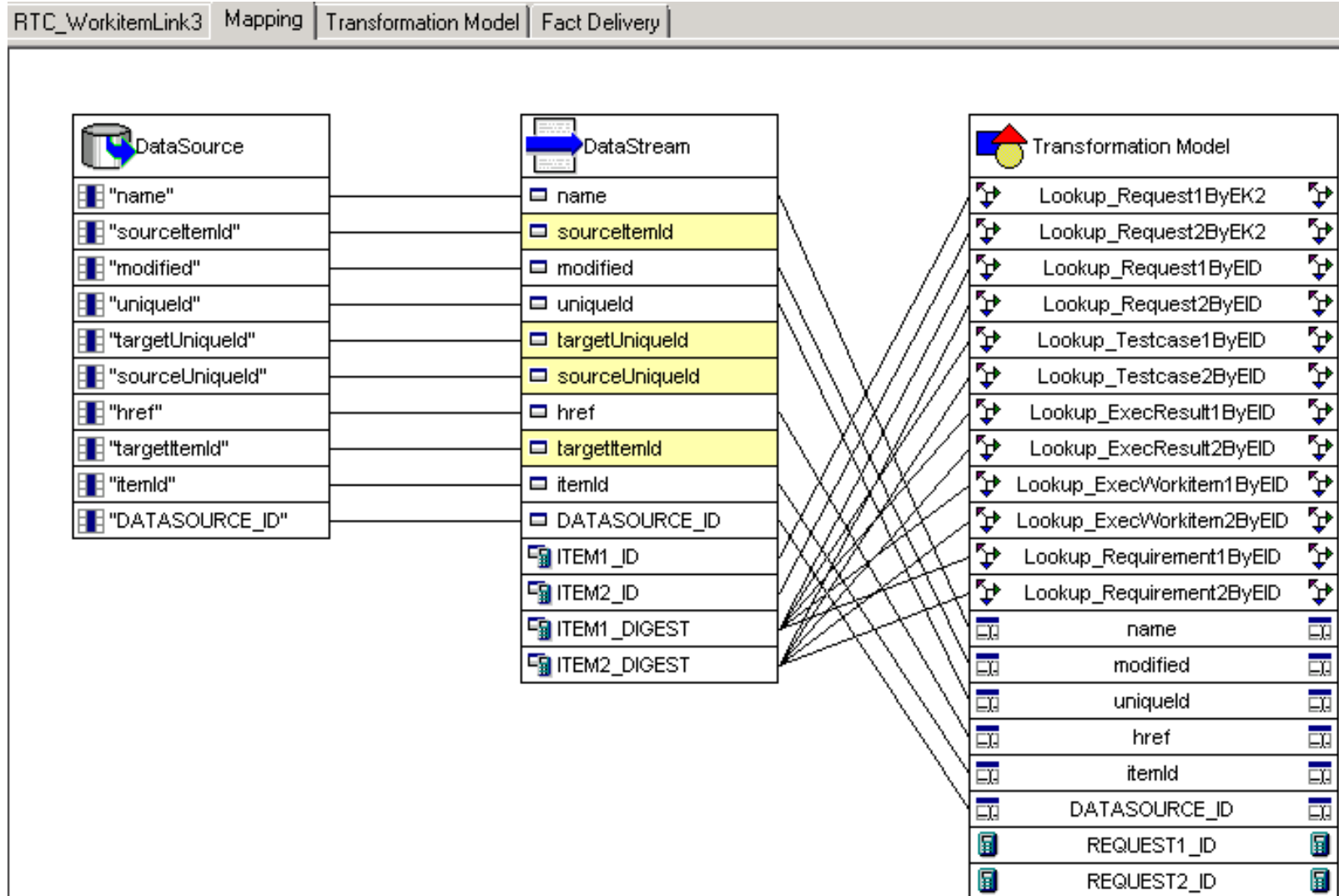
RTC_WorkitemCategory3	Mapping	Transformation Model	Fact Delivery
-----------------------	---------	----------------------	---------------



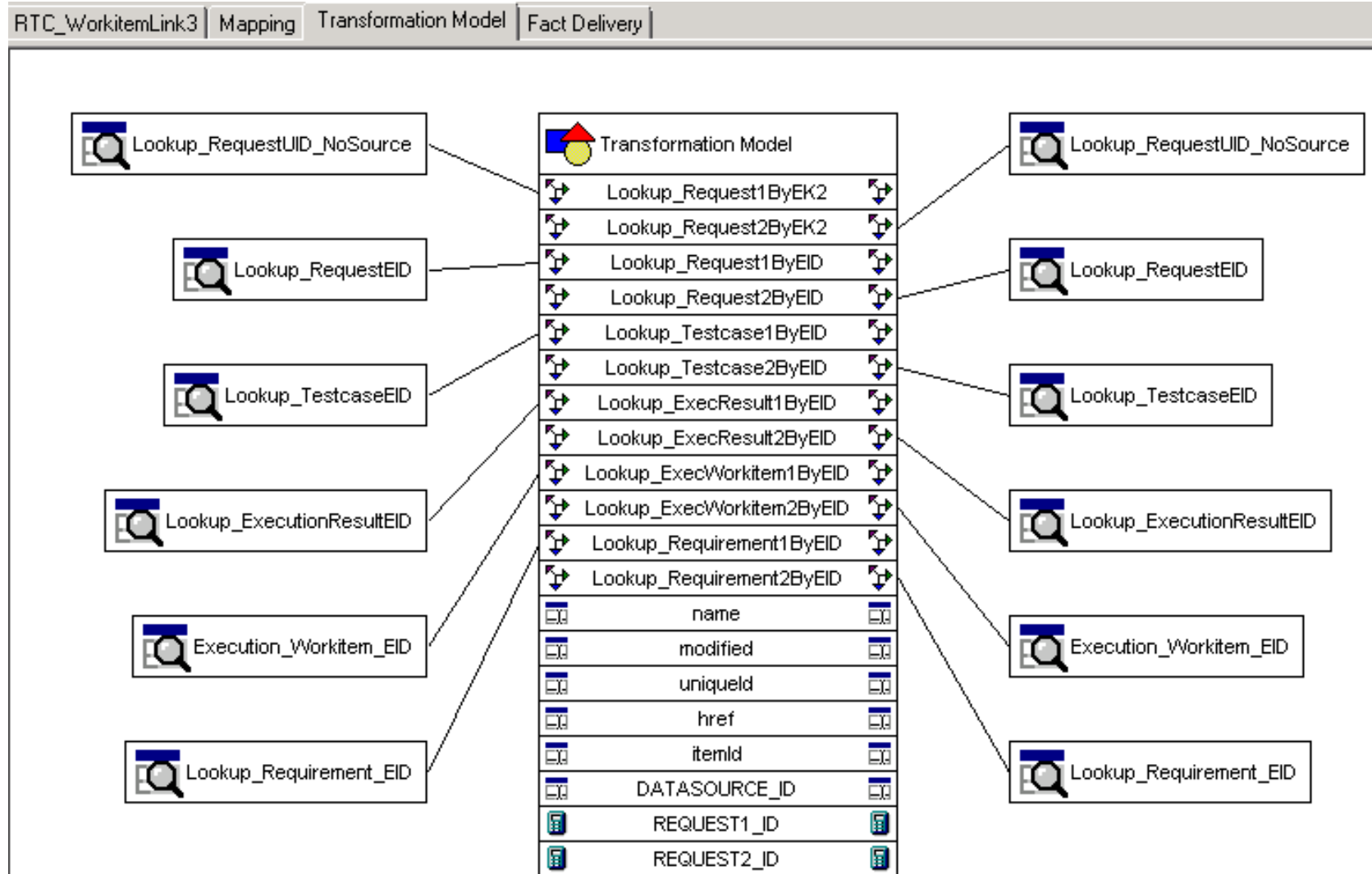
Example of a complex fact build



Fact build – Complex (continued)

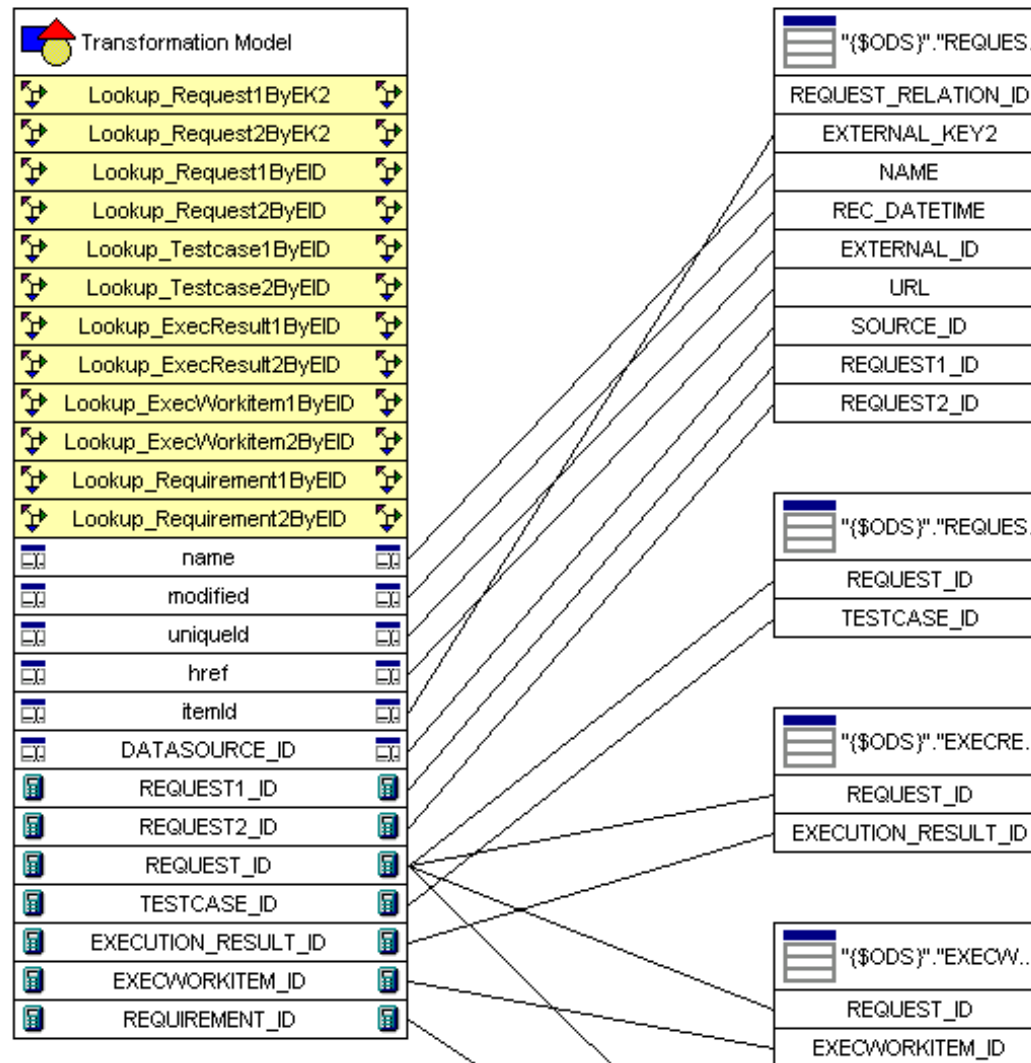


Fact build – Complex (continued)



Fact build – Complex (continued)

RTC_WorkitemLink3 | Mapping | Transformation Model | Fact Delivery



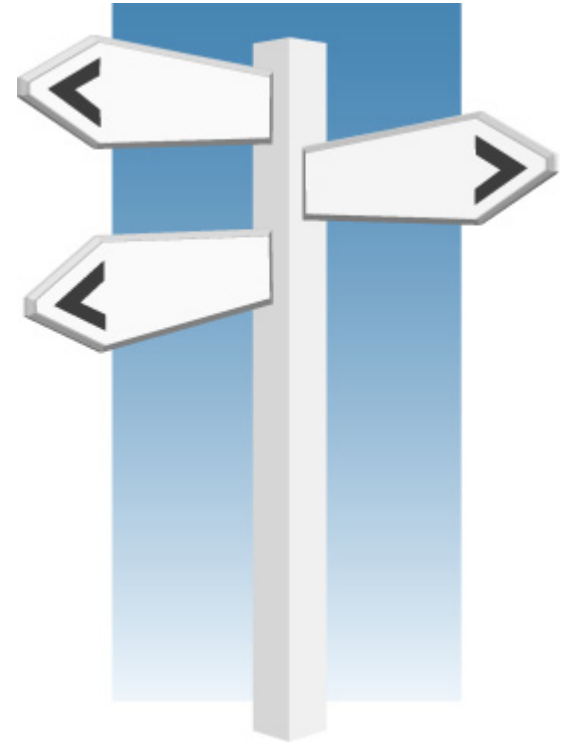


General steps to create a fact build

1. Right-click a folder and select “Insert Fact Build” from the menu.
2. Right-click the workspace and select “Insert Data Source” from the menu.
3. Define the connection and query.
4. Refresh the results columns.
5. Double-click the DataStream icon and define data stream items.
6. Right-click “Transformation Model.” Select “Mapping” from the menu and define transformation model mapping.
7. Right-click the workspace and select “Insert Relational Table Delivery.”
8. Define a connection for the delivery table and map fact items to table columns.

What's next?

- Module 3: IBM Rational Insight and IBM Cognos Data Manager
 - Understanding ETL and Data Manager concepts
 - Review data warehouse structure
 - Rational Insight ETL catalog structure
 - Fact builds
 - Job streams



Job streams

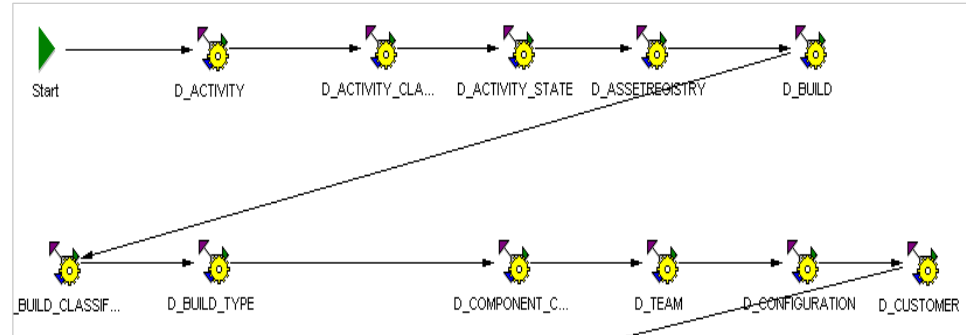
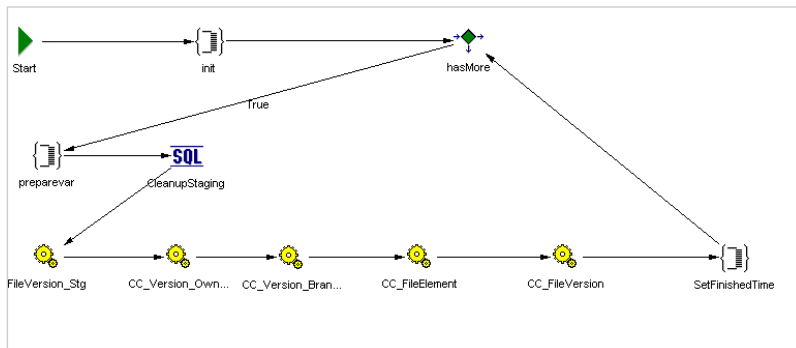
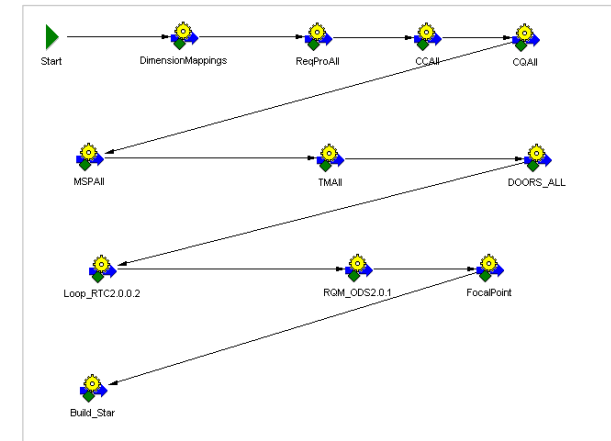
You can automate tasks that are related to managing the build process by creating a job stream to contain a series of steps.

A job stream contains a node for each step in a process:

- Fact build
- Dimension build
- SQL statement
- Condition
- Procedure
- Alert
- Email
- Other job stream

Each job stream can use any number of these nodes.

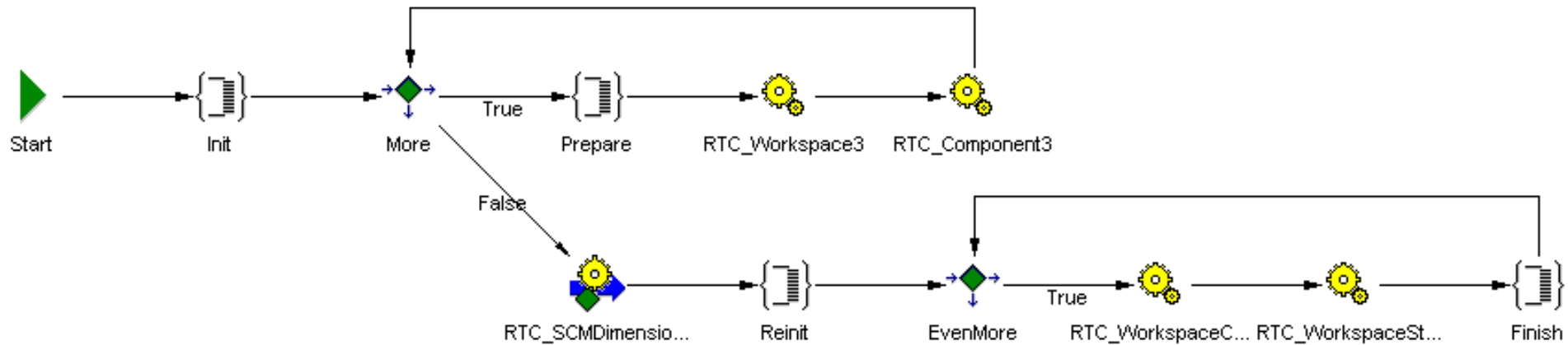
You can link nodes together to create a flow to instruct Data Manager in what order the nodes should be processed.



Screen capture showing a job stream with a series of fact build nodes:



Screen capture showing one fact build process containing procedures, conditions, and other job stream nodes:



Job streams can contain:

- **Fact build node**

- You can include any fact builds from the current catalog in a job stream.

- **Dimension build node**

- You can include any dimension builds from the current catalog in a job stream.

- **SQL node**

- An SQL node can contain a series of SQL statements that are processed when the SQL node is reached in the job stream flow.

- **Procedure node**

- A procedure node can contain one or more Data Manager functions or statements.

▪ **Condition node**

- A condition node provides branching between nodes for conditional execution. Each condition node can have many nodes linking to it, but only two output links, True and False.

▪ **Job stream node**

- A job stream node is a job stream within a job stream. This allows larger jobs to be broken down into a series of smaller jobs. When a job stream node is encountered, all the steps within the job stream node are processed before processing of the next node in the sequence starts.

▪ **Alert node**

- An alert node writes a user-defined audit record of the type ALERT into the Data Manager audit tables. You can use these records to record specific events that occur during job stream and build execution.

Job streams

- **Email node**

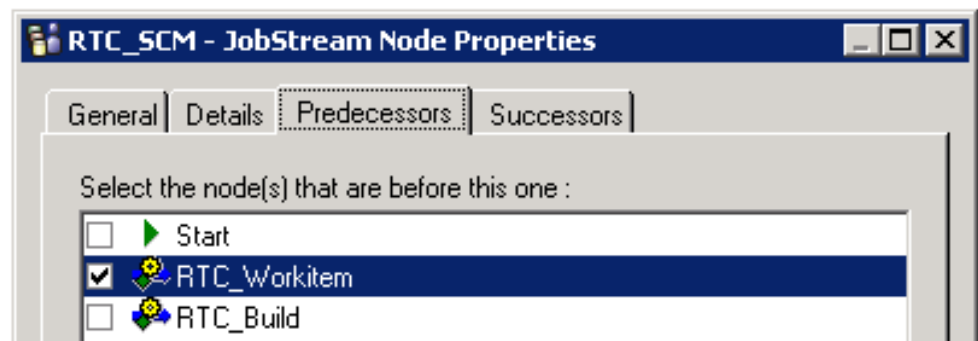
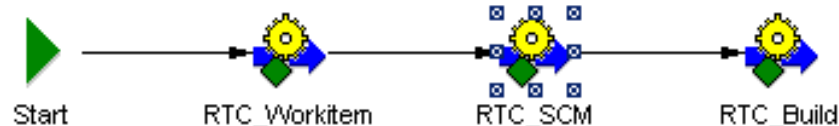
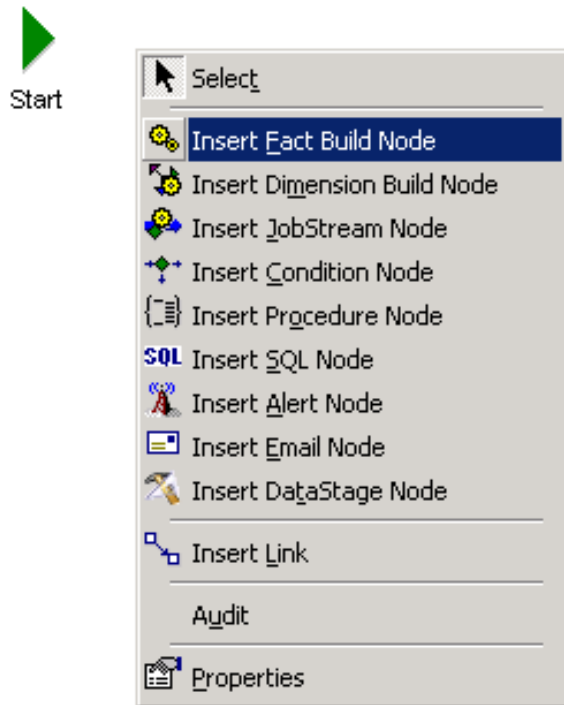
- Sends e-mail notifications.
- Useful for notifying key recipients of problems.
- Mail server needs to be configured in Cognos configuration tool.

- **Link node**

- Connects two components in a job stream.
- Dictates the sequence of execution.
- Is a component from which the link branching out is the predecessor. The component that the link targets is the successor in the sequence of execution.

General steps to create a job stream

1. Right-click a folder and select **Insert Job stream**.
2. Right-click the work area and choose the item to add.
3. Specify properties of the item, predecessors, successors, and so on.



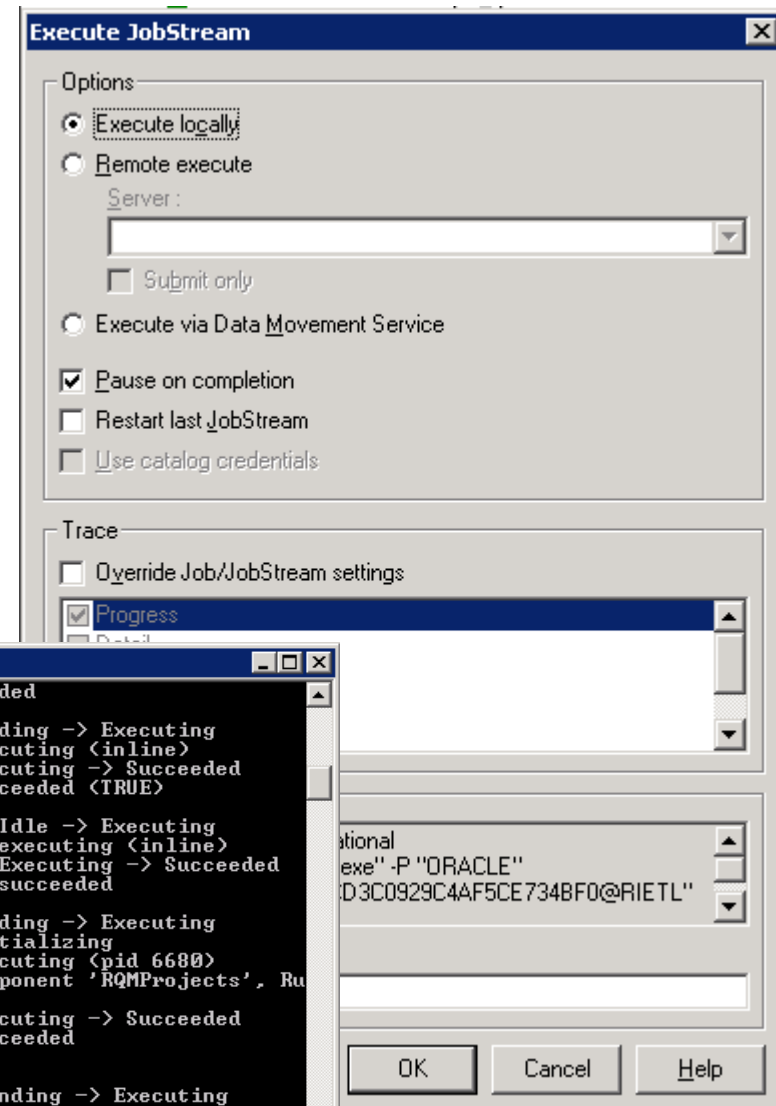
Job streams can be run in the following two ways:

- From the Cognos Data Manager interface
 - Manual run
 - User needs to be logged in

- From IBM Cognos Connections
 - Job stream needs to be published
 - Can be either manually started or automatically scheduled

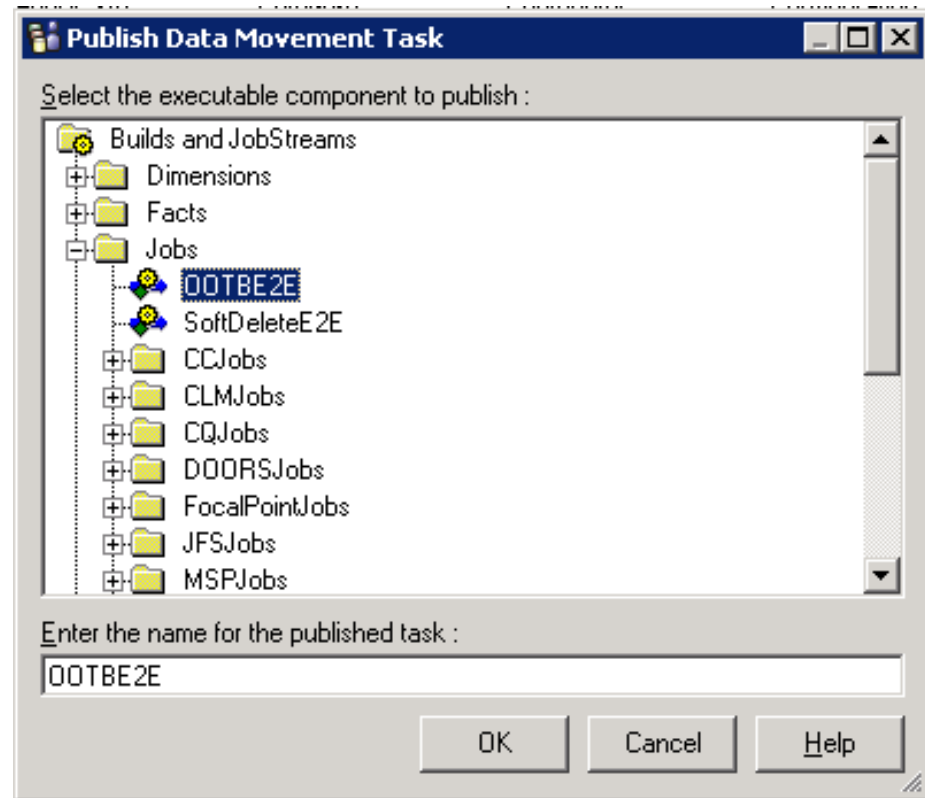
Running job streams from Data Manager

1. Right-click the job stream and click **Execute**.
2. Click **OK**.
3. A command window opens for the entire duration of the run, displaying ETL progress.



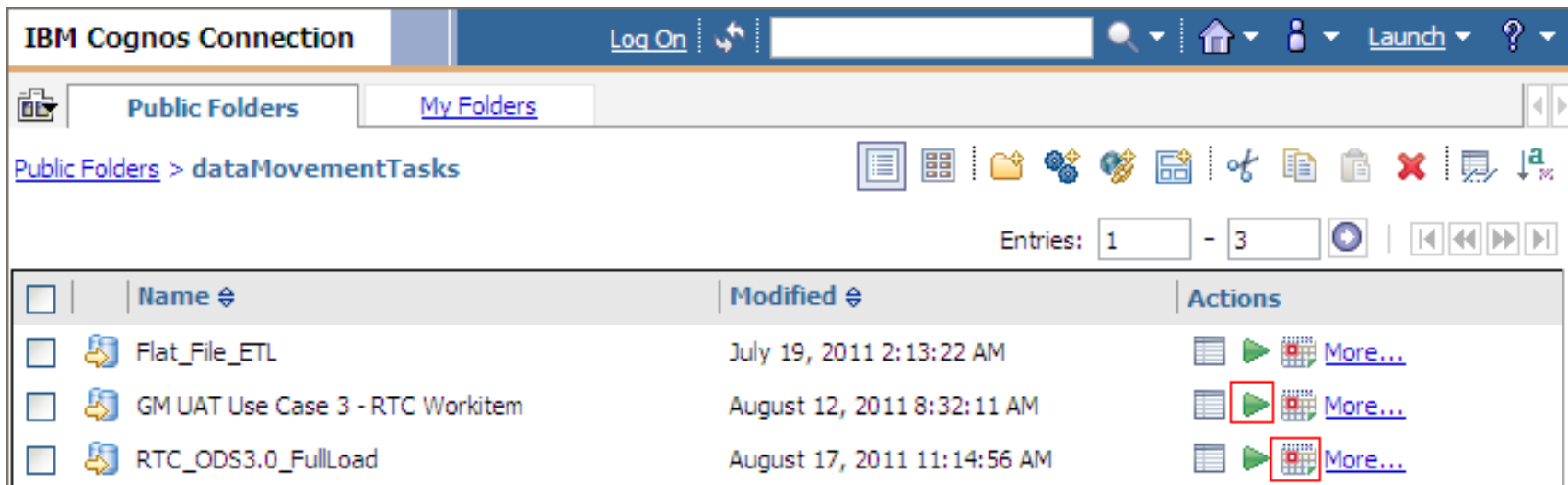
Job stream – publishing a job

1. Click **Actions** → **Publish Data Movement Task**.
2. Navigate to the job stream.
3. Click **OK**. The job is now published in the dataMovementTasks folder in the Public Folder in IBM Cognos Connections.



Running job streams from IBM Cognos Connections

- Click the green play button to run manually.
- Click the schedule button to set up a scheduled run.



The screenshot displays the IBM Cognos Connections web interface. At the top, there is a navigation bar with 'IBM Cognos Connection' and a 'Log On' button. Below this, a breadcrumb trail shows 'Public Folders > dataMovementTasks'. A table lists three tasks, each with a checkbox, a folder icon, a name, a 'Modified' date, and an 'Actions' column. The 'Actions' column contains a list icon, a green play button, a calendar icon, and a 'More...' link. Red boxes highlight the green play button and the calendar icon for the 'GM UAT Use Case 3 - RTC Workitem' and 'RTC_ODS3.0_FullLoad' tasks.

☐	Name	Modified	Actions
☐	Flat_File_ETL	July 19, 2011 2:13:22 AM	More...
☐	GM UAT Use Case 3 - RTC Workitem	August 12, 2011 8:32:11 AM	More...
☐	RTC_ODS3.0_FullLoad	August 17, 2011 11:14:56 AM	More...

Job stream – creating a schedule

Example of using the Schedule dialog box:

Schedule - GM UAT Use Case 3 - RTC Workitem

Schedule the entry to run at a recurring date and time. You can disable the schedule without losing any of its details.

☐ Disable the schedule

Priority:
3

Start:
Aug 22, 2011
5 : 00 PM

Frequency:
Select the frequency by clicking on a link.

☒ **By Day** [By Week](#) [By Month](#) [By Year](#) [By Trigger](#)

☐ Every 1 minute(s)

☐ Every 1 hour(s)

☒ Every 1 day(s)

End:
☒ No end date
☐ End by:
Aug 22, 2011
5 : 45 PM

OK Cancel

For more information

Rational Insight information center:

- <https://jazz.net/help-dev/rational-insight/index.jsp>
 - See “Reference” section for data models and report details.

Rational Insight Data Modeler training path:

- <https://www-304.ibm.com/jct03001c/services/learning/ites.wss/us/en?pageType=page&c=L806673M20298F27>
(or do a web search for “Rational Insight Data Modeler training path”)

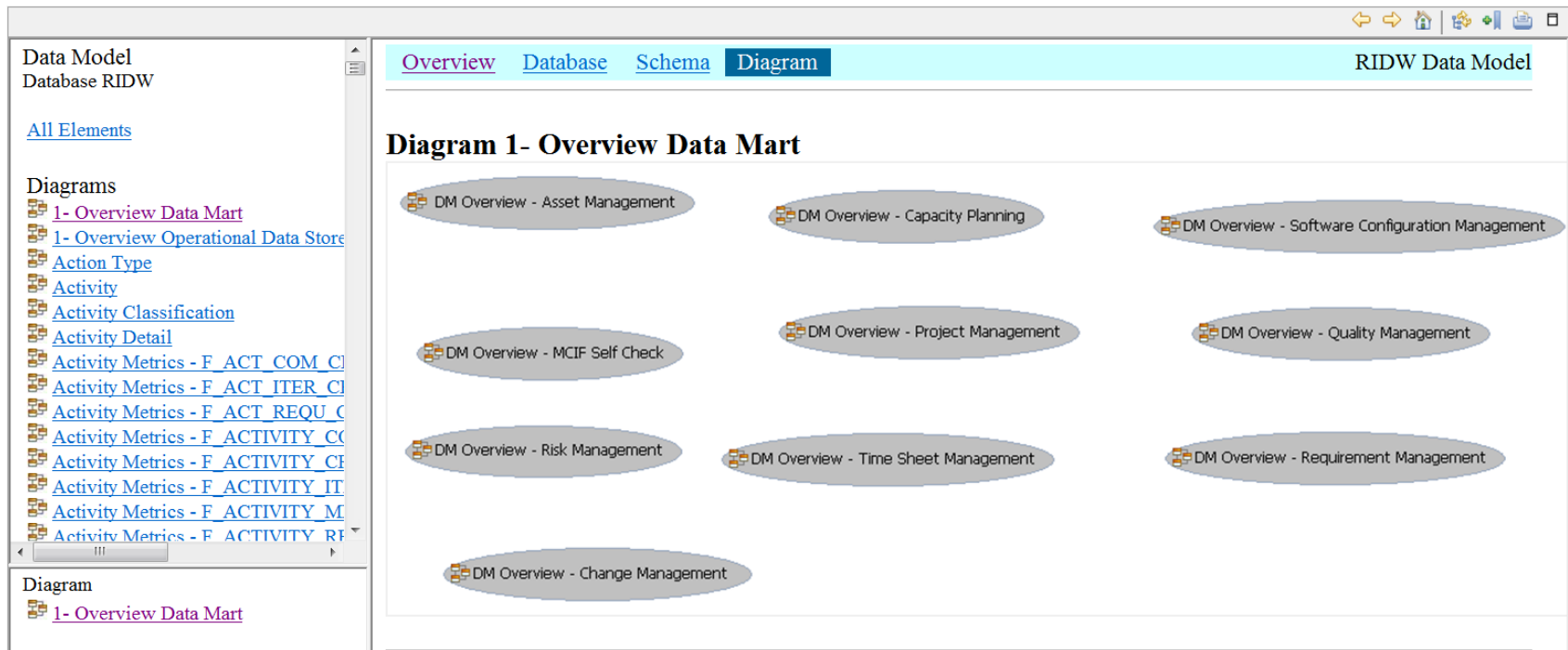
Cognos Data Manager 10.1 User Guide and Data Manager 10.1 Getting Started (see under Model and Design section)

- <http://publib.boulder.ibm.com/infocenter/cbi/v10r1m0/index.jsp>

For more information: Rational Insight data models

See the Rational Insight data models at

https://jazz.net/help-dev/clm/topic/com.ibm.rational.reporting.overview.doc/topics/c_reference_datadictionary.html



Thank
You

Legal notices

- **The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law:** INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions; therefore, this statement may not apply to you.
- This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.
- If you are viewing this information in softcopy, the photographs and color illustrations may not appear.
- Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.
- Any performance data contained herein was determined in a controlled environment. Therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.
- Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Legal notices

- This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

Trademarks and service marks

- IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at www.ibm.com/legal/copytrade.shtml.

- Other company, product, or service names may be trademarks or service marks of others.