

IICS: Cloud Data Integration Services

Student Guide

Version: IICS-R33-Cloud-DIS-202002



IICS: Cloud Data Integration Services

Version: IICS-R33-Cloud-DIS-202002

February 2020

Copyright (c) 1998–2020 Informatica LLC. All rights reserved.

This educational service, materials, documentation, and related software contain proprietary information of Informatica LLC and are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright law. Reverse engineering of the software is prohibited. No part of the materials and documentation may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC. The related software is protected by U.S. and/or international Patents and other Patents Pending.

Use, duplication or disclosure of the related software by the U.S. Government is subject to the restrictions set forth in the applicable software license agreement and as provided in DFARS 227.7202-1(a) and 227.7702-3(a) (1995), DFARS 252.227-7013(c)(1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19, or FAR 52.227-14 (ALT III), as applicable.

The information in this educational service, materials, and documentation is subject to change without notice. If you find any problems in this educational service, materials, or documentation, please report them to us in writing.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange, Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging, and Informatica Master Data Management are trademarks or registered trademarks of Informatica LLC in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

Portions of this educational service, materials, and/or documentation are subject to copyright held by third parties, including without limitation: Copyright © Adobe Systems Incorporated. All rights reserved. Copyright © Microsoft. All rights reserved. Copyright © Oracle. All rights reserved. Copyright @ the CentOS Project.

This Software is protected by U.S. Patent Numbers 5,794,246; 6,014,670; 6,016,501; 6,029,178; 6,032,158; 6,035,307; 6,044,374; 6,092,086; 6,208,990; 6,339,775; 6,640,226; 6,789,096; 6,820,077; 6,823,373; 6,850,947; 6,895,471; 7,117,215; 7,162,643; 7,243,110; 7,254,590; 7,281,001; 7,421,458; 7,496,588; 7,523,121; 7,584,422; 7,720,842; 7,721,270; and 7,774,791, international Patents and other Patents Pending.

DISCLAIMER: Informatica LLC provides this educational services, materials, and documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of non-infringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this educational service, materials, documentation, or related software is error free. The information provided in this educational service, materials, documentation, and related software may include technical inaccuracies or typographical errors. The information in this educational service, materials, documentation and related software is subject to change at any time without notice.

Document Conventions

This guide uses the following formatting conventions:

If you see...	It means...	Example
>	Indicates a sub menu to navigate to.	Click Repository > Connect. In this example, you should click the Repository menu or button and choose Connect.
boldfaced text	Indicates text you need to type or enter.	Click the Rename button and name the new source definition S_EMPLOYEE .
UPPERCASE	Database tables and column names are shown in all UPPERCASE.	T_ITEM_SUMMARY
<i>italicized text</i>	Indicates a variable you must replace with specific information.	Connect to the Repository using the assigned <i>login_id</i> .
Note:	The following paragraph provides additional facts.	Note: You can select multiple objects to import by using the Ctrl key.
Tip:	The following paragraph provides suggested uses or a Velocity best practice.	Tip: The m_ prefix for a mapping name is...

Other Informatica Resources

In addition to the student and lab guides, Informatica provides these other resources:

- Documentation and Knowledge Base
- Global Customer Support
- Professional Certification

Accessing Documentation and Knowledge Base

To get the latest documentation and Knowledge Base for your product, go to

<https://network.informatica.com>

Contacting Global Customer Support

You can contact a Customer Support Center by telephone or through the Online Support. Online Support requires a username and password. You can request a username and password at

<https://www.informatica.com/services-and-training/support-services/contact-us.html>

Obtaining Informatica Professional Certification

You can take and pass exams provided by Informatica to obtain Informatica Professional Certification. For more information, go to

<https://www.informatica.com/services-and-training/certification.html>

Table of Contents

- Module 1: Informatica Cloud Overview
- Module 2: Runtime Environments and Connections
- Module 3: Synchronization Task
- Module 4: Cloud Mapping Designer – Basic Transformations
- Module 5: Advanced Transformations and Mapping Tasks
- Module 6: Mapping Parameters
- Module 7: Expression Macro and Dynamic Linking
- Module 8: Replication Task
- Module 9: Masking Task
- Module 10: Mass Ingestion Task
- Module 11: Taskflows
- Module 12: Advanced Options
- Module 13: Hierarchical Connectivity
- Module 14: Intelligent Structure Model
- Module 15: IICS APIs
- Module 16: Exception Handling
- Module 17: Performance Tuning
- Module 18: Automating and Monitoring Tasks
- Module 19: Administration
- Module 20: SAML Setup
- Module 21: Discovery IQ



IICS: Cloud Data Integration Services

 Informatica[®]
University

Module 1

Informatica Cloud Overview



Module Objectives

After completing this module, you will be able to:

- Describe IICS as an iPaaS solution
- Define the key terminologies used in IICS
- Explore the IICS architecture
- List the Cloud Data Integration assets and components

3

© Informatica. Proprietary and Confidential.



Topic

Introduction to IICS



IICS as an iPaaS Solution

- IICS is a next-generation iPaaS solution that you can use to exchange data between applications or business partners
- Allows you to integrate, synchronize, and relate all data, applications, and processes that reside on-premise or in cloud
- Allows administrators, architects, and developers to easily process enterprise-ready data across Cloud, on-premise, big data, social, and mobile environments

5

© Informatica. Proprietary and Confidential.



IICS stands for Informatica Intelligent Cloud Services. It is a next-generation iPaaS solution that allows you to exchange data between applications or to exchange data externally with business partners.

Specifically, you can use IICS to integrate, synchronize, and relate all data, applications, and processes that reside on-premise or in your Cloud environment.

IICS is made up of several data management products that have a common user experience to accelerate productivity. You can access the IICS application via the Internet. Administrators, architects, and developers can use IICS to easily process enterprise-ready data across Cloud, on-premise, big data, social, and mobile environments.

Types of Business Processes

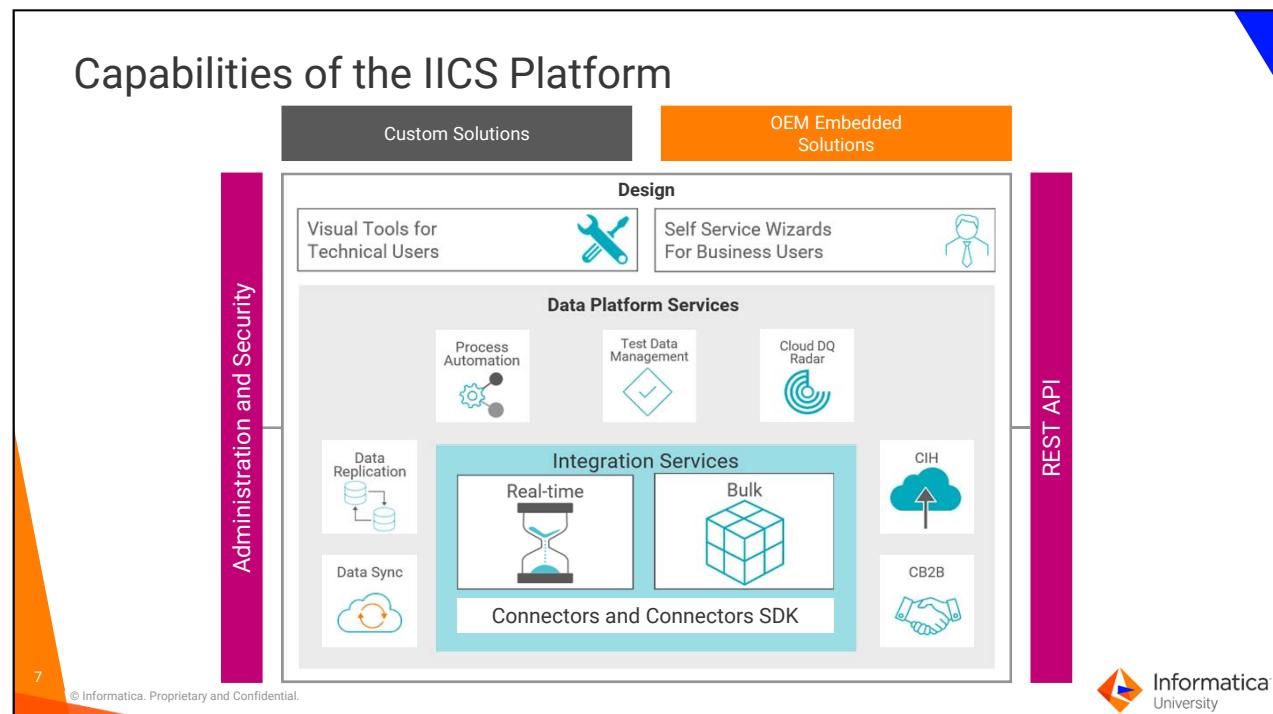
- You can use IICS to perform the following business processes:
 - Importing or migrating data
 - Object synchronization
 - Process integrations
 - Replication and archiving

6

© Informatica. Proprietary and Confidential.



- **Importing or Migrating Data** – Businesses often need to import data from an external system or migrate data from one contact management system to another. For example, you can import LEADS from a tradeshow or migrate data from an existing contact management system to Salesforce. Usually involves just a one-time push of data between the systems.
- **Object Synchronization** – Many organizations use IICS to synchronize copies of data in multiple systems. For example, if you have “Account” and “Contact” data in Salesforce, and the same information is in an Accounting or a Billing system, then you can use IICS to synchronize such information in each of these systems.
- **Process Integrations** – Process integrations help you connect one process with another and update both systems with appropriate information. For example, you have a Salesforce system where you track all sales-related activities such as opportunities. When an opportunity closes, you can update that information in your order management system.
- **Replication and Archiving** – This involves taking a backup of your data regularly or at scheduled intervals. You can also replicate the data to an on-premise database to run analytics.



The IICS platform provides real-time integration service and bulk integration service. It also provides connectors to connect to the data sources, and an SDK toolkit to create custom connectors. The core integration platform services include:

- Data Synchronization
- Data Replication
- Process automation
- Test data management
- Cloud DQ Radar
- Cloud Integration Hub (CIH)
- Cloud B2B

The platform holds all these services together with the help of visual tools for technical users and self-service wizards for business users. The platform also encompasses data governance and ensures that administration and security are seamless across the network. You can expose the dataset using REST API and utilize reusable codes from Informatica Market Place. The Informatica Marketplace is an open platform to host solutions that support all phases of the Data Integration lifecycle. Finally, the IICS platform allows you to use OEM Embedded Solutions.

Benefits of Using the IICS Platform

- Some of the benefits of using the IICS platform are:
 - Supports cloud, on-premise, and hybrid data management systems
 - Enterprise-class reliability and performance
 - End-to-end data management and governance
 - Supports both traditional data and big data platforms
 - Easy integration with on-premise and Cloud-based applications
 - Supports all major platforms

8

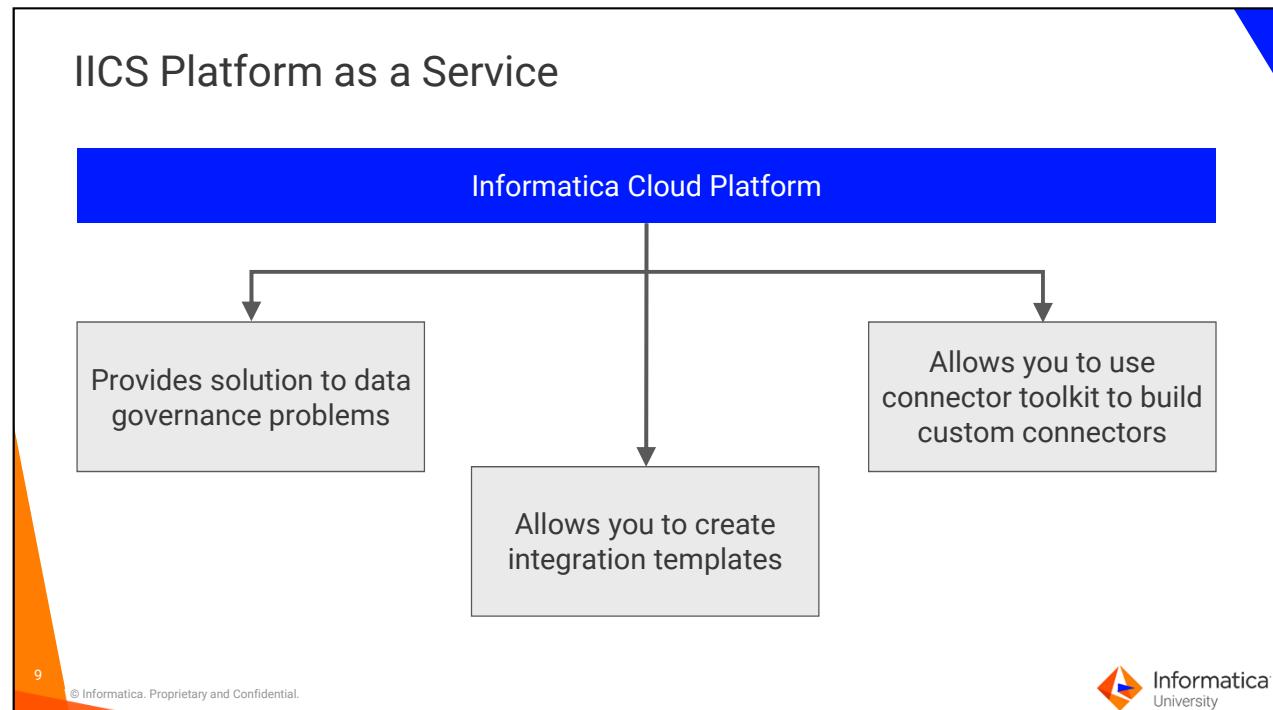
© Informatica. Proprietary and Confidential.



There are many benefits to using IICS platform that will give you a strategic advantage over your competitors.

IICS platform supports Cloud, on-premise, and hybrid data management systems. It also has an enterprise-class reliability and performance, with end-to-end data management and governance. It supports both traditional data and big data platforms. For example, IICS supports traditional data systems such as relational databases and data warehouses, as well as big data platforms such as Amazon Web Services and Hadoop.

You can easily integrate IICS with on-premise and Cloud-based applications. An important benefit of IICS is that it supports all major platforms such as Salesforce, Workday, Tableau, Microsoft Azure Blob, and so on.



IICS platform as a service provides solutions to data governance problems such as integrity, availability, and security of data. You can also use IICS platform to create templates and expose them to the business users using a hybrid approach. You can use the connector toolkit available with the platform to build custom connectors such as Workday, Eloqua, and Xactly.

So, with the help of these three services, you can aim for rapid development, configuration, and consumption of your SaaS application integration with the available on-premise data.

Building on the IICS Platform

- Hybrid Integration involves using the Cloud ICC model to utilize mapplets and PowerCenter services
- Re-use PowerCenter components
- Use a mapplet in Synchronization tasks, Mapping tasks, and Masking tasks
- Use integration templates to dynamically create PowerCenter-like mappings

10

© Informatica. Proprietary and Confidential.



Hybrid Integration involves using the Cloud Integration Competency Center (ICC) model to utilize mapplets and PowerCenter services. You can re-use PowerCenter components and leverage complex transformations from PowerCenter. You can also use a mapplet in Synchronization tasks, Mapping tasks, and Masking tasks.

As mentioned earlier, the IICS platform allows you to create integration templates. You can use these templates in the Cloud Mapping Designer to dynamically create PowerCenter-like mappings.

Topic

Key Terminologies



IICS Key Terminologies

Term	Definition
Source	Location from where you retrieve data
Target	Location to which you move the data
Task	Specifies the Data Integration job Task types: <ul style="list-style-type: none">• Synchronization Task• Replication Task• Mapping Task• Mass Ingestion Task• Masking Task
Connection	Provides the information that IICS needs to connect to an on-premise or a Cloud application
Org	An IICS organization. Generally one per company

12

© Informatica. Proprietary and Confidential.

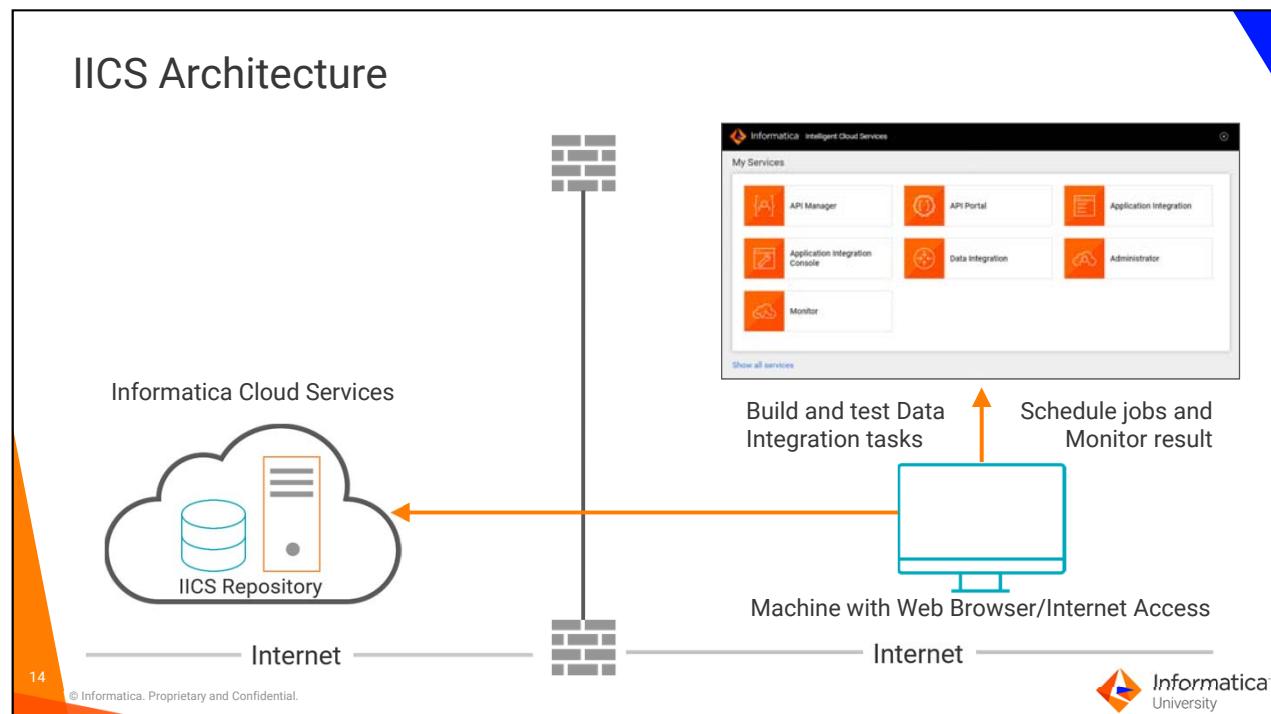


- **Source:** A source is the location from where you retrieve data. For example, if you load data from a CSV file into Salesforce, then your source is the CSV file.
- **Target:** A target is the location to which you move the data. If you load the data from a flat file into Salesforce, then your target is the object in Salesforce where you load the data.
- **Task:** A task specifies the Data Integration job, including the source and target, mappings, and any advanced options. Synchronization, Replication, Mapping, Mass Ingestion, and Masking are examples of the types of tasks.
- **Connection:** A connection provides the information that IICS needs to connect to your data sources and databases.
- **Org:** An Org is an IICS organization. Generally, you have one org per company where all users login and create objects. However, in some cases a company may have a production and a sandbox org. Large companies may have separate business units which require separate production orgs.

Topic

IICS Architecture

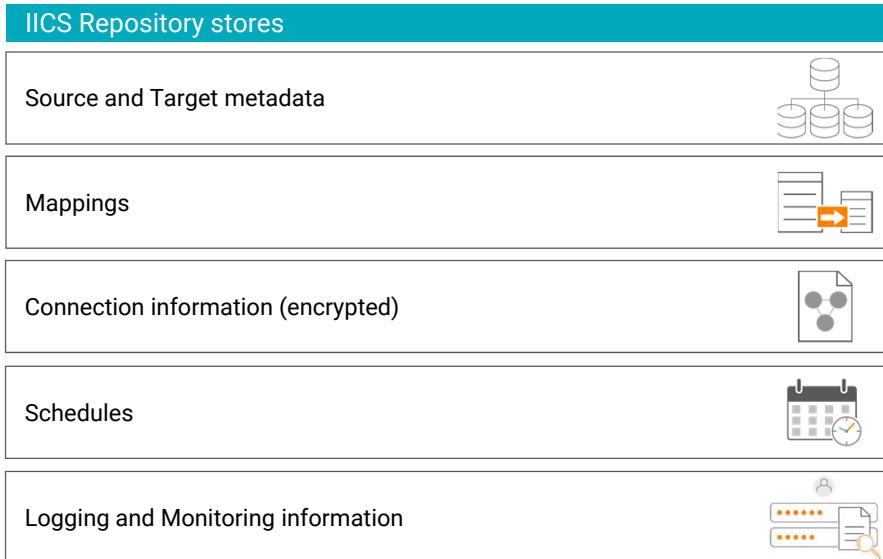




You have already seen that IICS is an iPaaS-based application. This means that you can access it from any machine with an Internet access and a web browser application. The application has multiple wizards that guide to build and test your Data Integration tasks. You can also use the application to schedule jobs and monitor the progress and result of the jobs.

When you access the application, your web browser connects to the Informatica Cloud Services through a secure HTTP. The Informatica Cloud Services includes the IICS repository that stores various information about the task.

IICS Repository



15

© Informatica. Proprietary and Confidential.



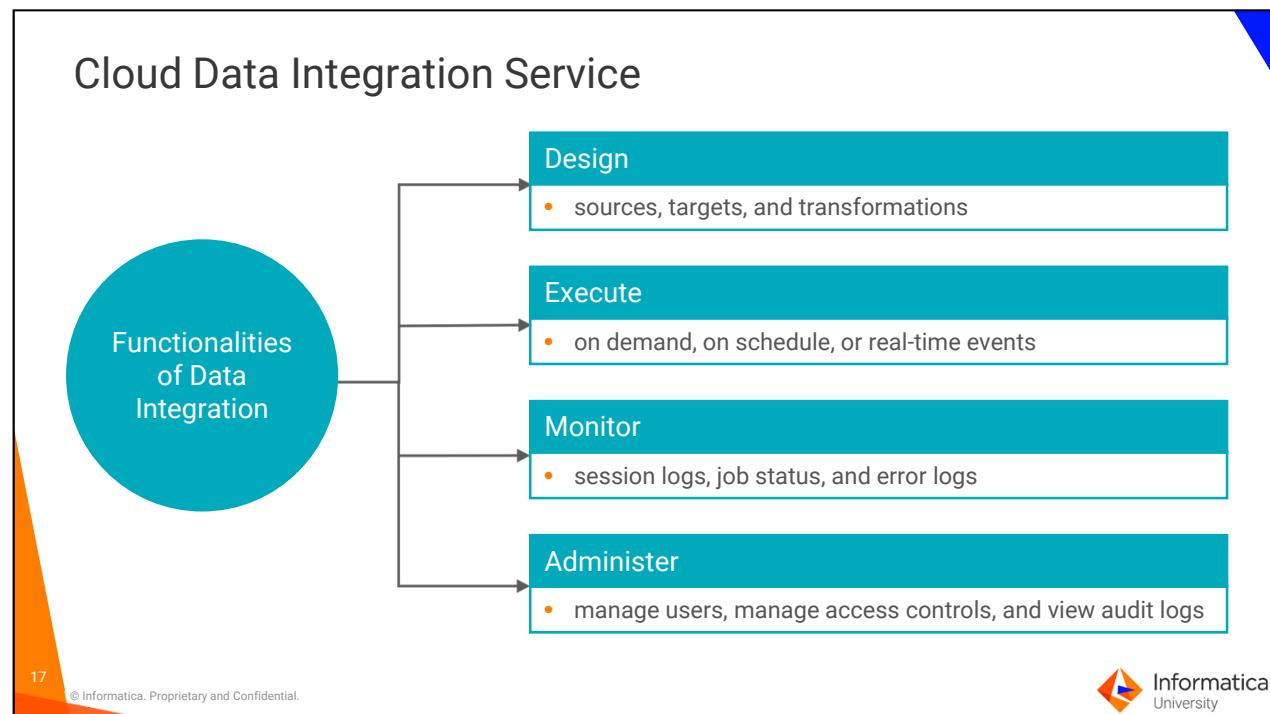
As you create, schedule, and run tasks, information is written to the IICS repository. This repository stores source and target metadata, mappings, connection information, schedules, and logging and monitoring information.

- **Source and Target metadata:** The repository stores metadata for each source and target object. This includes field names, data type, precision, and other information about the source and target object.
- **Mappings:** When you create a Data Integration task, the repository stores mappings and transformation rules.
- **Connection information:** The repository stores information that enables you to connect to specific source and target systems. The repository stores this information in an encrypted format.
- **Schedules:** You can configure tasks to run automatically using various scheduling options. The repository stores information regarding these schedules.
- **Logging and monitoring information:** The repository stores the results of all jobs. You can log in to IICS and view status details.

Topic

Cloud Data Integration Service





You can use the IICS Data Integration Service to design, execute, monitor, and administer tasks.

While designing a task, you can select from a variety of sources, targets, and transformations to transform and map data.

You can execute a task on demand, based on a schedule, or real-time events such as an outbound message triggered in Salesforce.

While monitoring a task, you can view session logs, job status, and error logs in the jobs section of the Data Integration Service.

As a part of the administrative tasks, you can manage users, manage their access rights and privileges, and view audit logs.

Assets and Components

- An asset is a task created in Informatica Cloud

Mapping Task

Synchronization Task

Masking Task

Replication Task

PowerCenter Task

Mass Ingestion Task

Mappings

Taskflows

- A component is used within an asset

- Saved Query

18

© Informatica. Proprietary and Confidential.



An asset is a task created in Informatica Cloud. Using the Data Integration service, you can create a Mapping task, Synchronization task, Masking task, Replication task, PowerCenter task, Mass Ingestion task, Mappings, and Taskflows.

A component is used within an asset. For example, to use a saved query in a synchronization task, you need to first create the saved query component, and then refer to it in the synchronization task asset.

Cloud Data Integration Assets



19

© Informatica. Proprietary and Confidential.



Mapping Task: A mapping task allows you to create custom tasks based on mappings or integration templates. This enables you to extend the capabilities of IICS and encapsulate repeatable business processes to create tasks.

Synchronization Task: A synchronization task allows you to load data and integrate applications, databases, and files. In a synchronization task, you can use advanced functionalities such as lookups, expressions, and multiple object sources. You can use the synchronization task for most of your integration jobs.

Masking Task: A masking task allows you to mask sensitive fields in source data with realistic test data for non-production environments.

Replication Task: A replication task is similar to a synchronization task; however, its focus is to move data out of an application and create a back-up.

PowerCenter task: You can import a PowerCenter workflow and run it as a Cloud Data Integration task.

Mass Ingestion Task: A mass ingestion task allows you to transfer, track, and monitor huge volumes of files between on-premise and cloud repositories.

Mapping: A mapping defines reusable data flow logic that you can use in mapping tasks. A mapping defines the flow of data from a source to a target.

Taskflows: A taskflow controls the sequence in which you execute the tasks. The execution sequence of tasks is based on the execution of the previous task. You must first create tasks and then add them to a taskflow.

Cloud Data Integration Components

Business Service

Maplets

Saved Query

Hierarchical Schema

Intelligent Structure Model

20

© Informatica. Proprietary and Confidential.



Business Service: A business service is a web service with configured operations. You can define a business service to add operations to the Web Services transformation in the Mapping Designer. You can use business service definitions in multiple mappings.

Maplets: A mapplet is a reusable transformation logic that you can use to transform source data before it is loaded to the target. You can create a mapplet in one of the following ways – create a mapplet in Data Integration, import a mapplet from PowerCenter, or generate a SAP BAPI or IDoc mapplet. After you create a mapplet, you can add it to a Mapplet transformation to use its transformation logic. Mapplets can be either active or passive. Passive mapplets contain a single input group, a single output group, and only passive transformations. Active mapplets contain at least one active transformation.

Saved Query: A saved query is a component that you can create to run SQL statements against a database. You can use a saved query as the source object in a synchronization task or as the query in a SQL transformation. Create a saved query when you want to use a database source that you cannot configure using the single or multiple object source options in a synchronization task.

Hierarchical Schema: A hierarchical schema is an asset that is based on a schema file or sample JSON file that you import to Data Integration. A hierarchical schema is required for Hierarchy Parser and Hierarchy Builder transformations. The schema defines the expected hierarchy of the output data.

Intelligent Structure Model: An intelligent structure model is an asset that is based on a sample file that contains data with little or no structure. Intelligent Structure Discovery determines the underlying patterns of the sample file and creates a model that can be used to transform, parse, and generate output groups. You can use an Intelligent Structure Model in a Structure Parser transformation in a Data Integration mapping.

Cloud Data Integration Components (continued)

Fixed-width File Format

File Listener

Shared Sequences

User-defined Functions

21

© Informatica. Proprietary and Confidential.



Fixed-width File Format: You can create and save fixed-width file formats that specify the formatting details for fixed-width flat files. You can use a fixed-width flat file as a source or target in mappings and mapping tasks. You can create multiple fixed-width file formats.

File Listener: A file listener listens to files on a defined location. IICS uses file listeners to monitor specific folders. A file listener receives notification through a call-back API when new files arrive at a monitored folder or when files in the folder are updated or deleted.

Shared Sequences: These are reusable sequences that you can use in multiple Sequence Generator transformations. When you use a shared sequence, the Sequence Generator transformation uses the properties of the shared sequence to generate values. You can use a shared sequence generator when you want to assign numeric values to your data in the same sequence in multiple mapping tasks. When you run the mapping task, Data Integration reserves a set of values in the sequence so that each mapping task generates unique values.

User-defined Functions: These are reusable functions that you can use in expressions. You can create user-defined functions to build complex expressions. User-defined functions use the same syntax and transformation language components as transformation and field expressions. You can include a user-defined function in a transformation expression in a mapping or mapplet, in a field expression in a mapping task, or in another user-defined function. You cannot use a user-defined function in an expression in a synchronization task.

Lab Activity

1-1 Navigating the IICS interface

In this lab, you will perform the following:

- Log in to the Informatica Cloud Org
- Access the Informatica Cloud online help
- Search the online help

Module Summary

This module showed you how to:

- Describe IICS as an iPaaS solution
- Define the key terminologies used in IICS
- Explore the IICS architecture
- List the Cloud Data Integration assets and components

23

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica[®]
University

Module 2

Runtime Environments and Connections



Module Objectives

After completing this module, you will be able to:

- Discuss Informatica Cloud runtime environments
- Explain the purpose of Informatica Cloud Secure Agent
- Explore the Secure Agent architecture
- View the Secure Agent log files
- List the steps to install the Secure Agent
- Define a connection
- Explore types of connectivity
- Discuss native and add-on connection types

3

© Informatica. Proprietary and Confidential.



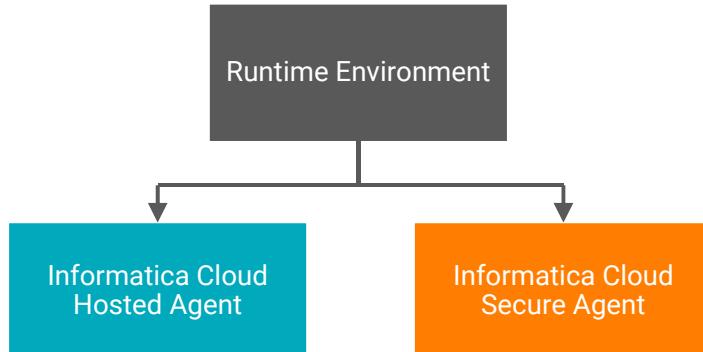
Topic

Informatica Cloud Runtime Environments



Runtime Environments

- An execution platform that runs a data integration or application integration task
- You must have at least one runtime environment in your Org



5

© Informatica. Proprietary and Confidential.



A runtime environment is the execution platform that runs a data integration or application integration task. To run tasks in your organization, you must have at least one runtime environment set-up in your organization.

Informatica Cloud supports two runtime environments – **Informatica Cloud Hosted Agent** and **Informatica Cloud Secure Agent**.

Informatica Cloud Hosted Agent

- You can use Informatica Cloud Hosted Agent to run tasks
- Run synchronization, mapping, and replication tasks
- Uses certain connectors such as Amazon S3, Google Analytics, Marketo, Microsoft Azure Blob Storage, and Workday
- Hosted Agent can process limited volumes of data
- You cannot add, delete, or configure a Hosted Agent

6

© Informatica. Proprietary and Confidential.



If your organization has the Cloud Runtime license, you can use the Informatica Cloud Hosted Agent to run the tasks.

The Hosted Agent can run synchronization, mapping, and replication tasks that use certain connectors such as Amazon S3, Google Analytics, Marketo, Microsoft Azure Blob Storage, Workday, and so on.

With Hosted Agent, you can only process limited volumes of data. Hence, if you want to handle huge volumes of data, you must use the Informatica Cloud Secure Agent.

The Hosted Agent runtime environment is managed by Informatica Cloud Data Integration. This means, you cannot add, delete, or configure a Hosted Agent.

Informatica Cloud Secure Agent

Lightweight and self-upgrading program that runs inside your network

Allows you to access all your local resources that reside behind your firewall

Your application data never gets staged or run through Informatica Cloud servers

7

© Informatica. Proprietary and Confidential.



The Informatica Cloud Secure Agent is another type of runtime environment. In addition to the Informatica Cloud application and repository, there is a local component called the Informatica Cloud Secure Agent.

The Secure Agent is a lightweight, self-upgrading program that runs on a machine inside your network. The Secure Agent is responsible for moving data directly from the source to the target. It allows you to access all your local resources, for example, databases or applications that reside behind the firewall.

Informatica Cloud Secure Agent is the local agent that runs the tasks. Therefore, your application data never gets staged or run through the Informatica Cloud servers. Your data remains completely secure and stays behind the firewall.

More about the Secure Agent

- Available for Windows and Linux platforms
- Run-time version of PowerCenter execution component
- Can install multiple agents within your network
- One agent per machine
- Secure agent is automatically linked to the IICS Org that you install it from

The Secure Agent is available for Windows and Linux platforms. If you are familiar with Informatica's PowerCenter, you will realize that the Secure Agent is the run-time version of the PowerCenter execution component. If you want to connect to multiple resources, you can install multiple agents within your network. However, you are restricted by the license agreement to the number of agents you can install within your network.

You must also note that you can install only one agent per machine. This means that you can have only one agent on your machine that communicates with a single IICS Org. When you install the agent on your machine, it is automatically linked to the IICS Org that you install it from. Different users can log in to the same IICS Org and install the agents on their machines.

When you install an agent in a production environment, you must ensure to install the agent on a machine that is always up and running and available to run the tasks.

Multiple Secure Agents

- In a Production environment, there is one Secure Agent installed in the Org
- You can install additional agents in the following scenarios:
 - when you want separate agent machines with different controls/permissions for different businesses/user groups
 - for failover purposes
 - for additional processing power
- Additional agent requires additional license fee

9

© Informatica. Proprietary and Confidential.

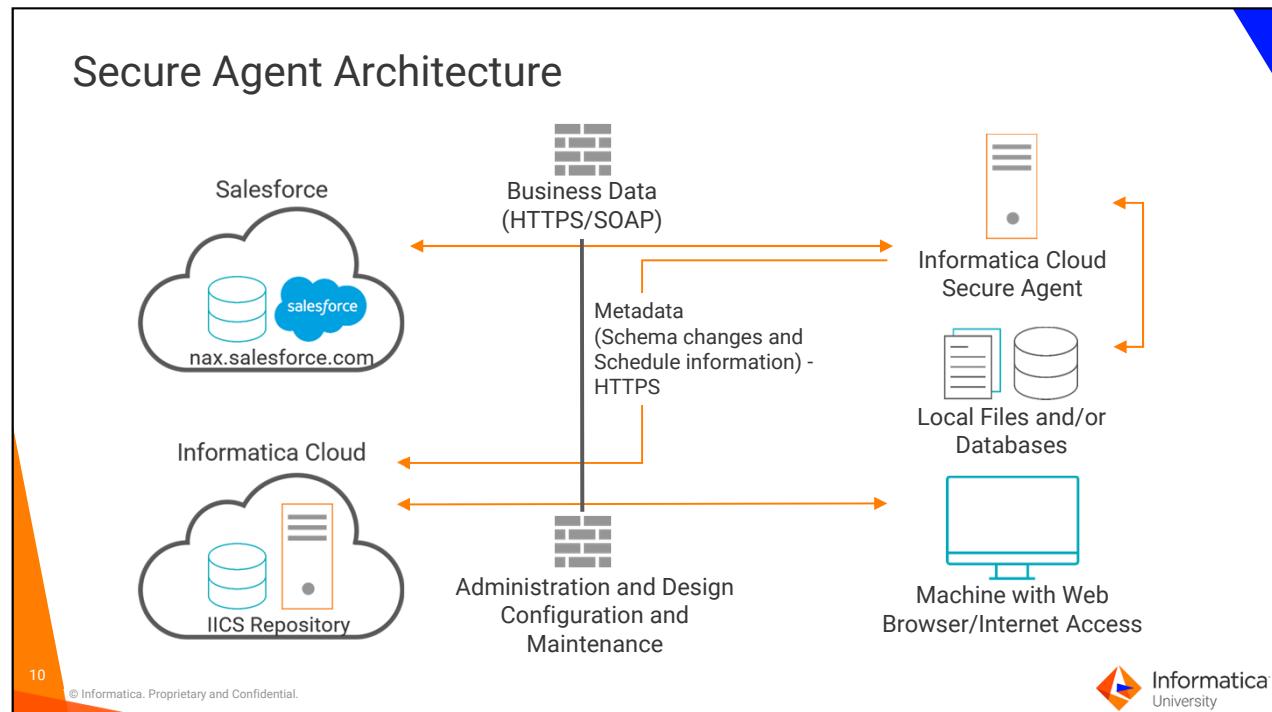


As you have just seen, you can install multiple Secure Agents within your network. In the production environment, there is usually one Secure Agent installed in the Org and all users run tasks using the same agent.

There are some instances where you will install additional agents.

- When you want separate agent machines with different controls and permissions for different businesses and user groups
- For Failover purposes
- For additional processing power. For example, when you want to run more than six tasks concurrently.

Again, you must note that additional agent requires additional license fee.



If you include the Secure Agent in the IICS architecture, you can see the agent running behind the firewall. The agent gives you access to any local files or on-premise databases or applications.

Once the task is initiated, the Secure Agent connects to the IICS Repository and downloads all metadata, including scheduling information, mappings, and so on. IICS performs the design and administration of tasks through a web browser.

When the agent wants to connect to a SaaS application, in this case Salesforce, it connects through the Business API.

Troubleshooting the Secure Agent

- If the agent status is inactive, communication with the Informatica Cloud Servers can be blocked by:
 - Windows Firewall
 - Virus Scanner
 - Content Filter
- If your company uses a Proxy Server, enter the following settings:
 - Proxy Server name
 - Proxy Port number
 - Proxy User ID (optional)
 - Proxy Password (optional)
- Visit Community site for more specific information

11

© Informatica. Proprietary and Confidential.



If the agent shows as inactive in the list of services, a probable reason is that the agent is not able to connect to the Informatica Cloud Server.

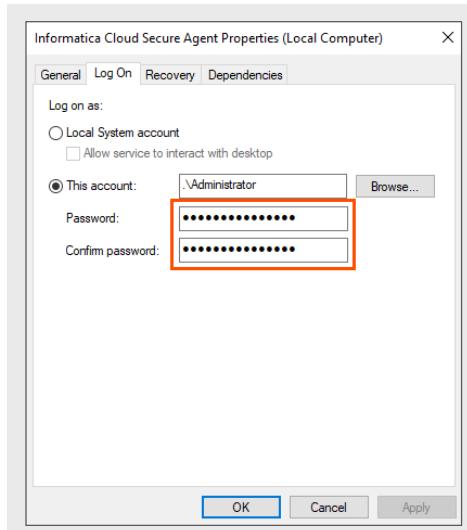
Sometimes, Windows firewall, virus scanner, or content filters can also block communication of the agent with the server. So, you can disable these applications and then check the status of the agent.

If your company uses a proxy server, you must configure proxy server settings such as the proxy server name, proxy port number, proxy user ID, and proxy password. You can get these settings from your IT department. You may also be able to obtain this information using your web browser or a detector application available on the Internet.

If you face issues while installing the agent, you can visit the Informatica Community site and search for probable solutions there. The site contains a lot of information on troubleshooting the agent.

Running Secure Agent as Local/Network User

- Agent inherits access privileges of Windows user to install the agent
- Secure agent needs permission to access directories on Windows
- Requires configuration of a new login for the Windows service



12

If you are a user who runs the agent on a Window's system, the agent inherits the access privileges of the Window's user. This happens automatically when you install the agent. The agent needs permission to access the directories on Windows. Therefore, you may have to reconfigure the Windows service.

The image shows the service properties that you can modify. This information is also documented on the Informatica Community site.

Topic

IICS Log Files



IICS Log Files

- The Secure Agent generates the following log files:
 - Session log
 - Error log
 - Success log
 - Infaagent log
 - Tomcat log

14

© Informatica. Proprietary and Confidential.



When the Secure Agent runs a task, it generates different log files such as the session log, error log, success log, infa agent log, and tomcat log.

Session Log

- Secure agent generates session logs for all tasks that run in the org
- Session logs provides technical details about the task
- Session logs are available at the following location:
 - C:\Program Files\Informatica Cloud Secure Agent\apps\Data_Integration_Server\logs

Local Disk (C:) > Program Files > Informatica Cloud Secure Agent > apps > Data_Integration_Server			
Name	Date modified	Type	Size
54.0.7	6/13/2019 10:08 AM	File folder	
conf	5/13/2019 3:37 PM	File folder	
data	5/13/2019 3:49 PM	File folder	
ext	5/13/2019 3:47 PM	File folder	
logs	5/13/2019 7:04 PM	File folder	

15

© Informatica. Proprietary and Confidential.



The Secure Agent generates session logs for all tasks that run in the org regardless of the status. The session logs provide all the technical details about the task. On the Secure Agent machine, the session logs are available at the location as shown in the image.

Error Log

- Error logs provides the list of all records that failed to process and the reason for the failure
- Error logs are available at the following location:
 - C:\Program Files\Informatica Cloud Secure Agent\apps\Data_Integration_Server\data\error

Local Disk (C:) > Program Files > Informatica Cloud Secure Agent > apps > Data_Integration_Server > data		
Name	Date modified	Type
AccessReport	5/13/2019 3:49 PM	File folder
cache	5/13/2019 4:53 PM	File folder
checksum	5/13/2019 3:49 PM	File folder
connections	6/13/2019 1:06 PM	File folder
error	5/13/2019 7:04 PM	File folder
logs	5/13/2019 3:49 PM	File folder

16

© Informatica. Proprietary and Confidential.

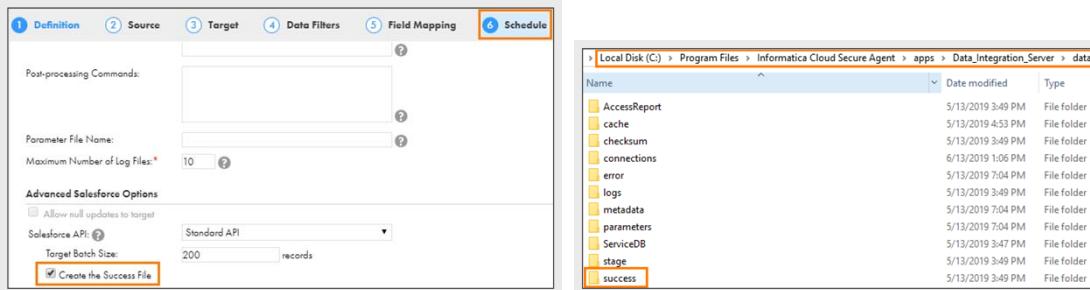


The error logs provide the list of all the bad records that failed to process and the reason for the failure. The error logs are available at the location as shown in the image.

Success Log

- Success logs provides a list of all success records
- Secure agent does not generate success logs by default
- You can create success logs if your task has a Salesforce target
- Success logs are available at the following location:

• C:\Program Files\Informatica Cloud Secure Agent\apps\Data_Integration_Server\data\success



17

© Informatica. Proprietary and Confidential.



The success logs provide a list of all records that were processed successfully. The Secure Agent does not generate success logs by default. However, you have the option to create success logs if your task has a Salesforce target. In step 6 of the synchronization task wizard, you can enable the option to create the success files for Salesforce targets. The Salesforce success files contains a record of all rows that you create, update, or delete. If you want to roll back the operation, you can use the success files to track the rows that were created.

The success logs are available at the location as shown in the image.

Infaagent Log

- Infaagent log provides all the details of the network connectivity
- Infaagent log is available at the following location:
 - **C:\Program Files\Informatica Cloud Secure Agent\apps\agentcore\infaagent.log**

Name	Date modified	Type	Size
54.2	5/13/2019 3:36 PM	File folder	
conf	5/13/2019 3:36 PM	File folder	
data	6/13/2019 1:08 PM	File folder	
logs	6/13/2019 1:07 PM	File folder	
abortAgentCore.bat	4/13/2019 1:42 AM	Windows Batch File	1 KB
abortAgentMgr.bat	4/13/2019 1:42 AM	Windows Batch File	1 KB
agent_start.bat	4/13/2019 1:42 AM	Windows Batch File	2 KB
agentcore.log	6/13/2019 10:16 AM	Text Document	2,481 KB
agentcore.log.1	6/4/2019 2:44 PM	1 File	10,241 KB
consoleAgentManager.bat	4/13/2019 1:42 AM	Windows Batch File	1 KB
customUninstallScript.bat	4/13/2019 1:42 AM	Windows Batch File	5 KB
FileLockDetect.exe	4/13/2019 1:42 AM	Application	66 KB
infaagent.exe	4/13/2019 1:42 AM	Application	92 KB
infaagent.log	6/11/2019 7:21 PM	Text Document	11 KB

18

© Informatica. Proprietary and Confidential.



The infa agent log provides all the details of the network connectivity such as success, failure, timing of the successful connection, and attempts made to reconnect.

The infa agent log is available at the location as shown in the image.

Tomcat Log

- Tomcat log provides details about the task such as, the start time, the request sent to the Integration Server, and response received from the Integration Server
- Tomcat log is available at the following location:
 - C:\Program Files\Informatica Cloud Secure Agent\apps\Data_Integration_Server\logs\tomcat



19

© Informatica. Proprietary and Confidential.



The tomcat log provides details about the task such as, the time when the task was started, the request sent to the Integration Server, and the response received from the Integration Server.

The tomcat log is available at the location as shown in the image.

Log File Naming Convention

- You can identify the type of task from the session log file names

 s_dss_0103A00I0000000000007_5_import.log	8/8/2018 3:48 PM	Text Document	2 KB
 s_dss_0103A00I0000000000007_7.log	8/8/2018 5:46 PM	Text Document	12 KB
 s_dss_0103A00I0000000000007_7_import.log	8/8/2018 5:45 PM	Text Document	2 KB
 s_mtt_0103A00Z00000000000B_2.log	8/30/2018 5:41 PM	Text Document	12 KB
 s_mtt_0103A00Z00000000000B_2_import.log	8/30/2018 5:41 PM	Text Document	2 KB
 s_mtt_0103A00Z00000000000U_1.log	8/8/2018 3:48 PM	Text Document	12 KB
 s_mtt_0103A00Z00000000000U_1_import.log	8/8/2018 3:47 PM	Text Document	2 KB
 s_mtt_0103A00Z00000000000U_2.log	8/8/2018 5:46 PM	Text Document	12 KB
 s_mtt_0103A00Z00000000000U_2_import.log	8/8/2018 5:46 PM	Text Document	2 KB
 s_mtt_0103A00Z00000000000W_1.log	9/4/2018 6:08 PM	Text Document	15 KB
 s_mtt_0103A00Z00000000000W_1_import.log	9/4/2018 6:08 PM	Text Document	2 KB
 s_mtt_0103A00Z00000000000X_3.log	9/6/2018 4:26 PM	Text Document	15 KB

20

© Informatica. Proprietary and Confidential.



You can identify the type of task from the session log file names.

In the file name, the prefix indicates the type of task. So, in this image, **dss** indicates a Synchronization task, **mtt** indicates a Mapping Task, and so on.

Every task in IICS has a unique ID. The ID is present in the next part of the log file name.

In the file names, you will also notice the **_1**, **_2**, and so on. This indicates the number of times the task was run. So, in the example image shown on your screen, the first log file, which indicates a Synchronization task, was run 5 times. Similarly, the last log file, which indicates a Mapping Task, was run 3 times.

Log File Storage

- IICS stores up to 10 log files per task
- When the task is run for the 11th time, the first log is overwritten
- Archive historical records of a task into another directory

21

© Informatica. Proprietary and Confidential.



IICS stores up to 10 log files per task.

When a task is run for the 11th time, the first log file is overwritten. Similarly, when the task is run for the 12th time, the second log file is overwritten, and so on.

If you want historical record of your tasks, then you must archive the log files in another directory.

Topic

Connection Types



What is a Connection?

- An Informatica Cloud object
- Provides access to data in cloud and on-premise applications, platforms, databases, and flat files
- Specifies the location of sources, lookup objects, and targets that are included in a task
- Can create a connection for any connector that is pre-installed in IICS
- Can create a connection by installing an add-on connector

23

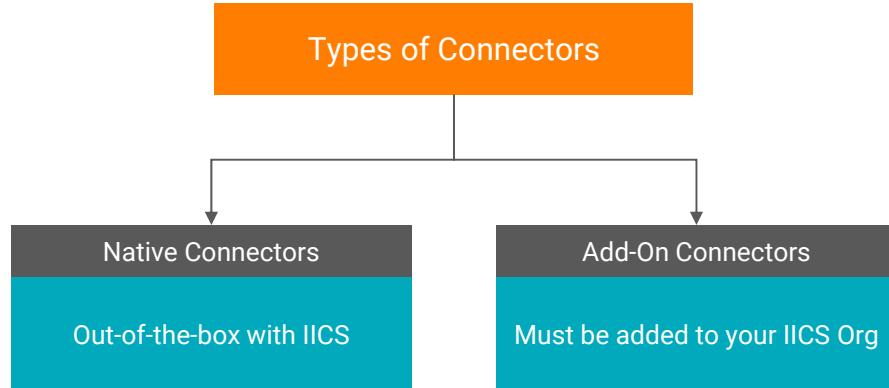
© Informatica. Proprietary and Confidential.



A connection is an Informatica Cloud object that provides access to data in the cloud and on-premise applications, platforms, databases, and flat files. Connections specify the location of sources, lookup objects, and targets included in a task.

You can create connections using connectors. You can create a connection for any connector that is pre-installed in IICS. You can also create a connection by installing an add-on connector created by Informatica or an Informatica partner.

Types of Connectivity



24

© Informatica. Proprietary and Confidential.



IICS provides two types of connectors to create a connection. They are – Native connectors and Add-on connectors.

Native connectors are provided out-of-the-box with IICS. This means they are pre-installed in IICS and do not require any additional setup.

Add-on connectors are additional connectors available to all Informatica Cloud customers. To use an Add-on connector, you must first add the connector to your IICS Org.

Connectivity Examples

Native	Add-On
<ul style="list-style-type: none">• Salesforce• Flat File• Oracle• FTP/SFTP• SQL Server• MySQL• ODBC• MS Access• SAP• Web Service• Microsoft Dynamics CRM	<ul style="list-style-type: none">• Amazon Redshift• Avature• Box• Eloqua• Chatter• Concur• Work Day• Amazon Web Services S3• Microsoft Azure Blob Storage V3• Marketo• Zuora

25

© Informatica. Proprietary and Confidential.



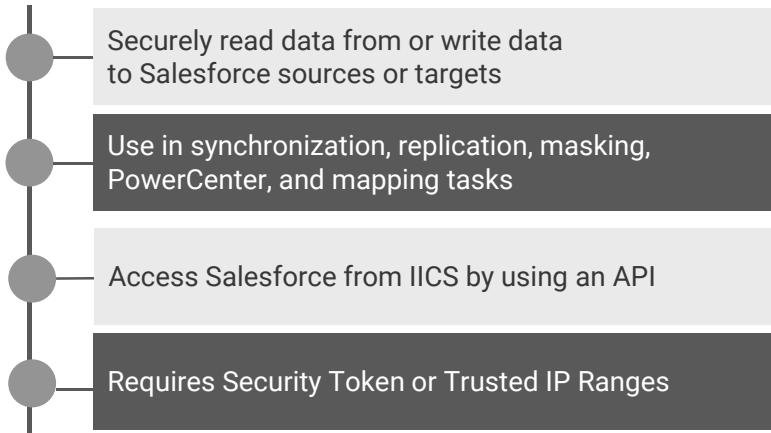
- File Transfer Protocol (FTP), and Secure File Transfer Protocol (SFTP) – These connections are similar to Flat File connections. However, the only difference with FTP and SFTP connections is that the file is on a remote machine.
- SQL Server – the supported versions for SQL server are 2000, 2005, 2008, 2012, and 2016.
- MySQL – The supported version is 5.0.x.

Topic

Native Connectors



Salesforce Connection



27

© Informatica. Proprietary and Confidential.



Salesforce connections allow you to securely read data from or write data to Salesforce sources or targets. You can use Salesforce connections in synchronization, replication, masking, PowerCenter, and mapping tasks.

You can also access Salesforce from IICS by using an API. When Salesforce is accessed by another application using APIs, it requires an extra layer of security. So, when you configure a Salesforce connection in IICS, you must provide the security token that is generated by Salesforce. As an alternative to using the security token, you can add the IICS server, as well as any machines that run the Secure Agent, as “Trusted IP Ranges” in your Salesforce account.

Salesforce Security Token



Salesforce automatically generates a security token

Security token provides improved security for the Salesforce account

Reset the security token via the Salesforce UI

Add the security token to Salesforce connection in IICS

28

© Informatica. Proprietary and Confidential.



Salesforce automatically generates a security token. The purpose of this token is to improve the security of the Salesforce account when it is accessed by a third-party application using an API.

You can get the security token, by using the “Reset Security Token” option in Salesforce. After you get the security token, you must enter it in the Salesforce connection properties in IICS.

The security token is valid until you reset it, change the Salesforce account password, or reset the account password.

Salesforce Trusted IP Range



Trusted IP Ranges define a list of IP addresses from which users can log in without receiving a login challenge

Alternative to using the security token

Add IICS Servers and any agent machine(s) IP addresses to Salesforce org

29

© Informatica. Proprietary and Confidential.



“Trusted IP Ranges” define a list of IP addresses from which users can log in without receiving a login challenge for verification of their identity.

Security tokens are tied to a user and can expire. This means that if you change your Salesforce account password, you will have to generate a new security token. So, instead of using the Salesforce security token, you can use the Salesforce Trusted IP Ranges. You must add IICS Servers and the IP addresses of the Secure Agent machines to the Salesforce Org.

To know more about adding “Salesforce Trusted IP Ranges”, as well as getting the IP address ranges for the IICS Servers, you can check the Informatica Cloud Online Help or visit the Informatica Community site.

Best Practice for Salesforce Connection

- Create a dedicated Salesforce user for Data Integration tasks
- Salesforce user specified in IICS connection shows up on all records as the Record Owner
- When records are modified through IICS, the user is listed in the 'Last Modified By' field

Account Detail		Edit Delete Include Offline
Account Owner	Informatica Training [Change]	
Account Name	NH Dneeds [View Hierarchy]	
Parent Account		
Account Number		
Account Site		
Type		
Industry		
Annual Revenue		
Billing Address	11 West Park Road Pasadena, CA 91001 USA	
Customer Priority		
SLA Expiration Date		
Number of Locations		
Active		
Created By	Informatica Training, 5/27/2019 1:07 AM	
Description		
Custom Links	Billing	
Rating		
Phone	754-3010	
Fax		
Website		
Ticker Symbol		
Ownership		
Employees	4	
SIC Code		
Shipping Address		
SLA		
SLA Serial Number		
Upsell Opportunity		
Last Modified By	Informatica Training, 5/27/2019 1:07 AM	

30

© Informatica. Proprietary and Confidential.



You must create a Salesforce user as the dedicated Data Integration user. When you create a Salesforce connection in IICS, you can use the dedicated user for the Salesforce connection.

When you create records in Salesforce through IICS, the Salesforce user specified in your IICS connection shows up on all records as the Record Owner. When you modify the records through IICS, the user is listed in the 'Last Modified By' field.

While you can change the Record Owner, you cannot change the 'Last Modified By' field, because it is set by Salesforce.

Salesforce Service URL

The screenshot shows the 'Connection Details' and 'Salesforce Connection Properties' sections. In the 'Standard Connection Properties' section, the 'Service URL' field is set to 'https://login.salesforce.com/services/Soap/u/40'. A callout bubble with an orange arrow points to this field, containing the text: 'Change the Service URL to "test.salesforce.com" to access a Sandbox'.

31 © Informatica. Proprietary and Confidential.

 Informatica University

When you create a Salesforce connection in IICS, you must configure the connection properties.

By default, the Service URL field is automatically populated as shown in the image. You must retain the default Service URL if you want to connect to a Production Org in Salesforce. However, if you want to connect to a Salesforce Sandbox instead of a Production Org, then in the Service URL field, you must replace 'login' with 'test'.

Flat File Connection

- Allows you to create, access, and store Flat Files
- Use Flat File connections in mapping tasks, PowerCenter tasks, replication tasks, and synchronization tasks
- Choose the formatting options for the Flat File
- Specify connection properties:

Runtime Environment	Directory	Date Format	Code Page
specifies the Secure Agent that IICS uses to access the Flat File in the local area network	specifies the location where you store the Flat File	specifies the format for date fields in the Flat File	specifies the code page of the system that hosts the Flat File

32

© Informatica. Proprietary and Confidential.



Flat File connections allow you to create, access, and store flat files. You can use Flat File connections in mapping tasks, PowerCenter tasks, replication tasks, and synchronization tasks. When you select a Flat File connection, you must choose the formatting options for the Flat File. When you choose the formatting options in a Source, Lookup, or Target transformation, you must specify whether the Flat File is a delimited Flat File or a fixed-width Flat File.

When you create a Flat File connection, you must specify some connection properties:

- **Runtime Environment:** Specifies the Secure Agent that IICS uses to access the Flat File in the local area network.
- **Directory:** This is the location where you store the Flat File. You must enter the full path to the directory or you can click the Browse button to locate and select the directory.
- **Date Format:** This is the format for date fields in the Flat File.
- **Code Page:** This is the code page of the system that hosts the Flat File.

Oracle Connection

- Allows you to read data from or write data to Oracle sources or targets
- Configure the following connection properties:
 - Runtime Environment
 - User Name
 - Password
 - Host
 - Port
 - Service Name/System ID
 - Schema
 - Code Page
 - Encryption Method
 - Crypto Protocol Version

33

© Informatica. Proprietary and Confidential.



Oracle connections allow you to read data from or write data to Oracle sources or targets. When you create an Oracle connection, you must configure few connection properties.

- **Runtime Environment:** This specifies the Secure Agent that IICS uses to access Oracle.
- **User Name:** This specifies the name for the database login user.
- **Password:** This is the password for the database login user.
- **Host:** This is the name of the machine that hosts the database server.
- **Port:** This is the network port number to connect to the database server. The default port number is 1521.
- **Service Name or System ID:** The service name or system ID uniquely identifies the Oracle database.
- **Schema:** This specifies the schema used for the Oracle connection.
- **Code Page:** This is the code page of the database server.
- **Encryption Method:** The encryption method specifies the method that the Secure Agent uses to encrypt the data exchanged between the Secure Agent and the database server.
- **Crypto Protocol Version:** This specifies the cryptographic protocols to use when you enable SSL encryption.

Oracle Connection (continued)

- Validate Server Certificate
- Trust Store
- Trust Store Password
- Host Name in Certificate
- Key Store
- Key Store Password
- Key Password
- Metadata Advanced Connection Properties
- Runtime Advanced Connection Properties

34

© Informatica. Proprietary and Confidential.



- **Validate Server Certificate:** Validates the certificate that is sent by the database server.
- **Trust Store:** Specifies the location and name of the trust store file.
- **Trust Store Password:** This is the password to access the contents of the trust store file.
- **Host Name in Certificate:** Indicates the host name of the machine that hosts the secure database.
- **Key Store:** Specifies the location and the file name of the key store.
- **Key Store Password:** This is the password for the key store file required for secure communication.
- **Key Password:** Specifies the password for the individual keys in the key store file required for secure communication.
- **Metadata Advanced Connection Properties:** These are the optional properties for the JDBC driver to fetch the metadata.
- **Runtime Advanced Connection Properties:** These are the optional properties for the ODBC driver to run the mappings.

ODBC Connection

- Allows you to connect to a database
- Use ODBC connections in synchronization tasks, mappings, and mapping tasks
- Standard interface for accessing Database Management System
- Uses ODBC driver as a translation layer
- Example:
 - QuickBooks



35

© Informatica. Proprietary and Confidential.



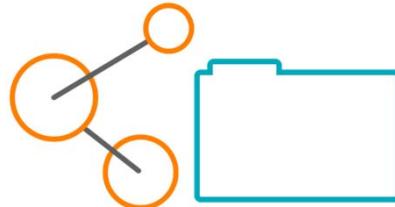
ODBC connections allow you to connect to a database. You can use ODBC connections in synchronization tasks, mappings, and mapping tasks.

ODBC drivers are available for different databases that you want to connect to. An application can use ODBC to query data from a Database Management System, regardless of the operating system it uses. ODBC accomplishes this independence by using an ODBC driver as a translation layer between the application and the Database Management System.

An example of an ODBC connection is “QuickBooks”. Although QuickBooks does not run on a true relational database, there is an ODBC driver made by a third-party company that connects to the QuickBooks file as a relational database.

FTP/SFTP Connection

- FTP and SFTP connection
 - allows you to extract data from or load data to a flat file on a remote machine
- Must define local and remote directories
- File structures of both directories must match
 - uses local file for data preview
- Commonly used in B2B scenarios
 - exchange data with partner at regular intervals
 - use FTP connection to access data file directly on your partner's machine



36

© Informatica. Proprietary and Confidential.



IICS supports FTP and SFTP connections. FTP connections allow you to extract data from or move data to a flat file on a remote machine. SFTP connections use secure protocols, such as Secure Shell or SSH to access source and target files.

When you configure FTP and SFTP connections, you must define a local and remote directory. It's very important to remember that the file structures of both the directories must match. You must define the local directory on the Secure Agent machine that contains a copy of the source or target files. The remote directory is the location of the files you want to use as source or target.

When you configure a task with FTP and SFTP connections, IICS uses the file structure of the local file to define the source or target for the task. The local file is used for metadata reference. IICS uses the local file to generate data preview. If the data in the local file does not match the data in the source or target file in the remote directory, data preview displays inaccurate results.

FTP and SFTP connections are commonly used in Business to Business scenarios where you exchange data with a business partner at regular intervals. Example: Your business partner sends you a data file daily and you load the file into Salesforce. To remove the dependency of your business partner sending you the files daily, you can use FTP and SFTP connections to access the data file directly on your partner's machine.

SAP Connection

- Use SAP connector to integrate with SAP systems in batch, asynchronous, or synchronous modes
- SAP is an application platform that integrates multiple business applications and solutions
- Developers can add business logic within SAP using J2EE or ABAP
- Data Integration supports iDoc read, iDoc write, or BAPI/RFC functions to integrate with SAP systems
- Use BAPI/RFC functions for object-level integration, and iDocs functions for message-level integration
- Use the SAP connection in synchronization tasks, mappings, and mapping tasks

37

© Informatica. Proprietary and Confidential.



You can use the SAP connector to integrate with SAP systems in batch, asynchronous, or synchronous modes based on your requirements.

SAP is an application platform that integrates multiple business applications and solutions, such as the Customer Relationship Management, Advanced Planner and Optimizer, and Bank Analyzer. Developers can add business logic within SAP using Java 2 Enterprise Edition or Advanced Business Application Programming (ABAP).

Data Integration supports IDoc read, IDoc write, or BAPI-RFC functions to integrate with SAP systems. BAPI stands for Business Application Programming Interface and RFC stands for Remote Function Call. You can use BAPI-RFC functions for object-level integration, and IDocs functions for message-level integration.

You can use the SAP connection in synchronization tasks, mappings, and mapping tasks.

Data Integration Using BAPI/RFC Functions

- BAPI functions allow third-party applications to synchronously integrate with SAP at the object-level
- Use BAPI functions to read, create, change, or delete data in SAP
- Define the functions in the SAP Business Objects Repository
- Call the functions as an ABAP program within SAP or from any external application
- SAP connector uses RFC protocol to call BAPI/RFC functions outside of SAP
- Import a BAPI/RFC function as a maplet to Data Integration
- Data Integration makes the RFC function calls to SAP to process data synchronously

38

© Informatica. Proprietary and Confidential.



BAPI functions allow third-party applications to synchronously integrate with SAP at the object-level. You can use these functions to read, create, change, or delete data in SAP.

You can define BAPI functions in the SAP Business Objects Repository. You can call the functions as an ABAP program within SAP or from any external application. SAP connector uses RFC protocol to call BAPI-RFC functions outside of SAP.

You can import a BAPI-RFC function as a maplet to Data Integration. You can then use the maplet in a mapping to read, create, change, or delete data in SAP. When you run the mapping or the mapping task, Data Integration makes the RFC function calls to SAP to process data synchronously.

Data Integration Using iDOC Functions

- iDoc functions electronically exchange data between SAP applications or between SAP applications and external programs
- iDoc is a message-based integration interface that processes data asynchronously
- iDoc is a component of ALE module that sends and receives iDocs over RFC protocol

39

© Informatica. Proprietary and Confidential.



Intermediate Document (iDoc) functions electronically exchange data between SAP applications or between SAP applications and external programs. iDoc is a message-based integration interface that processes data asynchronously.

iDoc is a component of Application Link Enabling (ALE) module within SAP that sends and receives iDocs over RFC protocol.

SAP iDocs and RFC/BAPI Connector Administration

- Verify that the SAP connector license is enabled for the IICS org
- Download and install the Microsoft Visual C ++ Redistributable
- Download and configure the SAP libraries for iDoc and BAPI/RFC
- Configure the **sapnrfc.ini** file stored on the Secure Agent machine
- Define SAP Connector as a logical system in SAP
- Configure SAP user authorizations
- Install and configure the SAP iDocs Metadata utility

40

© Informatica. Proprietary and Confidential.



Before you can use an SAP connection to process the data through iDocs or RFC-BAPI, the Administrator must perform tasks such as:

- Verify that the SAP connector license is enabled for the IICS org
- Download and install the Microsoft Visual C++ Redistributable
- Download and configure the SAP libraries for iDoc and BAPI-RFC
- Configure the **sapnrfc.ini** file that is stored on the Secure Agent machine
- Define SAP Connector as a logical system in SAP
- Configure SAP user authorizations, and
- Install and configure the SAP iDocs Metadata utility

After the administrator has performed the configuration, you can create and use SAP RFC-BAPI, iDoc Reader, and iDoc Writer connections in mappings.

SAP RFC/BAPI and iDoc Writer Connection Properties

- To access SAP data through the RFC/BAPI interface or to write SAP data through the iDoc Writer interface, you must configure the following connection properties:
 - User Name: specifies the name of the authorized SAP user
 - Password: specifies the user's password
 - Connection String: specifies the DEST entry that you specified in the sapnrfc.ini file for the SAP application server
 - Code Page: specifies the code page compatible with the SAP target
 - Language Code: specifies the language code that corresponds to the SAP language
 - Client Code: specifies the SAP client number

Connection Details	
Connection Name: [*]	SAP_Connection
Description:	<input type="text"/>
Type: [*]	SAP
SAP Connection Properties	
Runtime Environment: [*]	CDI_XX_FIRSTNAME
SAP Connection Type: [*]	SAP RFC/BAPI Interface
SAP RFC/BAPI Interface Connection Properties	
User Name: [*]	<input type="text"/>
Password: [*]	<input type="password"/>
Connection String: [*]	<input type="text"/>
Code Page: [*]	<input type="text"/>
Language Code: [*]	<input type="text"/>
Client Code: [*]	<input type="text"/>

41

© Informatica. Proprietary and Confidential.



SAP connections enable you to access SAP data through the iDoc or BAPI-RFC interfaces. To access SAP data through the RFC-BAPI interface or to write SAP data through the iDoc Writer interface, you must configure certain connection properties:

- User Name: specifies the name of the authorized SAP user.
- Password: specifies the user's password.
- Connection String: specifies the DEST entry that you specified in the **SAP nrfc.ini** file for the SAP application server.
- Code Page: specifies the code page that is compatible with the SAP target.
- Language Code: specifies the language code that corresponds to the SAP language.
- Client Code: specifies the SAP client number.

SAP iDoc Reader Connection Properties

- To read SAP data through the iDoc Reader interface, you must configure the Destination Entry and Code Page connection properties
- Destination Entry specifies the DEST entry that you specified in the **sapnrfc.ini** file for the RFC server program registered at a SAP gateway
- Code Page specifies the code page compatible with the SAP source

Connection Details	
Connection Name: [*]	SAP_Connection
Description:	<input type="text"/>
Type: [*]	SAP
SAP Connection Properties	
Runtime Environment: [*]	CDI_XX_FIRSTNAME
SAP Connection Type: [*]	iDoc Reader
iDoc Reader Connection Properties	
Destination Entry: [*]	<input type="text"/>
Code Page: [*]	<input type="text"/>

42

© Informatica. Proprietary and Confidential.



To read SAP data through the iDoc Reader interface, you must configure the Destination Entry and Code Page connection properties.

The Destination Entry specifies the DEST entry that you specified in the **sapnrfc.ini** file for the RFC server program registered at a SAP gateway. The Program ID for this destination entry must be the same as the Program ID for the logical system you defined in SAP to receive iDocs.

The Code Page specifies the code page compatible with the SAP source.

Topic

Add-on Connectors



Workday Connector

- Use Workday connector to connect to Workday from Data Integration
- Read data from or write data to Workday
- Workday exposes the web service API, which the Secure Agent uses to perform integration tasks through the SOAP protocol
- Use Workday connector in a Source transformation, Target transformation, or midstream in a Web Services transformation
- Interact with the Workday service to perform operations on non-relational hierarchical data
- To enable Workday connector, you must contact Informatica Support

44

© Informatica. Proprietary and Confidential.

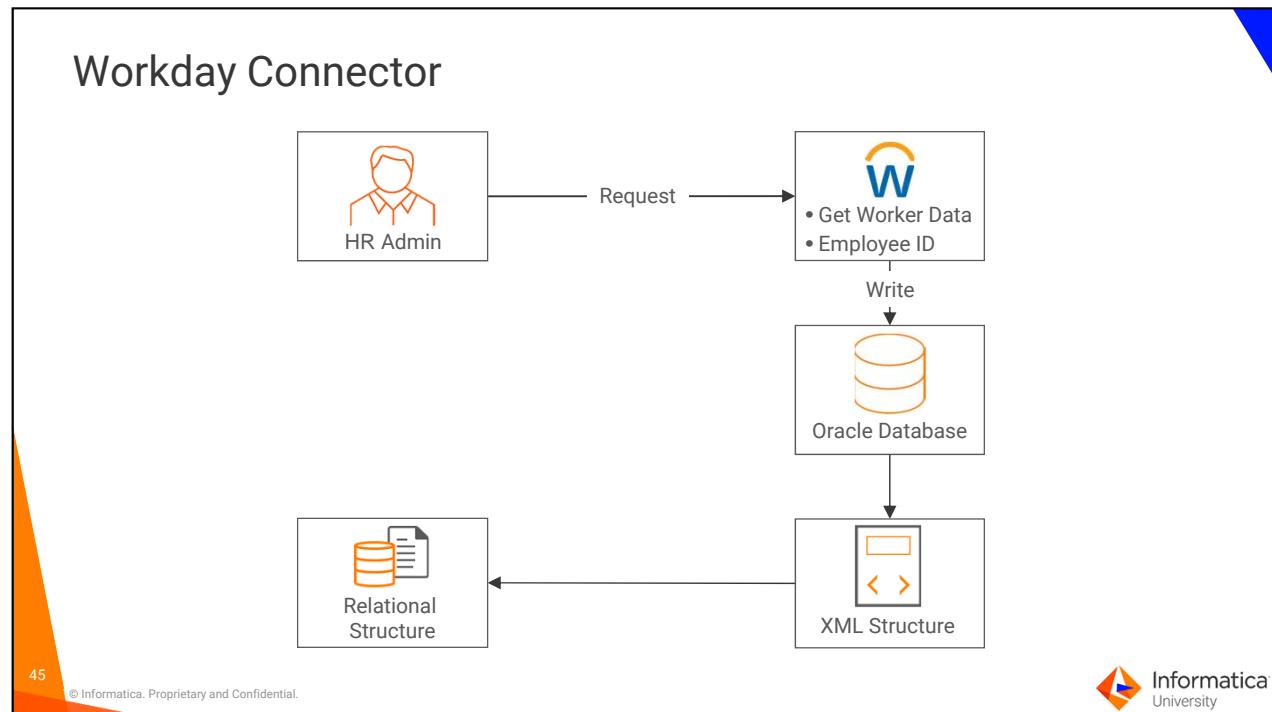


A Workday connector allows you to connect to Workday from Data Integration. You can read data from or write data to Workday.

Workday is an on-demand Cloud-based Enterprise Resource Application that includes financial management and human capital management applications.

You can use Workday connector in a Source transformation, Target transformation, or midstream in a Web Services transformation. You can interact with the Workday service to perform operations on non-relational hierarchical data.

To enable Workday connector, you must contact Informatica Support.



Here is an example.

You are an HR administrator and need to archive employees' details who left the organization in the past month. You can find the employee in Workday based on the employee ID, retrieve worker data through the “Get Workers” operation, and then write the details to an Oracle database target. With Workday connector, you can retrieve the worker data in an XML structure and then define a corresponding relational structure to write to the relational target.

Amazon Web Services S3 Connector

- Cloud-based store that stores many objects in one or more buckets
- You can connect to Amazon S3 buckets available in VPC, through VPC endpoints
- Read data from or write data to multiple Amazon S3 sources and targets
- Use Amazon S3 connection in synchronization tasks, mappings, and mapping tasks
- Configure the following connection properties:
 - Runtime Environment
 - Access Key
 - Secret Key
 - Folder Path
 - Master Symmetric Key
 - Code Page
 - Region Name

46

© Informatica. Proprietary and Confidential.



You can use an Amazon S3 connector to connect Informatica Cloud and Amazon S3. When you set up an Amazon S3 connection, you must configure few connection properties:

- The Runtime Environment specifies the name of the runtime environment where you want to run the tasks.
- The Access key specifies the access key ID to access the Amazon account resources.
- The secret key is used to access the Amazon account resources. The value is associated with the access key and uniquely identifies the account.
- The folder path specifies the complete path to the Amazon S3 objects and includes the bucket name and any folder name.
- The master symmetric key provides a 256-bit AES encryption key in the Base 64 format when you enable client-side encryption.
- The code page specifies the code page compatible with the Amazon S3 source.
- The region name specifies the name of the region where the Amazon S3 bucket is available. You have also generated the customer master key ID.

Microsoft Azure Blob Storage V3 Connector

- Reads data from or writes data to Microsoft Azure Blob Storage
- Use the connector to specify sources or targets in a Mass Ingestion task, Mapping, and Mapping task.
- Configure the following connection properties:
 - Runtime Environment
 - Account Name
 - Account Key
 - Container Name
 - Endpoint Suffix

47

© Informatica. Proprietary and Confidential.



You can use Microsoft Azure Blob Storage V3 connector to read data from or write data to Microsoft Azure Blob Storage. You can use this connector to specify sources or targets in a Mass Ingestion task, Mapping, and Mapping task.

When you create this connection, you must configure few connection properties, such as Runtime Environment, Account Name, Account Key, Container Name, and Endpoint Suffix.

- The runtime environment specifies the name of the runtime environment where you want to run the tasks.
- The account name specifies the Microsoft Azure Blob Storage account name.
- The account key specifies the Microsoft Azure Blob Storage access key.
- The container name specifies the Microsoft Azure Blob Storage container name.
- The endpoint suffix specifies the type of Microsoft Azure end-points.

Hive Connector

- Use Hive connector to connect to Hive from Data Integration
- Use a Hive object as a source or target in mappings and mapping tasks
- Use Hive connector on Kerberos and non-Kerberos clusters
- Use Hive to read data from and write data to partitioned and bucketed tables

48

© Informatica. Proprietary and Confidential.



You can use a Hive connector to connect to Hive from Data Integration. You can use a Hive object as a source or target in mappings and mapping tasks. You can use Hive Connector on Kerberos and non-Kerberos clusters. You can read data from and write data to partitioned and bucketed tables in Hive.

Hive Connector on Kerberos and Non-Kerberos Clusters

- On Kerberos and non-Kerberos clusters, install the Secure Agent outside the cluster and perform a read or write operation
- Connects to Hive to perform relevant data operations
- Supports all operators supported in HiveQL
- Supports the AND conjunction in simple filters
- Supports the AND and OR conjunction in advanced filters
- Supports filtering on all filterable columns in Hive tables

49

© Informatica. Proprietary and Confidential.



On Kerberos and non-Kerberos clusters, you can install the Secure Agent outside the cluster and perform a read or write operation. The Hive connector connects to Hive to perform relevant data operations. The connector supports all operators supported in HiveQL. It supports the **AND** conjunction in simple filters and the **AND** and **OR** conjunctions in advanced filters. The Hive connector also supports filtering on all filterable columns in Hive tables.

Hive Connector Distributions

- Hive Connector supports the following distributions for a read or write operation

Kerberos Cluster	Non-Kerberos Cluster
Cloudera 5.8 to Cloudera 5.13	Cloudera 5.8 to Cloudera 5.13
Hortonworks 2.5 and Hortonworks 2.6	Hortonworks 2.5 and Hortonworks 2.6
HDInsight 3.6	HDInsight 3.6

50

© Informatica. Proprietary and Confidential.



Hive Connection Properties

- Authentication Type
 - select Kerberos for a Kerberos cluster
 - select LDAP for an LDAP-enabled cluster
 - select None for a cluster that is not secure or not LDAP-enabled
- JDBC URL: The JDBC URL to connect to Hive
- JDBC Driver: The JDBC driver class to connect to Hive

51

© Informatica. Proprietary and Confidential.



When you set up a Hive connection, you must configure the connection properties such as:

- **Authentication Type:** For a Kerberos cluster, you must select the authentication type as Kerberos. For an LDAP-enabled cluster, select the authentication type as LDAP. For a cluster that is not secure or not LDAP-enabled, select the authentication type as None.
- **JDBC URL:** This specifies the JDBC URL to connect to Hive.
- **JDBC Driver:** This specifies the JDBC driver class to connect to Hive.

Hive Connection Properties (continued)

- **Username:** The username to connect to Hive in LDAP or None mode
- **Password:** The password to connect to Hive in LDAP or None mode
- **Principal Name:** The principal name to connect to Hive through Kerberos authentication
- **Impersonation Name:** The user name of the user that the Secure Agent impersonates to run mappings on a Hadoop cluster
- **Keytab Location:** The path and file name to the Keytab file for Kerberos login

52

© Informatica. Proprietary and Confidential.



- **Username:** This specifies the user name to connect to Hive in LDAP or None mode.
- **Password:** This specifies the password to connect to Hive in LDAP or None mode.
- **Principal Name:** This specifies the principal name to connect to Hive through Kerberos authentication.
- **Impersonation Name:** This specifies the name of the user that the Secure Agent impersonates to run mappings on a Hadoop cluster. You can configure user impersonation to enable different users to run mappings or connect to Hive. The impersonation name is required for the Hadoop connection if the cluster uses Kerberos authentication.
- **Keytab Location:** This specifies the path and file name to the Keytab file for Kerberos login.

Hive Connection Properties (continued)

- Configuration Files Path: The directory that contains the Hadoop configuration files for the client
- NameNode URI: The URI to access HDFS
- HDFS Staging Directory: The staging directory in the cluster where the Secure Agent stages the data before it writes to the target
- Hive Staging Database: The Hive database where external or temporary tables are created

53

© Informatica. Proprietary and Confidential.



- **Configuration Files Path:** This specifies the directory that contains the Hadoop configuration files for the client.
- **Name Node URI:** This specifies the URI to access HDFS.
- **HDFS Staging Directory:** This specifies the staging directory in the cluster where the Secure Agent stages the data before it writes to the target. You must have full permissions for the HDFS staging directory.
- **Hive Staging Database:** This specifies the Hive database where external or temporary tables are created. You must have full permissions for the Hive staging database to create and insert data.

CDM Folders Connector

- Use CDM Folders Connector to connect to the Microsoft Azure Data Lake Storage Gen2 storage and Power BI from Data Integration
- Read data from or write data in the .csv file format to the common data model folder
- Create an external dataflow on Power BI workspace to access the data from the common data model folder
- Create a CDM Folders connection and use the connection in mappings or mapping tasks

You can use CDM Folders Connector to connect to the Microsoft Azure Data Lake Storage Gen2 storage and Power BI from Data Integration. You can use a CDM Folders Connector to read data from or write data in the dot csv file format to the common data model folder present in the Microsoft Azure Data Lake Storage Gen2 storage. You can also use CDM Folders Connector to create an external dataflow on Power BI workspace to access the data from the common data model folder in the Microsoft Azure Data Lake Storage Gen2 storage.

You can create a CDM Folders connection and use the connection in mappings or mapping tasks.

Snowflake Cloud Data Warehouse V2 Connector

- Use Snowflake Cloud Data Warehouse V2 Connector to connect to Snowflake from Data Integration
- Create a Snowflake Cloud Data Warehouse V2 connection and use the connection in mass ingestion tasks, mappings, or mapping tasks
- Use a mass ingestion task to transfer files from any source that mass ingestion task supports to a Snowflake target
- To write data from Microsoft Azure Blob Storage to Snowflake, specify the external stage location on Snowflake to load the files

55

© Informatica. Proprietary and Confidential.



You can use a Snowflake Cloud Data Warehouse V2 Connector to connect to Snowflake from Data Integration.

You can use this connector to securely read data from and write data to Snowflake.

You can create a Snowflake Cloud Data Warehouse V2 connection and use the connection in mass ingestion tasks, mappings, or mapping tasks.

You can use a mass ingestion task to transfer files from any source that mass ingestion task supports to a Snowflake target. When you configure a mass ingestion task to load files to a Snowflake target, you must specify the file format and the copy options for the data files.

To write data from sources such as Microsoft Azure Blob Storage to Snowflake, you must specify the external stage location on Snowflake to load the files. You can choose to specify an external stage location for Amazon S3 on Snowflake.

Zendesk V2 Connector

- Use Zendesk V2 Connector to connect to Zendesk from Data Integration
- Uses REST call to connect to Zendesk
- Create a Zendesk V2 connection and use the connection in synchronization tasks, mapping, and mapping tasks
- Secure Agent uses the Zendesk API to read data from and write data to Zendesk

56

© Informatica. Proprietary and Confidential.



You can use Zendesk V2 Connector to connect to Zendesk from Data Integration. You can use this connector to read data from and write data to Zendesk.

Zendesk V2 Connector uses REST call to connect to Zendesk. You can create a Zendesk V2 connection and use the connection in synchronization tasks, mapping, and mapping tasks.

When you run a synchronization task or a mapping task, the Secure Agent uses the Zendesk API to read data from and write data to Zendesk. You can use Zendesk objects, such as Users, Tickets, or Organizations in a task.

Using Add-On Connectors

- Add connector to IICS Org
- Review documentation on Community site

The screenshot shows the Informatica Administrator interface with the 'Add-On Connectors' page open. The left sidebar includes options like Organization, SAML Setup, Licenses, Users, User Groups, User Roles, Runtime Environments, Connections, Add-On Connectors (which is selected and highlighted in dark grey), Schedules, Add-On Bundles, Swagger Files, and Logos. The main content area is titled 'Add-On Connectors' with the sub-instruction 'Install connectors published by Informatica Cloud and our partners, and then use the connectors to create connections.' Below this, two connectors are listed:

- Adobe Cloud Platform**
Publisher: Informatica
Version: 1
Updated On: Jul 11, 2018
Description: The Informatica Cloud Connector for Adobe Cloud Platform is a high volume data connector enabling users to move data between any cloud or on-prem sources and the Adobe Cloud Platform.
Free Trial
- Amazon Dynamo DB**
Publisher: Informatica Cloud
Version: 9
Updated On: Dec 7, 2017
Description: This connector enables users to extract and load data from DynamoDB. Users can also create a DynamoDB target.
Free Trial

57

© Informatica. Proprietary and Confidential.

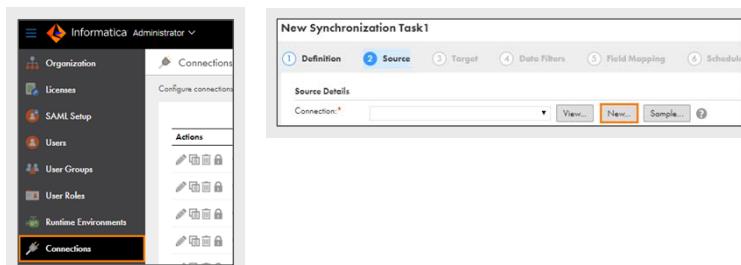


As mentioned earlier in the module, to use an add-on connector, you must first add the connector to your IICS Org.

You can visit the Informatica Community site to review the documentation for Add-on connectors.

Creating a Connection

- Create connection
 - from Administrator Service
 - on-the-fly via the wizard when configuring a task



- Connection becomes available to all users in the Org
- Secure agent must be installed in the Org before creating any connection

58

© Informatica. Proprietary and Confidential.



You can create connections in IICS by selecting “Connections” from the Administrator Service.

You can also create connections on-the-fly using the wizard when you configure a task. For example, you can create connections using the Synchronization task wizard.

When you create a connection, the connection becomes available to all the users in the Org.

Except for Software-as-a-Service (SaaS) applications, all connections depend on the Secure Agent. Therefore, you must first install the Secure Agent in your IICS Org before creating any connection.

Lab Activity

2-1 Creating a Salesforce Connection

In this lab, you will perform the following:

- Create a Salesforce connection

Lab Activity

2-2 Creating a Flat File Connection

In this lab, you will perform the following:

- Create a flat file connection

60

© Informatica. Proprietary and Confidential.



Lab Activity

2-3 Creating an Oracle connection

In this lab, you will perform the following:

- Create an oracle connection

Module Summary

This module showed you how to:

- Discuss Informatica Cloud runtime environments
- Explain the purpose of Informatica Cloud Secure Agent
- Explore the Secure Agent architecture
- View the Secure Agent log files
- List the steps to install the Secure Agent
- Define a connection
- Explore types of connectivity
- Discuss native and add-on connection types

62

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica[®]
University

Module 3 Synchronization Task



Module Objectives

After completing this module, you will be able to:

- Define Synchronization Task
- Describe Synchronization Task wizard
- Create a Synchronization Task
- Identify status of a Synchronization Task
- Discuss Activity Monitor and Activity Log

Synchronization Task Overview

- Synchronize the data between a source and a target
- Use case:

Read data from a flat file and write the data to Salesforce

Apply Filter on incoming data before writing it to the target

Transform data according to the business logic and update it on the target system

4

© Informatica. Proprietary and Confidential.



A Synchronization Task allows you to synchronize the data between a source and a target. In IICS, the supported source and target types include Database connection, Flat File connection, and Salesforce connection.

Some examples of business scenarios where you can use a synchronization task are:

- Reading data from a Flat File and writing the data to a Salesforce account
- Applying Filter on incoming data before writing it to the target
- Transforming the data according to the business logic and updating it on the target system

Topic

Synchronization Task Wizard



Step 1 – Definition Step

- Specify the Task Name, Location, Description, and the Task Operation
 - Task Name specifies the name of the synchronization task
 - Location specifies the location where you want to save the task
 - Description specifies a brief note about the task
 - Task Operation specifies the operation that the synchronization task performs
- Select one of the following task operations:
 - Insert
 - Update
 - Upsert
 - Delete

The screenshot shows the 'Definition' step of the Synchronization Task wizard. The 'Task Details' panel contains fields for 'Task Name' (set to 'Synchronization Task1'), 'Location' (set to 'Default'), and 'Description'. Below these, a dropdown menu for 'Task Operation' is shown. The top navigation bar includes tabs for 'Definition', 'Source', 'Target', 'Data Filters', 'Field Mapping', and 'Schedule'. The 'Definition' tab is currently selected.

© Informatica. Proprietary and Confidential.



The first step of the Synchronization task wizard is the Definition step. In this step, you must specify the task name, location, description, and the task operation.

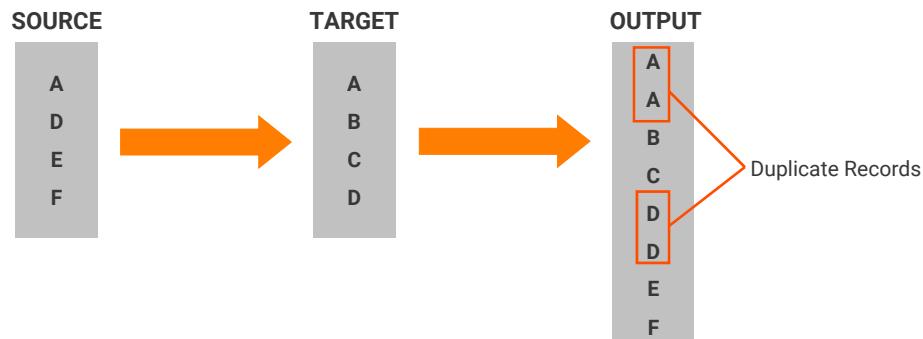
- **Task name** specifies the name of the synchronization task. It can contain alphanumeric characters, spaces, and some special characters such as underscore, dot, and hyphen.
- **Location** specifies the location where you want to save the task.
- **Description** specifies a brief note about the task.
- **Task operation** specifies the operation that the synchronization task performs. You can select one of the following task operations:
 - Insert
 - Update
 - Upsert
 - Delete

The list of available targets in a subsequent step depends on the operation that you select for the task.

Step 1 – Definition Step

Insert Operation

- Inserts all source rows into a target
- Ideal Use: one-time or initial data load
- Risk of duplicate records in the target



© Informatica. Proprietary and Confidential.



The first type of operation is the Insert operation. This operation inserts all rows from the source into the target. An ideal use case of the Insert operation in a synchronization task is to perform a one-time load of data into the target system. The Insert operation inserts all records in the source and does not take into account the existing records in the target. So, when you use the Insert operation, there is a risk of creating duplicate records in the target system.

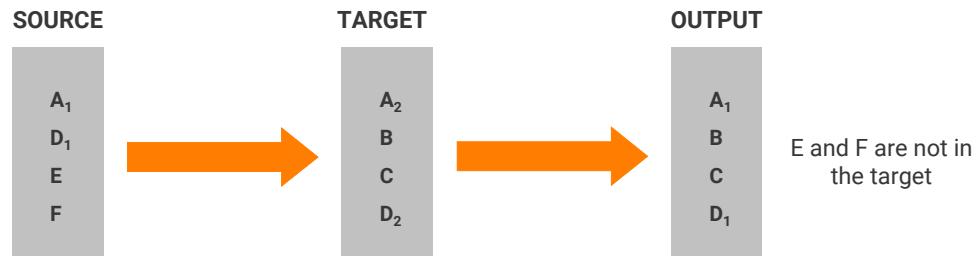
Example

The source contains records A, D, E, and F. The target contains records A, B, C, and D. When you run a synchronization task with an Insert operation, duplicate records are created for A and D, as these two records already exist in the target.

Step 1 – Definition Step

Update Operation

- Updates rows in the target that exist in the source
- Does not update source rows that do not exist in the target



8

© Informatica. Proprietary and Confidential.



The Update operation updates only those rows in the target that exist in the source. If there is a row in the source that does not exist in the target, then that row gives an error when the task runs. However, rest of the task continues to run.

Example

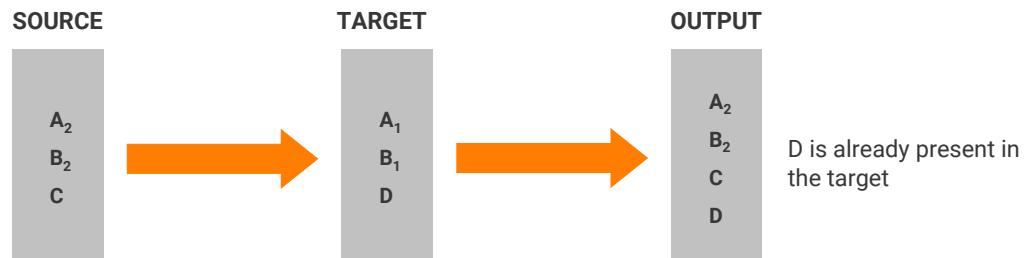
The source contains version one of record A, version one of record D, along with records E and F. The target contains version two of record A, version two of record D, and records B and C.

When you run a synchronization task with an Update operation, the task updates records to A₁ and D₁. The records E and F return an error because they do not exist in the target. Also, note that the records B and C remain unchanged as they do not exist in the source.

Step 1 – Definition Step

Upsert Operation

- Updates existing records and insert new records in the target
- Does not support flat file target



9

© Informatica. Proprietary and Confidential.



The Upsert operation allows you to use a single task to update existing records and insert new records in the target. Ideally, you must use the Upsert operation to synchronize data between two systems.

The Upsert operation does not support a flat file target.

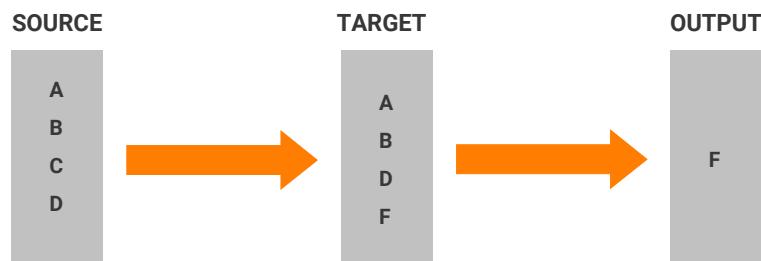
Example

The source contains version two of record A, version two of record B, and another record C. The target system contains records A₁, B₁, and D. When you run a synchronization task with an Upsert operation, the task updates the records to A₂, B₂, and C in the target system.

Step 1 – Definition Step

Delete Operation

- Deletes all rows from target that exists in the source



10 © Informatica. Proprietary and Confidential.



The Delete operation allows you to delete all rows from the target that exists in the source.

Example

The source system contains records A, B, C, and D. The target contains records A, B, D, and F. When you run a synchronization task with the Delete operation, the task deletes the records A, B, and D from the target.

Step 2 – Source Step

- The source for a synchronization task can be:

Single Flat File

Database Table

Salesforce Object

Saved Query

- Single object
 - perform operation on a single source object
- Multiple object
 - configure multiple database tables or Salesforce objects as the source
- Saved Query
 - create the Saved Query component
 - create a Saved Query from one or more database tables
 - enter a valid SQL SELECT statement to select the columns you want to use in the task

11

© Informatica. Proprietary and Confidential.



Step 2 of the synchronization task wizard allows you to define the source information. The source for a synchronization task can be a single flat file, a database table, a Salesforce object, or a customer source object, also called a Saved Query.

You can select the source type to be a Single object, Multiple object, or Saved Query. When you specify the source type as single, you can perform operation on a single source object. For multiple source types, you can configure multiple database tables or Salesforce objects as the source for the synchronization task.

To use a Saved Query in a Synchronization Task, you must first create the Saved Query component. You can create a Saved Query from one or more database tables. To create a Saved Query, you must enter a valid SQL SELECT statement to select the columns you want to use in the task. The data integration then uses that SQL statement to retrieve the information from the source. You can edit the data type, precision, or scale of each column before you save the Saved Query.

Step 3 – Target Step

- The target for a synchronization task can be:

Single Flat File

Database Table

Salesforce Object

- Target connections that you can use also depend on the task operation you select
 - Example: If you select the Upsert operation, you cannot use a flat file target connection
- Target Object can be an existing object, or you can create a new target at runtime

12

© Informatica. Proprietary and Confidential.



Step 3 of the synchronization task wizard is to set up the target information. For a synchronization task, you can write data to a single flat file, a database table, or a Salesforce object. The target connections that you can use also depend on the task operation you select. For example, if you select the Upsert operation, you cannot use a flat file target connection because you cannot Upsert records into a flat file target.

After you select the target connection, you must select a target object. You can use an existing object, or you can create a new target at runtime. You can create a new target at runtime only for Flat File and relational database connections.

Step 4 – Data Filters Step

- Limits the data retrieval from the source
- You can apply two types of data filters – simple and advanced
- Use a simple data filter when all the filter conditions can be joined together using the AND operator
- Use advanced data filter when your source type is a flat file

13

© Informatica. Proprietary and Confidential.



Step 4 of the synchronization task wizard is the Data Filters step. A data filter allows you to limit the data that you retrieve from the source. The data filters act as a WHERE clause of the query that retrieves records from the source. For a synchronization task, you can apply two types of data filters - simple and advanced. The filter type is determined based on the source connection type and the filter condition.

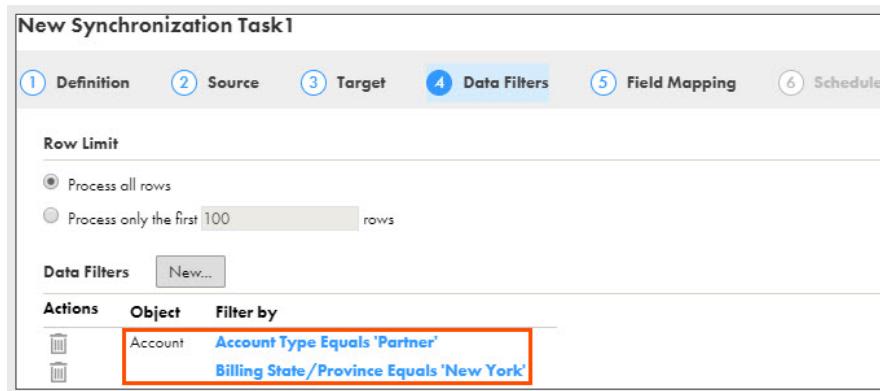
You can use a simple data filter when all the filter conditions can be joined together using the AND operator. Therefore, only those records that meet all the filter conditions will be passed through to the target.

You must note that you can use a simple data filter only if your source type is not a flat file. If your source type is a flat file, then you must use the advanced data filter.

Step 4 – Data Filters Step

Simple Filters – Example

- Source: Salesforce Accounts
- Scenario: Only load Partner accounts in New York



14

© Informatica. Proprietary and Confidential.



In this scenario, you want to load records from the Salesforce Account object. However, you must ensure that the Account Type is Partner, and the Billing State is New York. To meet these conditions, you must create two simple data filters as shown in the image.

The two filter conditions are joined together using the WHERE clause along with the AND operator.

Step 4 – Data Filters Step

Advanced Filters

- Use an advanced data filter:
 - when the filter conditions are complex
 - when the connection is a flat file connection
 - when it is necessary to use the OR operator in the data filter
- In an advanced filter, one expression contains all the filter conditions

15

© Informatica. Proprietary and Confidential.



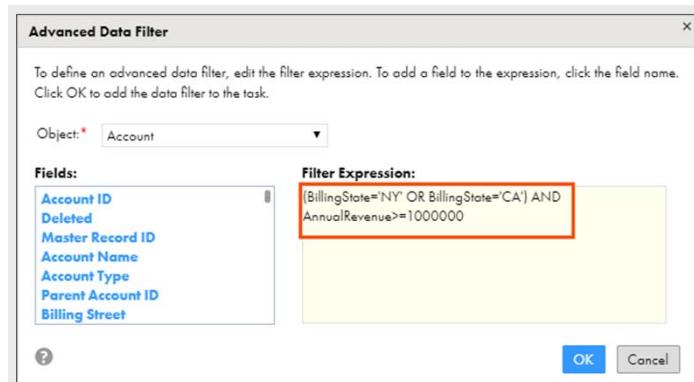
You can use an advanced data filter when the filter conditions are complex, when the connection is a flat file connection, or when it is necessary to use the OR operator in the data filter.

One of the main differences between a simple filter and an advanced filter is that, in an advanced filter, one expression contains all the filter conditions.

Step 4 – Data Filters Step

Advanced Filters – Example

- Source: Salesforce Accounts
- Scenario: Only load accounts where Billing State is either New York or California, and Annual Revenue is greater than or equal to 1,000,000



16

© Informatica. Proprietary and Confidential.



Consider an example of an advanced data filter. Here, you want to load records from the Salesforce Account object. However, you want to load only those accounts where the Billing State is either New York or California and the Annual Revenue is greater than or equal to one million dollars.

You can meet these conditions by creating an advanced data filter as shown on your screen.

Step 4 – Data Filters Step

Data Filter Variables

- You can use System Variables to filter new or modified records
- IICS provides access to the following System Variables:
 - \$LastRunDate
 - \$LastRunTime
 - \$ErrorFileName
 - \$SuccessFileName

17

© Informatica. Proprietary and Confidential.



You can use the data filter variables to filter newly inserted records. For Upsert operation, you can use the system variables to filter records that have changed after a successful task run.

IICS provides access to the following system variables:

- **\$LastRunDate** returns the last date on which the task ran successfully.
- **\$LastRunTime** returns the last time when the task ran successfully.
- **\$ErrorFileName** returns the name of the error file that gets generated.
- **\$SuccessFileName** returns the name of the success file that gets generated.

Step 5 – Field Mapping Step

Automatch and Clear Mapping

- Define the mapping between the source fields and the target fields
- The Automatch feature matches the source field with the target field in the following ways:
 - the Exact Field name option matches fields with the exact same name
 - the Smart Match option matches fields with similar names
- Use the Clear Mapping option to clear or rematch all the field mappings

The screenshot shows the 'Field Mapping' step of the synchronization task wizard. It displays a mapping grid where source fields from 'Source: Employee.csv' are mapped to target fields in 'Target: Contact'. The grid includes columns for 'Status', 'Name', 'Actions', and 'Expression/Lookup'. The 'Actions' column contains various mapping expressions, such as 'fx [] Emp_ID' and 'fx [] lookup[Outlet_Name = Name] Account.Id'. The 'Expression/Lookup' column provides detailed logic for the mappings. The 'Automatch' dropdown menu is open, showing options like 'Exact Field Name' and 'Smart Match'. Buttons for 'Clear Mapping', 'Validate Mapping', 'Edit Types...', and 'Related Objects...' are also visible.

Step 5 of the synchronization task wizard is the Field Mapping step. To complete the synchronization task configuration, you must define the mapping between the source fields and the target fields.

The Automatch feature matches the source field with the target field in the following ways:

- The Exact Field name option matches fields with the exact same name
- The Smart Match option matches fields with similar names

An example of Smart Match is when you have a source field "Cust_Name" and a target field "Customer_Name". The Smart Match function automatically links the "Cust_Name" field with the "Customer_Name" field.

If you want to clear or rematch all the field mappings, you can use the Clear Mapping option available in the field mapping step.

Step 5 – Field Mapping Step

Field Properties

- The following icons identify the field properties:

Primary Key



Indicates that field is primary key for object

External ID



Indicates that field is external ID for object (Salesforce.com only)

Non-null



Indicates that field cannot contain a null value

19

© Informatica. Proprietary and Confidential.



The golden key icon indicates that the field is the primary key for the object.

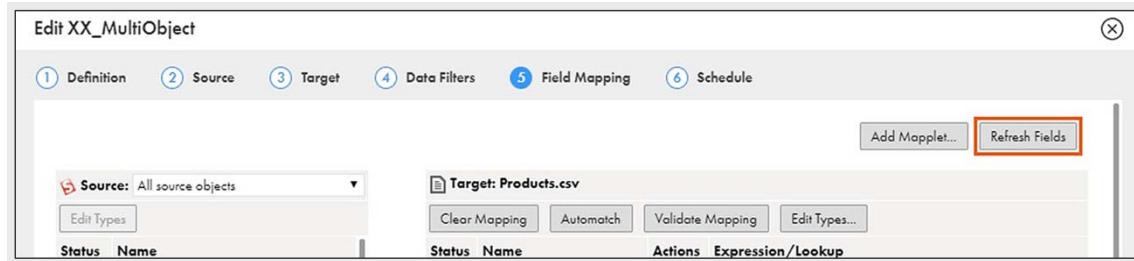
The white key in a golden box indicates that the field is the external ID field for the object. The External ID field applies only to Salesforce objects.

The white star in a grey circle indicates that the field cannot contain null values. Sometimes this field is automatically populated by Salesforce.

Step 5 – Field Mapping Step

Refresh Fields

- When you create a synchronization task, the Data Integration service stores field metadata for all the source and target fields
- When you change the source or target of an existing task, you must use the Refresh Fields option



20

© Informatica. Proprietary and Confidential.



Another important feature of the field mapping step is “Refresh Fields”. When you create a synchronization task, the Data Integration service stores field metadata for all the source and target fields. When you change the source or target of an existing task, you must use the “Refresh Fields” option to update the cache and view the latest field attributes.

Step 5 – Field Mapping Step

Edit Types and Validate Mapping

- Map the fields with compatible data types
- To configure or edit the field data types, IICS provides the Edit type option
- Edit type option is not available for all target types
- After you map the source and target fields, you must always validate the mapping

21

© Informatica. Proprietary and Confidential.



When you map the source fields with target fields, you must map the fields with compatible data types. To configure or edit the field data types, IICS provides the Edit type option.

You must note that the Edit type option is not available for all target types. If the synchronization task contains multiple sources or targets, you must first select the source or target you want to edit and then edit the data type. After you map the source and target fields, you must always validate the mapping.

Step 5 – Field Mapping Step

Field Expression

- Transforms source data before loading to target
- You can use field expressions in the following scenarios:
 - mapping multiple source fields to a single target field
 - converting data values (to/from date and string)
 - performing data clean-up (trimming leading or trailing blank spaces and removing unnecessary characters)

22

© Informatica. Proprietary and Confidential.



You can use Field expressions to transform the source data before you load it to the target.

A few instances where you can use field expressions are:

- when you want to map multiple source fields to a single target field
- when you want to convert data values
- when you want to perform data clean-up

For example, trimming leading or trailing blank spaces, or removing unnecessary characters from data, and so on.

Step 5 – Field Mapping Step

Field Expression – Example

- When you drag and drop two source fields onto a single target field, IICS automatically writes the field expression



23

© Informatica. Proprietary and Confidential.



When you drag and drop two source fields onto a single target field, IICS automatically writes the field expression.

In this example, when you drag phone and area code fields from the source and drop it onto the Phone Number field in the target, IICS automatically writes the field expression.

You can see that IICS writes the expression to concatenate Area Code and Phone number. You can also edit this expression to perform additional formatting on the source fields, like adding space or adding parentheses around the area code, and so on.

Step 5 – Field Mapping Step

Field Lookup

- Use the Field Lookup option to look up some fields for integration purposes
- Field Lookup allows you to retrieve information from any lookup connection
- Field lookup retrieves the data based on a lookup condition

24

© Informatica. Proprietary and Confidential.



When you synchronize data between a source and a target, you can use the Field Lookup option to look up some fields required for integration.

The Field Lookup feature allows you to retrieve information from any lookup connection such as Salesforce object, database table, or a flat file connection. The Field Lookup retrieves the data, based on the lookup condition defined for the task. You can use this feature when you have missing or inaccurate data.

Step 5 – Field Mapping Step

Field Lookup – Example

Name	Address	City	State	Zip
DHL	34 West Street	Mckinney		75070

1	A	B	C	D	E
	Zip	lat	long	City	State County
32899	75063	32.767268	-96.777626	IRVING	TX DALLAS
32900	75065	33.131089	-97.008563	LAKE DALLAS	TX DENTON
32901	75067	33.047871	-97.050897	LEWISVILLE	TX DENTON
32902	75068	33.17679	-96.958293	LITTLE ELM	TX DENTON
32903	75069	33.141438	-96.588295	MC KINNEY	TX COLLIN
32904	75070	33.230381	-96.627018	MC KINNEY	TX COLLIN
32905	75074	33.109044	-96.578819	PLANO	TX COLLIN
32906	75075	33.162417	-96.71546	PLANO	TX COLLIN

Name	Address	City	State	zip
DHL	34 West Street	Mckinney	TX	75070

25

© Informatica. Proprietary and Confidential.



Consider an example where you can use “field lookup” to update the missing information.

Observe the sample data from the source. Notice that the state field contains no value. However, you have the zip code value available in the source.

Now, if you have a resource, such as a table or a file of zip codes and their associated states, you can perform a lookup using the value in the Zip field.

When the Synchronization task writes the record to the target, it inserts the return value that is defined in the resource. So, in this case, the value “Texas” is written to the “State” field in the target.

Rules and Guidelines for Lookup

- If the lookup is on a flat file:
 - must be comma-delimited
 - make sure Secure Agent has access to flat file directory (Windows OS)
- Minimize number of lookups per task
- Source field and lookup field in the lookup condition must have compatible data types

When the lookup object is a flat file, then the file must be a comma-delimited flat file. You must also ensure that the Secure Agent has access to the flat file directory.

To improve processing efficiency, you must minimize the number of lookups per task.

Finally, you must ensure that the source field and lookup field in the lookup condition have compatible data types.

Field Lookup Configuration

Step 1

Select lookup connection and object

Step 2

Select lookup fields

Step 3

Select lookup return value

27

© Informatica. Proprietary and Confidential.



Configuring a Field lookup is a three-step process.

Step 1: You must select a connection and an object to perform the lookup. The lookup connection does not need to be the same as the source or the target connection for the task.

Step 2: You must specify the lookup to compare.

Step 3: You must select the return value. You get the return value when the lookup finds a match in the lookup object.

When you configure a field lookup, you must also configure how the Data Integration service handles multiple matching return values. The Data Integration service can randomly choose a matching value or return an error. When the lookup returns an error, the Data Integration service writes the row to the error rows file.

Step 6 – Schedule Step



Allows you to run tasks at a specific time or at regular intervals

Create a schedule from the Administrator service or from the Schedule step of the synchronization task wizard

- To create a schedule, enter the following details:
 - Schedule Name
 - Starts
 - Time Zone
 - Repeat Frequency

28

© Informatica. Proprietary and Confidential.



The last step of the synchronization task wizard is the Schedule step.

A schedule allows you to run tasks at a specific time or at regular intervals. You can create a schedule from the Administrator service or from the Schedule step of the synchronization task wizard.

To create a schedule, enter the following details:

- **Schedule Name:** This field specifies the name of the schedule.
- **Starts:** This field sets the date and time for the schedule to start.
- **Time Zone:** This field sets the time zone for the schedule. The time zone can differ from the organization time zone or user time zone.
- **Repeat Frequency:** The Repeat frequency field determines how often the tasks run. You can choose the repeat frequency for a schedule as - Does not repeat, Every N minutes, Hourly, Daily, Weekly, and Monthly.

Step 6 – Schedule Step

- Email notifications allow you to monitor the status of the task
- When you configure email notifications at the Org level, the notification is applicable to all tasks in the Org
- When you configure email notifications at the individual task level, the notification is applicable only to that individual task
- Can use SQL commands to perform database level tasks
 - pre-processing commands
 - post-processing commands
- Can use Operating System commands to perform Operating System level tasks

29

© Informatica. Proprietary and Confidential.



The schedule step also allows you to configure email notification options for the task. Email notifications allow you to monitor the status of the task. You can configure email notifications at the Org level or at individual task level.

When you configure email notifications at the Org level, the notification is applicable to all tasks in the Org. When you configure email notifications at the individual task level, the notification is applicable only to that individual task.

In the advanced options section of the schedule step, you can also use SQL commands to perform database level tasks. The task runs pre-processing commands before it reads the data from the source. The task can also run post-processing commands after it writes the data to the target. You can also use Operating System commands to perform Operating System level tasks. You must note that if the pre or post-processing commands fail, the synchronization task also fails.

Topic

Activity Monitor



Activity Monitor

- Displays the status of the task
- Shows current and past tasks in the org
- You can view job details and download a session log while a job is running
- You use the Activity Monitor to stop a running task, restart a previous task, or to refresh the status of currently running task

Instance Name	Location	Subtasks	Start Time	End Time	Rows Processed	State
XX_FirstName_Openweathermap_34149762969-1	Default	2 Tasks	Aug 1, 2019, 3:05...	Aug 1, 2019,...	View Subtasks	Success
XX_FirstName_PhoneNumber_Masking-1	Default	1 Tasks	Aug 1, 2019, 5:13...	Aug 1, 2019,...	964	Success
XX_FirstName_PhoneNumber_Masking-2	Default	1 Tasks	Aug 1, 2019, 12:4...	Aug 1, 2019,...	0	Failed
XX_FirstName_PKChunking-1	Default	1 Tasks	Aug 1, 2019, 2:29...	Aug 1, 2019,...	48	Success
XX_FirstName_Proj-1	Default	1 Tasks	Aug 1, 2019, 11:1...	Aug 1, 2019,...	10	Success
XX_FirstName_Pushdown-1	Default	1 Tasks	Aug 1, 2019, 4:51...	Aug 1, 2019,...	0	Failed

31

© Informatica. Proprietary and Confidential.



When you start a task, the Activity Monitor displays the status of the task. The Activity Monitor also displays the status of previously run and currently running tasks in your Org.

You can view job details and download a session log while a job is running so that you can more easily monitor long running jobs.

You can use the Activity Monitor to stop a running task, restart a previous task, or to refresh the status of a currently running task.

It is important to note that there is no rollback feature. So, if you stop a task that has already processed some rows, then that data remains in the target.

Activity Monitor

Task Status

Starting	Job is starting
Running	Job is in progress
Success	Job ran successfully without errors
Failed	<ul style="list-style-type: none">An error caused the job to failNo rows were moved from source to target
Warning	<ul style="list-style-type: none">Job completed but some rows failedReview the Error Rows file to identify failed rows and review error messages

32 © Informatica. Proprietary and Confidential.

 Informatica
University

A synchronization task can have one of the following status: **Starting**, **Running**, **Success**, **Failed**, and **Warning**.

- **Starting** indicates that the job is starting.
- **Running** indicates that the job is in progress.
- **Success** indicates that the task completed without any errors.
- **Failed** indicates that the job failed due to an error. For failed status, no data is written to the target system.
- **Warning** indicates that the task completed but with errors. When the status of the task is warning, IICS generates an error rows file. You must review the file to analyze the failed rows and the corresponding error messages.

Lab Activity

3-1 Creating a Synchronization Task

In this lab, you will perform the following:

- Create and configure a synchronization task
- Run the task and validate the results in Salesforce

Lab Activity

3-2 Using Filter, Expression, and Lookup in a Synchronization Task

In this lab, you will perform the following:

- Create data filter
- Create field expressions
- Use a lookup to relate outlet name and account name

Lab Activity

3-3 Creating a Synchronization Task with Multiple Object Source Types

In this lab, you will perform the following:

- Create a Synchronization task to load data from multiple Salesforce objects into a Flat File

35

© Informatica. Proprietary and Confidential.



Lab Activity

3-4 Using Pre and Post SQL Commands in a Synchronization Task

In this lab, you will perform the following:

- Use pre and post SQL commands in a Synchronization task

Module Summary

This module showed you how to:

- Define Synchronization Task
- Describe Synchronization Task wizard
- Create a Synchronization Task
- Identify status of a Synchronization Task
- Discuss Activity Monitor and Activity Log

37

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 4

Cloud Mapping Designer – Basic Transformations



Module Objectives

After completing this module, you will be able to:

- Discuss Cloud Mapping Designer
- List the mapping designer terminologies
- Define mappings
- Discuss basic transformations in the Cloud Mapping Designer
- Explain field rules
- State best practices for creating mappings

Topic

Cloud Mapping Designer



Overview of Cloud Mapping Designer

Allows creation of flexible mappings to address more advanced use cases such as:

perform aggregations

create additional fields using an expression

use a lookup to return multiple values

execute the logic in a specific order

Parameterize elements of the mapping and create reusable mappings or templates

5

© Informatica. Proprietary and Confidential.



The Cloud Mapping Designer is a simple web-based interface that allows you to create end-to-end mappings. You can use the Cloud Mapping Designer to handle specific use cases that cannot be handled using the Synchronization task. The Cloud Mapping Designer allows you to perform transformation logic on the data. For example, you can perform aggregations, create additional fields using an expression, or use a lookup to return multiple values. The Cloud Mapping Designer can also execute the logic in a specific order. For example, you can filter data using the return value from a lookup.

The Cloud Mapping Designer allows you to create a mapping that includes multiple sources or targets. It also allows you to join data from heterogeneous sources. For example, you can join the data from a Sequel Server database table with the data from Salesforce. You can also write data to multiple targets.

The Cloud Mapping Designer also enables you to parameterize elements of the mapping, which allows you to create reusable mappings or templates. You can parameterize all aspects of the mapping, including sources and targets, filter criteria, lookup conditions, and so on.

Mapping Designer Terminologies

Term	Description
Canvas	Area where you can create a mapping
Shapes	Represent sources, targets, and data transformations
Mapping	Defines end-to-end integration that you create using Mapping Designer
Parameterized Mapping	Mapping that contains one or more parameters
Mapping Task	Allows you to configure a mapping
Deploy	Apply changes made to a mapping across all dependent mapping configuration tasks

6

© Informatica. Proprietary and Confidential.



Canvas refers to the area where you create a mapping.

Shapes represents sources, targets, and data transformations. You can drag and drop shapes on to the canvas.

Mapping defines the end-to-end integration that you create using the Mapping Designer.

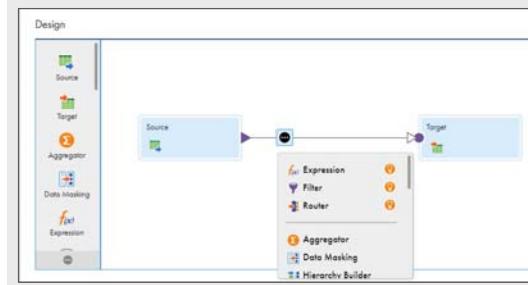
Parameterized Mapping refers to a mapping that contains one or more parameters.

Mapping Task is an automatically generated wizard that allows you to configure a mapping, including schedules and parameter value assignments.

When you make changes to a mapping, you can promote the changes to any dependent mapping task by deploying them.

CLAIRE Transformation Recommendations

- Enable Informatica's AI engine CLAIRE for transformation recommendations during mapping design
- Click the **Add Transformation** icon to add transformations to a mapping directly on the mapping canvas
- The **Add Transformation** icon appears when you hover over the link between transformations or when you select an unconnected transformation



What are CLAIRE transformation recommendations?

You can enable Informatica's AI engine CLAIRE for transformation recommendations during mapping design. The CLAIRE engine uses metadata from IICS organizations to recommend transformations to include in a mapping flow.

You can click the **Add Transformation** icon to add transformations to a mapping directly onto the mapping canvas. The **Add Transformation** icon appears when you hover over the link between transformations or when you select an unconnected transformation.

So, if your organization has CLAIRE recommendations enabled, you can see recommended transformations in the Add Transformation menu.

The image shows the Add Transformation icon and the Add Transformation menu.

Mapping

- Defines the flow of data from the source to the target
- Add transformation shapes to perform data transformation tasks
- Define rules for fields that are part of the transformations
- Links visually represent the flow of data in a mapping



8

© Informatica. Proprietary and Confidential.



A mapping is an object that you create in the Cloud Mapping Designer. It defines the flow of data from the source to the target.

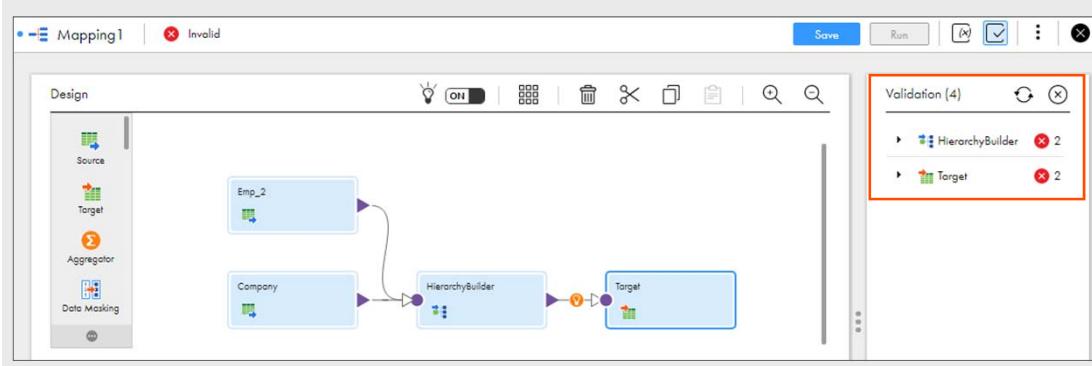
To create a mapping, you must add transformation shapes to the mapping flow. The transformations perform various data transformation tasks.

You can also define rules for fields that come into the transformations. The rules are based on certain criteria and allows you to create flexible mappings. For example, you can choose to include or exclude fields of a certain data type. This allows the mapping to remain valid even when you add new fields to the source.

The links in the mapping visually represent the flow of data in the mapping.

Validating a Mapping

- When you save a mapping, the Mapping Designer validates the mapping
- The status of the mapping can be valid or invalid
- Use the Validation panel to view the location and details of mapping errors



When you save the mapping, the Mapping Designer automatically validates the mapping. The status of the mapping is displayed in the header of the Mapping Designer. The status can be either valid or invalid.

If the mapping is invalid, you can use the Validation panel to view the location and details of the errors. The Validation panel displays a list of transformations in the mapping. An error icon is displayed next to the transformation that includes errors.

As you can see in the image, the sample mapping has three errors in the Target Transformation.

Running a Mapping

- When the mapping is valid, you can perform a test run to verify the results of the mapping
- Test run creates a temporary mapping task
- Data Integration service deletes the temporary mapping task after the test run is complete

10

© Informatica. Proprietary and Confidential.



After you check the validity of a mapping, and ensure that it is free of errors, you can perform a test run to verify the results of the mapping.

In the test run, you run a temporary mapping task. The task reads data from the source, writes data to the target, and performs all transformation logics in the data flow. The Data Integration service deletes the temporary task after the test run is complete.

Mapping Lifecycle

1

Design a Mapping

- Must contain at least one source and one target
- Optionally create parameters for connections, objects, and transformations

2

Run and Test the Mapping

- Verify that mapping works
- Enter values for parameters

3

Create Mapping Tasks

- Run on schedule or as part of task flow
- Enter values for parameters

11

© Informatica. Proprietary and Confidential.



There are many stages in the lifecycle of a mapping.

In the first stage, you design a mapping using the Mapping Designer. For a mapping to be valid, it must contain at least one source and one target. In the mapping, you can optionally create parameters for connections, objects, transformations, filters, lookup conditions, and so on.

In the second stage, run and test the mapping to verify that the mapping works. If you have parameters in the mapping, you must specify a value for each parameter.

In the final stage of the mapping lifecycle, create mapping tasks. You can configure the mapping tasks to run on a schedule or use the mapping tasks in a taskflow. You can also include pre or post-processing commands in a mapping task. If the mapping contains parameters, you must specify values for the parameters in the mapping tasks.

Topic

Basic Transformations



What is a Transformation?

- A mapping object that modifies or passes data
- Can be active or passive
- Active transformation changes the number of rows that pass through it
- Passive transformation does not change the number of rows that pass through it

13

© Informatica. Proprietary and Confidential.



A transformation is a mapping object that modifies the data in the mapping or passes it on to the next step of the mapping. A transformation can either be active or passive.

An active transformation changes the number of rows that pass through it. For example, a Filter transformation is an active transformation because it passes data only when the criteria specified in the filter condition is fulfilled.

A passive transformation does not change the number of rows that pass through it. For example, an Expression transformation is a passive transformation because it transforms the data based on the expression that you specify. However, it does not change the number of rows that pass through it.

Source Transformation

- Active transformation
- Reads data from a source
- Defines connection and object
- Configure advanced options based on the connection type
- Add or remove fields from source

14

© Informatica. Proprietary and Confidential.



The Source transformation is an active transformation that reads data from a source. You can add one or more Source transformations to a mapping.

When you configure a Source transformation, you must specify a connection and an object. You can also configure advanced options based on the connection type. For example, when you select a Salesforce connection, you can use multiple related source objects and configure the Salesforce API advanced source option.

You can also add or remove fields from the Source transformation.

Target Transformation

- Defines the target connection and object for the mapping
- Specify whether you want to use the Insert, Update, Upsert, or Delete operation
- Configure advanced options based on the connection type

15

© Informatica. Proprietary and Confidential.



The Target transformation defines the target connection and object for the mapping. When you configure the target, you must specify whether you want to use the Insert, Update, Upsert, or Delete operation. You must also map incoming fields to the target fields.

In the Target transformation, you can configure advanced options based on the connection type. For example, when you select a Salesforce connection, you can configure options for success and error log details.

Filter Transformation

- Filters data based on the filter condition
- Place the Filter transformation close to the mapping sources
- Returns either a TRUE or FALSE value
- Can include multiple conditions
- Can only link a single transformation to the Filter transformation

16

© Informatica. Proprietary and Confidential.



The Filter transformation filters data based on the filter condition that you define. To improve job performance, you must place the Filter transformation close to the mapping sources. This way, you can remove unnecessary data from the data flow.

A filter condition is an expression that returns either a TRUE or a FALSE value. When the filter condition returns a TRUE value for a row, the Filter transformation passes the row to the rest of the data flow. When the filter condition returns a FALSE value, the Filter transformation drops the row.

You can filter data based on one or more conditions. For example, to work with data within a specified date range, you can create conditions to remove data before and after the specified dates.

You must note that you can only link a single transformation to the Filter transformation. You cannot merge multiple transformations into it.

Filter Condition

- Expression that returns either a TRUE or FALSE value
- Create one or more simple filter conditions
- A simple filter condition includes a field name, an operator, and a value
- Can use the following operators:
 - = (equals)
 - < (less than)
 - > (greater than)
 - <= (less than or equal to)
 - >= (greater than or equal to)
 - != (not equals)
- Use advanced filter condition to define complex expression

17

© Informatica. Proprietary and Confidential.



You have just seen that the filter condition is an expression that returns either a TRUE or a FALSE value. You can create one or more simple filter conditions. A simple filter condition includes a field name, an operator, and a value. You must note that filter conditions are case sensitive.

A simple filter includes operators such as – equals, less than, greater than, less than or equal to, greater than or equal to, and not equals.

When you define multiple simple filter conditions in a mapping task, the task evaluates the conditions in the order that you specify. The task evaluates the filter conditions using the AND logical operator and returns rows that match all the filter conditions.

You can also use an advanced filter condition to define a complex expression. When you configure a complex expression, you can incorporate multiple conditions using the AND or the OR logical operators.

Joiner Transformation

Joins two related heterogeneous sources

Joins data based on a join condition

Performs inner join operation

18

© Informatica. Proprietary and Confidential.



A Joiner transformation joins two related heterogeneous sources. The sources can reside in different source systems. For example, you can join data from the Salesforce Account object with the data in the customer's database table.

The Joiner transformation joins data based on a join condition that you define. You can also create multiple join conditions.

The Joiner transformation performs an inner join, which results in rows from both sources that match all the conditions. Source rows that do not match the join conditions are dropped from the data flow.

Joiner Transformation (continued)

- Select Master or Detail group to connect a source to Joiner transformation
- Connect source with smaller data set to Master group
- Use multiple Joiner transformations to join more than two sources in a mapping
- To avoid field name conflicts:
 - rename matching fields in upstream transformation
 - rename matching fields in Source transformation
 - pass data through an Expression transformation and rename fields

19

© Informatica. Proprietary and Confidential.



In a Joiner transformation, there are two sources used for joins. These two sources are called, Master Source and Detail Source.

In the properties of a Joiner transformation, you can select which data source can be Master and which source can be Detail source. When the task runs, the Master Source is cached into the memory for joining purpose. So, it is recommended to select the source with a smaller number of records as the Master Source.

You can join only two sources at a time to a Joiner transformation. If you want to connect more than two sources in the mapping, you must use multiple Joiner transformations.

Sometimes, when you join sources with matching field names, there can be a field name conflict. To avoid these conflicts, you must:

- Rename matching fields in an upstream transformation
- Rename matching fields in the Source transformation, and
- Pass data through an Expression transformation and rename fields

Expression Transformation

- Allows you to create new fields within the mapping
- Allows you to create an expression field and a variable field
- Expression field defines the calculations that you perform on an incoming field and also acts as the output field for results
- Variable field holds a variable in the expression or in other Expression transformations within the mapping
- Use an Expression transformation to:
 - perform non-aggregate calculations
 - concatenate or split incoming field values
 - insert a hard coded value into a field

20

© Informatica. Proprietary and Confidential.



An Expression transformation allows you to create new fields within the mapping. When you create a new field, you must specify the field name, type, precision, and scale. The Expression transformation allows you to create an expression field and a variable field.

An expression field defines the calculations that you perform on an incoming field and also acts as the output field for results. You can then use the field in the data flow. You can use multiple expression fields to perform calculations on incoming fields.

A variable field holds a variable in the expression or in other Expression transformations within the mapping. A variable field is not available for use downstream in the data flow.

You can use an Expression transformation to perform non-aggregate calculations, concatenate or split incoming field values, or insert a hard-coded value into a field.

Lookup Transformation

- Retrieves data from a lookup object based on a condition
- Configure the following properties:
 - Select lookup connection and object
 - Specify how the lookup handles multiple matching rows
 - Configure lookup condition
 - Specify return fields

21

© Informatica. Proprietary and Confidential.



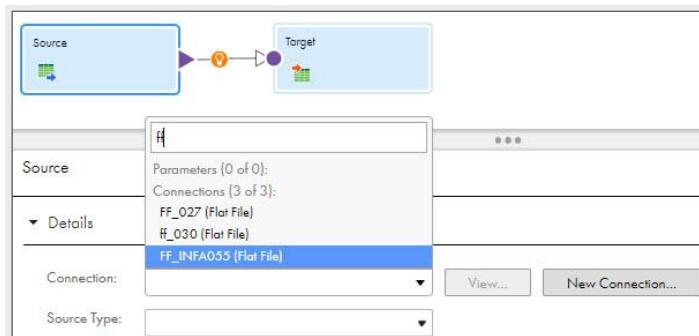
A Lookup transformation retrieves the data from a lookup object based on a condition that you define. When you configure a Lookup transformation, there are several options that you must configure.

You must select the lookup connection and object. Then, you must specify how the lookup must behave if multiple matches are found. For example, you can choose to return any row that matches, return the first row, return the last row, or report an error, and so on.

You must also configure the lookup condition and specify how incoming rows must match with rows in the lookup object. Finally, you must also specify the fields that you want to return.

Selection List Search in Transformation

- Enter a search string in the selection list to search for a particular connection



22

© Informatica. Proprietary and Confidential.



When you select a connection or data field in a transformation, you can enter a search string in the selection list. You can search by name or type.

For example, when you configure a connection that contains "ff" in a Source transformation, you can enter "ff" in the connection list to find all the connections that include that string.

The example displayed on the slide shows a search in the Connection field for a Source transformation.

Topic Field Rules



Field Rules

Define how data enters a transformation from the upstream transformation

By default, includes all fields in the mapping

You can define rules for all transformations, except Source transformation

Evaluates multiple rules in a specific order

24

© Informatica. Proprietary and Confidential.



Field rules define how data enters a transformation from the upstream transformation.

If you do not define any field rules in the mapping, then by default, all fields are included.

You can define field rules for all transformations, except the Source transformation, because it is always at the start of the mapping.

If you define multiple field rules, then the rules are evaluated in the specified order. For example, if you want to include all incoming fields, except date fields, you can use the default field rule that includes all fields. You must then create a second field rule that excludes fields by data type, with the data type set to date-time.

Field Rules – Renaming Fields

- Renaming fields avoids naming conflicts
- Rename fields individually or in bulk
- Field naming conflicts propagate throughout the data flow
- Rename fields before the transformation where the error occurred
- Can use a prefix, suffix, or pattern to rename fields in bulk

25

© Informatica. Proprietary and Confidential.



Renaming fields avoids naming conflicts and helps in identifying the origin of the field. When you configure field rules, you can rename fields either individually or in bulk.

Field naming conflicts propagate throughout the data flow. If you get a field naming conflict error, you must rename fields before the transformation where the error occurred.

There are three options to rename fields in bulk. You can add a prefix, a suffix, or use a pattern.

Renaming fields allows you to bring data from multiple sources into the mapping. For example, you can use “bulk rename in a prefix” to avoid naming conflicts and easily identify the original source of all fields.

Field Rules – Selection Criteria

Criteria	Definition
All Fields	<ul style="list-style-type: none"> Includes or excludes all fields Can rename all fields in bulk
Named Fields	<ul style="list-style-type: none"> Includes or excludes specific fields Can create parameters to represent fields Can rename fields individually or in bulk
Fields by Data Types	<ul style="list-style-type: none"> Includes or excludes fields of a selected data type Can rename fields in bulk
Fields by Text or Pattern	<ul style="list-style-type: none"> Includes or excludes fields based on a prefix, suffix, or pattern Can rename fields in bulk

26

© Informatica. Proprietary and Confidential.



When you configure a field rule, you must specify the field selection criteria to determine which incoming fields apply to the field rule.

- **All Fields:** Includes or excludes all fields. When you choose the “All Fields” criteria, you can rename all fields in bulk.
- **Named Fields:** Includes or excludes fields that you specify. You can also create a parameter to represent a field to include or exclude. When you choose the “Named Fields” criteria, you can rename fields individually or in bulk.
- **Fields by Data Types:** Includes or excludes fields with the data types that you specify. When you choose this criterion, you can rename all fields in bulk.
- **Fields by Text or Pattern:** Includes or excludes fields based on a prefix, suffix, or pattern. When you select the prefix or suffix option, you can enter the text to use as the prefix or suffix. When you select the pattern option, you can enter a regular expression or use a parameter for the pattern. When you choose “Fields by Text or Pattern” criteria, you can rename all fields in bulk.

Best Practices for Creating Mappings

- Create a single data flow between the source and the target
- Connect all transformations to the data flow
- Use a Joiner transformation to join heterogeneous sources
- Resolve field naming conflicts before fields come into a transformation
- If you use a parameter for an object, you must use parameters for all conditions or field mappings in the data flow

27

© Informatica. Proprietary and Confidential.



You must create a single data flow between the source and the target. You must then connect all the transformations to the data flow.

To join heterogeneous sources, you must use a Joiner transformation.

You must resolve field naming conflicts before fields come into a transformation.

Finally, if you use a parameter for an object, you must use parameters for all conditions or field mappings in the data flow.

Lab Activity

4-1 Creating a Mapping Using Basic Transformations

In this lab, you will perform the following:

- Use Joiner, Lookup, Filter, and Expression transformations in a mapping

Module Summary

This module showed you how to:

- Discuss Cloud Mapping Designer
- List the mapping designer terminologies
- Define mappings
- Discuss basic transformations in the Cloud Mapping Designer
- Explain field rules
- State best practices for creating mappings

29

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 5

Advanced Transformations and Mapping Tasks



Module Objectives

After completing this module, you will be able to:

- Discuss Mapplets
- Explain advanced transformations in the Cloud Mapping Designer
- Describe Mapping Tasks
- Discuss mapping updates and deployment

3

© Informatica. Proprietary and Confidential.



Topic

Aggregator Transformation



Aggregator Transformation Overview



Aggregator Transformation

Allows you to perform aggregate calculations on group of data

Active transformation

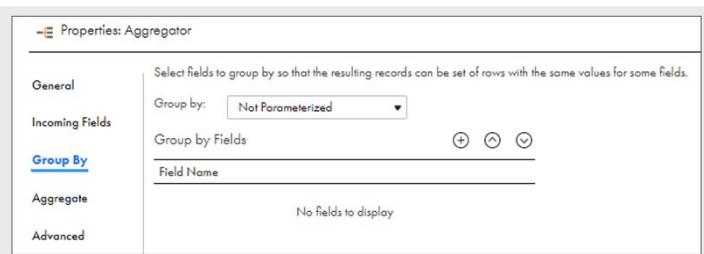
Data Integration Service stores the data temporarily in an aggregate cache, until it completes the aggregation

The Aggregator transformation allows you to perform aggregate calculations, such as average and sum on groups of data. The Aggregator transformation is an active transformation. This means, the number of rows that enter the transformation changes after you apply the transformation.

When you run a mapping that uses an Aggregator transformation, the Data Integration Service stores the data temporarily in an aggregate cache, until it completes the aggregation.

Group By Field

- Use Group By fields to group data for aggregate expressions
- When you configure a Group By field, the mapping task groups rows with the same data in the field
- The task performs aggregate calculations on each group and writes the result to the last row in the group
- When you select more than one Group By field, the task creates a group for each unique combination of data in the Group By fields



© Informatica. Proprietary and Confidential.



The Group By field groups data for aggregate expressions. When you configure a Group By field, the mapping task groups rows with the same data in the field. The task performs aggregate calculations on each group and writes the result to the last row in the group.

The result of an aggregate expression varies based on the Group By fields that you configure. When you select more than one Group By field, the task creates a group for each unique combination of data in the Group By fields. You can configure Group By fields in the Group By tab of the Properties panel.

Aggregator Transformation – Source

- Distributed annual sales across four quarters for each store:

Store Name	Q1	Q2	Q3	Q4	Year
Walmart	30	50	48	80	2011
BestBuy	120	100	88	150	2011
Kellogs	80	108	123	134	2011
Walmart	40	60	88	100	2012
BestBuy	110	120	98	140	2012
Kellogs	100	98	133	234	2012
Walmart	40	56	68	80	2013
BestBuy	125	200	98	150	2013
Kellogs	90	128	123	134	2013
Walmart	90	86	88	180	2014
BestBuy	125	200	148	190	2014
Kellogs	90	138	126	154	2014

7

© Informatica. Proprietary and Confidential.



Consider this sample source data on which you can apply the Aggregator transformation. As shown in this example, each store distributes the annual sales across four quarters. Now, let's calculate the annual sales and quarterly average of each store per year.

Aggregator Transformation – Target

- To know the stores' annual sales and the quarterly average for each year, use the Aggregator transformation

Store Name	Annual_Sales	Quaterly_Avg	Year
BestBuy	458	115	2011
BestBuy	468	117	2012
BestBuy	573	143	2013
BestBuy	573	143	2014
Kellogs	445	111	2011
Kellogs	565	141	2012
Kellogs	475	119	2013
Kellogs	475	119	2014
Walmart	208	52	2011
Walmart	288	72	2012
Walmart	244	61	2013
Walmart	244	61	2014

8

© Informatica. Proprietary and Confidential.



As you can see, after applying the Aggregator transformation, you can easily obtain the store's annual sales and the quarterly average for each year.

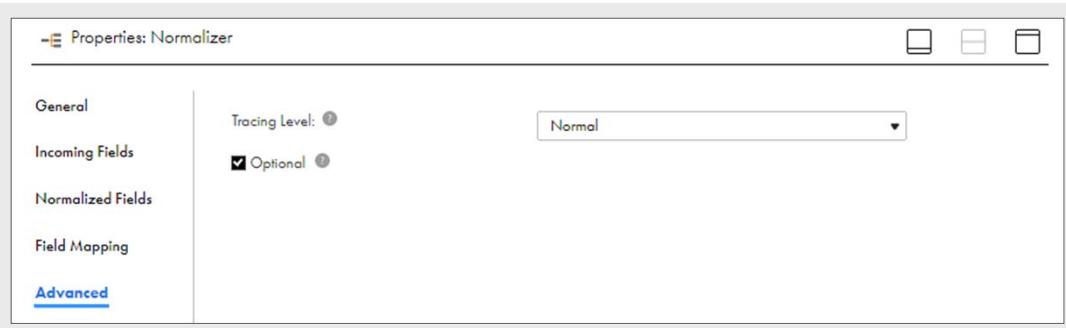
Topic

Normalizer Transformation



Normalizer Transformation Overview

- An active transformation
- For a row that contains multiple-occurring field, it returns a row for each instance of the multiple-occurring field



10

© Informatica. Proprietary and Confidential.



The Normalizer transformation is an active transformation that transforms one incoming row into multiple output rows. When the Normalizer transformation receives a row that contains multiple-occurring fields, it returns a row for each instance of the multiple-occurring field.

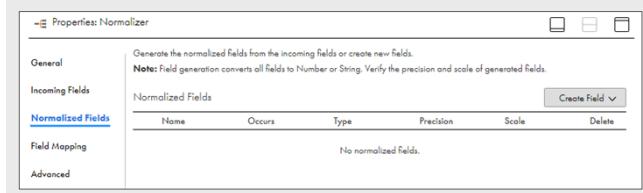
When you configure a Normalizer transformation, you must define the normalizer properties.

In the Normalized Fields tab, you must define the multiple-occurring fields and specify additional fields that you want to use in the mapping.

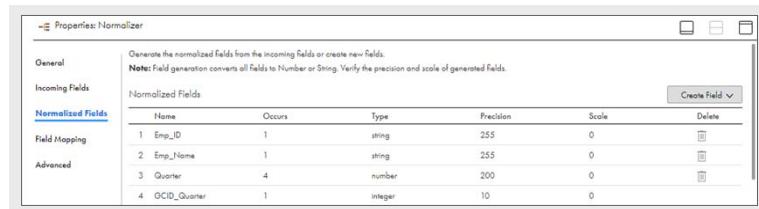
In the Field Mapping tab, you must connect the incoming fields to the normalized fields.

Normalized Fields

- Normalized fields are the fields that occurs multiple times and holds different data value for a single row



- Set the Occurs value to an integer greater than 1



11

© Informatica. Proprietary and Confidential.



Normalized fields are the fields that occur multiple times and holds different data value for a single row. The Normalized Fields tab represents normalized fields. In the Normalized Fields tab, you can define additional incoming fields in a mapping. To define multiple occurring fields, you must set the Occurs value to an integer greater than 1 for that field. The Normalizer transformation generates a column ID for every field that has the Occurs value greater than 1.

Generated Key

- Normalizer transformation generates key values for the normalized data
- The mapping task generates following keys for the normalized data:
 - Generated Key (GK)
 - Generated Column ID (GCID)

Name	Occurs	Type	Precision	Scale	Delete
1 Emp_ID	1	string	255	0	[Delete]
2 Emp_Name	1	string	255	0	[Delete]
3 Quarter	4	number	200	0	[Delete]
4 GCID_Quarter	1	integer	10	0	[Delete]
5 GK_Quarter	1	bigint	19	0	[Delete]

12

© Informatica. Proprietary and Confidential.



The Normalizer transformation generates key values for normalized data. These generated keys appear on the Normalized Fields tab, when you configure the field to have more than one occurrence.

The mapping task generates the following fields for normalized data:

- Generated Key is a key value that the task generates each time it processes an incoming row. When a task runs, it begins with the Generated Key having value 1, and increments the value by 1 for each processed row. The naming convention for the Normalizer generated key is GK_<redefined field name>.
- Generated Column ID is a column ID value that represents the instance of multiple-occurring data. The Normalizer transformation uses a generated column ID for each field that is configured to occur more than once. The naming convention for the Normalizer generated key is GCID_<redefined field name>.

The image shows that the Generated Key is GK_Department and the Generated Column ID is GCID_Department.

Topic

Java Transformation



Java Transformation

- Provides a simple, native programming interface to define transformation functionality
- You do not need advanced knowledge of the Java programming language
- Can be an active or a passive transformation

14

© Informatica. Proprietary and Confidential.



You can extend Data Integration functionality with the Java transformation. The Java transformation provides a simple, native programming interface to define transformation functionality with the Java programming language.

You can use the Java transformation to quickly define simple or moderately complex transformation functionality without having advanced knowledge of the Java programming language. The Java transformation can be an active or a passive transformation.

Java Transformation (continued)

- Secure Agent requires a JDK to compile the Java code and generate byte code for the transformation
- Azul Open JDK is installed with the Secure Agent
- Azul Open JDK includes the JRE
- Secure Agent uses the JRE to execute the byte code, process input rows, and generate output rows
- Define transformation behavior for a Java transformation based on the following events:
 - Transformation receives an input row
 - Transformation processes all input rows
 - Transformation receives a transaction notification

15

© Informatica. Proprietary and Confidential.



The Secure Agent requires a Java Development Kit (JDK) to compile the Java code and generate byte code for the transformation. Azul Open JDK is installed with the Secure Agent, so you don't have to install a separate JDK. Azul Open JDK includes the Java Runtime Environment (JRE).

The Secure Agent uses the JRE to execute generated byte code at run time. When you run a mapping or mapping task that includes a Java transformation, the Secure Agent uses the JRE to execute the byte code, process input rows, and generate output rows.

To create a Java transformation, you must write Java code snippets that define the transformation logic. You must also define the transformation behavior for a Java transformation based on events such as:

- The transformation receives an input row
- The transformation processes all input rows, and
- The transformation receives a transaction notification

Note: You cannot invoke expressions in a Java transformation.

Topic

SQL Transformation



SQL Transformation

- Use the SQL transformation to call a stored procedure or function, process a saved query, or process a SQL query
- The SQL transformation can process the following types of SQL statements:
 - Stored procedure or stored function
 - Saved query or user-entered query

17

© Informatica. Proprietary and Confidential.



You can use the SQL transformation to call a stored procedure or function, process a saved query, or process a query that you create in the transformation SQL editor.

The types of SQL statements that the SQL transformation can process are:

- Stored procedure or stored function, and
- Saved query or user-entered query

Now let's see how you can use these SQL statements in the SQL transformation.

Stored Procedure or Stored Function

- A stored procedure is a pre-compiled collection of database procedural statements and optional flow control statements
- A stored function is similar to a stored procedure, except that a function returns a single value
- The SQL transformation passes input parameters to the stored procedure or function
- The stored procedure or function passes the return value or values to the output fields of the transformation

18

© Informatica. Proprietary and Confidential.



A stored procedure is a precompiled collection of database procedural statements and optional flow control statements, similar to an executable script. Stored procedures reside in the database and run within the database. A stored function is similar to a stored procedure, except that a function returns a single value.

When the SQL transformation processes a stored procedure or function, it passes input parameters to the stored procedure or function. The stored procedure or function passes the return value or values to the output fields of the transformation.

Stored Procedure or Stored Function (continued)

- Use stored procedures to perform the following tasks:
 - Check the status of a target database before loading data into it
 - Determine if enough space exists in a database
 - Perform a specialized calculation
 - Retrieve data by a value
 - Drop and re-create indexes
- You can perform calculations using a stored procedure instead of a mapping

19

© Informatica. Proprietary and Confidential.



You can use a stored procedure to perform tasks such as:

- Check the status of a target database before loading data into it
- Determine if enough space exists in a database
- Perform a specialized calculation
- Retrieve data by a value, and
- Drop and re-create indexes

You can also use a stored procedure to perform a calculation that you usually perform in a mapping. For example, if you have a stored procedure to calculate sales tax, you can perform the calculation in a SQL transformation instead of re-creating the calculation in an Expression transformation.

Saved Query or User-Entered Query

- SQL transformation can process a saved query or process a query that you enter in the SQL editor
- You can pass strings or parameters to the query
- SQL transformation outputs multiple rows when the query has a SELECT statement

20

© Informatica. Proprietary and Confidential.



You can also configure the SQL transformation to process a saved query that you create in Data Integration or you can enter a query in the SQL editor. The SQL transformation processes the query and returns rows and database errors.

You can pass strings or parameters to the query to define dynamic queries or change the selection parameters. You can output multiple rows when the query has a SELECT statement.

Saved Query or User-Entered Query (continued)

- In a Static SQL query, you cannot change the query statement
 - Data Integration prepares the SQL query once and runs the query for all input rows
- In a Dynamic SQL query, you can change the query statements and the data
 - Data Integration prepares the SQL query for each input row

21

© Informatica. Proprietary and Confidential.



You can create a Static SQL query or a Dynamic SQL query.

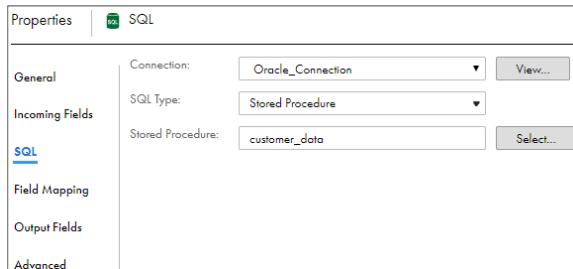
In a Static SQL query, you cannot change the query statement. However, you can use query parameters to change the data. Data Integration prepares the SQL query once and runs the query for all input rows.

In a Dynamic SQL query, you can change the query statements and the data. Data Integration prepares the SQL query for each input row.

SQL Transformation – Use Case

• Scenario:

- You have a mapping that includes user IDs in the data flow. You want to include user names in addition to user IDs.



• Solution:

- You have a stored procedure that matches user IDs with user names in the database.
- Add a SQL transformation to the mapping.
- Select the stored procedure.
- Map the userId incoming field with the userId input field in the stored procedure.
- Check the Output Fields tab of the SQL transformation to confirm that it includes the username field.
- Run the mapping and observe that the username value is returned with the user ID.

22

© Informatica. Proprietary and Confidential.



Use Case:

Consider that you have a mapping that includes user IDs in the data flow. You want to include user names in addition to user IDs.

You have a stored procedure that matches user IDs with user names in the database. You add a SQL transformation to your mapping, select the stored procedure, and map the userId incoming field with the userId input field in the stored procedure. You check the Output Fields tab of the SQL transformation to confirm that it includes the username field. When you run the mapping, the username value is returned with the user ID.

Points to Note When Using SQL Transformation

- When you use a stored procedure in a SQL transformation, you must define input/output parameters in the stored procedure
- The input/output parameters appear as input/output fields in the SQL transformation

23

© Informatica. Proprietary and Confidential.



When you use a stored procedure in a SQL transformation, you must define input/output parameters in the stored procedure. The input/output parameters appear as input/output fields in the SQL transformation. If you do not define input/output parameters, the mapping becomes invalid.

Topic

Union Transformation



Union Transformation

- An active transformation
- For Data Integration patterns, it is common to combine two or more data sources into a single stream that includes the union of all rows
- The Union transformation enables you to make the metadata of the streams alike
- You can add, change, or remove specific fields when you merge data sources
- At run time, the mapping task processes input groups in parallel
- When the mapping runs, it merges data into a single output group based on the field mappings

25

© Informatica. Proprietary and Confidential.



The Union transformation is an active transformation that you can use to merge data from multiple pipelines into a single pipeline.

For Data Integration patterns, it is common to combine two or more data sources into a single stream that includes the union of all rows. The data sources often do not have the same structure, so you cannot freely join the data streams. The Union transformation enables you to make the metadata of the streams alike, so you can combine the data sources in a single target.

The Union transformation merges data from multiple sources similar to the UNION ALL SQL statement. For example, you can use the Union transformation to merge employee information from **ADP Workforce Now** with data from a **Workday** employee object.

With a Union transformation, you can add, change, or remove specific fields when you merge data sources.

At run time, the mapping task processes input groups in parallel. It concurrently reads the sources connected to the Union transformation and pushes blocks of data into the input groups of the transformation. When the mapping runs, it merges data into a single output group based on the field mappings.

Union Transformation – Notes

- Add all Source transformations and include the other upstream transformations that you want to use
- You can use a Sequence Generator transformation upstream from a Union transformation

26

© Informatica. Proprietary and Confidential.



Remember a few important points when you use a Union transformation in a mapping.

- Before you add a Union transformation to a mapping, add all Source transformations and include the other upstream transformations that you want to use.
- You can use a Sequence Generator transformation upstream from a Union transformation, if you connect both the Sequence Generator and a Source transformation to one input group of the Union transformation.

Topic

Lookup Transformation



Lookup Transformation Overview



Lookup
Transformation

Passive transformation

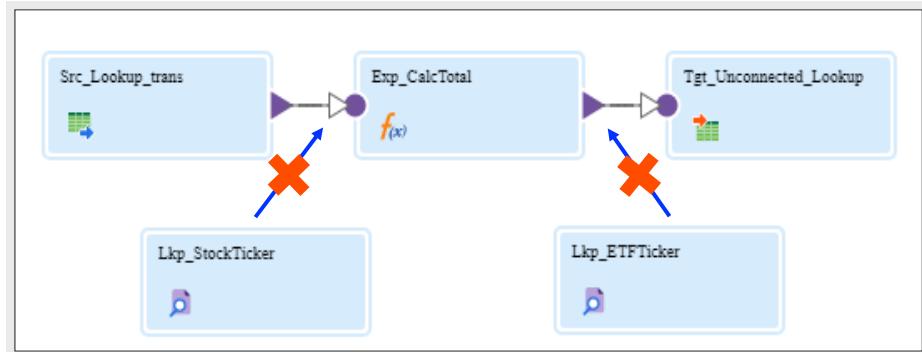
Allows you to perform lookup on relational tables, flat files, synonyms, and views

Extract the data from the lookup table or file based on the lookup condition

The Lookup transformation is a passive transformation. It allows you to perform lookup on relational tables, flat files, synonyms, and views. You must use a lookup condition to extract data from lookup tables or flat files.

Unconnected Lookup

- Neither connected to the source nor to the target
- Use when you want the lookup to return only one value
- Reuse the lookup multiple times in a mapping



29

© Informatica. Proprietary and Confidential.



An Unconnected Lookup transformation is neither connected to the source nor to the target. You can use the Unconnected Lookup transformation when you want the lookup to return only one value. You can reuse the lookup multiple times in a mapping.

You can use an unconnected lookup only with a database or flat file data source.

Dynamic Lookup

- When you enable lookup caching, a mapping task builds the lookup cache when it processes the first lookup request
- Cache can be static or dynamic
- If the cache is static, the data in the lookup cache does not change as the mapping task runs
- If the cache is dynamic, the task updates the cache based on the actions defined in the task

30

© Informatica. Proprietary and Confidential.



When you enable lookup caching, a mapping task builds the lookup cache while it processes the first lookup request. The cache can be static or dynamic.

If the cache is static, the data in the lookup cache does not change as the mapping task runs. If the task uses the cache multiple times, the task uses the same data.

If the cache is dynamic, the task updates the cache based on the actions defined in the task. So, if the task uses the lookup multiple times, downstream transformations can use updated data.

You can use a dynamic lookup cache to keep the lookup cache synchronized with the target. You can also use a dynamic lookup cache if the source data contains duplicate primary keys.

Dynamic Lookup (continued)

- The mapping task performs one of the following actions on the dynamic lookup cache when it reads a row from the source:
 - Inserts the row into the cache
 - Updates the row in the cache
 - Makes no change to the cache
- The dynamic Lookup transformation includes the return field and the New Lookup Row

31

© Informatica. Proprietary and Confidential.



Now, based on the results of the lookup query, the row type, and the Lookup transformation properties, the mapping task performs one of the following actions on the dynamic lookup cache when it reads a row from the source:

- Inserts the row into the cache: The mapping task inserts the row when it is not in the cache. The task flags the row as insert.
- Updates the row in the cache: The mapping task updates the row when it exists in the cache. It updates the row in the cache based on the input fields, and flags the row as an update row.
- Makes no change to the cache: The mapping task makes no change when the row is in the cache, and nothing changes. The task flags the row as unchanged.

The dynamic Lookup transformation includes the return field and the New Lookup Row, which describes the changes that the task makes to each row in the cache.

Note that you cannot use a parameterized source, target, or lookup with a Lookup transformation that uses a dynamic cache.

Topic

Rank Transformation



Rank Transformation

- Selects the top or bottom range of data
- Use the Rank Transformation to:
 - Return the largest or smallest numeric values in a group
 - Return strings at the top or bottom of the mapping sort order
- Rank transformation differs from the transformation functions MAX and MIN

33

© Informatica. Proprietary and Confidential.



The Rank transformation selects the top or bottom range of data. You can use the Rank transformation to return the largest or smallest numeric values in a group. You can also use the Rank transformation to return strings at the top or bottom of the mapping sort order. For example, you can use a Rank transformation to select the top 10 customers by region, or you can identify the three departments with the lowest expenses in salaries and overhead.

The Rank transformation differs from the transformation functions MAX and MIN because the Rank transformation returns a group of values, not just one value. The SQL language provides many functions that are designed to handle groups of data. However, identifying top or bottom groups of values within a set of rows is not possible with the use of standard SQL functions.

Rank Transformation (continued)

- An active transformation
- Example:
 - Select the top 10 rows from a source that contains 100 rows
- When you run a mapping that contains a Rank transformation, Data Integration caches input data until it can perform the rank calculations

34

© Informatica. Proprietary and Confidential.



The Rank transformation is an active transformation because it can change the number of rows that pass through it. For example, you can configure the transformation to select the top 10 rows from a source that contains 100 rows. In this case, 100 rows pass into the transformation but only 10 rows pass from the Rank transformation to the downstream transformation or target.

When you run a mapping that contains a Rank transformation, Data Integration caches input data until it can perform the rank calculations.

Topic

Sequence Generator Transformation



Sequence Generator Transformation

- A passive and connected transformation that generates numeric values
- Use the Sequence Generator transformation to:
 - Create unique primary key values
 - Replace missing primary keys or
 - Cycle through a sequential range of numbers
- Contains pass-through fields and two output fields – NEXT VAL and CURR VAL that you can connect to one or more downstream transformations
- Mapping task generates a numeric sequence of values each time the mapped fields enter a connected transformation
- After the task completes, you can see the last value generated for a Sequence Generator transformation

36

© Informatica. Proprietary and Confidential.



The Sequence Generator transformation is a passive and connected transformation that generates numeric values. You can use the Sequence Generator transformation to create unique primary key values, replace missing primary keys, or cycle through a sequential range of numbers.

The Sequence Generator transformation contains pass-through fields and two output fields, NEXT VAL and CURR VAL, that you can connect to one or more downstream transformations.

The mapping task generates a numeric sequence of values each time the mapped fields enter a connected transformation. You can set the range of numbers in the Mapping Designer. You can change the initial number in the sequence when you run the task.

After the task completes, you can see the last value generated for a Sequence Generator transformation.

Sequence Generator Transformation – Notes

- Cannot connect a Sequence Generator transformation to an upstream transformation
- Sequence Generator transformation cannot be connected alone
- When you map the NEXT VAL and CURR VAL output fields, ensure that the data type of the mapped field is appropriate
- When you run the mapping, the current value is not saved
- When you run the task, you can edit the current value to start the sequence with a specified value

37

© Informatica. Proprietary and Confidential.



You cannot connect a Sequence Generator transformation to an upstream transformation.

The Sequence Generator transformation cannot be connected alone. You must use at least one field from the other connected transformation in the mapping. For example, if you connect a Sequence Generator transformation and a Source transformation to a Target transformation, you must map at least one field from the Source transformation to the target.

When you map the NEXT VAL and CURR VAL output fields, ensure that the data type of the mapped field is appropriate.

When you run the mapping in the Mapping Designer, the current value is not saved. So, each time you run the mapping, it begins with the initial value.

When you run the task in the Mapping Task wizard, you can edit the current value to start the sequence with a specified value.

Topic
Data Masking Transformation



Data Masking Transformation

- Use the Data Masking transformation to change sensitive production data to realistic test data for non-production environments
- Modifies source data based on the masking rules that you configure for each column
- A passive transformation
- Provides masking rules based on the source data type and masking type that you configure for a port
- For strings, you can replace the characters with the characters that you want to apply in the mask
- For numbers and dates, you can provide a range of numbers for the masked data
- Integration Service replaces characters based on the locale that you configure with the masking rules

39

© Informatica. Proprietary and Confidential.



You can use the Data Masking transformation to change sensitive production data to realistic test data for non-production environments. The Data Masking transformation modifies source data based on the masking rules that you configure for each column.

You can maintain data relationships in the masked data and maintain referential integrity between database tables. The Data Masking transformation is a passive transformation.

The Data Masking transformation provides masking rules based on the source data type and masking type that you configure for a port. For strings, you can replace the characters with the characters that you want to apply in the mask. For numbers and dates, you can provide a range of numbers for the masked data. You can configure a range that is a fixed or percentage variance from the original number. The Integration Service replaces characters based on the locale that you configure with the masking rules.

Masking Techniques

- Credit Card Masking
- Email Masking
- IP Address Masking
- Key Masking
- Phone Masking
- Random Masking
- Social Insurance Number (SIN) Masking
- Social Security Number (SSN) Masking
- Custom Substitution Masking
- Substitution Masking
- URL Masking

40

© Informatica. Proprietary and Confidential.



The masking technique is a type of data masking that you want to apply to a selected column. You can select one of the following masking techniques:

- Credit Card masking applies a credit card mask format to columns of string data type that contain credit card numbers.
- Email masking applies an email mask format to columns of string data type that contain email addresses.
- IP Address masking applies an IP address mask format to columns of string data type that contain IP addresses.
- Key masking produces deterministic results for the same source data and seed value. You can apply key masking to datetime, string, and numeric data types.
- Phone masking applies a phone number mask format to columns of string data type that contain phone numbers.
- Random masking produces random results for the same source data and mask format. You can apply random masking to datetime, string, and numeric data types.

Topic
Cleanse Transformation



Cleanse Transformation

- Adds a cleanse asset to a mapping
 - A cleanse asset is a set of data transformation operations that standardize the form and content of your data
- You can add a single cleanse asset to a Cleanse transformation
- A cleanse asset can perform the following operations:
 - Change the character case of the input data
 - Remove leading and trailing spaces from input data
 - Remove values from the input data
 - Find and replace values in the input data

42

© Informatica. Proprietary and Confidential.



The Cleanse transformation adds a cleanse asset to a mapping. You can create the cleanse asset in Data Quality. A cleanse asset is a set of data transformation operations that standardize the form and content of your data.

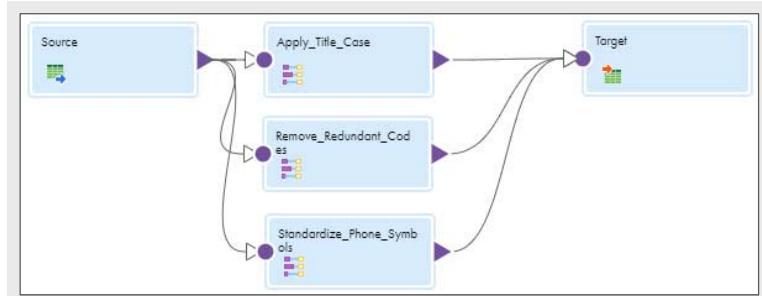
You can add a single cleanse asset to a Cleanse transformation. Each cleanse asset maps to a single input field in the mapping.

A cleanse asset can perform one or more of the following operations:

- Change the character case of the input data.
- Remove leading and trailing spaces from input data.
- Remove values from the input data. You can enter the values that you want the mapping to remove, or use a dictionary to specify the values.
- Find and replace values in the input data. You can enter the values that you want the mapping to find and replace, or use a dictionary to specify the values.

Cleanse Transformation (continued)

- You can configure multiple operations in a cleanse asset and add any type of operation to the asset multiple times
- You can add multiple Cleanse transformations to a mapping
- A Cleanse transformation shows incoming and outgoing fields



43

© Informatica. Proprietary and Confidential.



You can configure multiple operations in a cleanse asset. You can also add any type of operation to the asset multiple times. The mapping performs the operations on an input data field in a sequence that you define, so that a single cleanse asset can specify multiple changes to the input field data. You can add multiple Cleanse transformations to a mapping and apply standardization operations to multiple data fields.

A Cleanse transformation is similar to a Mapplet transformation, as it allows you to add data transformation logic that you designed elsewhere to a mapping. Like mapplets, cleanse assets are reusable assets.

A Cleanse transformation shows incoming and outgoing fields. It does not display the logic that the cleanse asset contains or allow you to edit the cleanse asset. To edit the cleanse asset, you must open it in Data Quality.

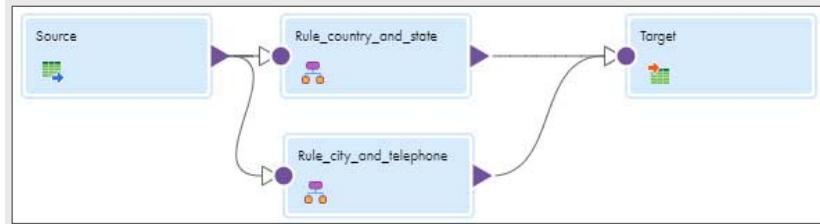
Topic

Rule Specification Transformation



Rule Specification Transformation

- Adds a rule specification asset to a mapping
 - A rule specification asset is a set of one or more logical operations that analyze data according to the business criteria that you define
- Each Rule Specification transformation can contain a single rule specification
- You can add multiple Rule Specification transformations to a mapping



45

© Informatica. Proprietary and Confidential.



The Rule Specification transformation adds a rule specification asset to a mapping. You can create a rule specification asset in Data Quality.

A rule specification asset is a set of one or more logical operations that analyze data according to the business criteria that you define. The rule specification generates an output that indicates whether the data satisfies the business criteria. The rule specification can also update the data that it analyzes. You can define the logical operations as IF/THEN/ELSE statements in Data Quality.

Each Rule Specification transformation can contain a single rule specification. You can add multiple Rule Specification transformations to a mapping.

Rule Specification Transformation (continued)

- Use a Rule Specification transformation to define:
 - Types of data that a business data set contains
 - Set of conditions that the business data must satisfy
 - Actions to take when the data satisfies the conditions of the business rule
 - Actions to take when the data fails to satisfy the conditions of the business rule
- A Rule Specification transformation is similar to a Maplet transformation
- A Rule Specification transformation shows incoming and outgoing fields

46

© Informatica. Proprietary and Confidential.



You can use a Rule Specification transformation to perform the following tasks:

- Define the types of data that a business data set contains.
- Define a set of conditions that the business data must satisfy.
- Define the actions to take when the data satisfies the conditions of the business rule.
- Define the actions to take when the data fails to satisfy the conditions of the business rule.

A Rule Specification transformation is similar to a Maplet transformation, as it allows you to add data analysis and data transformation logic that you designed elsewhere to a mapping. Like mapplets, rule specifications are reusable assets.

A Rule Specification transformation shows incoming and outgoing fields. It does not display the logic that the rule specification contains or allow to you edit the rule specification. To edit the rule specification, you must open it in Data Quality.

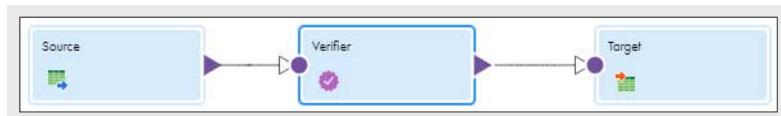
Topic

Verifier Transformation



Verifier Transformation

- Adds a verifier asset to a mapping
 - A verifier asset defines a template for input and output address data that you can connect to the input and output fields on the Verifier transformation
- Connect fields in source data or in upstream transformations to input ports on the Verifier transformation
- Connect output ports on the Verifier transformation to downstream transformations in the mapping or to the mapping target



48

© Informatica. Proprietary and Confidential.



The Verifier transformation adds a verifier asset to a mapping. You create a verifier asset in Data Quality.

A verifier asset defines a template for input and output address data that you can connect to the input and output fields on the Verifier transformation.

You must connect the fields in your source data or in upstream transformations to the input ports on the Verifier transformation. You must connect the output ports on the Verifier transformation to downstream transformations in the mapping or to the mapping target.

Verifier Transformation (continued)

- Use the Verifier transformation to perform the following operations on input address data:
 - Compare address records in the source data with address definitions in the address reference data
 - Fix errors and complete partial address records
 - Write output addresses in the format that the verifier asset specifies
 - Report on the deliverable status of each address and the nature of error in the address
 - Provide suggestions for any ambiguous or incomplete address

49

© Informatica. Proprietary and Confidential.



The Verifier transformation performs the following operations on the input address data:

- The transformation compares address records in the source data with address definitions in the address reference data.
- It fixes errors and completes partial address records. To fix an address, the transformation must find a positive match with an address in the reference data. The transformation copies the required data elements from the address reference data to the address records.
- It writes output addresses in the format that the verifier asset specifies. You can define a verifier asset in Data Quality to create address records that suit your business needs. You can also create addresses with the structure that the mail carrier requires.
- It can report on the deliverable status of each address and the nature of any error or ambiguity in the address.
- It can provide suggestions for any ambiguous or incomplete address.

A Verifier transformation is similar to a Mapplet transformation, as it allows you to add address verification logic that you created elsewhere to a mapping. Like mapplets, verifiers are reusable assets. A Verifier transformation shows incoming and outgoing fields. It does not display the address data that the verifier contains or allow to you edit the verifier. To edit the verifier, you must open it in Data Quality.

Topic Mapplets



Mapplets

- Reusable transformation logic that you can use to transform source data before it is loaded to the target
- Use the Maplet Designer to create a mapplet or upload a mapplet that you exported from PowerCenter
- Add the mapplet to a Maplet transformation to use its transformation logic
- Mapplets can be either active or passive

51

© Informatica. Proprietary and Confidential.



A mapplet is a reusable transformation logic that you can use to transform source data before it is loaded to the target. You can use the Maplet Designer to create a mapplet or upload a mapplet that you exported from PowerCenter. After you create a mapplet, you can add it to a Maplet transformation to use its transformation logic. Mapplets can be either active or passive.

Mapplet Input

- Mapplet input can be an Input transformation, a Source transformation, or both
- Use an Input transformation when you want the mapplet to receive input data from one or more upstream transformations
- You can use multiple Input transformations in a mapplet
- You can include one or more Source transformations in a mapplet to provide source data
- A mapplet must contain at least one Input transformation or Source transformation

52

© Informatica. Proprietary and Confidential.



To use a mapplet in a Mapplet transformation, you must configure the mapplet input and output.

Mapplet input can be an Input transformation, a Source transformation, or both. Use an Input transformation when you want the mapplet to receive input data from one or more upstream transformations.

You can use multiple Input transformations in a mapplet. When you use the mapplet in a Mapplet transformation, each Input transformation becomes an input group. Use multiple Input transformations when you have multiple pipelines in a mapplet, or when you want the mapplet to receive input from multiple upstream transformations.

You can include one or more Source transformations in a mapplet to provide source data. When you use only Source transformations for mapplet input, the mapplet is the first object in the mapping pipeline and contains no input groups.

A mapplet must contain at least one Input transformation or Source transformation.

Maplet Output

- Maplet output can be an Output transformation, a Target transformation, or both
- Use an Output transformation when you want the maplet to pass data to one or more downstream transformations
- Use a Target transformation when you want the maplet to write data to a target
- A maplet must contain at least one Output transformation or Target transformation

53

© Informatica. Proprietary and Confidential.



Maplet output can be an Output transformation, a Target transformation, or both.

Use an Output transformation when you want the maplet to pass data to one or more downstream transformations. When you use the maplet in a Maplet transformation, each Output transformation becomes an output group. Each output group can pass data to one or more pipelines in a mapping.

Use a Target transformation when you want the maplet to write data to a target. When you use only a Target transformation for maplet output, the maplet is the last object in the mapping pipeline.

A maplet must contain at least one Output transformation or Target transformation.

Parameters in Mapplets

- You can use input parameters in a mapplet
- Data Integration renames the parameters when you use the mapplet in a Mapplet transformation

54

© Informatica. Proprietary and Confidential.



You can use input parameters in a mapplet. You can specify the value of the parameters when you configure the mapping task.

When you include parameters in a mapplet, Data Integration renames the parameters when you use the mapplet in a Mapplet transformation. The parameter names are prefixed with the name of the Mapplet transformation.

PowerCenter Mapplets

- Use a PowerCenter mapplet to create a mapplet in Data Integration
- Create a mapplet in PowerCenter and export the mapplet to an XML file
- Upload the XML file to Data Integration
- PowerCenter mapplet can contain one or more Source transformations, but it cannot contain a Target transformation
- Use PowerCenter mapplets in the following Data Integration tasks:
 - Synchronization tasks
 - Mapping tasks
 - Masking tasks

55

© Informatica. Proprietary and Confidential.



You can use a PowerCenter mapplet to create a mapplet in Data Integration. To use a PowerCenter mapplet, you create a mapplet in PowerCenter and export the mapplet to an XML file. Then you upload the XML file to Data Integration. When you upload a PowerCenter mapplet to Data Integration, you must specify whether the mapplet is active or passive.

A PowerCenter mapplet can contain one or more Source transformations, however it cannot contain a Target transformation.

You can use PowerCenter mapplets in the following Data Integration tasks:

- Synchronization tasks: You can use one mapplet in a synchronization task.
- Mapping tasks: You can use multiple mapplets in a mapping task.
- Masking tasks: You can use passive mapplets in a masking task to mask target fields.

PowerCenter XML Files for Mapplets

- If the mapplet includes a transformation that uses a connection, then the PowerCenter XML file must contain only one workflow, one session task, one mapping, and one mapplet
- If the mapplet does not include a transformation that uses a connection, then the PowerCenter XML file must include one mapplet
 - The session can use any type of connection
 - You do not have to map all source and target fields in the PowerCenter mapping

You must note the following rules when you use a PowerCenter XML file to create a Data Integration mapplet:

- If the mapplet includes a transformation that uses a connection, then the PowerCenter XML file must contain only one workflow, one session task, one mapping, and one mapplet.
- If the mapplet does not include a transformation that uses a connection, then the PowerCenter XML file must include one mapplet. The workflow, Session task, and mapping are optional.
- The session can use any type of connection.
- You do not have to map all source and target fields in the PowerCenter mapping.

PowerCenter XML Files for Mapplets (continued)

- If you use a mapplet in a synchronization task, the PowerCenter mapplet cannot contain multiple Input transformations
- If you use a mapplet in a mapping task, the PowerCenter mapplet can contain multiple Input transformations
- Data Integration flattens PowerCenter mapplets with multiple input groups into mapplets with one input group
- PowerCenter mapplet cannot contain reusable objects such as shortcuts

- If you use a mapplet in a synchronization task, the PowerCenter mapplet cannot contain multiple Input transformations.
- If you use a mapplet in a mapping task, the PowerCenter mapplet can contain multiple Input transformations.
- Data Integration flattens PowerCenter mapplets with multiple input groups into mapplets with one input group. Therefore, the ports in each input group in the PowerCenter mapplet must have unique names. If the names are not unique, rename the input ports in PowerCenter before you export the PowerCenter XML file that contains the mapplet.
- The PowerCenter mapplet cannot contain reusable objects such as shortcuts because Data Integration does not use a repository to store reusable objects. Export the mapplet without reusable objects.

Topic

Maplet Transformation



Mapplet Transformation

- Mapplet transformation inserts a mapplet that you created in Data Integration, imported from PowerCenter, or generated from a SAP asset into a mapping
- Each Mapplet transformation can contain one mapplet
- Mapplet transformation can be active or passive based on the transformation logic within the mapplet
- An active mapplet includes at least one active transformation
- A passive mapplet includes only passive transformations

59

© Informatica. Proprietary and Confidential.



The Mapplet transformation inserts a mapplet that you created in Data Integration, imported from PowerCenter, or generated from a SAP asset into a mapping. Each Mapplet transformation can contain one mapplet. You can add multiple Mapplet transformations to a mapping or mapplet.

The Mapplet transformation can be active or passive based on the transformation logic within the mapplet.

An active mapplet includes at least one active transformation. An active mapplet can return a number of rows that is different from the number of source rows passed to the mapplet.

A passive mapplet includes only passive transformations. A passive mapplet returns the same number of rows that are passed from the source.

Mapplet Transformation (continued)

- Use the Mapplet transformation to accomplish the following goals:
 - extend the data transformation capabilities of Data Integration
 - reuse transformation logic in different mappings
 - hide complex transformation logic

60

© Informatica. Proprietary and Confidential.



You can use the Mapplet transformation to accomplish the following goals:

Extend the data transformation capabilities of Data Integration: For example, you want to create a mapping that passes customer records to a target if the customers pass a credit check. You create a Web Services transformation to run a credit check on each customer. You include the Web Services transformation in a mapplet and use the mapplet in a mapping to perform the credit check.
Note: Web Services transformation is discussed in a later module.

Reuse transformation logic in different mappings: For example, you have different fact tables that require a series of dimension keys. You create a mapplet that contains a series of Lookup transformations to find each dimension key. You include the mapplet in different fact table mappings instead of re-creating the lookup logic in each mapping.
Note: Lookup transformation is discussed later in this module.

Hide complex transformation logic: The Mapplet transformation shows the mapplet incoming and outgoing fields. It does not display the transformations that the mapplet contains.

Topic Mapping Task



Mapping Task

- Allows you to process data based on the data flow logic defined in a mapping
- You can define parameters that associate with the mapping
- Can configure the task to run on a schedule
- Can add pre and post-processing commands to the task

A Mapping task allows you to process data based on the data flow logic defined in a mapping. When you create a Mapping Task, you must select a mapping to use in the task. You can define parameters that associate with the mapping. You can also configure the Mapping Task to run on a schedule. You can also add pre and post-processing commands to a Mapping Task.

Mapping Task Features

- Download a Mapping Task and import it into PowerCenter
- Invoke a Mapping Task via the REST API or the Salesforce outbound message

63

© Informatica. Proprietary and Confidential.



A Mapping task has two important features:

- You can download a Mapping Task and import it into PowerCenter, and
- You can invoke a Mapping Task via the REST API or the Salesforce outbound message

Mapping Updates and Deployment

- Deploy updated mapping to Mapping Task
- After you make changes to a mapping, you can perform the following actions:
 - Save changes as a new mapping
 - Save changes and deploy the mapping to Mapping Task

64

© Informatica. Proprietary and Confidential.



When you make changes to a mapping, you must deploy the changes so that the Mapping Tasks can use the updated mapping. After you update a mapping, you can perform one of two tasks:

- Save changes as a new mapping: You must use this option if you want to keep the previous version of the mapping and the related Mapping Tasks.
- Save changes and deploy the mapping to Mapping Tasks: You must use this option if the changes are compatible with the existing tasks. You must note that you may have to edit the Mapping Tasks to verify that they are valid with the new version of the mapping. You can view the last deployed version of the mapping. If needed, you can also revert to the last deployed version of the mapping.

Lab Activity

5-1 Using Normalizer, Aggregator, and Rank Transformations in a Mapping

In this lab, you will perform the following:

- Configure a mapping in Informatica Cloud
- Use Normalizer, Rank, and Aggregator transformations in the mapping

65

© Informatica. Proprietary and Confidential.



Lab Activity

5-2 Creating a Mapping using Unconnected Lookup Transformation

In this lab, you will perform the following:

- Use Unconnected Lookup transformations in the mapping

Lab Activity

5-3 Creating a Mapping Task

In this lab, you will perform the following:

- Create a Mapping Task

Lab Activity

5-4 Using Maplet Transformation in a Mapping

In this lab, you will perform the following:

- Use Maplet transformation in a mapping

Module Summary

This module showed you how to:

- Discuss Mapplets
- Explain advanced transformations in the Cloud Mapping Designer
- Describe Mapping Tasks
- Discuss mapping updates and deployment

69

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 6

Mapping Parameters



Module Objectives

After completing this module, you will be able to:

- Explain parameters
- View the use cases of parameters
- List the types of parameters
- Describe a parameter file
- Discuss best practices for creating parameters

What are Parameters?

- Parameters are placeholders that represent values in a mapping
- You can use parameters to hold values that
 - you want to define at run-time
 - change between task runs
- You can override complete source queries in relational database connections

4

© Informatica. Proprietary and Confidential.



Parameters are placeholders that represent values in a mapping. You can use parameters to hold values to define at run-time, such as a source connection, a target object, or the join condition for a Joiner transformation. You can also use parameters to hold values that change between task runs, such as a time stamp that increments each time you run a mapping.

You can also use parameters to override complete source queries in relational database connections.

Use Cases of Parameters

- Parameterize source and target connections
- Reuse parameterized mapping in Development, Test, and Production environments
- Parameterize filter conditions and expressions

5

© Informatica. Proprietary and Confidential.



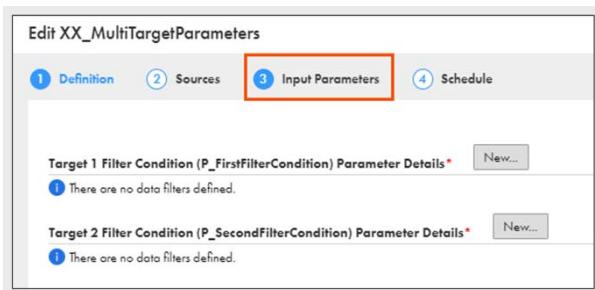
There are various use cases where you add a parameter to a mapping.

You can use parameters when you want to parameterize source and target connections for a mapping and reuse it in your Development, Test, and Production environments. You can create multiple mapping tasks using the same mapping.

You can also parameterize filter conditions and expressions to allow business users to easily specify values without editing the mapping.

Adding Parameters to a Mapping

- Allows you to create flexible mapping templates
- Any part of a mapping can be parameterized:
 - Source connection and source object
 - Target connection, target object, and field mapping
 - Join condition
 - Filter criteria
 - Lookup connection, lookup object, and lookup condition
 - Expression



6

© Informatica. Proprietary and Confidential.



You can add parameters to a mapping in order to create flexible mapping templates. Business users can use the mapping templates to create multiple mapping tasks.

Any part of a mapping can be parameterized. For example,

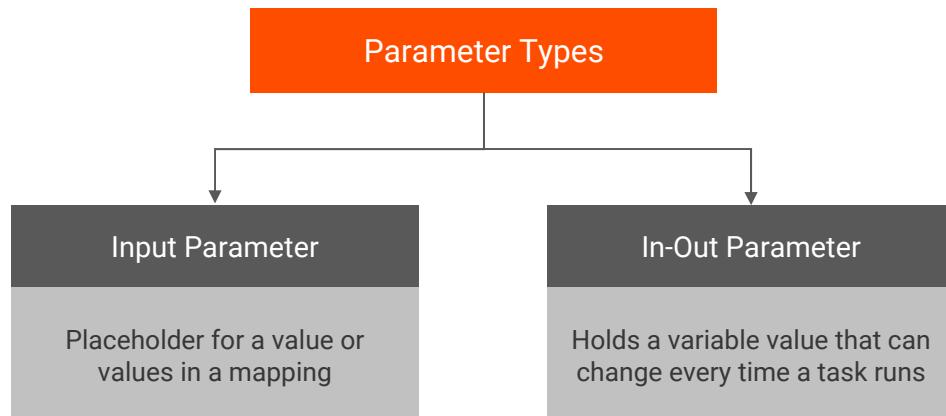
- The Source connection and source object.
- The Target connection, target object, and field mapping. Target supports full or partial parameterization of the field mapping.
- You can also parameterize a Join condition, or a filter criteria.
- A Lookup connection, lookup object, lookup condition, as well as Expressions.

Topic

Parameter Types



Parameter Types



8

© Informatica. Proprietary and Confidential.



You can create the following types of parameters in a mapping – Input Parameters and In-Out Parameters.

An input parameter is a placeholder for a value or values in a mapping. Input parameters help you to control the logical aspects of a data flow or to set other variables that you can use to manage different targets. You can define an input parameter in a mapping and set the value of the parameter when you configure a mapping task.

An In-Out parameter holds a variable value that can change every time a task runs. When you define an In-Out parameter, you can set a default value in the mapping. However, you would typically set the value at run time using an Expression transformation. You can also change the value in the mapping task.

Input Parameter

- You can use input parameters in the following transformations:

Source	Lookup
Target	Mapplet
All transformations with incoming fields	Rank
Aggregator	Router
Data Masking	Sorter
Expression	SQL
Filter	Structure Parser
Joiner	Union

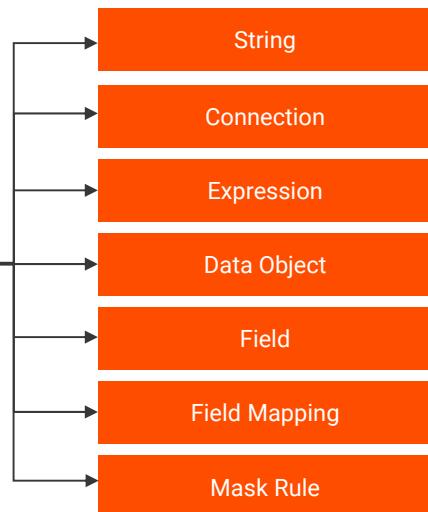
9

© Informatica. Proprietary and Confidential.



Types of Input Parameters

Types of Input Parameters



10

© Informatica. Proprietary and Confidential.



- **String:** A String parameter represents a string value. In the task, the string parameter displays as a textbox in most instances.
- **Connection:** A Connection parameter represents a connection. You can specify the connection type for the parameter or allow any connection type. In the task, the connection parameter displays a list of connections.
- **Expression:** An Expression parameter represents an expression. In the task, the expression parameter displays the Field Expression dialog box to configure an expression.
- **Data Object:** A Data Object parameter represents a data object, such as a source table or a source file. In the task, the data object parameter appears as a list of available objects from the selected connection.
- **Field:** A Field parameter represents a field. In the task, the field parameter displays as a list of available fields from the selected object.
- **Field Mapping:** A Field Mapping parameter represents field mappings for the task. You can create a full or partial field mapping. A full field mapping parameter displays all fields for configuration, and a partial field mapping parameter displays the unmapped fields.
- **Mask Rule:** A mask rule parameter represents a masking technique. In the task, the mask rule parameter displays a list of masking techniques.

In-Out Parameters

- Act as persistent task variables
- Can store a date value for the last record that loads from a data warehouse
- Can help you manage the update process for a slowly changing dimension table
- For example, you can:
 - update values after each task execution
 - handle incremental data loading for a data warehouse

11

© Informatica. Proprietary and Confidential.



In-Out parameters act as persistent task variables. The parameter values update during task execution. The parameter can store a date value for the last record that loads from a data warehouse, or it can help you manage the update process for a slowly changing dimension table.

Some examples where you can use an In-Out parameter:

- You can use an In-Out parameter to update values after each task execution. So, you can use the SetVariable, SetMaxVariable, SetMinVariable, or SetCountVariable functions in an Expression transformation to update parameter values each time you run a task.
- You can also use an In-Out parameter to handle incremental data loading for a data warehouse. In this case, you set a filter condition to select records from the source that meets the load criteria. When the task runs, you include an expression to increment the load process.

In-Out Parameters (continued)

- Use In-Out parameters in the following transformations:
 - Source
 - Target
 - Aggregator
 - Expression
- You cannot use In-Out parameters in expression macros and in an at-scale mapping
- Unlike Input parameters, an In-Out parameter can change each time a task runs
- You can reset the In-Out parameters in a mapping task

12

© Informatica. Proprietary and Confidential.



You cannot use In-Out parameters in expression macros and in an at-scale mapping.

For each In-Out parameter, you need to configure the variable name, datatype, default value, aggregation type, and retention policy. You can also use a parameter file that contains the value to be applied at run time. For a specific task run, you can change the value in the mapping task.

Unlike input parameters, an In-Out parameter can change each time a task runs. The latest value of the parameter displays in the job details when the task completes successfully. Next time the task runs, the mapping task compares the In-Out parameter to the saved value.

You can also reset the In-Out parameters in a mapping task, and then view the saved values in the job details.

Parameter File

- A list of user-defined parameters and their associated values
- Defines values that you want to update without having to edit the task
- Save the parameter file in the Secure Agent directory
- Parameter values are treated as String values
- You cannot use a parameter file if the mapping task is based on an at-scale mapping

13

© Informatica. Proprietary and Confidential.



A parameter file is a list of user-defined parameters and their associated values. You can use user-defined parameters in data filters, expressions, and lookup expressions in a Synchronization task or a Mapping task.

You can use a parameter file to define values that you want to update without having to edit the task. For example, you can use a parameter file for a Sales quota that changes quarterly.

You must save the parameter file in the Secure Agent directory. The parameter values are applied when the task runs.

Parameter values are treated as String values. So, when you use a parameter in an expression, you must use the appropriate function to convert the value to the necessary data type.

Note: You cannot use a parameter file if the mapping task is based on an at-scale mapping.

Parameter Best Practices

- You must first configure the mapping with a specific connection to select source, target, or lookup object
- Use parameters for all conditions or field mappings in the data flow that use fields from the object
- Enter the appropriate label for the parameter
- Provide a description for the parameter

14

© Informatica. Proprietary and Confidential.



You must first configure the mapping with a specific connection to select source, target, or lookup object. When the mapping is complete, you can replace the connection with a parameter.

When you use a parameter for an object, use parameters for all conditions or field mappings in the data flow that use fields from the object.

When you create a parameter, you must enter the appropriate label for the parameter. The label displays in the mapping task and helps users to enter the right data.

You must also provide a description for the parameter. The value that you enter in the description field displays as a tool tip in the mapping task wizard.

Lab Activity

6-1 Performing Complete Parameterization

In this lab, you will perform the following:

- Create a completely parameterized mapping

15

© Informatica. Proprietary and Confidential.



Lab Activity

6-2 Using Parameter File in a Mapping task

In this lab, you will perform the following:

- Build a fully parameterized mapping

Lab Activity

6-3 Using In-Out parameters for Incremental Data Loading

In this lab, you will perform the following:

- Create a mapping using Input-Output Parameters

Module Summary

This module showed you how to:

- Explain parameters
- View the use cases of parameters
- List the types of parameters
- Describe a parameter file
- Discuss best practices for creating parameters



IICS: Cloud Data Integration Services

 Informatica[®]
University

Module 7

Expression Macro and Dynamic Linking



Module Objectives

After completing this module, you will be able to:

- Describe Expression Macro
- List types of Expression Macro
- Explain Dynamic Linking
- Discuss Flat File target time stamps

Topic

Expression Macro



Expression Macro Overview

- An expression macro creates repetitive or complex expressions in mappings
- Use an expression macro to perform calculations across a set of fields or constants
- In an expression macro:
 - One or more input fields represent source data for the macro
 - An expression represents the calculations that you want to perform
 - An output field represents the results of the calculations
- At run time, the task expands the expression to include all the input fields and constants, and then writes the results to output fields
- You can create expression macros in Expression and Aggregator transformations

5

© Informatica. Proprietary and Confidential.



An expression macro allows you to create repetitive or complex expressions in mappings.

You can use an expression macro to perform calculations across a set of fields or constants. For example, you can use an expression macro to replace null values in a set of fields or to label items based on a set of sales ranges.

In an expression macro, one or more input fields represent source data for the macro. An expression represents the calculations that you want to perform, and an output field represents the results of the calculations. At run time, the task expands the expression to include all the input fields and constants, it then writes the results to the output fields.

You can create expression macros in Expression and Aggregator transformations. However, you cannot combine an expression macro and an in-out parameter in an Expression transformation.

Types of Expression Macros

Vertical Macro

- Generates a set of similar expressions to perform the same calculation on multiple incoming fields

Horizontal Macro

- Generates one extended expression that includes a set of fields or constants

Hybrid Macro

- Generates a set of vertical expressions that also expand horizontally

6

© Informatica. Proprietary and Confidential.



Vertical Macro: A vertical macro expands an expression vertically. A vertical macro generates a set of similar expressions to perform the same calculation on multiple incoming fields.

Horizontal Macro: A horizontal macro expands an expression horizontally. A horizontal macro generates one extended expression that includes a set of fields or constants.

Hybrid Macro: A hybrid macro expands an expression both vertically and horizontally. A hybrid macro generates a set of vertical expressions that also expands horizontally.

Vertical Macro

- Macro input field represents the incoming fields
- Expression represents the calculations you want to perform on all incoming fields
- Macro output field represents the results of the calculations
- The names of the output fields are not explicitly defined in the mapping
- Configure a field rule in the downstream transformation to include the output fields that the macro generates
- Link the output fields to target fields in the Target transformation
- When the task runs, it performs the following actions:
 - Generates multiple expressions based on the macro input fields
 - Replaces the macro output fields with the actual output fields
 - Uses the output fields to pass the results of the calculations to the rest of the mapping

7

© Informatica. Proprietary and Confidential.



You can use a vertical macro to apply a macro expression to a set of incoming fields. The macro input field in a vertical macro represents the incoming fields. The expression represents the calculations that you want to perform on all incoming fields. The macro output field represents a set of output fields that passes the results of the calculations to the rest of the mapping. You can configure the macro expression in the macro output field.

The macro output field represents the output fields of the macro. However, the names of the output fields are not explicitly defined in the mapping. To include the results of a vertical macro in the mapping, you must first configure a field rule in the downstream transformation to include the output fields that the macro generates.

To write the results of a vertical macro to the target, you must link the output fields to target fields in the Target transformation. When the task runs, it generates multiple expressions to perform calculations on each field that the macro input field represents. The task also replaces the macro output fields with the actual output fields. The task then uses the output fields to pass the results of the calculations to the rest of the mapping.

Note: The macro output field does not pass any data.

Vertical Macro – Example

Scenario

Remove leading and trailing spaces from the customer's address fields

Solution

```
LTRIM(RTRIM(%Addresses%))
```

Output

```
LTRIM(RTRIM(Street))
LTRIM(RTRIM(City))
LTRIM(RTRIM(State))
LTRIM(RTRIM(ZipCode))
```

8

© Informatica. Proprietary and Confidential.



Look at an example of a vertical macro expression.

Consider that you want to remove leading and trailing spaces from the customer's address fields.

You can use the vertical macro expression to trim leading and trailing spaces from the address fields.

When the task runs, it generates a set of expressions to trim spaces from the address fields. As you can see, the expression removes the leading and trailing spaces from the Street, City, State, and Zip Code address fields.

Horizontal Macro

- Use a horizontal macro to generate a single complex expression
- Macro input field represents a set of incoming fields or a set of constants
- Expression represents the calculations that you want to perform on the incoming fields or constants
- Horizontal macro expression produces one result, and the transformation output field passes the results to the rest of the mapping
- Results of the expression pass to the downstream transformation with the default field rule
- To write the results to the target, connect the transformation output field to a target field in the Target transformation

9

© Informatica. Proprietary and Confidential.



You can use a horizontal macro to generate a single complex expression that includes a set of incoming fields or a set of constants.

The macro input field represents a set of incoming fields or a set of constants. The expression represents the calculations that you want to perform on the incoming fields or constants. You must include a horizontal expansion function in the expression.

The horizontal macro expression produces one result, and the transformation output field passes the results to the rest of the mapping. You can configure the horizontal macro expression in the transformation output field.

The results of the expression passes to the downstream transformation with the default field rule. Unlike a vertical macro, you don't have to configure additional field rules to include the results of a horizontal macro in the mapping. To write the results of a horizontal macro to the target, you must connect the transformation output field to a target field in the Target transformation.

Horizontal Macro – Example

Scenario

Check for null values in all the fields of a customer record

Solution

```
%OPR_SUM{IIF(ISNULL(%AllFields%),1,0]}%
```

Output

```
IIF(ISNULL(AccountID, 1,0)+IIF(ISNULL(AccountName, 1, 0)+IIF(ISNULL(ContactName, 1, 0)+IIF(ISNULL(Phone, 1, 0)+IIF(ISNULL>Email, 1, 0)...
```

10

© Informatica. Proprietary and Confidential.



Assume that you want to check for null values in all the fields of a customer record.

You can use the horizontal macro expression to check for any null values in a field. The expression returns the value 1 when a field is null, and the horizontal expansion function **%OPR_SUM%** returns the total number of null fields.

When the task runs, it expands the expression horizontally to include all the incoming fields. As you can see, the expression checks for null values in the customer's AccountID, AccountName, ContactName, Phone, and Email fields.

Hybrid Macro

- Expands an expression both vertically and horizontally
- Configure a hybrid macro based on your business requirements

11

© Informatica. Proprietary and Confidential.



A hybrid macro expands an expression both vertically and horizontally. You can configure a hybrid macro based on your business requirements.

Hybrid Macro – Example

Scenario

Format the date fields in a customer's record to the 'mm-dd-yyyy' format

Solution

```
%OPR_IIF[ IsDate(%dateports%,%fromdateformat%),To_String(To_Date(%dateports%,%fromdateformat%),  
'mm-dd-yyyy'),%dateports%]
```

Output

```
IIF(IsDate(StartDate,'mm/dd/yy'),To_String(To_Date(StartDate,'mm/dd/yy'),'mm-dd-yyyy'),  
IIF(IsDate(StartDate,'mm/dd/yyyy'),To_String(To_Date(StartDate,'mm/dd/yyyy'),'mm-dd-yyyy'),  
StartDate))  
  
IIF(IsDate(EndDate,'mm/dd/yy'),To_String(To_Date(EndDate,'mm/dd/yy'),'mm-dd-yyyy'),  
IIF(IsDate(END_DT,'mm/dd/yyyy'),To_String(To_Date(EndDate,'mm/dd/yyyy'),'mm-dd-yyyy'),  
EndDate))
```

12

© Informatica. Proprietary and Confidential.



Assume that you want to format a date field in the customer's record to a specific format.

You can use the hybrid macro expression to convert the date fields to the required format. In the expression, the **%fromdateformat%** macro input field defines the different date formats used in the date fields.

When the task runs, it expands the expression vertically and horizontally. The expression expands vertically to create an expression for the **Start Date** and **End Date** fields that the **%dateports%** represents. The expression also expands horizontally to use the constants that the **%fromdateformat%** represents to evaluate the incoming fields.

Topic

Dynamic Linking



Dynamic Linking Overview

- IICS allows you to create a new target file at runtime
- Use this feature when you don't know the field names and the nature of the data that comes from the source
- You can create a new target file at runtime only in mappings

14

© Informatica. Proprietary and Confidential.



IICS allows you to use dynamic linking to create a new target file at runtime.

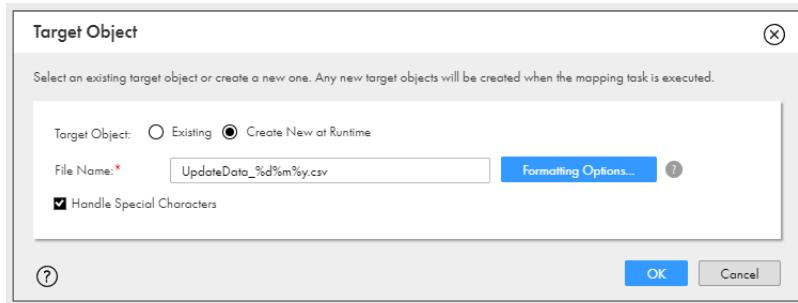
You can use this feature when you don't know the field names and the nature of the data that comes from the source. When you choose to create a new target file at runtime, you don't have to manually write codes to create a table and populate it.

It is important to know that you can create a new target file at runtime only in mappings, and not in Synchronization, Replication, and other tasks.

Creating New Target at Runtime

To create a new target file, you must perform the following steps:

1. Select **Create New at Runtime** option in the Target Tab
2. Specify the file name for the new target file
3. Use the Formatting Options to configure the format of the target file



15

© Informatica. Proprietary and Confidential.



To create a new target file, you must perform the following steps:

1. Select **Create New at Runtime** option in the Target Tab.
2. Specify the file name for the new target file.
3. Use the Formatting Options to configure the format of the target file.

Flat File Target Time Stamps

- You can append the time-stamp information to the file name to show when the file is created
- Some common function formats are:

Special Character	Description
%d	Day as a two-decimal number, with a range of 01-31
%m	Month as a two-decimal number, with a range of 01-12
%y	Year as a two-decimal number without the century, with a range of 00-99
%Y	Year including the century, for example 2019
%T	Time in 24-hour notation, equivalent to %H:%M:%S
%H	Hour in 24-hour clock notation, with a range of 00-24
%I	Hour in 12-hour clock notation, with a range of 01-12
%M	Minute as a decimal, with a range of 00-59
%S	Second as a decimal, with a range of 00-60
%p	Either AM or PM

16

© Informatica. Proprietary and Confidential.



When you create a Flat File target at run time, you can append time stamp information to the file name to show when the file was created. When you specify the file name for the target file, you can include special characters based on Linux function formats. The Mapping Task uses these function formats to include time stamp information in the file name.

Here are some of the common function formats:

- %d, represents the day as a two-decimal number, with a range of 1 to 31.
- %m, represents the month as a two-decimal number, with a range of 1 to 12.
- %y in lower case, represents the year as a two-decimal number without the century, and with a range of 0 to 99.
- %Y in upper case, represents the year including the century, for example 2019.
- %T, represents the time in 24-hour notation.
- %H, represents the hour in 24-hour clock notation, with a range of 0 to 24.
- %I, represents the hour in 12-hour clock notation, with a range of 1 to 12.
- %M, represents the minute as a decimal, with a range of 0 to 59.
- %S, represents the second as a decimal, with a range of 0 to 60.
- %p, represents the time as either AM or PM.

Lab Activity

7-1 Using an Expression Macro in a Mapping

In this lab, you will perform the following:

- Use an Expression Macro in a mapping

Lab Activity

7-2 Using Dynamic Linking in a Mapping

In this lab, you will perform the following:

- Use Dynamic Linking by creating a Flat File at runtime
- Append time stamp in the name of the file

Module Summary

This module showed you how to:

- Describe Expression Macro
- List types of Expression Macro
- Explain Dynamic Linking
- Discuss Flat File target time stamps



IICS: Cloud Data Integration Services

 Informatica
University

Module 8

Replication Task



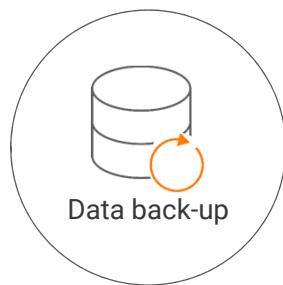
Module Objectives

After completing this module, you will be able to:

- Explain the purpose of a Replication task
- List the features of a Replication task
- Discuss load type options in a Replication task
- Describe source and target options in a Replication task

Replication Task Overview

- Replicates data from a cloud-based application or a relational database table to a target
 - Replicates all rows
 - Replicates only changed rows
- Use cases:



4

© Informatica. Proprietary and Confidential.



A Replication task allows you to replicate data from a cloud-based application or a database table, to a target. You can replicate all rows of a source object each time the task runs, or replicate only those rows that changed since the last time the task ran. You can also use a Replication task to reset target tables and create target tables.

You can create and use a Replication task when you need to perform a regular back-up of data, when you need to archive data for compliance purposes, or when you need to move data to a data warehouse to perform offline reporting.

Replication Task Features

Automatically replicates the data and schema to the target

You can schedule the Replication task

In-built incremental processing

5

© Informatica. Proprietary and Confidential.



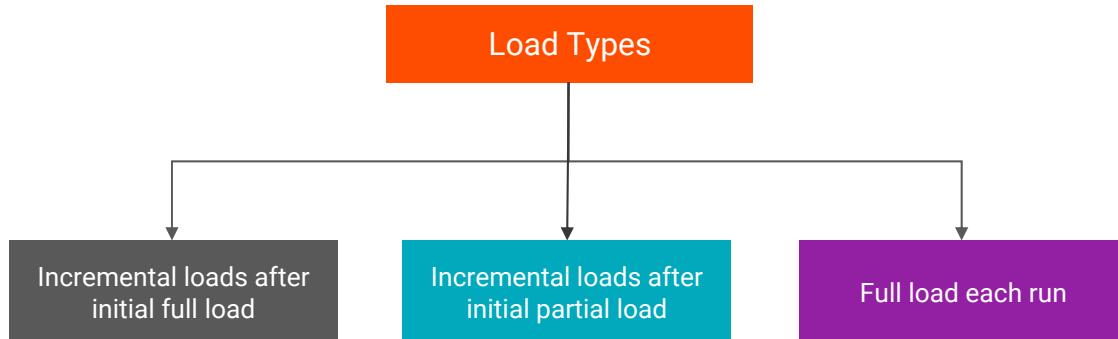
A Replication task automatically replicates data and schema to the target. One of the major differences between a synchronization task and a Replication task is that, in a synchronization task you must have a target to integrate the data. However, a Replication task can create a target for you.

You can configure a Replication task to run on a schedule.

The Replication task has in-built incremental processing. This means that when you replicate data from a cloud-based application to a database table, you can choose to capture only the changed data.

Load Types

- Determines the type of operation that you can use when you replicate data from the source to the target



6

© Informatica. Proprietary and Confidential.



The load type determines the type of operation you can use when you replicate data from a source to the target. When you replicate data, you can use one of three load types:

- Incremental loads after initial full load:** The first time the Replication task runs, it performs a full load, replicating all rows of the source. For each subsequent run, the Replication task performs an incremental load. In an incremental load, the Replication task uses an Upsert operation to replicate rows that changed since the last time the task ran. You can specify the load type when the task uses a Salesforce source and a database target.
- Incremental loads after initial partial load:** The Replication task always performs an incremental load with this load type. The first time the Replication task runs, the Replication task processes rows created or modified after a specified point in time. For each subsequent run, the Replication task replicates rows that changed since the last time the task ran. You can specify the load type when the task uses a Salesforce source and a database target.
- Full load each run:** The Replication task replicates all rows of the source objects in the task during each run. You can specify this load type when the task uses a Salesforce or database source and a database or flat file target.

Replication Task – Source Options

- You can replicate a single object or multiple objects
- If an error occurs while replicating multiple objects, you can choose to either cancel or continue with the processing of the remaining objects
- For Salesforce sources only, you can include archived and deleted rows

7

© Informatica. Proprietary and Confidential.



When you configure a Replication task, you have the following source options:

- You can replicate a single object or multiple objects.
- If an error occurs while replicating multiple objects, you can choose to either cancel or continue with the processing of the remaining objects.
- For Salesforce sources only, you can include archived and deleted rows.

Replication Task – Target Options

- Specify the load type to be incremental or full
- In the target, retain, or delete rows that are deleted from the source
- A prefix is added to name the resulting table or flat file
 - can change the default prefix and assign a unique prefix

- When you replicate data from a Salesforce source to a database table, you can specify the load type to be incremental or full. The incremental load type loads only new or changed source rows to the target. The full load type loads all source rows to the target.
- In the target, you can retain or delete the rows that are deleted from the source.
- By default, a prefix is added to the name of the resulting table or flat file. For example, if you use the Replication task to replicate data from the Salesforce Account object to a database table, the default name of the resulting table would be SF_ACCOUNT.

You can change the default prefix and assign a unique prefix. Assigning a unique prefix allows users to replicate data using the same database connection, without the risk of overwriting data.

Other Replication Task Options

Exclude fields for each source object

For Salesforce sources, can apply row limits

Apply data filters to retrieve only a certain subset of data

Enable high precision calculations for Salesforce sources

You can exclude fields for each source object. For example, if you have fields in Salesforce that the business is not using, then you can exclude those fields.

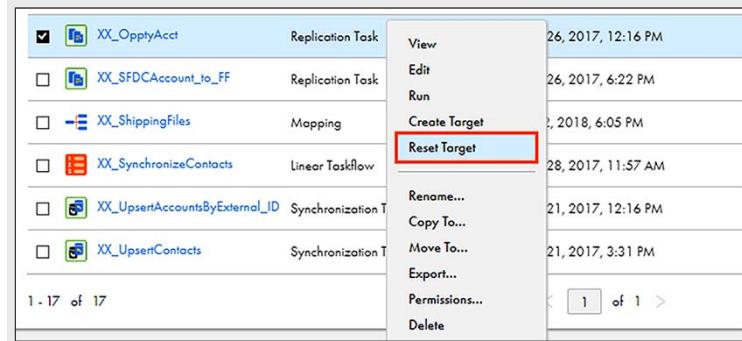
If your source is a Salesforce object, then you can apply row limits. For example, if you want to perform a test run of a task, you can limit the number of rows you export. For non-Salesforce sources, this option is disabled.

You can also apply data filters to retrieve only a certain subset of data. For example, if you want to replicate data to a data warehouse for reporting, you can filter and replicate only records of a certain type.

Finally, you can enable high precision calculations for Salesforce sources. When you can enable high precision calculations, the Replication task reads data with a precision of up to 28 in Salesforce calculated fields and writes the data to the target.

Resetting the Target Table

- While using Replication task, IICS provides 'Reset Target' option to drop the target table
- Use this option:
 - when the task load type is incremental, and
 - if any one of the following changes are made to a source field – datatype, precision, and scale



© Informatica. Proprietary and Confidential.



How can you reset a target table using a Replication task?

Usually, to drop a target table, you have to go to the database that stores the table, and then drop the table there. When you use a Replication task, IICS provides a Reset Target option that enables you to drop the target table within the IICS interface.

You can use this option when the task type is incremental and if you change the data type, precision, or scale of a source field.

Assume that you have a Replication task scheduled to run every day, with an incremental load type. If you change the data type, precision, or scale of a source field, you will get an error when the task runs again because the data type, precision, or scale for the source field and target field are inconsistent. In this scenario, you can use the Reset Target option to drop the target table after you make changes to the source fields. The next time the task runs, it will run a full load and create the target table.

Generating Non-Unique Index

- Replication task generates non-unique indexes for target tables that do not exist
- Replication task also generates an index when you use the 'Create Target' option on the Replication Tasks page
- Non-unique index is based on the Salesforce ID field

11

© Informatica. Proprietary and Confidential.



When you replicate Salesforce sources to target database tables that do not exist, the Replication task generates a non-unique index for each target table. The Replication task also generates an index when you use the **Create Target** option on the Replication Tasks page and replicate the Salesforce sources to target database tables.

The index is generated based on the Salesforce ID field. Indexes are not generated for Salesforce sources that do not include a Salesforce ID field.

Lab Activity

8-1 Replicating Data to a Flat file

In this lab, you will perform the following:

- Create a Replication task to replicate data to a CSV file

Module Summary

This module showed you how to:

- Explain the purpose of a Replication task
- List the features of a Replication task
- Discuss load type options in a Replication task
- Describe source and target options in a Replication task

13

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica[®]
University

Module 9

Masking Task



Module Objectives

After completing this module, you will be able to:

- Explain the purpose of a masking task
- Discuss masking task source and target options
- Define masking rule types
- Refresh masking task metadata
- Reset a masking task
- Discuss guidelines for masking data
- Create a masking task

Masking Task Overview

- Allows you to mask sensitive fields in the source data with realistic test data for non-production environments
- Allows you to migrate data from one Salesforce org to another Salesforce org
- Allows you to choose the source and the target and then select a masking rule
- Allows you to mask data “inplace” to overwrite existing data

4

© Informatica. Proprietary and Confidential.



You can use a masking task to mask the sensitive fields in source data with realistic test data for non-production environments. A masking task allows you to migrate data from one Salesforce org to another. When you migrate data, you can mask sensitive fields with realistic test data.

When you configure a masking task, you must choose the source and the target and then select a masking rule for each field in the source that you want to mask.

You can also use the “in place” masking option to mask the data in the same system from which the masking task reads the data. In simpler terms, the “in place” masking option allows you to overwrite the existing data in a Salesforce org.

The Masking Task Wizard

The screenshot shows the 'The Masking Task Wizard' interface. On the left, a vertical navigation bar lists six steps: STEP 1 (Definition), STEP 2 (Configure source), STEP 3 (Configure target), STEP 4 (Configure data filters), STEP 5 (Define masking rules), and STEP 6 (Schedule). The first step, 'Definition', is highlighted in teal. To the right of the navigation bar is a detailed configuration panel titled 'Task Details'. It includes fields for 'Task Name' (with a red asterisk indicating it's required), 'Location' (set to 'REG\Default'), and 'Description'. A 'Browse' button is also present. The top of the configuration panel has tabs for 'Definition', 'Source', 'Target', 'Data Filters', 'Masking', and 'Schedule', with 'Definition' being the active tab.

STEP 1 Definition

STEP 2 Configure source

STEP 3 Configure target

STEP 4 Configure data filters

STEP 5 Define masking rules

STEP 6 Schedule

5 © Informatica. Proprietary and Confidential.

 Informatica University

Using a six-step wizard, you can configure a masking task.

- In the first step, you can define the masking task.
- In the second step, you can configure the source options.
- In the third step, you can configure the target options.
- In the fourth step, you can configure the data subset or filters.
- In the fifth step, you can define data masking rules.
- In the sixth step, you can configure the scheduling options for the masking task.

Masking Source Options

- You can add a single object or multiple related objects
- A single object does not contain any related objects
- Multiple objects have an explicit relationship defined in Salesforce
 - Salesforce Opportunity object is related to the Campaign object
- If you select multiple source objects, you can choose an object and add the related parent, child, and self-reference objects manually

The screenshot shows the 'Source' tab of the masking task configuration. The 'Source Details' section includes fields for Connection, Source Type (radio buttons for Single and Multiple, with Single selected), Source Object (dropdown set to Account), and two checkboxes: 'Display technical names instead of labels' (checked) and 'Display source fields in alphabetical order' (unchecked). Below this is a 'Data Preview' section.



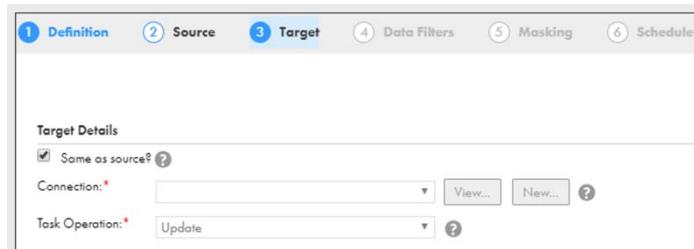
You can add a single object or multiple related objects in a masking task.

You can add a single object that does not contain any related objects. You can also add multiple objects that have an explicit relationship defined in Salesforce. For example, if you use the **Opportunity** object in Salesforce as a source, you can add the related **Campaign** object as well. All Salesforce objects in a multiple-object source must have a predefined relationship in Salesforce.

If you select multiple source objects, you can choose an object and add the related parent, child, and self-reference objects manually. A self-reference relationship is the one in which a source object references to itself within a task.

Masking Target Options

- For a Salesforce source, you can select another Salesforce connection for your target
- Select the task operation that you want to perform in the target
- Select the ‘Same as Source’ option in the target step if you want to mask data “inplace”
- When the target is different from the source, select one the following task operations:
 - Insert
 - Update
 - Upsert



© Informatica. Proprietary and Confidential.



If you have a Salesforce source, you can select another Salesforce connection for your target. This means that the source and target type must be the same in a masking task.

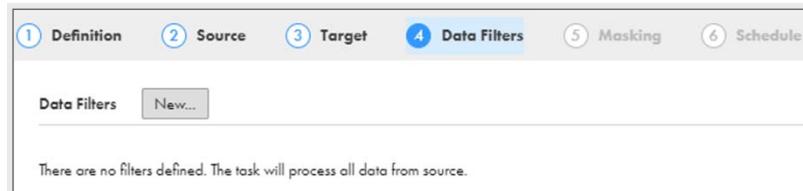
When you configure the target options, you can select the task operation that you want to perform in the target. If you want to mask the data “in place”, you must select the ‘Same as Source’ option in the target step. When the target is same as the source, you can perform only an Update operation.

When the target is different from the source, you can select one the following task operations:

- Insert: This operation ignores the existing target data and inserts all the source data.
- Update: This operation updates data in the target location based on the source data.
- Upsert: This operation updates existing target data. If data does not exist in the target, the masking task inserts the data.

Data Subset

- Allows you to specify row limit or data filter options
- Available only if the source and target are not from the same connection
- For Salesforce sources, create a filter on a single Salesforce object



8

© Informatica. Proprietary and Confidential.



A Data Subset allows you to specify row limit or data filter options in a masking task. You must note that these options are available only if the source and target connections are different. You cannot use the filter if they are from the same connection or account. Therefore, you cannot use these options if you want to mask the data “in place”.

For Salesforce sources, you can create a filter on a single Salesforce object.

Defining Field Masking Rules

- Specify masking rules for columns that you want to mask
- Column data type determines the available masking rules
- If you do not specify a masking rule for a column, the column is copied unmasked

The screenshot shows the 'Masking' step of the Informatica Data Masking wizard. The interface has tabs at the top: 1 Definition, 2 Source, 3 Target, 4 Data Filters, 5 Masking (which is selected), and 6 Schedule. Below the tabs is a button bar with 'Default Assignment', 'Clear Assignment', 'Add Maplet...', and 'Configure'. The main area is titled 'Define Masking Rules' and shows a table for 'Source Object: Account'. The table has columns for 'Status' (checkboxes), 'Name' (AccountNumber, AccountSource, Active__c, AnnualRevenue, BillingCity), and 'Masking Rule' (dropdown menus). The 'Masking Rule' column for AccountNumber, AccountSource, Active__c, and AnnualRevenue each has a dropdown menu with three options: 'None', 'Masked', and 'Unmasked'. The 'Masking Rule' column for BillingCity also has a dropdown menu with three options: 'None', 'Masked', and 'Unmasked'.

9

© Informatica. Proprietary and Confidential.



In the Masking step of the wizard, you can select masking rules for columns that you want to mask.

The column data type determines the available masking rules. For example, integer and date fields have fewer masking options.

If you do not specify a masking rule for a column, then the column is copied ‘unmasked’ to the target.

Masking Rule Types

- Credit card masking
- Date masking
- Email masking
- IP address masking
- Key masking
- Nullification masking

10

© Informatica. Proprietary and Confidential.



A masking rule defines the logic that masks the data. As we saw earlier, the type of masking rule that you can apply depends on the data type of the field that you want to mask.

- **Credit card masking:** A credit card masking rule applies a built-in mask format to mask credit card numbers.
- **Date masking:** A date masking rule applies a date mask format to columns of string data type that contain dates.
- **Email masking:** An email masking rule applies an email mask format to columns of string data type that contains email addresses.
- **IP address masking:** An IP address masking rule applies an IP address mask format to columns of string data type that contain IP addresses.
- **Key masking:** A key masking rule produces repeatable results for the same source data. You can apply key masking to datetime, string, and numeric data types.
- **Nullification masking:** A nullification masking rule transfers a null value from the source to the target.

Masking Rule Types (continued)

- Phone masking
- Random masking
- Social Insurance Number (SIN) masking
- Social Security Number (SSN) masking
- Substitution masking
- URL masking

11

© Informatica. Proprietary and Confidential.

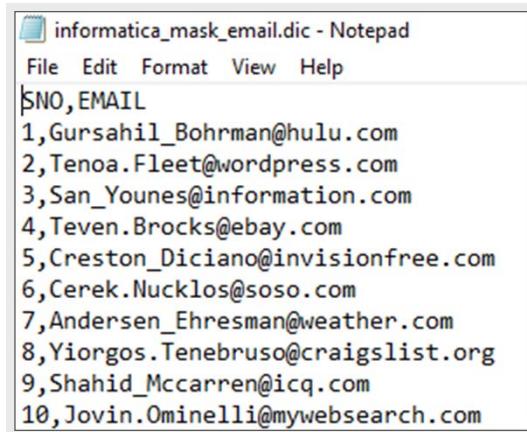


- **Phone masking:** This rule applies a phone number mask format to columns of string data type that contain phone numbers.
- **Random masking:** This rule produces random, non-repeatable results for the same source data and masking rules. You can apply random masking to datetime, string, and numeric data types.
- **Social Insurance Number (SIN) masking:** This rule applies a built-in mask format to modify the Social Insurance numbers according to the specified format.
- **Social Security Number (SSN) masking:** This rule applies a built-in mask format to modify the Social Security numbers.
- **Substitution masking:** This rule replaces a column of data with similar but unrelated data from a default dictionary. You can apply substitution masking to columns with string data type.
- **URL masking:** This rule applies a URL mask format to columns of string data type that contain URLs.

Remember, for some of the masking rules, you can configure additional options.

Dictionary Files

- Masking task uses a set of built-in dictionary files or the custom dictionary files
- Masking task performs a lookup on the selected dictionary and replaces the source data with data from the dictionary
- You can find dictionary files in the following directory:
 - C:\Program Files\Informatica Cloud Secure Agent\apps\Data_Integration_Server\data
- You cannot edit or rename dictionary files



The screenshot shows a Notepad window with the title 'informatica_mask_email.dic - Notepad'. The menu bar includes File, Edit, Format, View, and Help. The content of the file is a list of email addresses, each preceded by a number from 1 to 10, followed by the header 'SNO,EMAIL'. The list includes various email addresses such as Gursahil_Bohrman@hulu.com, Tenoa.Fleet@wordpress.com, San_Younes@information.com, Teven.Brocks@ebay.com, Creston_Diciano@invisionfree.com, Cerek.Nucklos@soso.com, Andersen_Ehresman@weather.com, Yiorgos.Tenebruso@craigslist.org, Shahid_Mccarren@icq.com, and Jovin.Ominelli@mywebsearch.com.

SNO	EMAIL
1	Gursahil_Bohrman@hulu.com
2	Tenoa.Fleet@wordpress.com
3	San_Younes@information.com
4	Teven.Brocks@ebay.com
5	Creston_Diciano@invisionfree.com
6	Cerek.Nucklos@soso.com
7	Andersen_Ehresman@weather.com
8	Yiorgos.Tenebruso@craigslist.org
9	Shahid_Mccarren@icq.com
10	Jovin.Ominelli@mywebsearch.com

12

© Informatica. Proprietary and Confidential.



The masking task uses a set of built-in dictionary files or custom dictionary files that you create. When you configure a substitution masking operation, you can select a dictionary that contains substitute values.

The masking task performs a lookup on the selected dictionary and replaces the source data with data from the dictionary.

The dictionary files are located in the secure agent installation directory. While you cannot edit or rename the dictionary files, you can change the content within the specified file structure.

The image shows a sample dictionary file for email masking.

Refresh Metadata

- Masking task imports the source and target metadata
- If you make changes to Salesforce objects or objects in the task, the metadata imported when you created the task can get outdated
- Masking task requires the latest metadata to define relationships between objects and to determine fields that you can mask
- Refresh the metadata before you run a masking task to ensure that the source and target metadata in the task is up to date

13

© Informatica. Proprietary and Confidential.



When you create a masking task, the task imports the source and target metadata. Over time, you might update the Salesforce objects and add or delete objects. You might also add or delete objects in the masking task.

When there are changes to Salesforce objects or objects in the task, the metadata imported when you created the task can get outdated. If you run the same masking task at regular intervals, the metadata imported in the task will not be the latest. The masking task requires the latest metadata to define relationships between objects and to determine fields that you can mask.

A masking task can fail if it does not use the updated metadata in the Salesforce source and target. You can refresh the metadata before you run a masking task to ensure that the source and target metadata in the task is up to date.

There are two ways in which you can refresh the metadata in a masking task, as shown in the next few pages.

Refresh Metadata Without Editing the Task

- Refresh runs as a separate job
- You cannot run an instance of a masking task and a metadata refresh of the task at the same time
- If the refresh job fails, the metadata does not update

You can refresh the metadata without editing the task.

When you refresh the metadata without editing the task, the refresh runs as a separate job. You cannot run an instance of a masking task and a metadata refresh of the task at the same time. If the refresh job fails at any point, the metadata does not update. So the source metadata and target metadata remain consistent.

Refresh Metadata From Within the Task

- Refresh the source and target fields from within a masking task
- You cannot view the progress of the refresh or perform other tasks during the refresh
- The refresh process can take some time, based on the number of objects and the size of the metadata

15

© Informatica. Proprietary and Confidential.



You can refresh the source and target fields from within a masking task when you create or update a masking task. You cannot view the progress of the refresh or perform other tasks during the refresh. You can continue to create or update and save the masking task after the refresh finishes. The refresh process can take some time, based on the number of objects and the size of the metadata.

Best Practices For Refreshing Metadata

- If you want to update many objects, refresh the metadata without editing the task
- If you want to update fewer objects, refresh the metadata from within the task

16

© Informatica. Proprietary and Confidential.



You can choose how you want to refresh the metadata based on the number of objects to refresh. As a best practice, if you want to update many objects, it is recommended that you refresh the metadata without editing the task. To update fewer objects or less metadata, you can refresh the metadata from within the task.

Reset Task

- You can reset a masking task that has a different source and target and contains data filters
- Masking task with data filters performs different steps
- If the task fails at any of the steps, it continues from the point of failure when you restart the task
- Reset returns the task status to Start
- When you restart the task, the task starts from the first step
- Tasks that use the same source and target or do not include data filters do not require subset computation or staging tables

17

© Informatica. Proprietary and Confidential.



You can reset a masking task that has a different source and target and contains data filters.

A masking task with data filters performs different steps including staging data, subset computation, load to target, and drop staging tables. If a task fails at any of the steps, it continues from the point of failure when you restart the task.

You can choose to reset the task before you restart the task. The reset returns the task status to Start. When you restart the task, the task starts from the first step. It performs all steps of staging, subset computation, load to target, and drop staging tables, based on how you configure the task.

Tasks that use the same source and target or do not include data filters do not require subset computation or staging tables.

Masking Guidelines

- Target must contain all source fields
- Field names in the target must match the field names in the source
- Target can contain additional fields

18

© Informatica. Proprietary and Confidential.



When you mask the data from a source org to a target org, you must ensure that the target org contains all the source fields. Additionally, the field names in the target org must match the field names in the source org.

The target org can contain additional fields. However, the additional fields will not have masked data.

Lab Activity

9-1 Creating a Masking Task

In this lab, you will perform the following:

- Create a masking task to mask phone number

Module Summary

This module showed you how to:

- Explain the purpose of a masking task
- Discuss masking task source and target options
- Define masking rule types
- Refresh masking task metadata
- Reset a masking task
- Discuss guidelines for masking data
- Create a masking task

20

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 10

Mass Ingestion Task



Module Objectives

After completing this module, you will be able to:

- Discuss mass ingestion task
- Describe mass ingestion task sources and targets
- List file processing actions
- Create a mass ingestion task

Mass Ingestion Task Overview

- Enables you to transfer, track, and monitor huge volumes of files between on-premise and cloud repositories
- Define source and target for the task
- To improve performance, define the number of files the task must transfer in a batch
- Can schedule the task

4

© Informatica. Proprietary and Confidential.



The mass ingestion task enables you to transfer, track, and monitor huge volumes of files between on-premise and cloud repositories.

When you create a mass ingestion task, you must define the source from which you want to transfer files and the target to which you want to transfer the files. To improve the performance of a mass ingestion task, you can define the number of files the task must transfer in a batch.

You can configure the task to run on a schedule. You can also configure the task to perform actions, such as compression, decompression, encryption, or decryption of files.

Note: To use the mass ingestion task feature, your organization must have the Mass Ingestion and Mass Ingestion Runtime licenses.

Mass Ingestion Task Sources

- **Supported source types:**

- Local folder
- Advanced FTP V2
- Advanced FTPS V2
- Advanced SFTP V2
- Amazon S3 V2
- Google Cloud Storage V2

- **Supported source types:**

- Hadoop Files V2
- Microsoft Azure Blob Storage V3
- Microsoft Azure Data Lake Store Gen2
- Microsoft Azure Data Lake Store V3
- File Listener – Use a file listener component as a source

Mass Ingestion Task Targets

- **Supported target types:**

- Local folder
- Advanced FTP V2
- Advanced FTPS V2
- Advanced SFTP V2
- Amazon S3 V2
- Amazon Redshift V2
- Google BigQuery V2

- **Supported target types:**

- Google Cloud Storage V2
- Hadoop Files V2
- Microsoft Azure Blob Storage V3
- Microsoft Azure Data Lake Store Gen2
- Microsoft Azure Data Lake Store V3
- Microsoft Azure SQL Data Warehouse V3
- Snowflake Cloud Data Warehouse V2

File Processing Actions

- Encryption
- Decryption
- Compression
- Decompression
- Flatten File Structure
- Virus Scan

7

© Informatica. Proprietary and Confidential.



You can apply the following actions on the files that the mass ingestion task transfers:

Encryption: Uses the Pretty Good Privacy (PGP) method to encrypt files. PGP is an encryption program that provides cryptographic privacy and authentication for data communication. The mass ingestion task encrypts files and flattens the file structure in the target directory.

Decryption: Uses the PGP method to decrypt files. The mass ingestion task decrypts files and flattens the file structure in the target directory.

Compression: Uses Zip, Tar, or Gzip compression methods to compress files. The mass ingestion task compresses files and flattens the file structure in the target directory.

Decompression: Uses Unzip, Untar, or Gunzip decompression methods to decompress files. The mass ingestion task decompresses files and flattens the file structure in the target directory.

Flatten File Structure: Moves the files from multiple folders to a single folder in the target directory.

Virus Scan: Reviews and identifies viruses in the files that the mass ingestion task transfers. Mass ingestion uses the Internet Content Adaptation Protocol (ICAP) to scan files and detect malwares. The ICAP server scans the files and sends a response code as 200 when the scan does not identify any virus in the files. The mass ingestion task fails when the scan detects virus.

Note: The mass ingestion task performs the file processing actions that you configure in a sequential order. The mass ingestion task retains the file structure if you do not configure any action.

Configuring a Mass Ingestion Task

Step 1

Define the task

Step 2

Configure the source

Step 3

Configure the target

Step 4

Configure a schedule (Optional)

8

© Informatica. Proprietary and Confidential.



Configuring a mass ingestion task is a four-step process.

Step 1, you must define the task. Here, you must provide information such as the task name and the secure agent that will run the task.

Step 2, you must configure the source connection from where you want to transfer the files.

Step 3, you must configure the target connection to which you want to transfer the files.

Step 4 is an optional step. If needed, you can configure the task to run on a schedule.

Note: When you configure a local folder as a source connection, you can use the Batch Size option to define the number of files the mass ingestion task must transfer in a batch. The default batch size is set to 5. Defining the batch size improves the performance of the task.

Lab Activity

10-1 Creating a Mass Ingestion Task

In this lab, you will perform the following:

- Create a mass ingestion task

Module Summary

This module showed you how to:

- Discuss mass ingestion task
- Describe mass ingestion task sources and targets
- List file processing actions
- Create a mass ingestion task

10

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 11

Taskflows



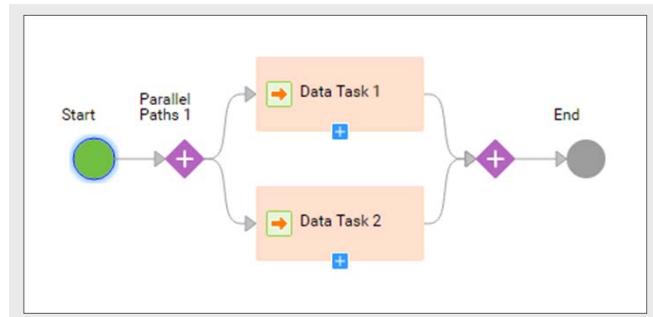
Module Objectives

After completing this module, you will be able to:

- Define a taskflow
- List the steps in a taskflow
- Define Linear taskflow
- List the task types in a Linear taskflow
- Discuss the taskflow templates
- Explain the use of parameters in a taskflow
- Discuss the use of REST API to run taskflows
- Explain the use of a file listener to invoke a taskflow

Taskflow Overview

- A taskflow controls the execution sequence of the tasks
- Create tasks and add them to a taskflow
- Can configure email notifications for a taskflow
- Can configure a taskflow to run on a schedule
- Taskflow allows you to:
 - run parallel tasks
 - use advanced decision making
 - perform other advanced orchestrations



4

© Informatica. Proprietary and Confidential.



A taskflow controls the execution sequence of a mapping task or a synchronization task based on the output of the previous task. You must first create tasks and then add them to a taskflow. You can configure email notifications for a taskflow. Email notifications allows users to receive notifications about the status of the taskflow. You can also configure a taskflow to run on a schedule. A sample taskflow is shown in the image.

With the various tasks and taskflow templates available in IICS, you can run tasks in parallel, use advanced decision making criteria, and perform other advanced orchestrations.

You will see more about taskflow templates later in the module.

Taskflow Steps

- Use taskflow steps to add and orchestrate Data Integration tasks

	Assignment Step	Sets a value to a field
	Data Task Step	Adds a task to a taskflow
	Notification Task Step	Sends an email notification to specified recipients
	Command Task Step	Runs shell scripts or batch commands from a file on the Secure Agent machine
	Subtaskflow Step	Embeds and reuses an existing taskflow
	Decision Step	Enables a taskflow to take different paths based on the value of a specific field
	Parallel Paths Step	Enables a taskflow to run multiple tasks at the same time
	Jump Step	Jumps from one part of the taskflow to another
	End Step	Defines the HTTP status code that must be used when a taskflow completes
	Wait Step	Pauses the taskflow execution for a specific duration

5

© Informatica. Proprietary and Confidential.



- **Assignment Step:** An Assignment step allows you to set a value to a field. With this step, you can use input, output, and temporary fields.
- **Data Task Step:** A Data Task step allows you to add a task to a taskflow. You can configure how the taskflow handles errors and warnings, perform actions based on a schedule, and override runtime parameters.
- **Notification Task Step:** A Notification Task step allows you to send an email notification to specified recipients.
- **Command Task Step:** A Command Task step allows you to run shell scripts or batch commands from a file on the Secure Agent machine. You can use the Command Task outputs to orchestrate subsequent tasks in the taskflow.
- **Subtaskflow Step:** A Subtaskflow step allows you to embed and reuse an existing taskflow. You can configure the input fields to provide the input when you run the taskflow.
- **Decision Step:** You can use a Decision step when you want a taskflow to take different paths based on the value of a specific field.
- **Parallel Paths Step:** You can use a Parallel Paths step when you want a taskflow to run multiple tasks at the same time.
- **Jump Step:** A Jump step allows you to jump from one part of the taskflow to another.
- **End Step:** You can use an End step to define the HTTP status code that must be used when a taskflow completes.
- **Wait Step:** A Wait step allows you to pause the taskflow execution for a specific duration.

Topic

Linear Taskflow



Linear Taskflow Overview

- Simplified version of the Data Integration taskflow
- Cannot control the execution sequence of tasks based on the previous task in the taskflow
- Groups multiple Data Integration tasks and runs them serially
- Can configure email notifications for a linear taskflow
- Can also configure a linear taskflow to run on a schedule

7

© Informatica. Proprietary and Confidential.



A linear taskflow is a simplified version of the Data Integration taskflow. It cannot control the execution sequence of tasks based on the previous task in the taskflow.

A linear taskflow groups multiple Data Integration tasks and runs them serially in the order that you specify.

You can configure email notifications for a linear taskflow. You can also configure the taskflow to run on a schedule.

Linear Taskflow Example

Scenario

Update a list of contacts on a monthly basis

Solution

- Create a linear taskflow with a Synchronization task to Upsert accounts followed by a Synchronization task to Upsert contacts for the accounts
- Schedule the linear taskflow to run each month

Assume that you need to update a list of contacts on a monthly basis. To do this, you can Upsert recent account information and then Upsert contact information for each account. So, you can create a linear taskflow with a Synchronization task to Upsert accounts followed by a Synchronization task to Upsert contacts for the accounts. You can then schedule the linear taskflow to run each month.

Tasks in a Linear Taskflow

- Linear taskflow can include the following task types:
 - Synchronization Task
 - Replication Task
 - Mapping Task
 - Masking Task
 - PowerCenter Task
- You can edit a linear taskflow
- You cannot delete a taskflow that was published, previously run from the taskflow designer, or associated with a schedule

9

© Informatica. Proprietary and Confidential.



A linear taskflow can include Synchronization Task, Replication Task, Mapping Task, Masking Task, and PowerCenter Task.

You can also edit a linear taskflow. If you add a task to a linear taskflow that is currently running, Data Integration does not run the new task until the next time the linear taskflow runs.

Note: You cannot delete a taskflow that was published, previously run from the taskflow designer, or associated with a schedule. You must first unpublish the taskflow, and then delete it.

Topic

Taskflow Templates



Taskflow Templates

The screenshot shows the Informatica Data Integration Service interface. On the left, there is a vertical sidebar with six categories: Basic, Parallel Tasks, Parallel Tasks with Decision, Sequential Tasks, Sequential Tasks with Decision, and Single Tasks. The 'Basic' category is highlighted. To the right of the sidebar is a main content area. At the top of this area, there are three tabs: Mappings, Taskflows (which is selected and highlighted in blue), and Components. Below these tabs is a list of taskflow types with their descriptions:

- Taskflow**: Create control flow logic to run multiple Data Integration tasks.
- Parallel Tasks**: Creates a taskflow with multiple Data Integration tasks that run concurrently.
- Parallel Tasks with Decision**: Creates a taskflow with multiple Data Integration tasks that run concurrently and a Decision step that directs further processing.
- Sequential Tasks**: Creates a taskflow with multiple Data Integration tasks that run consecutively.
- Sequential Tasks with Decision**: Creates a taskflow with multiple Data Integration tasks that run consecutively and a Decision step that directs further processing.
- Single Task**: Creates a taskflow with a single Data Integration task.

At the bottom left of the content area, there is a small orange triangle containing the number '11'. At the bottom right, there is a logo for 'Informatica University' with the text '© Informatica. Proprietary and Confidential.'

You can use a taskflow template instead of creating a taskflow from scratch. IICS provides pre-created templates such as Basic, Parallel Tasks, Parallel Tasks with Decision, Sequential Tasks, Sequential Tasks with Decision, and Single Task, that you can use as per your business requirements. You can find these templates when you create a new asset from the Data Integration Service home page.

Basic Template

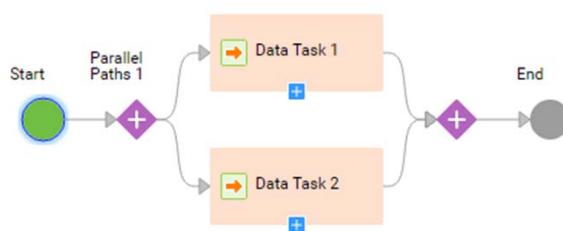
- Provides a canvas with a Start and an End step



Basic Template: The basic template provides a canvas with a Start and an End step. You can create and configure a taskflow based on your business requirements.

Parallel Tasks Template

- Can be used to run two or more Data Integration tasks in parallel



13

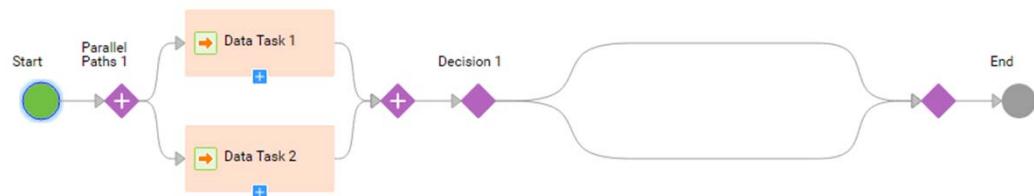
© Informatica. Proprietary and Confidential.



Parallel Tasks Template: You can use this template when you want to run two or more Data Integration tasks in parallel.

Parallel Tasks With Decision Template

- Can be used to run two or more Data Integration tasks in parallel, and then make a decision based on the outcome of any task



14

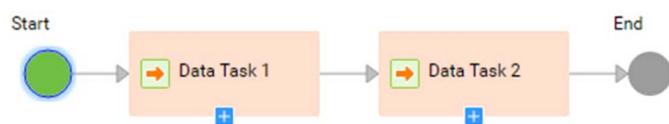
© Informatica. Proprietary and Confidential.



Parallel Tasks with Decision Template: You can use this template when you want to run two or more Data Integration tasks in parallel, and then make a decision based on the outcome of any task.

Sequential Tasks Template

- Can be used to run two Data Integration tasks consecutively



15

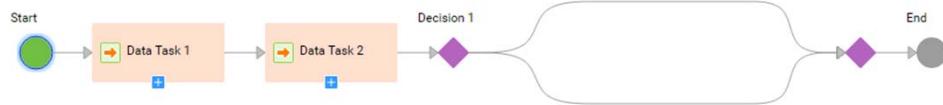
© Informatica. Proprietary and Confidential.



Sequential Tasks Template: You can use this template when you want to run two Data Integration tasks consecutively.

Sequential Tasks With Decision Template

- Can be used to run two Data Integration tasks consecutively, and then make a decision based on the output of either of the two tasks



16

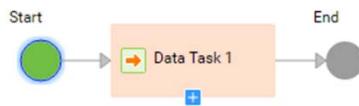
© Informatica. Proprietary and Confidential.



Sequential Tasks with Decision Template: You can use this template when you want to run two Data Integration tasks consecutively, and then make a decision based on the output of either of the two tasks.

Single Task Template

- Can be used to run a Data Integration task on a daily or weekly basis



17

© Informatica. Proprietary and Confidential.



Single Task Template: You can use this template when you want to run a Data Integration task on a daily or weekly basis.

Topic

Parameters in a Taskflow



Parameters in a Taskflow

- Use a taskflow to pass Input and In-Out parameters to a task
- An Input parameter is a placeholder for a value or values in a mapping
- An In-Out parameter is a placeholder for a value that you can pass in to or out of a mapping
- When you add a mapping task to a taskflow, you can override parameter values with the Data Task step or with the Assignment step

19

© Informatica. Proprietary and Confidential.



You can use a taskflow to pass Input and In-Out parameters to a task. An Input parameter is a placeholder for a value or values in a mapping. An In-Out parameter is a placeholder for a value that you can pass in to or out of a mapping.

You can design a mapping with Input or In-Out parameters.

When you add a mapping task to a taskflow, you can override the parameter values with the Data Task step or with the Assignment step. The mapping task passes the parameters to the mapping.

Input Parameters

- You can define the value of the input parameter when you configure the mapping task
- You can override the following subset of mapping input parameters:
 - **Source object:** You can change the object from which the mapping task reads the data
 - **Source connection:** You can change the connection that the mapping task uses to read data from the source
 - **Target connection:** You can change the connection that the mapping task uses to write data to the target
 - **Target object:** You can change the object to which the mapping task writes the data

20

© Informatica. Proprietary and Confidential.



You can define the value of the input parameter when you configure the mapping task. When you use the mapping task in a taskflow, you can override the parameter values for the source object, source connection, target connection, and target object.

When you override the parameter value for the source object, you can change the object from which the mapping task reads the data.

When you override the parameter value for the source connection, you can change the connection that the mapping task uses to read the data from the source.

When you override the parameter value for the target connection, you can change the connection that the mapping task uses to write the data to the target.

When you override the parameter value for the target object, you can change the object to which the mapping task writes the data.

In-Out Parameters

- You can define the value of the in-out parameter when you configure the mapping task
- An in-out parameter can change each time a task runs

21

© Informatica. Proprietary and Confidential.



You can also define the value of the in-out parameter when you configure the mapping task. Unlike input parameters, an in-out parameter can change each time a task runs. You can use the taskflow to override any type of in-out parameters that the mapping task supports.

Guidelines for Using Parameters in Taskflows

- Use the Data Task step to override input or in-out parameters
- For advanced use cases, use the Assignment step to override the parameters
- If you override the same parameter in both the Assignment step and the Data Task step, the taskflow uses the value assigned in the Data Task step
- You can get the current value of the in-out parameter only after the taskflow executes the Data Task step

22

© Informatica. Proprietary and Confidential.



You can use the Data Task step to override input or in-out parameters. However, for advanced use cases, you can override the input or in-out parameters with an Assignment step.

If you define a field to override the same parameter in both the Assignment step and the Data Task step, the taskflow uses the value assigned in the Data Task step.

You can get the current value of the in-out parameter only after the taskflow executes the Data Task step.

Topic

Running a Taskflow Using REST API



Running a Taskflow Using REST API

- You can invoke a taskflow as an API by publishing the taskflow as a service
- Data Integration generates the service URL and the SOAP service URL
- When you invoke a taskflow as an API, you can dynamically provide input parameters for the tasks in the taskflow

24

© Informatica. Proprietary and Confidential.



You can invoke a taskflow as an API by publishing the taskflow as a service. When you publish a taskflow, Data Integration generates the service URL and the SOAP service URL. You can use these endpoint URLs to invoke the taskflow as an API.

When you invoke a taskflow as an API, you can dynamically provide input parameters for the tasks in the taskflow.

Generating the Service URL

1. In the Start tab, enter the user details in the Allowed Users field

Demo_ParallelTaskflow Properties

General	Binding: REST/SOAP
Start	Allowed Groups:
Input Fields	Allowed Users: apiuser@infa.com
Temp Fields	

25

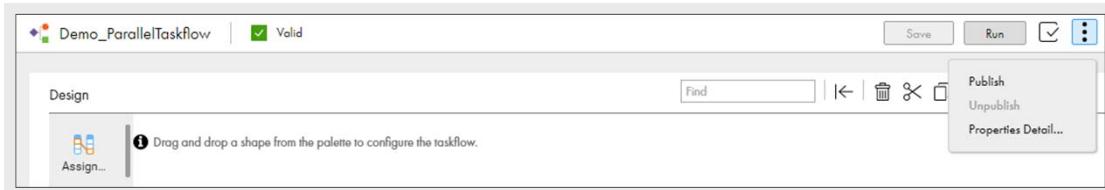
© Informatica. Proprietary and Confidential.



After you create a taskflow, navigate to the **Start** tab under the taskflow properties. In the **Allowed Users** field, enter the user details that can start the taskflow.

Generating the Service URL

2. Publish the taskflow



After saving the taskflow, you must publish it so that the Data Integration can generate the service API for the taskflow.

Generating the Service URL

3. Navigate to the Properties Detail to get the service URL



27

© Informatica. Proprietary and Confidential.



After the taskflow is published, navigate to the Properties Detail to get the service URL.

Generating the Service URL

4. Copy the Service URL

Properties Detail for Demo_ParallelTaskflow

Basic

Unique Name: Demo_ParallelTaskflow
Publication Status: Published
Published On: 2019-06-25 15:54
Published By:
Applies To: * Any *

Endpoints

Service URL: https://na1.dm-us.informaticacloud.com/active-bpel/r/Demo_ParallelTaskflow [View Swagger File](#) [Copy](#)
SOAP Service URL: https://na1.dm-us.informaticacloud.com/active-bpel/soap/Demo_ParallelTaskflow [View WSDL File](#) [Copy](#)

28

© Informatica. Proprietary and Confidential.



The Service URL is specified under the Endpoints section on the Properties Detail page.

You can use the Service URL to start the taskflow from a REST client such as the Postman.

Topic

Invoking a Taskflow Through a File Listener



Invoking a Taskflow Through a File Listener

- Define the binding type as event and select the file listener
- When you publish the taskflow, the taskflow subscribes to the file listener
- When a file event occurs, the file listener invokes the taskflow
- You can monitor the execution of the file listener and the events

30

© Informatica. Proprietary and Confidential.



You can invoke a taskflow through a file listener.

Within a taskflow, you can define the binding type as event and select the file listener. When you publish the taskflow, the taskflow subscribes to the file listener that is defined in it. When a file event occurs, the file listener invokes the taskflow.

For example, if you configure the file listener to listen for new files on a folder, the file listener invokes the associated taskflow each time a new file arrives in the specified folder.

You can monitor the execution of the file listener and the events that occur on each run job of the file listener.

Lab Activity

11-1 Creating a Parallel Taskflow

In this lab, you will perform the following:

- Configure Taskflow using a template

Lab Activity

11-2 Passing In-out Parameters in a Taskflow

In this lab, you will perform the following:

- Create a Taskflow

Lab Activity

11-3 Invoking a Taskflow Through a File Listener

In this lab, you will perform the following:

- Create a Synchronization Task
- Create a File Listener
- Create a Taskflow

Module Summary

This module showed you how to:

- Define a taskflow
- List the steps in a taskflow
- Define Linear taskflow
- List the task types in a Linear taskflow
- Discuss the taskflow templates
- Explain the use of parameters in a taskflow
- Discuss the use of REST API to run taskflows
- Explain the use of a file listener to invoke a taskflow



IICS: Cloud Data Integration Services

 Informatica
University

Module 12

Advanced Options



Module Objectives

After completing this module, you will be able to:

- Explain Primary Key (PK) chunking
- Discuss a PK chunking use case
- Define optimal chunk size
- List the benefits of using PK chunking
- Explain Lookup SQL override
- Discuss use cases for Lookup SQL override
- Define rules and guidelines for overriding the Lookup query

3

© Informatica. Proprietary and Confidential.



Topic

PK Chunking



PK Chunking Overview

- Advanced option that you can configure for Salesforce objects
- In IICS, PK chunking is available for bulk API tasks that use Salesforce API version 32 or above
- Enable PK chunking on the Schedule tab of the task
- Splits bulk queries into small chunks
- Enables a query to fetch huge volumes of records for each batch job run
- Uses Salesforce Object Query Language (SOQL) queries to chunk records

5

© Informatica. Proprietary and Confidential.



PK chunking is an advanced option that you can configure for Salesforce objects. Salesforce enables PK chunking for its bulk API jobs. In IICS, this feature is available for bulk API tasks that use Salesforce API version 32 or above. You can enable PK chunking on the Schedule tab of the task.

PK chunking enables you to split bulk queries into small chunks based on the primary keys of the queried records. Each chunk is processed as a separate batch and you must download the results of each batch separately. PK chunking enables a query to fetch huge volumes of records for each batch-job run. Salesforce recommends that you must enable PK chunking for objects that have more than 10 million records.

PK chunking uses Salesforce Object Query Language to chunk records.

PK Chunking – Use Case

- Organization has more than 10 million records to transfer from a Salesforce org
- Organization faces connection issues and Synchronization task times out while transferring records
- Organization has Salesforce connection with API version 32 or higher in IICS
- Organization enables PK chunking with a chunk size of 25,000
- Each query fetches a maximum of 25,000 records
- Using PK chunking, organization quickly transfers records to their on-premise database tables

6

© Informatica. Proprietary and Confidential.



Here is a scenario where you can use PK chunking.

Assume that an organization wants to transfer more than 10 million records from a Salesforce org to their on-premise database tables. Due to the large number of records, the organization frequently faces connection issues and the Synchronization task times out while transferring records. The organization has a Salesforce connection with API version 32 or higher in IICS.

This is an ideal situation where the organization can enable PK chunking in their IICS org, with the chunk size configured to 25 thousand. In this case, the bulk queries sent to Salesforce has a WHERE clause based on the primary key of the table. Each query fetches a maximum of 25 thousand records. Therefore, by using PK chunking, the organization can quickly transfer records to their on-premise database tables.

Optimal Chunk Size

- Smaller chunk sizes increase performance and speed of queries
 - Results in more Bulk API batches
- Larger chunk sizes results in fewer Bulk API batches
 - Affects the performance if the chunk size is too large
- Experiment with chunk sizes to find the optimal chunk size for each data set
 - Chunk size cannot be larger than 250,000 records

7

© Informatica. Proprietary and Confidential.



Smaller chunk sizes increase the performance and speed of the queries. However, smaller chunk sizes result in more Bulk API batches and require more time to complete the job. Larger chunk sizes results in fewer Bulk API batches. However, if the chunk size is too large, it can cause the job to time-out frequently.

To find the optimal chunk size for each data set, you must experiment with the chunk sizes. Based on the number of records you want to fetch, you can specify a smaller or larger chunk size. Therefore, you must create different chunk sizes and check the optimal chunk size.

It is also important to note that the maximum chunk size cannot be larger than 250 thousand records.

Benefits of PK Chunking

- Allows you to use it for most standard objects and all custom objects in Salesforce
- Improves job performance as smaller chunks of data run with each Bulk API batch
- Restarts job from the point of failure

You can use PK chunking for most standard objects and all custom objects in Salesforce. PK chunking improves job performance as smaller chunks of data run with each Bulk API batch. When a job fails, PK chunking can restart the job from the point of failure.

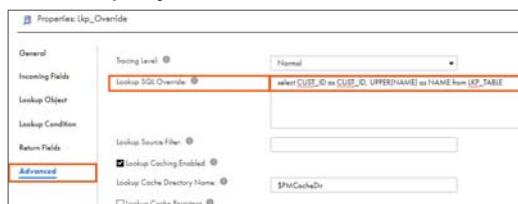
Topic

Lookup SQL Override



Lookup SQL Override

- When a mapping includes a Lookup transformation, the mapping task queries the lookup object
- The mapping task runs a default lookup query when the first row of data enters the Lookup transformation
- If the Lookup transformation performs a relational lookup, you can override the default query



10

© Informatica. Proprietary and Confidential.

- The default query contains a SELECT statement that includes all lookup fields in the mapping
- The SELECT statement also contains an ORDER BY clause that orders all columns in the same order in which they appear in the Lookup transformation
- If you want to change the ORDER BY clause, you must add a WHERE clause or transform the lookup data before it is cached
- You can override the default query on the **Advanced** tab of the Lookup transformation



When a mapping includes a Lookup transformation, the mapping task queries the lookup object based on the fields and properties that you configure in the Lookup transformation.

The mapping task runs a default lookup query when the first row of data enters the Lookup transformation. If the Lookup transformation performs a relational lookup, you can override the default query.

The default query contains a SELECT statement that includes all lookup fields in the mapping. The SELECT statement also contains an ORDER BY clause that orders all columns in the same order in which they appear in the Lookup transformation. To view the default query, you must run the mapping task. The default query appears in the log file.

If you want to change the ORDER BY clause, you must add a WHERE clause or transform the lookup before it caches the data.

You can override the default query on the **Advanced** tab of the Lookup transformation.

Lookup SQL Override – Example

- A Lookup transformation returns the following fields from Microsoft SQL Server table ALC_ORDER_DETAILS:

Name	Type	Precision	Scale	Origin
1 ORDERID	integer	10	0	ALC_ORDER_DETAILS
2 PRODUCTID	integer	10	0	ALC_ORDER_DETAILS
3 UNITPRICE	decimal	10	4	ALC_ORDER_DETAILS
4 QUANTITY	integer	10	0	ALC_ORDER_DETAILS
5 DISCOUNT	double	15	0	ALC_ORDER_DETAILS

- The transformation uses the following lookup condition:
 - ORDERID=in_ORDERID

11

© Informatica. Proprietary and Confidential.



Lookup SQL Override – Example (continued)

- When you run the mapping task, the following default query appears in the log file:

```
LKPDP_1> DBG_21097 [2020-01-31 14:11:33.509] Lookup Transformation [1kp_ALC_ORDER_DETAILS]: Default sql to create lookup cache:  
SELECT PRODUCTID,UNITPRICE,QUANTITY,DISCOUNT,ORDERID  
FROM "icsauto"."ALC_ORDER_DETAILS" ORDER BY ORDERID,PRODUCTID,UNITPRICE,QUANTITY,DISCOUNT
```

- To override the ORDER BY clause and sort by PRODUCTID, enter the following query in the **Lookup SQL Override** field on the **Advanced** tab:

```
SELECT PRODUCTID AS PRODUCTID, UNITPRICE AS UNITPRICE, QUANTITY AS QUANTITY, DISCOUNT AS DISCOUNT, ORDERID AS ORDERID  
FROM "icsauto"."ALC_ORDER_DETAILS" ORDER BY PRODUCTID --
```

- When you run the mapping task again, the following query appears in the log file:

```
LKPDP_1> DBG_21312 [2020-01-31 14:14:36.734] Lookup Transformation [1kp_ALC_ORDER_DETAILS]: Lookup override sql to create cache:  
SELECT PRODUCTID AS PRODUCTID, UNITPRICE AS UNITPRICE, QUANTITY AS QUANTITY, DISCOUNT AS DISCOUNT, ORDERID AS ORDERID  
FROM "icsauto"."ALC_ORDER_DETAILS" ORDER BY PRODUCTID -- ORDER BY ORDERID,PRODUCTID,UNITPRICE,QUANTITY,DISCOUNT
```

Lookup SQL Override – Use Cases

1

Customize the SQL queries in a Lookup transformation

2

Meet complex selection or join criteria

3

Modify data types or formats in the lookup table

13

© Informatica. Proprietary and Confidential.



Following are a few scenarios where you can use the Lookup SQL override option.

You can use Lookup SQL Override:

- When you want to customize the SQL queries in a lookup transformation
- When you want to meet complex selection or join criteria, or
- When you want to modify the data types or formats in the lookup table. You can use the database functions to adjust the data types or formats in the lookup table to match the data types and formats of fields in the mapping.

Overriding the Lookup Query

Rules and Guidelines

- You can override the lookup SQL query for relational lookups
- If you override the lookup query, enable lookup caching for the transformation
- Enter the entire SELECT statement using the syntax required by the database
- Enclose all database reserved words in quotes
- Include all lookup and return fields in the SELECT statement
- Use an alias for each column in the query

14

© Informatica. Proprietary and Confidential.



Overriding the Lookup Query (continued)

Rules and Guidelines

- If the ORDER BY clause contains multiple columns, enter the columns in the same order as the fields in the lookup condition
- If the multiple Lookup transformations share a lookup cache, use the same lookup SQL override for each Lookup transformation
- You cannot include parameters in the lookup SQL override
- If you configure a lookup SQL override and a lookup source filter in the same transformation, the mapping task ignores the filter

Module Summary

This module showed you how to:

- Explain Primary Key (PK) chunking
- Discuss a PK chunking use case
- Define optimal chunk size
- List the benefits of using PK chunking
- Explain Lookup SQL override
- Discuss use cases for Lookup SQL override
- Define rules and guidelines for overriding the Lookup query

16

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 13

Hierarchical Connectivity



Module Objectives

After completing this module, you will be able to:

- Explain Web Service
- List the types of web services
- Discuss REST Web Services
- Compare JSON and XML response types
- Discuss REST V2 connector
- Describe Web Services transformation
- Discuss Hierarchy Parser transformation
- Explain Hierarchical Schemas
- Discuss Hierarchy Builder transformation

3

© Informatica. Proprietary and Confidential.



Topic

REST Web Services



Web Service Overview

- A web service integrates applications and uses open standards, such as SOAP, WSDL, and XML
- SOAP is the communication protocol for web services
- WSDL is an XML schema that describes the protocols, formats, and signatures of the web service operations
- Web service operations include requests for information, requests to update data, and requests to perform tasks
- Web service uses transport protocols such as HTTP, FTP, and SMTP to send messages between network applications

5

© Informatica. Proprietary and Confidential.



A web service integrates applications and uses open standards, such as Simple Object Access Protocol (SOAP), Web Services Description Language (WSDL), and Extensible Markup Language (XML).

SOAP is the communication protocol for web services.

WSDL is an XML schema that describes the protocols, formats, and signatures of the web service operations.

Web service operations include requests for information, requests to update data, and requests to perform tasks.

Web service uses transport protocols such as HTTP, FTP, and SMTP to send messages between network applications.

The World Wide Web (WWW) uses HTTP protocol to send data.

Web Service Payload

- A Payload is the content that sends the message using the transport protocol
- Types of XML payload messages:

Atom

MIME

Binary Files

- JSON is a non-XML type of payload
 - It is lightweight and consumes less data for file transfers

6

© Informatica. Proprietary and Confidential.



A payload is the content that sends the message using the transport protocol. The payload of most web services is in XML format. Atom, MIME, and Binary Files are some of the different types of XML payload messages.

Atom is an XML based file format that contains certain tags. Atom is a standard way of describing the data.

Multipurpose Internet Mail Extension (MIME), transports multimedia information through synchronous internet mail.

Binary Files can be BASE 64 encoded or text files. A web service uses Binary files as payloads for FTP transfers.

JavaScript Object Notation (JSON) is a non-XML type of payload. JSON is a self-describing language, which is similar to XML. JSON is lightweight and consumes less data for file transfers. While IICS supports JSON file formats, Informatica PowerCenter currently does not support this format.

Types of Web Services

REST	SOAP	XML over HTTP
<ul style="list-style-type: none"> • Considers data and functionality as resources • Sends a query string to the web service and generates an XML or JSON format file as output 	<ul style="list-style-type: none"> • Standard protocol to exchange messages between web service and client in XML format • Requires WSDL to generate code • SOAP web services are of two types – RPC Encoded and Document Style <ul style="list-style-type: none"> • RPC Encoded: asks for the data by the name of the procedure and for certain data types • Document Style: SOAP body contains an XML document that validates against a pre-defined XML schema document 	XML message is posted over HTTP to access web service resources

7

© Informatica. Proprietary and Confidential.



REST web service: The REST web service considers data and functionality as resources. You can access the resources with the help of Uniform Resource Identifier (URI). REST sends a query string to the web service and generates an XML or JSON format file as output. REST web service is simple and easy to test.

SOAP web service: SOAP is the standard protocol to exchange messages between the web service and the client in XML format. It requires WSDL to generate code. WSDL is an XML file that defines the process to implement a web service. SOAP web services are complex and difficult to manage manually. SOAP web services are of two types. The first type is known as Remote Procedure Call (RPC) encoded and the second type is the Document Style. The RPC encoded type is procedural in nature. This means that it asks for the data by the name of the procedure and for certain data types. In the Document Style type, the SOAP body contains an XML document that validates against a pre-defined XML schema document. An XML message creates all the required information. You send the request and get the response in XML format.

The third type of webservice is XML over HTTP. In this type of web service, an XML message is posted over HTTP to access web service resources and it returns the response in XML format.

REST Web Service

- Uses HTTP protocol for data exchange
- Every component in REST web service is a resource
- You can access the resources using HTTP standard methods like GET, POST, PUT, or DELETE
- Request message consists of an end point URL, type of operation it performs, request header, and a request body

The REST web service uses HTTP protocol for data exchange.

Every component in the REST web service is a resource and you can access these resources by using HTTP standard methods like GET, POST, PUT, or DELETE.

To access a resource, you must send a message or request to the server that contains the resource. Each message consists of an end point URL, type of operation it performs, request header, and a request body.

Components of REST Web Service Message

URL	<ul style="list-style-type: none">URL or URI identifies each resource in REST architectureParameters in the URL that helps to process a request are known as URL request parameters
Method	<ul style="list-style-type: none">Defines HTTP operation to performOperations can be GET, POST, PUT, or DELETE
Request Header	<ul style="list-style-type: none">Contains metadata of HTTP requestParameters that are passed in the header are called request parameters
Request Body	<ul style="list-style-type: none">Contains actual message contentParameters that are passed in the body are called request parameters

9

© Informatica. Proprietary and Confidential.



URL: URL or URI identifies each resource in the REST architecture. The purpose of a URL is to locate a resource on the server that hosts the web service. The parameters in the URL that helps to process a request are known as “URL request parameters”.

Method: Method represents the HTTP operation that you want to perform. The operations that you can perform are GET, POST, PUT, or DELETE.

Request Header: Header contains the metadata for the HTTP message request. Some examples of metadata are, client type, format supported by the client, format of the message body, cache settings, and so on. To post a request, you can pass parameters in the header. These parameters are called header request parameters.

Request Body: Body is the actual message content. In a REST web service, the representation of resources is present in the body of a message. You can pass parameters in the request body. The parameters passed in the request body are called request parameters.

Example of REST Web Service

```
GET http://del.icio.us/api/peej/bookmarks/?start=1&end=2
<?xml version="1.0"?>
<bookmarks start="1" end="2"
next="http://del.icio.us/api/peej/bookmarks?start=3&end=4">
<bookmark url="http://www.example.org/one" tags="example,test"
href="http://del.icio.us/api/peej/bookmarks/a211528fb51
08cddaa4b0d3aecdbdcf" time="2005-10-21T19:07:30Z">Example of a bookmark
</bookmark>
</bookmarks>
```

10

© Informatica. Proprietary and Confidential.

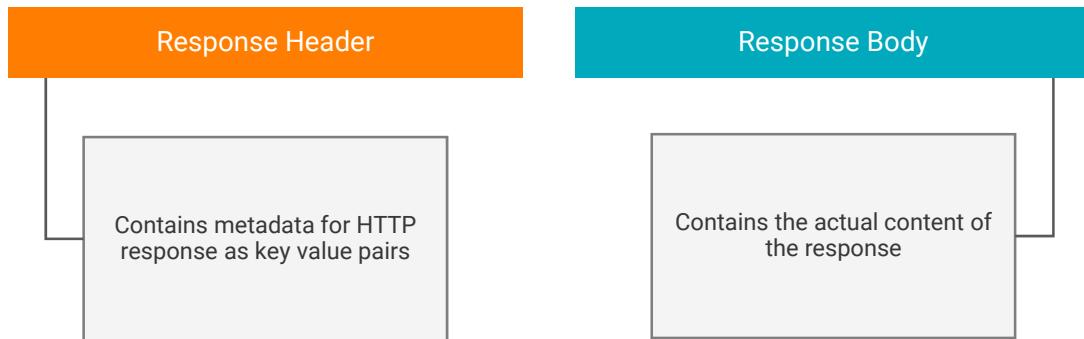


Here is an example of a REST web-service request message in XML format.

The first line uses the HTTP method 'GET' and the end point URL. The complete XML payload message starts from the next line.

This example message retrieves all the bookmarks and tags of a user.

REST Web Services Response



11

© Informatica. Proprietary and Confidential.



The REST web service response is an HTTP response that contains a response header and a response body.

The Response Header contains metadata for HTTP response as key value pairs. For example, content length, type, response date, server type, and so on.

The Response Body contains the actual content of the response message. The response message can be in JSON or XML format.

REST Web Service (JSON Response)

- Query string contains the input to the web service
- REST API returns the response in JSON format
- JSON is lightweight and transfers less information about the data over the internet



```
[{"coordinates": null, "favorited": false, "truncated": false, "created_at": "Wed Aug 29 17:12:58 +0000 2012", "id": "2340859602684612608", "entities": {"urls": [{"expanded_url": "https://dev.twitter.com/blog/twitter-certified-products", "url": "https://t.co/Wj38xAnT", "indices": [92, 113]}], "display_url": "dev.twitter.com/blog/twitter-c\u2026"}}
```

12

© Informatica. Proprietary and Confidential.



This example demonstrates how REST API sends a request and receives a response in JSON format.

The highlighted section represents the query string. The query string contains the input to the web service to process the data and return the response.

In most cases, REST API returns the response in JSON format.

JSON is lightweight and transfers less information about the data over the internet. For security of data transfer, JSON is a better choice over XML.

REST Web Service (XML Response)

- Advantages of XML response:
 - Easy to understand
 - Easy to test
- Disadvantage of XML response:
 - Not flexible
- Applications that return XML response:
 - LinkedIn
 - Facebook

13

© Informatica. Proprietary and Confidential.



REST web services can also return a response in XML format.

There are few advantages of a REST web service returning an XML response. An XML response is easy to understand and test.

However, the disadvantage of an XML response is that it is not flexible. So, it is difficult to generate code from a REST web service if the HTTP returns a response in an XML format.

LinkedIn and Facebook are some of the applications that returns an XML response.

JSON versus XML

JSON	XML
Extremely lightweight	Extremely verbose
Used commonly in JavaScript enabled web clients	Better suited for Adobe Flash, Oracle ADF, Microsoft InfoPath, or open source solutions such as NetBeans
Deals well with atomic values or lists or hashes of atomic values	Deals well with extremely complex unstructured data
Use for schema-less and entity-free data	Use for namespace and well-formed mixed content documents

14

© Informatica. Proprietary and Confidential.



REST V2 Connector

- You can use REST V2 Connectors:
 - To interact with web service applications that support REST API
 - In a Source transformation, Target transformation, or midstream in a Web Services transformation
 - Midstream in a mapping to pass a single or multiple requests to a web service application and process the response data
- In a source, target, and midstream transformation, you can use the REST methods such as GET, PUT, POST, or DELETE
- Configure one of the REST authentication types such as BASIC, DIGEST, or OAUTH Version 1
- Use REST V2 Connector to process XML and JSON data

15

© Informatica. Proprietary and Confidential.



You can use REST V2 Connectors to interact with web service applications that support REST API. You can also use them in a Source transformation, Target transformation, or midstream in a Web Service transformation. You can use REST V2 Connectors midstream in a mapping to pass a single or multiple requests to a web service application and process the response data. You can also pass data obtained from multiple transformations in the mapping pipeline and process the data. In a source, target, and midstream transformation, you can use the REST methods such as GET, PUT, POST, or DELETE.

When you create a REST V2 connection, you can configure one of the REST authentication types such as BASIC, DIGEST, or O AUTH Version 1.

Finally, you can use REST V2 Connectors to process XML and JSON data.

REST Web Services – Best Practices

- Set the URL request parameter, form request parameter, and header request parameter in such a way that it fetches maximum data from the REST endpoint server
- When the endpoint supports both XML and JSON based responses, use JSON based response
- Use sample response XML or JSON file field to manually control the metadata
- When you receive a response, you must download raw response data from REST web service in a separate folder

16

© Informatica. Proprietary and Confidential.



You must set the URL request parameter, form request parameter, and header request parameter in such a way that it fetches maximum data from the REST endpoint server. This allows the connector to analyze and store more metadata.

When the endpoint supports both XML and JSON based responses, use JSON based response, as it is more stable.

You must use a sample response XML or JSON file field to manually control the metadata.

When you receive a response, you must download raw response data from REST web service in a separate folder. The raw data helps you to understand if the response is minified or not. In computer programming languages, 'minification' is the process of removing all unnecessary characters from source code without changing its functionality.

Topic

Web Services Transformation



Web Services Transformation

- Web Services transformation connects to a web service as a web service client to access, transform, or deliver data
- Web service client request and web service response are SOAP messages
- Mapping task processes SOAP messages with document or literal encoding
- Web Services transformation does not support RPC encoded or document encoded WSDL files
- SOAP request and response messages can contain hierarchical data

18

© Informatica. Proprietary and Confidential.



A Web Services transformation connects to a web service as a web-service-client, to access, transform, or deliver data. The web-service client-request and the web-service response are SOAP messages. The Mapping task processes SOAP messages with document or literal encoding. The Web Services transformation does not support RPC encoded or document encoded WSDL files.

SOAP request messages and response messages can contain hierarchical data, such as data that follows an XML schema.

Using Web Services Transformation

Create a Web Services Consumer connection and use a WSDL URL and an endpoint URL



Define a business service



Use the Cloud Mapping Designer to configure the Web Services transformation in a mapping

19

© Informatica. Proprietary and Confidential.



To use a Web Services transformation, you must first create a Web Services Consumer connection and use a WSDL URL and an endpoint URL. Then, you must define a business service. A business service is a web service with configured operations. Finally, you must use the Cloud Mapping Designer to configure the Web Services transformation in a mapping.

Topic

Hierarchy Parser Transformation



Hierarchy Parser Transformation

- Converts hierarchical input into relational output
- Processes XML or JSON input from the upstream transformation and provides relational output to the downstream transformation
- Configures a hierarchical schema that defines the expected hierarchy of the output data
- Converts a hierarchical input based on the hierarchical schema that you associate with the transformation

21

© Informatica. Proprietary and Confidential.



The Hierarchy Parser transformation converts hierarchical input into relational output. The transformation processes XML or JSON input from the upstream transformation and provides relational output to the downstream transformation.

You can configure a hierarchical schema that defines the expected hierarchy of the output data from a sample file or schema file. The Hierarchy Parser transformation converts hierarchical input based on the hierarchical schema that you associate with the transformation. You can use an existing hierarchical schema or create a new schema.

Hierarchical Schemas

- A hierarchical schema is based on a schema file or sample file that you import to Data Integration
- The schema defines the expected hierarchy of the input data
- You can create hierarchical schema in two ways:
 - Create a standalone hierarchical schema and associate it with any transformation
 - Create a hierarchical schema within a specific transformation
- When you create a standalone hierarchical schema, you can import a JSON sample file or .xsd file as the basis of the schema
- You can create, edit, or delete a hierarchical schema

22

© Informatica. Proprietary and Confidential.



A hierarchical schema is based on a schema file or sample file that you import to Data Integration. If you import a sample file, Data Integration generates a schema based on the structure of the sample file. The schema defines the expected hierarchy of the input data.

You can create a hierarchical schema in two ways. You can either create a standalone hierarchical schema that can be associated with any transformation that you choose; or, you can create the schema within a specific transformation.

When you create a standalone hierarchical schema, you can import a JSON sample file or .xsd file as the basis of the schema.

You can create, edit, or delete a hierarchical schema. However, if you have used the hierarchical schema in a transformation, you cannot edit or delete it.

Points to Remember

- Ensure that you have the Hierarchy Parser transformation on the Transformation palette in the mapping designer
- Confirm that the Data Transformation Package is assigned to your org
- When you create a mapping with Hierarchy Parser transformation, you must provide a text file as the source

23

© Informatica. Proprietary and Confidential.



You must first ensure that you have the Hierarchy Parser transformation on the Transformation palette in the mapping designer. If the transformation is not available, you can contact Informatica Cloud Support to enable the feature. After the transformation is assigned to your org, you can configure the privileges for the Hierarchy Parser transformation.

You must also confirm that the Data Transformation Package is assigned to your org. If the package is not assigned, you must restart the secure agent and then check if the package is assigned.

Finally, when you create a mapping with Hierarchy Parser transformation, you must provide a text file as the source. The text file contains the physical directory location of the XML file.

Topic

Hierarchy Builder Transformation



Hierarchy Builder Transformation

- Converts relational input into hierarchical output
- Processes relational input from the upstream transformation and provides JSON or XML output to the downstream transformation
- Configures a hierarchical schema that defines the expected hierarchy of the output data
- Produces hierarchical output based on the hierarchical schema that you associate with the transformation and the way that you map the data

25

© Informatica. Proprietary and Confidential.



The Hierarchy Builder transformation converts a relational input into hierarchical output.

The transformation processes relational input from the upstream transformation and provides JSON or XML output to the downstream transformation.

You can configure a hierarchical schema that defines the expected hierarchy of the output data from a sample file or schema file. The Hierarchy Builder transformation produces hierarchical output based on the hierarchical schema that you associate with the transformation and the way that you map the data. You can use an existing hierarchical schema or create a new schema.

Hierarchy Builder Transformation – Field Mapping

- Configure the field mapping in a Hierarchy Builder transformation
- Field mapping defines the link between relational elements and schema elements

The screenshot shows the 'Properties: HierarchyBuilder' dialog box with the 'Field Mapping' tab selected. On the left, under 'Relational Fields: (6 of 7 fields mapped)', there are two groups: 'Company' and 'Emp_2'. Under 'Company', 'CompanyName' is mapped to 'Key'. Under 'Emp_2', 'Name' is mapped to 'Key'. On the right, under 'Hierarchy Fields: (6 of 6 fields mapped)', there are three groups: 'Employees*', 'Employee*', and 'Subdivision*'. 'Employees*' contains 'Name' and 'Address', both mapped to 'Company'. 'Employee*' contains 'EmployeeID', 'Department', 'Building', and 'Room', all mapped to 'Emp_2'. 'Subdivision*' contains 'Building' and 'Room', both mapped to 'Emp_2_Building'.



26

© Informatica. Proprietary and Confidential.



After you select a schema, you must configure the field mapping in a Hierarchy Builder transformation. Field mapping defines the link between relational elements and schema elements to provide the hierarchical output. You can configure the field mapping in the Field Mapping tab of the Properties panel.

If you have more than one group, you must define primary and foreign keys in the Field Mapping editor. The Field Mapping editor displays the relational fields on the left side and schema elements on the right side. To link a relational field to schema element, drag the relational element to the schema element. The Mapped Field column shows the relational field to which the schema element is mapped.

If the input relational fields constitute just one group, it treats the data as denormalized input and there is no need to define primary or foreign keys.

The image shows that the transformation has two relational groups. The "CompanyName" field is the primary key in the "Company" group and the "Name" field is the foreign key that links to the "Employee" group. The "Name" field is the primary key in the Employee group.

Lab Activity

13-1 Creating a Mapping using a REST V2 Connector

In this lab, you will perform the following:

- Create a REST connection using the REST V2 connector
- Get the JSON message and write it to a Flat File

Lab Activity

13-2 Using Web Services Transformation in a Mapping

In this lab, you will perform the following:

- Use Web Services transformation in a mapping

Lab Activity

13-3 Creating a mapping using Hierarchy Parser Transformation

In this lab, you will perform the following:

- Import a hierarchical schema
- Create a mapping using Hierarchy Parser transformation

Lab Activity

13-4 Creating a mapping using Hierarchy Builder Transformation

In this lab, you will perform the following:

- Import a hierarchical schema
- Create a mapping using Hierarchy Builder transformation

30

© Informatica. Proprietary and Confidential.



Module Summary

This module showed you how to:

- Explain Web Service
- List the types of web services
- Discuss REST Web Services
- Compare JSON and XML response types
- Discuss REST V2 connector
- Describe Web Services transformation
- Discuss Hierarchy Parser transformation
- Explain Hierarchical Schemas
- Discuss Hierarchy Builder transformation

31

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 14

Intelligent Structure Model



Module Objectives

After completing this module, you will be able to:

- Discuss Intelligent Structure Model
- Explain Intelligent Structure Discovery process
- List the steps to create an Intelligent Structure Model
- Refine the discovered structure
- Edit the Intelligent Structure Model
- Use the Intelligent Structure Model in a Structure Parser transformation

Overview of Intelligent Structure Model

- An Intelligent Structure Model is an asset based on a sample file that contains data with little or no structure
- Intelligent Structure Discovery:
 - Determines the patterns of the sample file and creates a model that can be used to transform, parse, and generate output groups
 - Automatically interprets input data and discover the patterns, repetitions, relationships, and types of data in unstructured files
 - Creates a model that defines the expected output data

4

© Informatica. Proprietary and Confidential.



An Intelligent Structure Model is an asset that is based on a sample file that contains data with little or no structure. Intelligent Structure Discovery determines the underlying patterns of the sample file and creates a model that can be used to transform, parse, and generate output groups.

Normally, long, complex files with little or no structure can be difficult to parse. Intelligent Structure Discovery can automatically interpret input data and discover the patterns, repetitions, relationships, and types of data in unstructured files.

Intelligent Structure Discovery creates a model that defines the expected output data. You can use an Intelligent Structure Model in mappings to parse unstructured, semi-structured, or structured data.

Using Intelligent Structure Discovery

- You can use sample JSON files to enrich an existing intelligent structure
- You can view, edit, refine, and export the Intelligent Structure Model
- You can create models for:
 - Microsoft Excel
 - Microsoft Word tables
 - PDF forms
 - CSV files
 - Unstructured text files

5

© Informatica. Proprietary and Confidential.



You can use sample JSON files to enrich an existing intelligent structure that is based on the sample JSON file. Intelligent Structure Discovery adds new data that it finds in the sample files to the structure.

After you create an Intelligent Structure Model you can view, edit, refine, and export it.

You can create models for Microsoft Excel, Microsoft Word tables, PDF forms, CSV files, and unstructured text files. You can also create models for structured data such as XML and JSON files.

Intelligent Structure Discovery Process

- Create an intelligent structure using Intelligent Structure Discovery
- After you provide a sample file:
 - Intelligent Structure Discovery determines the underlying and repeating patterns of the data
 - Creates a structure that represents the data fields and their relationships
- You can associate an Intelligent Structure Model with a Structure Parser transformation and use it in a mapping

6

© Informatica. Proprietary and Confidential.



You can create an intelligent structure using Intelligent Structure Discovery.

After you provide a sample file, Intelligent Structure Discovery determines the underlying and repeating patterns of the data. It then creates a structure that represents the data fields and their relationships. You can quickly model data for files whose structure is complex and takes time to interpret.

After you save an Intelligent Structure Model, you can associate it with a Structure Parser transformation, and use it in a Data Integration mapping.

Selecting a Sample File

- Use a simple sample file similar to the files used in production
- Data types of fields in the sample file must match the data types of fields in the production file
- Use a simplified sample file to generate the model:
 - If the input data has tables provide a table with just a few sample rows of data
 - If the input is a JSON file that contains repeating groups of data, limit the number of repetitions
- If the intelligent structure does not match the input file, there may be a large amount of unidentified data and data loss
- Intelligent Structure Discovery can parse files that have some variations in the date format

7

© Informatica. Proprietary and Confidential.



When you select a sample file to base an intelligent structure, it needs to be similar to the files used in production. The data types of fields in the sample file must match the data types of fields in the production file.

You must use a simplified sample file to generate the model. For example, if the input data has tables, you must provide a table with just a few sample rows rather than many rows of data. If you use a JSON input file that contains repeating groups of data, you must limit the number of repetitions.

If the intelligent structure does not match the input file that you plan to use, or only partially matches the input file, there may be a large amount of unidentified data and data loss. However, Intelligent Structure Discovery can parse files that have some variations in the date format.

Selecting a Sample File

- The model can address data drift in certain cases
- Sample data that is used to create the model

96E8FA61FD3B02CA.dxTomcat	[18/Apr/2019:14:54:24 +0300]	GET /dx-console/ HTTP/1.1' 302
96E8FA61FD3B02CA.dxTomcat	[18/Apr/2019:14:54:25 +0300]	GET /dx-console/com.informatica.b2b.dx.main.jsp/HTTP/1.1'302
96E8FA61FD3B02CA.dxTomcat	[18/Apr/2019:14:54:25 +0300]	GET /dx-console/login.jsp HTTP/1.1' 200 4472

- Data that you parse with the model

96E8FA61FD3B02CA.dxTomcat this_is_new_version_data	[4-Jun-2019 05:14:24WIT]	GET /dx-console/ HTTP/1.1' 302
96E8FA61FD3B02CA.dxTomcat this_is_new_version_data	[4-Jun-2019 05:14:24WIT]	GET /dx-console/com.informatica.b2b.dx.main.jsp/HTTP/1.1'302
96E8FA61FD3B02CA.dxTomcat this_is_new_version_data	[4-Jun-2019 05:14:24WIT]	GET /dx-console/login.jsp HTTP/1.1' 200 4472

The model can also address data drift in certain cases.

The first image is of a sample data used to create the model.

The second image is of the data that you parse with the model.

As you can see, some of the data has drifted and is in a different location with respect to the other data. However, the model can parse this variation.

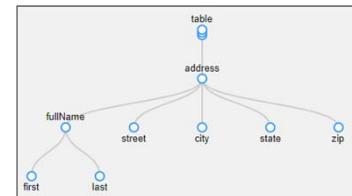
Intelligent Structure Example

first	last	street	city	state	zip
Carrine	Stone	17 Torrent	Livingston	PA	10173
Poona	Tillkup	52 Perez A	Livingston	PA	10256
Tasha	Herrera	158 Shiraz	Kensington	WA	33823

CSV Input File



Intelligent Structure Discovery



Intelligent Structure Model

Look at an example of an Intelligent Structure Discovery process.

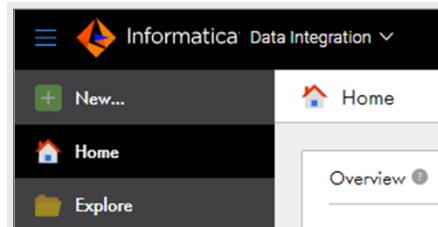
Consider that you want to create an intelligent structure for a CSV input file that contains customer data. As you can see, the input file contains the customer's name and address information. When you run an Intelligent Structure Discovery on the input file, it creates an Intelligent Structure Model. In the model you can see that Intelligent Structure Discovery creates nodes that represent the fields in the input file, such as first, last, street, city, state, and zip.

The structure represents the data fields and also defines the relationships between the fields. In this case, Intelligent Structure Discovery recognizes that 'Carrine' is the first name and 'Stone' is the last name of a person. So, it groups the nodes 'first' and 'last' together under the 'fullName' node, to represent the relationship of the data with each other.

Intelligent Structure Discovery also recognizes that the entire data in the input file represents addresses. So, it groups the data under a parent node 'address'.

Creating an Intelligent Structure Model

1. In the Data Integration home page, click **New**.



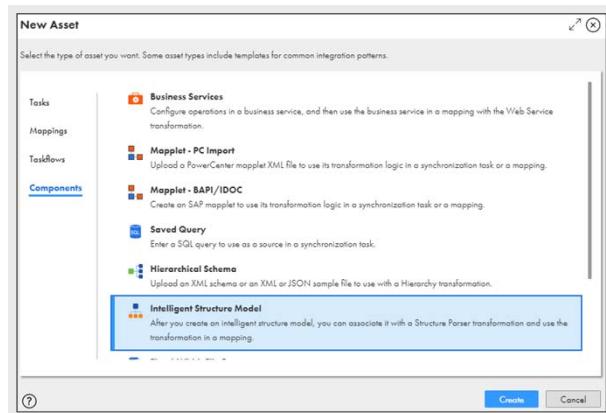
10

© Informatica. Proprietary and Confidential.



Creating an Intelligent Structure Model

1. In the Data Integration home page, click New.
2. In the New Asset window, click Components, select Intelligent Structure Model, and click Create.



11

© Informatica. Proprietary and Confidential.



Creating an Intelligent Structure Model

1. In the Data Integration home page, click New.
2. In the New Asset window, click Components, select the Intelligent Structure Model, and click Create.
3. On the Intelligent Structure Model page, provide a name for the model, and select the location where you want to save the model.



The screenshot shows a software interface titled "Intelligent Structure Model Details". It has three input fields: "Name" containing "Customer_Intelligent_Structure_Model", "Location" set to "Default", and an empty "Description" field. A "Save" button is located in the top right corner of the dialog box.

Creating an Intelligent Structure Model

1. In the Data Integration home page, click New.
2. In the New Asset window, click Components, select the Intelligent Structure Model, and click Create.
3. On the Intelligent Structure Model page, provide a name for the model and select the location where you want to save the model.
4. In the Sample File field, click the file icon to browse for and select a file on which to base the model.



13

© Informatica. Proprietary and Confidential.



Creating an Intelligent Structure Model

1. In the Data Integration home page, click New.
2. In the New Asset window, click Components, select the Intelligent Structure Model, and click Create.
3. On the Intelligent Structure Model page, provide a name for the model and select the location where you want to save the model.
4. In the Sample File field, click the file icon to browse for and select a file on which to base the model.
5. After you select the file, click Discover Structure, and save the model.



14

© Informatica. Proprietary and Confidential.



Refining a Discovered Structure

- After you discover the file structure:
 - View the expected output
 - Refine the nodes and output groups
 - Save the model
- Use the **Visual Model** tab and the **Table** tab to refine the output

15

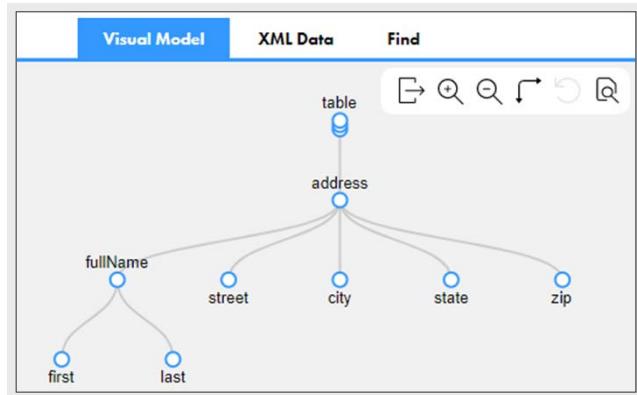
© Informatica. Proprietary and Confidential.



After you discover the file structure, you can view the expected output, refine the nodes and output groups, and then save the model. You can use the Visual Model tab and the Table tab to understand and refine the output.

Refining a Discovered Structure – Visual Model Tab

- Displays the output in a graphical, tree-like structure
- You can use the Visual Model tab to:
 - Trace how the input data is mapped to a node
 - Rename, combine, or exclude nodes from the output



16

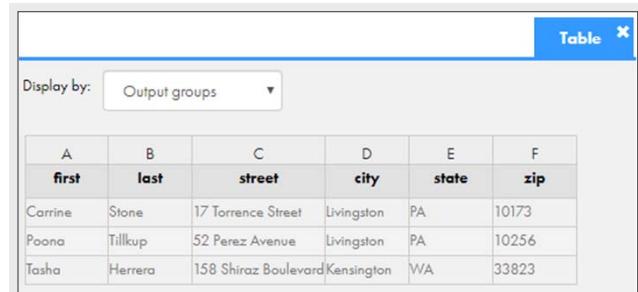
© Informatica. Proprietary and Confidential.



The Visual Model tab displays the output in a graphical, tree-like structure. The intelligent structure shows the discovered types of data as nodes and displays their relationship to each other in a graphical format. You can use the Visual Model tab to trace how the input data is mapped to a node. You can also perform actions on nodes, such as renaming, combining, or excluding nodes from the output.

Refining a Discovered Structure – Table Tab

- Displays the relational output that the intelligent structure produces
- Output is organized in one or more output groups
- You can use the Table tab to:
 - Know how a node is mapped to an output group
 - Rename nodes or exclude nodes from the output



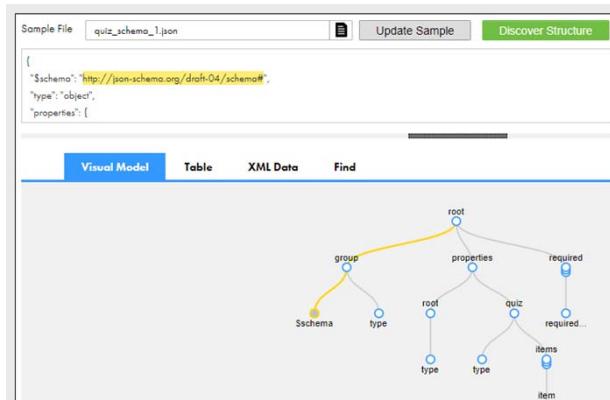
The screenshot shows a software interface titled "Table" with a "Display by:" dropdown set to "Output groups". The main area contains a grid with columns labeled A through F. Column A is "first", B is "last", C is "street", D is "city", E is "state", and F is "zip". The data rows are: Carrine, Stone, 17 Torrence Street, Livingston, PA, 10173; Poona, Tillkup, 52 Perez Avenue, Livingston, PA, 10256; and Tasha, Herrera, 158 Shiraz Boulevard, Kensington, WA, 33823.

A	B	C	D	E	F
first	last	street	city	state	zip
Carrine	Stone	17 Torrence Street	Livingston	PA	10173
Poona	Tillkup	52 Perez Avenue	Livingston	PA	10256
Tasha	Herrera	158 Shiraz Boulevard	Kensington	WA	33823

The Table tab displays the relational output that the intelligent structure produces. The output is organized in one or more output groups. An output group contains one or more nodes. You can use the Table tab to determine how a node is mapped to an output group. You can also use the Table tab to rename nodes or exclude nodes from the output.

Enriching an Existing Structure with New Samples

- You can use additional sample files to enrich the structure with the new fields
- You can only enrich an existing structure based on a JSON sample file
- Intelligent Structure Discovery creates nodes for new data in the sample file



18

© Informatica. Proprietary and Confidential.



After you create an intelligent structure from one sample file, you can use additional sample files to enrich the structure with new fields that exist in the new samples.

You must note that you can only enrich an existing structure based on a JSON sample file.

To add data to the structure, Click **Update Sample**, and select the new sample file. Intelligent Structure Discovery creates nodes for new data in the sample file.

Editing the Intelligent Structure Model

- You can edit the basic details of the model

Name	Type	Uploaded On	Location	Description	Tags	Status
XX_FirstName_FullName_Opt	Mapping	Jun 6, 2019 5:24 AM	Default			Valid
XX_FirstName_GoalSchema	Hierarchical Schema	Jun 5, 2019 11:59 PM	Default			Valid
XX_FirstName_ShoppingFiles	Mapping	May 20, 2019 3:53 AM	Default			Valid
XX_FirstName_StructureModel	Intelligent Structure Model	Jun 5, 2019 7:42 AM	Default			Valid
XX_FirstName_StructureParser	Mapping	Jun 5, 2019 10:55 AM	Default			Valid
XX_FirstName_UncorrectedMapping	Mapping	Jun 1, 2019 3:27 AM	Default			Valid

- You can change the name, location, and description for the model

XX_FirstName_StructureModel

Intelligent Structure Model Details

Name: XX_FirstName_StructureModel

Location: Default

Description:

Sample File StructureModel.txt Discover Structure

19

© Informatica. Proprietary and Confidential.



After you save an Intelligent Structure Model, you can edit the basic details of the model.

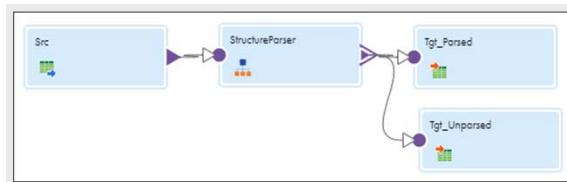
On the **Explore** page, navigate to the project and folder to access the saved Intelligent Structure Model.

Select the row that contains the Intelligent Structure Model. In the **Actions** menu, select **Edit**.

In the Intelligent Structure Model page, you can change the name, description, and location for the model.

Using the Model in Structure Parser Transformation

- You can use an Intelligent Structure Model in a Structure Parser transformation
- Structure Parser transforms input data into a user defined structured format based on an Intelligent Structure Model
- Use the Structure Parser transformation to analyze the data
- When you create a mapping with a Structure Parser transformation, you can select:
 - Intelligent Structure Model that the Structure Parser uses
 - Type of input for the transformation
 - Output for downstream transformations



20

© Informatica. Proprietary and Confidential.



As seen earlier, you can use an Intelligent Structure Model in a Structure Parser transformation. A Structure Parser transforms input data into a user defined structured format, based on an intelligent structure model. You can use the Structure Parser transformation to analyze data such as log files, clickstreams, XML or JSON files, Word tables, and other unstructured or semi-structured formats.

When you create a mapping with a Structure Parser transformation, you can select the intelligent structure model that the Structure Parser uses, the type of input that the transformation expects to receive, and the output that you pass to downstream transformations. An Intelligent Structure Model is required for Structure Parser transformations.

Lab Activity

14-1 Creating an Intelligent Structure Model

In this lab, you will perform the following:

- Create Intelligent Structure Model

Lab Activity

14-2 Using Structure Parser Transformation in a Mapping

In this lab, you will perform the following:

- Create a mapping using Structure Parser transformation

Module Summary

This module showed you how to:

- Discuss Intelligent Structure Model
- Explain Intelligent Structure Discovery process
- List the steps to create an Intelligent Structure Model
- Refine the discovered structure
- Edit the Intelligent Structure Model
- Use the Intelligent Structure Model in a Structure Parser transformation

23

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 15

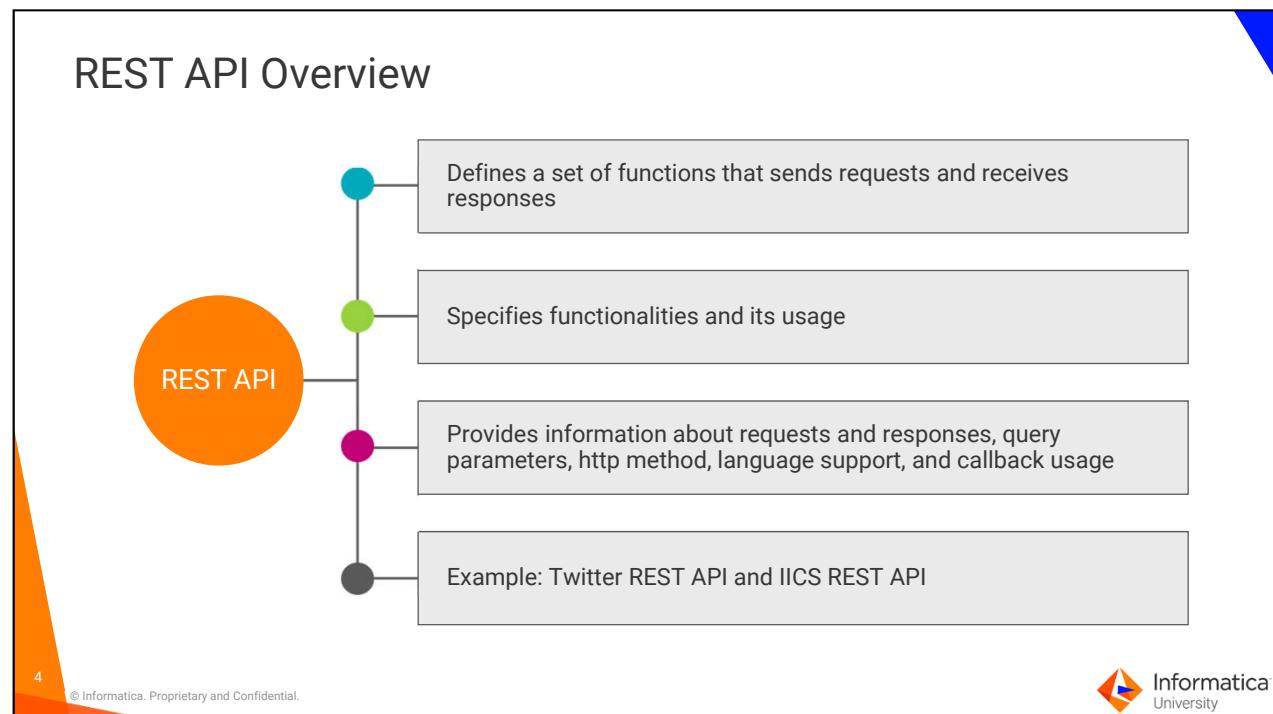
IICS APIs



Module Objectives

After completing this module, you will be able to:

- Explain REST API
- Discuss IICS REST API
- Describe IICS REST API versions
- Discuss request header and request body configurations
- Describe return lists
- Explain the RunAJob utility



A REST API defines a set of functions that allows you to send requests and receive responses using the HTTP protocol.

The REST API specifies the functionalities it can provide and also specifies how to use those functionalities. It also provides information about request and response formats, query parameters, HTTP method to be used – for example, GET, POST, PUT, or DELETE, language support, callback usage, and so on.

Some common examples of REST APIs include, Twitter REST API, and IICS REST API.

IICS REST API

- Allows you to access information from the Informatica Cloud Org
- Enables you to create new objects, update existing objects, delete objects, run tasks and taskflows, and update connection, and schedule information
- Allows you to access information from the Activity Log, Activity Monitor, and Audit Log
- Allows you to access details about the tasks, taskflows, secure agents, connections, schedules, and users

5

© Informatica. Proprietary and Confidential.



IICS REST API allows you to access information from the Informatica Cloud Org using a third-party application or service. You can use the IICS REST API to create new objects, update existing objects, delete objects, run tasks and taskflows, update connection, and schedule information for the Org.

The REST API includes a complete set of resources that allows you to access information from the Activity Log, Activity Monitor, and Audit Log. You can also access details about tasks, taskflows, secure agents, connections, schedules, users, and so on.

When you use the Informatica Cloud REST API, you do not have to manually log in to the Org to perform the tasks.

Using IICS REST API

To use IICS REST API you must have:

Valid Informatica Cloud login credentials

Knowledge of REST API guidelines

To perform a task using IICS API:

Use appropriate resource and method to configure a request

Use applicable objects

6

© Informatica. Proprietary and Confidential.



To use IICS REST API, you must have valid Informatica Cloud login credentials and knowledge of REST API guidelines.

To perform a task, you must first configure a request using the IICS REST API. To configure a request, you must use the appropriate resource and method. You must also use all the applicable objects.

Informatica Cloud returns the requested information, performs the requested task, or returns an error object, and related messages.

IICS Platform REST APIs

- IICS includes common platform functionality that is applicable to all services
 - Platform Resource task enables you to list all tasks in your organization
- Some functionalities are specific to a service
 - Mapping task is applicable only to the Data Integration service

7

© Informatica. Proprietary and Confidential.



IICS includes common platform functionality that is applicable to all the services. For example, you can use the platform resource task to list all the tasks in your organization.

It is also important to know that apart from the platform functionalities, some functionalities are specific to a service. For example, a mapping task is applicable only to the Data Integration service.

IICS REST API Versions

- Supports the platform REST API version 2 and version 3 resources and service-specific resources
- You can log in to IICS using the platform REST API version 2 or version 3 login resource

IICS supports the platform REST API versions 2 and 3 resources and service-specific resources.

You can log in using the REST API versions 2 or 3 platforms.

IICS REST API Versions – A Comparison

- Format:

- version 2 supports XML and JSON calls
- version 3 supports JSON calls

- Login URL:

- for version 2, use <https://dm-<POD region>.informaticacloud.com/ma/api/v2/user/login>
- for version 3, use <https://dm-<POD region>.informaticacloud.com/saas/public/core/v3/login>

North America

Europe

Asia

us

eu

ap

9

© Informatica. Proprietary and Confidential.



The format that you can use depends on the API version. Version 2 supports XML and JSON calls. Version 3 supports only JSON calls.

The login URLs that you can use also depends on the API version. In the login URL, the Point of Deployment or POD region is based on the location of the Informatica Cloud data center.

- For the North America data center, the POD region is **us**
- For Europe, the POD region is **eu**
- For Asia, the POD region is **ap**

The URL that you receive when you register with IICS includes the POD region.

IICS REST API Versions – A Comparison (continued)

- Base URL:
 - for version 2 resources, use <serverUrl>/api/v2/<API name>
 - for version 3 resources, use <baseApiUrl>/public/core/v3/<API name>
- Request URL:
 - for version 2 resources, use <serverUrl>/api/v2/<API name>
 - for version 3 resources, use <baseApiUrl>/public/core/v3/<API name>
- Session ID:
 - for version 2 resources, use **icSessionId** in the header
 - for version 3 resources, use **INFAS-SESSION-ID** in the header

10

© Informatica. Proprietary and Confidential.



The login response includes the base URL that you must include in subsequent calls. The name of the base URL attribute and the URL that you use after login depends on the API version. The URL that you use in requests differs between the version 2 and version 3 resources.

The login response includes a session ID that you must include in headers during the session. You can use the same session ID for versions 2 and version 3 resources. The name of the attribute for session ID also depends on the API version.

For version 2 resources, you can use **icSessionID** in the header.

For version 3 resources, you can use **INFASESSIONID** in the header.

Request Header – Version 2

For version 2 calls, use the following format in the REST API request header:



```
<METHOD> <serverUrl>/<URI> HTTP/<HTTP version>
Content-Type: application/<json | xml>
Accept: application/<json | xml>
icSessionId: <SessionId>
```

Note

If you use the Postman tool, requests automatically include the HTTP version.

11

© Informatica. Proprietary and Confidential.



The request header is slightly different for versions 2 and 3 resources.

- **METHOD** indicates the method you want to use, such as GET, POST, or DELETE.
- **Server URL** indicates the Base URL for all version 2 resources, except login and register.
- **URI** indicates the Resource URI.
- **HTTP version** indicates the HTTP version that you use.
- **Content-Type** indicates the Format of the request.
- **Accept** indicates the Request format that you want to receive.
- **IcSessionID** indicates the IICS session ID.

If you use a tool such as Postman, requests automatically include the HTTP version. So, if you enter the HTTP version in the URL, the request will not be successful because the HTTP version occurs twice in the URL.

Request Header – Version 3

For version 3 calls, use the following format in the REST API request header:



```
<METHOD> <baseApiUrl>/<URI> HTTP/<HTTP version>
Content-Type: application/json
Accept: application/json
INFA-SESSION-ID: <SessionId>
```

Note

If you use the Postman tool, requests automatically include the HTTP version.

12

© Informatica. Proprietary and Confidential.



For version 3 calls, you can use the format displayed above, in the REST API request header.

In this format:

- **METHOD** indicates the method you want to use, such as GET, POST, or DELETE.
- **base API URL** indicates the Base URL for all version 3 resources except login.
- **URI** indicates the Resource URI.
- **HTTP version** indicates the HTTP version that you use.
- **Content-Type** indicates the Format of the request.
- **Accept** indicates the Request format that you want to receive.
- **INFA SESSION ID** indicates the IICS session ID.

If you use a tool such as Postman, requests automatically include the HTTP version. So, if you enter the HTTP version in the URL, the request will not be successful because the HTTP version occurs twice in the URL.

Request Body

- Request body passes additional attributes for the resource
- Passes attributes as part of an object
- If a request includes sub-objects for attributes, you must declare the sub-objects before listing the related attributes

13

© Informatica. Proprietary and Confidential.



You can use the request body to pass additional attributes for the resource. When you pass attributes in a request body, you pass the attributes as part of an object.

For example, to log in with the login resource, you pass the required username and password attributes in a login object.

Some requests include sub-objects for attributes. You must declare the sub-objects before listing the related attributes.

Request Body – JSON Format

When you use JSON format for version 2 REST API calls, you can optionally define a request object with the @type attribute.

```
{  
  "@type": "<request object>",  
  "<attribute1>": "<value1>",  
  "<attribute2>": "<value2>",  
}
```

When an attribute includes an object, state the attribute and use the object name.

```
{  
  "@type": "<request object>",  
  "<attribute1>": "<value1>",  
  "<attribute2>": {  
    "@type": "<attribute object>",  
    "<attributeA>": "<valueA>",  
    "<attributeB>": "<valueB>",  
    "@type": "<attribute object>",  
    "<attributeD>": "<valueD>",  
    "<attributeE>": "<valueE>",  
    "<attribute3>": "<value3>",  
  }  
}
```

14

© Informatica. Proprietary and Confidential.



When you use the JSON format for version 2 REST API calls, you can optionally define a request object with the **@type** attribute, as shown in the example.

When an attribute includes an object, you must state the attribute and use the object name.

For version 3 REST API calls, do not use the **@type** attribute.

Request Body – XML Format

When you use XML format, define a request object as an enclosing set of tags.

```
<request object>
  <attribute1>value1</attribute1>
  <attribute2>value2</attribute2>
</request object>
```

When an attribute includes an object, enclose the attribute object within the attribute tags.

```
<request object>
  <attribute1>value1</attribute1>
  <attribute2>
    <attribute object>
      <attributeA>valueA</attributeA>
    </attribute object>
    <attribute object>
      <attributeB>valueB</attributeB>
    </attribute object>
  </attribute2>
  <attribute3>value3</attribute3>
</request object>
```

15

© Informatica. Proprietary and Confidential.



When you use the XML format, you must define a request object as an enclosing set of tags.

When an attribute includes an object, you must enclose the attribute object within the attribute tags.

Return Lists – JSON Format

JSON does not use additional attributes. The REST API encloses the list in square brackets ([]).

```
[  
  {  
    "<attribute1>": "<value1>",  
    "<attribute2>": "<value2>",  
  }{  
    "<attribute1>": "<value1>",  
    "<attribute2>": "<value2>",  
  }]  
]
```

In the JSON format, it does not use additional attributes. The REST API encloses the list in square brackets.

Return Lists – XML Format

When the REST API returns a series of objects in XML, it encloses the list in the root tag.

```
<root>
  <return object 1>
    <attribute1>value1</attribute1>
    <attribute2>value2</attribute2>
  </return object 1>
  <return object 2>
    <attribute1>value1</attribute1>
    <attribute2>value2</attribute2>
  </return object 2>
</root>
```

When the REST API returns a series of objects in XML format, it encloses the list in the root tag.

RunAJob Utility

- Runs a JAR file that calls an IICS REST API to run a job
- The utility provides the following job details:
 - User who initiated the job
 - Time the job was initiated
 - Run ID for the job
- Use the utility to run published taskflows and the following tasks:

Mapping Task

Synchronization Task

Replication Task

Masking Task

PowerCenter Task

Workflow

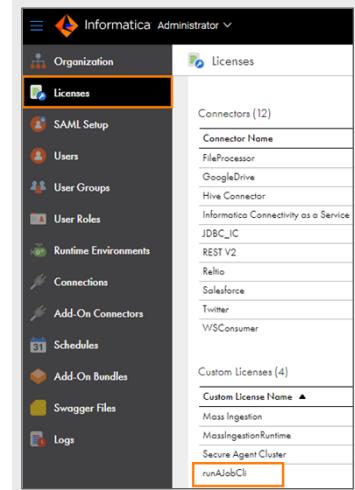
When you use the IICS REST API, you can use the RunAJob utility instead of the job resource to run Data Integration tasks or published taskflows.

The RunAJob utility runs a JAR file that calls an IICS REST API to run a job. After the job completes, the utility provides the job details such as the user who initiated the job, the time the job was initiated, and the run ID for the job.

You can use the RunAJob utility to run published taskflows and tasks such as a mapping task, synchronization task, replication task, masking task, PowerCenter task, and a workflow.

RunAJob Utility (continued)

- To use the RunAJob utility, you must have the **runAJobCli** package enabled in your Org
- If the package is not enabled in your Org, contact Informatica Global Customer Support
- The RunAJob utility can be found in the following location:
 - **C:\Program Files\Informatica Cloud Secure Agent\apps\runAJobCli**
- To use the RunAJob utility, the secure agent host must have Java version 1.8 or higher installed



 Informatica
University

19

To use the RunAJob utility, you must have the **runAJobCli** package enabled in your Informatica Cloud Org.

To see if your organization is licensed to use the utility, log in to your organization and in the Administrator Service, click Licenses. Then scroll down to the bottom of the page and look for the **runAJobCli** package.

If you don't see the package in your Org, you must contact Informatica Global Customer Support to enable it.

When the package is enabled, the utility can be found in the location as shown in the image.

To use the RunAJob utility, the secure agent host must have Java version 1.8 or higher installed.

RunAJob Utility Setup

- Create copies of the RunAJob properties template files that are included with the utility and configure the new files
- The Restenv_default.properties file specifies the Informatica Cloud login credentials and the job polling behavior
- The Log4j_default.properties file specifies the level of detail to return in log files
- You can find the template files in the following location:
 - C:\Program Files\Informatica Cloud Secure Agent\apps\runAJobCli

Name	Date modified	Type
lib	6/13/2019 1:02 PM	File folder
cli.bat	6/13/2019 1:02 PM	.symlink
cli.sh	6/13/2019 1:02 PM	.symlink
log4j_default.properties	6/13/2019 1:02 PM	.symlink
restenv_default.properties	6/13/2019 1:02 PM	.symlink
runAJob_backup.bat	6/13/2019 1:02 PM	.symlink
runAJob_backup.sh	6/13/2019 1:02 PM	.symlink
runAJobCli.jar	6/13/2019 1:02 PM	.symlink

20

© Informatica. Proprietary and Confidential.



To set up the RunAJob utility, create copies of the RunAJob properties template files that are included with the utility and configure the new files.

The RunAJob utility includes the **restenv_default.properties** template file and the **log4j_default.properties** template file.

The **restenv_default.properties** file specifies the Informatica Cloud login credentials and the job polling behavior.

The **log4j_default.properties** file specifies the level of detail to return in log files.

To customize the RunAJob properties, you can copy the template files to create a **restenv_default.properties** file and a **log4j_default.properties** file and then configure the properties. You can use the template files that are included with the utility as a reference.

You can find the template files in the location as shown in the image.

Login Properties

- Specify the Informatica Cloud login credentials in the Restenv.properties file
- Alternatively, you can include the login parameters as arguments in a task command

Parameter	Description
baseUrl	Base URL. Default is https://dm-us.informaticacloud.com/ma.
username	Informatica Cloud user name
password	Informatica Cloud password

21

© Informatica. Proprietary and Confidential.



You must specify the Informatica Cloud login credentials in the **restenv.properties** file. Alternatively, you can include the login parameters as arguments in a task command.

Job Status

- Specify the frequency at which the RunAJob utility polls for status in the **Restenv.properties** file

Parameter	Description
ACTIVITYMONITORWAIT	The amount of time the utility waits before retrying if an internal exception occurs, such as a login failure or network problem. Default is 5000 milliseconds.
TOTALWAIT	The maximum amount of time the utility waits for a job to complete before polling the activity monitor and activity log again for status. Default is 5000 milliseconds.
RETRYCOUNT	The number of times the utility polls for status. This parameter is used for polling the activity monitor and activity log for job status and for internal exceptions such as login failure or network problems. Default is 3.

22

© Informatica. Proprietary and Confidential.



To configure the job status, you must specify the frequency at which the RunAJob utility polls for status in the **restenv.properties** file.

The parameters that you can use in the **restenv.properties** file are:

ACTIVITY MONITOR WAIT: This parameter specifies the amount of time the utility waits before retrying, if an internal exception occurs, such as a login failure or network problem. The default value is 5000 milliseconds.

TOTAL WAIT: This parameter specifies the maximum amount of time the utility waits for a job to complete before polling the activity monitor and activity log again for status. The default value is 5000 milliseconds.

RETRY COUNT: This parameter specifies the number of times the utility polls for status. This parameter is used for polling the activity monitor and activity log for job status and for internal exceptions such as login failure or network problems. The default value is 3.

Log File Detail

- Specify the level of detail to return in log files in the `log4j.properties` file
- To return basic information about the job, set the level of detail to **Info**
- To return all the job details for debugging purposes, set the level of detail to **Debug**

23

© Informatica. Proprietary and Confidential.



You can specify the level of detail to return in log files in the **log4.properties** file.

If you want the log to return basic information about the job such as user ID, job ID, and time the task was initiated, set the level of detail to **Info**.

If you want the log to return all the job details for debugging purposes, set the level of detail to **Debug**. You can also set this property as an argument in a task command.

Using the RunAJob Utility

- Type the RunAJob utility command, `cli.bat runAJobCli` followed by arguments
- For each job, specify the task or taskflow to run
- Syntax to run a task:
`cli.bat runAJobCli -t <tasktype> -n <task name> -fp <folder path to the task>`
- Syntax to run a synchronization task
`cli.bat runAJobCli -t DSS -n dss_Arch_2308 -fp myproject/folder1`
- Syntax to run a taskflow:
`cli.bat runAJobCli -t TASKFLOW -un <taskflow name>`

24

© Informatica. Proprietary and Confidential.



To use the RunAJob utility, you must type the RunAJob utility command, which is **cli.bat runAJobCli**, followed by the arguments.

For each job, you must specify the task or taskflow to run. The syntax that you use to run a taskflow is slightly different from the syntax you use to run a task.

RunAJob Utility Arguments

- You can use the following arguments in a RunAJob command:

Parameter	Argument
username	-u
password	-p
baseUrl	-bu
taskId	-i
folderPath	-fp
frsId	-fi
taskName	-n
taskType	-t
waitFlag	-w
debug	-d
insecure	-k

25

© Informatica. Proprietary and Confidential.



Job Status Codes

- If the job is successful, the RunAJob utility returns a SUCCESS value of 0
- For failed jobs, the utility returns errors
- If any required parameters are missing or are invalid in a command, an error message displays and the REST API call fails

Code	Description
-1	Exception
0	Success
1	Warning
2	No wait
3	Failure
4	Timeout
5	Error
6	Running

26

© Informatica. Proprietary and Confidential.



If the job is successful, the RunAJob utility returns a SUCCESS value of 0. If the task fails, the utility returns errors.

If any required parameters are missing or are invalid in a command, an error message displays and the REST API call does not run.

Demonstration

RunAJobUtility

- View the video **Module15_Video1_RunAJob_Utility.mp4**

Lab Activity

15-1 Running a Mapping Task Using REST API

In this lab, you will perform the following:

- Use a REST client application to call the login resource and obtain a session ID
- Start a Mapping task
- Log out of the Informatica Cloud API session

Module Summary

This module showed you how to:

- Explain REST API
- Discuss IICS REST API
- Describe IICS REST API versions
- Discuss request header and request body configurations
- Describe return lists
- Explain the RunAJob utility



IICS: Cloud Data Integration Services

 Informatica
University

Module 16

Exception Handling



Module Objectives

After completing this module, you will be able to:

- Explain user-defined, non-fatal, and fatal exceptions
- Define exception handling techniques
- Describe a reject or bad file

Types of Exceptions

User-defined Exceptions

- User-defined exceptions occur because of improper handling of data by business users

Non-fatal Exceptions

- Informatica Cloud Server ignores non-fatal exceptions and causes the records to dropout from the target

Fatal Exceptions

- Fatal exceptions occur when Informatica Cloud Server cannot access the source, target, or repository

4

© Informatica. Proprietary and Confidential.



Informatica Cloud Server can encounter user-defined, non-fatal, or fatal exceptions while running a task.

User-defined Exception: This kind of exception occurs due to improper handling of data by business users. For example, if a user enters an incorrect date for a credit card transaction.

Non-Fatal Exception: A non-fatal exception does not force the session to stop on its first occurrence. Informatica Cloud Server ignores non-fatal exceptions and causes the records to dropout from the target table. For example, a data conversion transformation error prevents the record from loading to the target table.

Fatal Exception: This exception occurs when Informatica Cloud Server cannot access the source, target, or repository. A fatal exception results in stopping the session. This can include connection failures or target database errors, such as unavailability of database space to load the data.

User Defined Exceptions – Error Handling Functions

ERROR()	ABORT()
<ul style="list-style-type: none">Causes Data Integration service to skip a row and issue an error messageData Integration service writes the error message to the session log file or the error log tables	<ul style="list-style-type: none">Causes Data Integration service to stop the session and issue an error messageData Integration service writes the error message to the session log file or the error log tables

Use the ERROR and ABORT functions in an Expression transformation to validate the data

5

© Informatica. Proprietary and Confidential.



IICS provides the ERROR and ABORT functions to handle user-defined exceptions.

The ERROR function causes the Data Integration service to skip a row and issue an error message. The Data Integration service writes the error message to the session log file or the error log tables, based on the error logging configuration for the session.

The ABORT function causes the Data Integration service to stop the session and issue an error message. The Data Integration service writes the error message to the session log file or the error log tables, based on the error logging configuration for the session. When the Data Integration service encounters an ABORT function, it stops processing data at the row it encounters the error.

You can use the ERROR and ABORT functions in an Expression transformation to validate the data.

User Defined Exceptions – Error Tables

- Create error tables and add error records to the table
- Configure the session properties to check for error records and move them to error tables
- Use record flags to identify records that need reprocessing
- Error table example:

Table Columns	Description
ERROR_SEQ_ID	Error row sequence identifier
COLUMN_1	Source data element 1
COLUMN_2	Source data element 2
.....
COLUMN_n	Source data element n
ERROR_FIXED_FLAG	Value is set to Y, if data issue is fixed and ready to reprocess, else set to N
PROCESSED_FLAG	Value is set to Y, if the record is reprocessed, else set to N

6

© Informatica. Proprietary and Confidential.



You can create error tables and add error records to the table.

You can configure the session properties to check for error records and move them to error tables.

A typical ETL design reads error records from the error table.

In the error tables, you can use record flags to identify records that need reprocessing.

Above is an example of an error table that includes all the columns from the source table, and additional columns to identify the status of the error records.

Non-Fatal Exceptions

- Non-fatal exceptions do not force the session to stop on its first occurrence
- Configure the 'Stop on Error' option to indicate the number of non-fatal errors the task can encounter before the session stops
- Types of non-fatal errors:

Reader Error	Writer Error	Transformation Error
Occurs when the Data Integration Service reads data	Occurs when the Data Integration Service writes data	Occurs when the Data Integration Service performs data transformation

7

© Informatica. Proprietary and Confidential.



As discussed earlier, a non-fatal exception does not force the session to stop on its first occurrence. You can configure the 'Stop on Error' option in the session properties, to indicate the number of non-fatal errors the task can encounter before the session stops. If you specify the number as zero, non-fatal errors do not cause the session to stop.

There are three types of non-fatal errors – Reader Error, Writer Error, and Transformation Error.

A Reader Error can occur when the Data Integration Service reads data from database sources, Flat Files, or other types of source systems.

A Writer Error can occur when the Data Integration Service writes data to targets or databases.

A Transformation Error can occur when the Data Integration Service performs data transformation.

Non-Fatal Exception – Example

- The error message indicates that the datatype of one of the columns used in the Expression transformation is invalid

```
Message Code: TT_11132
Message: Transformation [exp_pct_lt_exp] had an error evaluating output column [FGT_LT_EXP].
Error message is [<
    > [TO_DATE]: invalid string for converting to Date
    ... t:TO_DATE(s:'9999999T',s:'YYYYMMDD')].
```

Example

The error message indicates that the datatype of one of the columns used in the Expression transformation is invalid. You can resolve this issue by assigning appropriate datatypes for the columns in the Expression transformation.

Handling Non-Fatal Exceptions

- Non-fatal exceptions causes the records to drop out of the ETL process
- You can handle non-fatal exceptions using the following techniques:
 - Default Field Value Setting
 - Row Error Logging
 - Error Handling Setting

9

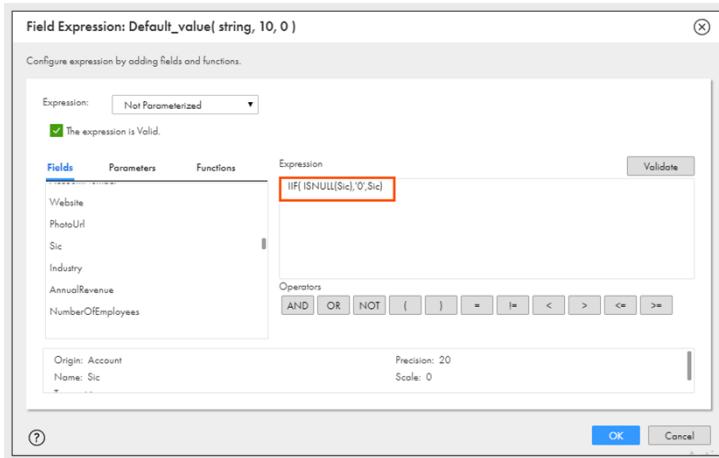
© Informatica. Proprietary and Confidential.



Non-fatal exceptions cause the records to drop out of the ETL process. This can cause quality issues. So, you must handle non-fatal exceptions to save important data. There are three techniques to handling non-fatal exceptions – by using the Default Field Value Setting, by providing Row Error Logging information, or by using Error Handling Settings.

Default Field Value Setting

- Use the default value property to handle null value exceptions and unexpected transformation errors



10

© Informatica. Proprietary and Confidential.

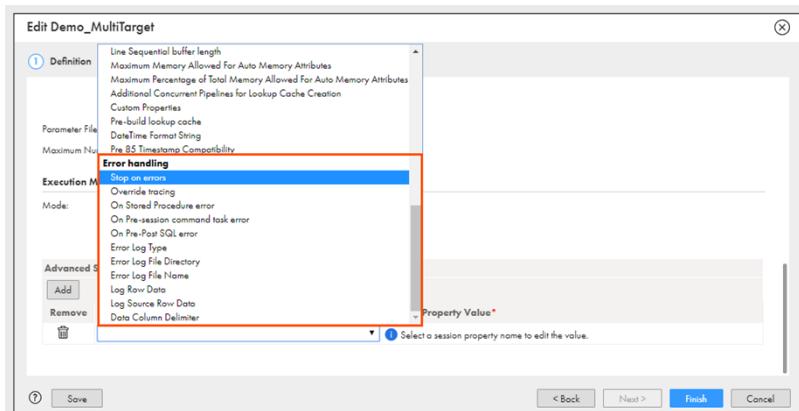


You can use the default value property to handle null value exceptions and unexpected transformation errors.

The image shows an expression that validates the Salary field. If the Salary field for a record is null, the expression transformation returns the default value “zero”.

Row Error Logging

- Does not require error configuration in the mapping creation stage
- In the mapping task, under the Schedule tab, you can set the session properties for error handling



11

© Informatica. Proprietary and Confidential.



For row error logging, you do not require any error configuration in the mapping creation stage. After you create the mapping, you can use the mapping to create a mapping task. When you configure the session for the mapping task, you can set the session properties for error handling under the Schedule tab.

Error Handling Settings

Error Handling Options	Description
Stop on Errors	Indicates how many non-fatal errors the task can encounter before it stops the session.
Override Tracing	Overrides tracing levels set on an object level.
On Stored Procedure Error	Determines the behavior when a task based on a Visio template encounters pre-session or post-session stored procedure errors.
On Pre-Session Command Task Error	Determines the behavior when a task that includes pre-session shell commands encounters errors.
On Pre-Post SQL Error	Determines the behavior when a task that includes pre-session or post-session SQL encounters errors.
Error Log Type	Specifies the type of error log to create. You can specify flat file or no log. Default is none.

12

© Informatica. Proprietary and Confidential.



Error Handling Settings (continued)

Error Handling Options	Description
Error Log File Directory	Specifies the directory where errors are logged.
Error Log File Name	Specifies error log file name.
Log Row Data	Specifies whether or not to log transformation row data.
Log Source Row Data	Specifies whether or not to log source row data.
Data Column Delimiter	Specifies the Delimiter for string type source row data and transformation group row data.

Fatal Exceptions

- All the read and write processes stop and the Data Integration Service rolls back all the data that is not committed to the target database
- Fatal Exceptions occur when there is a loss of connection and the Data Integration Service cannot access the source, target, or repository
- When the session encounters a fatal error, the Data Integration Service terminates the session
- To handle fatal errors, use a re-startable ETL design for your mapping or use the mapping recovery features of IICS

14

© Informatica. Proprietary and Confidential.



You saw earlier that fatal exceptions stop any ongoing sessions. The read and write processes stop, and the Data Integration Service rolls back all the data that is not committed to the target database. Fatal exceptions occur when there is a loss of connection and the Data Integration Service cannot access the source, target, or repository. This can also include target database errors, such as lack of database space to load data.

When the session encounters a fatal error, the Data Integration Service terminates the session. To handle fatal errors, you can either use a re-startable ETL design for your mapping or use the mapping recovery features of IICS.

Fatal Exception – Example

- The error message indicates that the Oracle database table 'TEST_DATA' has insufficient space to load the data

```
ERROR    6/12/2019 6:20:08 AM   DIRECTOR      REP_12400      Repository Error ( 
ORA-30032: the suspended (resumable) statement has timed out
ORA-01536: space quota exceeded for tablespace 'TEST_DATA'
```

Example

The error message indicates that the Oracle database table 'TEST DATA' has insufficient space to load the data. You can resolve this issue by increasing the table space of the 'TEST DATA' table.

Bad Files or Reject Files

- Bad or Reject files holds the data for the entire row that the target rejects
- Bad files extend the scope of error tracking
- Allows you to perform detailed analysis to track the exact error and take corrective actions

Bad Files, which are also known as Reject Files holds the data for the entire row that the target rejects. When you create a session with a target, the session creates bad files.

Bad files extend the scope of error tracking. They allow you to perform detailed analysis to track the exact error and take corrective actions.

Lab Activity

16-1 Creating a Mapping to Handle Non-fatal Errors

In this lab, you will perform the following:

- Configure a mapping to handle non-fatal errors

Module Summary

This module showed you how to:

- Explain user-defined, non-fatal, and fatal exceptions
- Define exception handling techniques
- Describe a reject or bad file

18

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 17

Performance Tuning



Module Objectives

After completing this module, you will be able to:

- Describe partitions
- Explain types of partitions
- List partition rules and guidelines
- Discuss pushdown optimization
- List pushdown optimization types
- Discuss secure agent groups
- Discuss Data Transformation Manager (DTM) process and its configuration

Topic Partitions



Partitions Overview

- Partitions enable you to optimize performance for mapping tasks
 - Reduces execution time of the task by processing partitions of data concurrently
- Enable partition while configuring the Source transformation
 - When you configure partitions in the Source transformation, partitioning occurs throughout the mapping

5

© Informatica. Proprietary and Confidential.

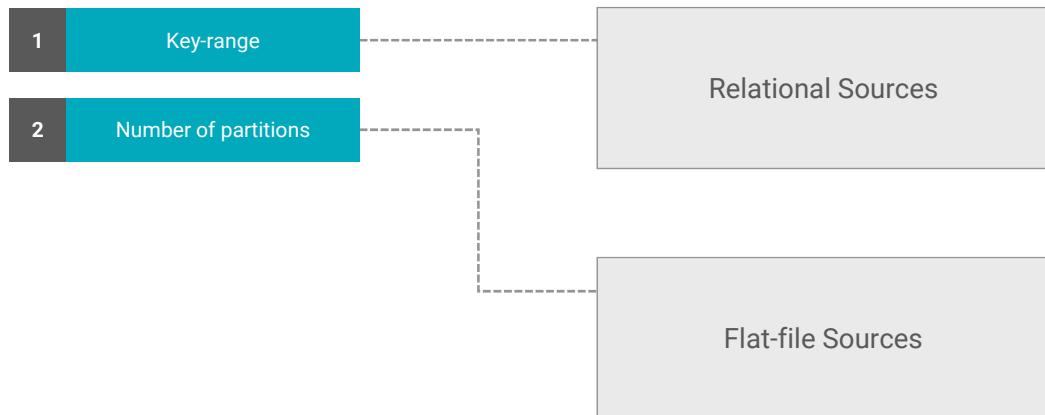


You can use partitions to optimize performance for mapping tasks. It reduces the execution time of the task by processing partitions of data concurrently.

If a mapping task processes large data sets or includes transformations that perform complex calculations, the task can take a long time to process. When you use multiple partitions, the mapping task divides data into partitions and processes the partitions concurrently, thereby reducing the execution time of the task.

You can enable partitions when you configure the Source transformation in the Mapping Designer. When you configure partitions in the Source transformation, partitioning occurs throughout the mapping.

Types of Partitioning



© Informatica. Proprietary and Confidential.



There are majorly two types of partitioning methods – One is based on a key-range, and the other is based on the number of partitions.

The “key range partitioning” method is typically used for relational source types. The “number of partitioning” method is used for a source type that does not allow key range partitioning such as a flat file source, or when the mapping includes a transformation that does not support key-range partitioning.

Types of Partitioning – Specify a Key Range

- Use key range partitioning method for a mapping with a relational source
- Mapping task distributes rows of data based on a field that you define as a partition key
- Specify one field in the source as the partition key
- Define a range of values for the partition key
- Key ranges can be of the following datatypes:
 - String
 - Number
 - Date/time (MM/DD/YYYY HH24:MI:SS)
- For a mapping with multiple sources, use the same number of key ranges for each source

7

© Informatica. Proprietary and Confidential.



You can use the key range partitioning method for a mapping with a relational source. When you enable partitioning using a key range, the mapping task distributes rows of data based on a field that you define as a partition key. You must select one field in the source as the partition key, and then define a range of values for it.

Key ranges can be of String, Number, or Date-Time data type. You must note that if the key range is of Number data type, you cannot use decimals in the key range values. If the key range is of Date-Time data type, you must use the default date-time format as shown above.

If the mapping includes multiple sources, you must use the same number of key ranges for each source.

Key Range Partitioning – Example

- Partition the source data into three partitions based on postal codes
- Specify the key ranges as follows:
 - First partition: Minimum value to 30000
 - Second partition: 30001 to 50000
 - Third partition: 50001 to maximum value
- On the Partitions tab for the Source transformation, select the BILLINGPOSTALCODE field for the partition key



8

© Informatica. Proprietary and Confidential.



Example

Consider that you have customer names, addresses, and purchasing history in a relational database source. You decide to partition the source data into three partitions based on postal codes. For the first partition, you specify the key range as 'Minimum value to 30000'. For the second partition, you specify the key range as '30001 to 50000'. Finally, for the third partition, you specify the key range as '50001 to maximum value'.

On the Partitions tab of the Source transformation, you select the BILLINGPOSTALCODE field for the partition key. As shown in the image, you add three key ranges to create three partitions.

Types of Partitioning – Specify the Number of Partitions

- Specify the number of partitions for a source type that does not allow key range partitioning
- Can also use this method when the mapping includes a transformation that does not support key range partitioning
- Can specify up to 64 partitions
- Consider the number of records you want to pass in the mapping to determine an appropriate number of partitions
- For a mapping with multiple sources, specify same number of partitions for each source

9

© Informatica. Proprietary and Confidential.



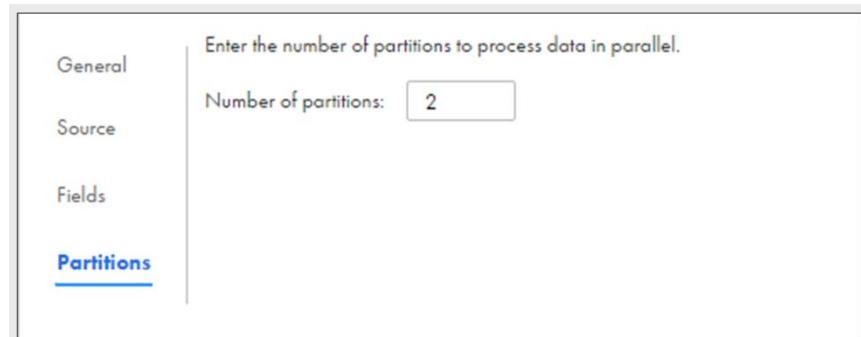
You can specify the number of partitions for a source type that does not allow key range partitioning, such as a flat file source. You can also use this partitioning method when the mapping includes a transformation that does not support key range partitioning. When you enable partitioning based on the number of partitions, you can specify up to 64 partitions.

You must consider the number of records you want to pass in the mapping to determine an appropriate number of partitions for the mapping. For a small number of records, partitioning may not be helpful.

If the mapping includes multiple sources, you must specify the same number of partitions for each source.

Number of Partitions – Example

- Mapping task uses 1GB flat file source
- Specify two partitions in the Source transformation
- On the Partitions tab of the Source transformation, enter the number of partitions



10

© Informatica. Proprietary and Confidential.



Here is an example where you can specify the number of partitions.

Consider that you have a mapping task that uses a large, 1GB flat file source. You decide to specify two partitions in the Source transformation to optimize performance.

On the Partitions tab of the Source transformation, enter the number of partitions, as shown in the image.

Partitioning Restrictions

- You cannot partition a mapping in the following situations:

01

When the mapping uses a parameterized source or source query

02

When the mapping includes Web Services or Hierarchy Parser transformation

03

When the mapping includes multiple sources that use custom relationships or advanced relationships

11

© Informatica. Proprietary and Confidential.



There are certain types of mappings in which you cannot use partitions.

You cannot use partitions in mappings that uses a parameterized source or source query. Partitioning is also not supported in mappings that includes Web Services or Hierarchy Parser transformation. You cannot partition mappings that includes multiple sources that use custom relationships or advanced relationships.

When you configure partitions, you must save and run the mapping in order to validate the partition settings.

Partitioning Rules and Guidelines

- For Flat File partitioning, session performance is optimal with large source files
- Set up caching in the Sequence Generator transformation
- The sequence numbers that the Normalizer and Sequence Generator transformations generate may not be sequential
- The Sorter transformation sorts data in each partition separately
- Place a Sorter transformation before any Joiner or Aggregator transformation
- You cannot use parameters for key range values

12

© Informatica. Proprietary and Confidential.



For Flat File partitioning, session performance is optimal with large source files.

When you enable partitioning in a mapping that has a Sequence Generator transformation, you must ensure that you set up caching in the Sequence Generator transformation.

The sequence numbers that the Normalizer and Sequence Generator transformations generate may not be sequential for a partitioned source, however they are unique.

When you enable partitioning in a mapping that has a Sorter transformation, the task sorts data in each partition separately.

You must place a Sorter transformation before any Joiner or Aggregator transformation.

You cannot use parameters for key range values.

Topic

Pushdown Optimization



Pushdown Optimization Overview

- Use pushdown optimization to push transformation logic to source databases or target databases for execution
- Task converts the transformation logic into a SQL query
- The amount of transformation logic that you can push to the database depends on the database, transformation logic, and task configuration

14

© Informatica. Proprietary and Confidential.



You can use the pushdown optimization technique to push transformation logic to source databases or target databases for execution. Using pushdown optimization on database resources can improve the task performance.

When you run a task that is configured for pushdown optimization, the task converts the transformation logic into a SQL query. The task sends the query to the database and the database executes the query.

The amount of transformation logic that you can push to the database depends on the database, the transformation logic, and the task configuration. The task processes all transformation logic that it cannot push to a database.

Pushdown Optimization Types

Source pushdown optimization

The task analyzes the mapping from source to target or until it reaches the transformation logic that it cannot push to the source database.

The task generates and executes a Select statement based on the transformation logic for each transformation that it can push to the database.

The task reads the results of the SQL query and processes the remaining transformations.

15

© Informatica. Proprietary and Confidential.



When you configure a task with a Source pushdown optimization, the task analyzes the mapping from the source to the target or until it reaches a transformation logic that it cannot push to the source database. The task generates and executes a Select statement based on the transformation logic for each transformation that it can push to the database. The task then reads the results of the SQL query and processes the remaining transformations.

Pushdown Optimization Types

Target pushdown optimization

The task analyzes the mapping from target to source or until it reaches the transformation logic that it cannot push to the target database.

The task generates an Insert, Delete, or Update statement based on the transformation logic for each transformation that it can push to the target database.

The task processes the transformation logic up to the point where it can push the transformation logic to the database. The task then executes the generated SQL on the target database.

16

© Informatica. Proprietary and Confidential.



When you configure a task with a Target pushdown optimization, the task analyzes the mapping from the target to the source or until it reaches a transformation logic that it cannot push to the target database. The task generates an Insert, Delete, or Update statement based on the transformation logic for each transformation that it can push to the target database. The task processes the transformation logic up to the point where it can push the transformation logic to the database. The task then executes the generated SQL query on the target database.

Pushdown Optimization Types

Full pushdown optimization

The task analyzes the mapping from source to target or until it reaches the transformation logic that it cannot push to the target database.

The task generates and executes SQL statements against the source or target, based on the transformation logic that it can push to the database.

You can use full pushdown optimization when the source and target databases are in the same relational database management system.

17

© Informatica. Proprietary and Confidential.



When you configure a task with a Full pushdown optimization, the task analyzes the mapping from the source to the target or until it reaches a transformation logic that it cannot push to the target database. The task generates and executes SQL statements against the source or target based on the transformation logic that it can push to the database. You can use full pushdown optimization when the source and target databases are in the same relational database management system.

Cross-schema Pushdown Optimization

- Enable cross-schema pushdown optimization for tasks that use source or target objects associated with different schemas within the same database
- To use cross-schema pushdown optimization, create a connection for each schema
- Cross-schema pushdown optimization is enabled by default

What is a cross-schema pushdown optimization?

You can enable cross-schema pushdown optimization for tasks that use source or target objects associated with different schemas within the same database.

To use cross-schema pushdown optimization, you must create a connection for each schema. The database and the database user name and password must be the same for both connections.

Cross-schema pushdown optimization is enabled by default in a task.

Pushdown Optimization User-defined Parameters

- Use a user-defined parameter to perform pushdown optimization based on the parameter value defined in the parameter file
- Use a pushdown optimization user-defined parameter when you want to perform different pushdown optimization options
- Example:
 - Use source or target pushdown optimization during the peak hours of the day
 - Use full pushdown optimization from midnight until 2 a.m.

19

© Informatica. Proprietary and Confidential.



You can use a user-defined parameter to perform pushdown optimization based on the parameter value that you define in the parameter file. You can use these parameters when you want to perform different pushdown optimization options at different times.

For example, you can use source or target pushdown optimization during the peak hours of the day and use full pushdown optimization from midnight until 2 a.m. when the database activity is low.

Pushdown Optimization – Connections

- When you run a pushdown optimization session that involves multiple database connection objects, IICS selects only one connection as the active connection
- IICS uses the active connection to execute the pushdown SQL query
- When the source and target reside in separate databases, enable session property 'Allow Pushdown for User Incompatible Connections'

The screenshot shows the 'Advanced Session Properties' dialog box. It has tabs for 'Add', 'Remove', 'Session Property Name*', and 'Session Property Value*'. Under 'Session Property Name*', the value 'Allow Pushdown for User Incompatible Connections' is selected. Under 'Session Property Value*', there are two radio buttons: 'Yes' (unchecked) and 'No' (checked). A small trash can icon is also visible.

20

© Informatica. Proprietary and Confidential.



When you run a pushdown optimization session that involves multiple database connection objects, IICS selects only one connection as the active connection. The selected connection is used to execute the pushdown SQL query.

When the source and target reside in separate databases, you must enable the flag 'Allow Pushdown for User Incompatible Connections' in the Advanced Session Properties of the task.

Pushdown Optimization – Error Handling

- Some functionalities available in Data Integration Service are not available in Database processing
- If an error occurs in a pushdown optimization session, the database handles the error
- You cannot use IICS error handling features for pushdown optimization session failures
- For failed pushdown optimization sessions, IICS cannot perform incremental recovery

21

© Informatica. Proprietary and Confidential.



What happens when you encounter an error in a task that is enabled for pushdown optimization?

Some functionalities available in Data Integration Service are not available in Database processing.

When you enable pushdown optimization, the database executes the SQL query. If any error occurs, the database handles the errors. You cannot use IICS error handling features for pushdown optimization session failures. So, if you configure a session for full pushdown optimization and the session fails, IICS cannot perform incremental recovery because the database processes the transformations.

Benefits of Using Pushdown Optimization

- It avoids reading large volumes of data to and from:
 - the database and across the network
 - the database and the runtime environment where the secure agent DIS is running
- Pushdown optimization is most efficient when you deal with large volumes of data

Note

Data transformation is faster using the Informatica secure agent Data Integration Service, rather than using the database engine. However, when you use pushdown optimization, you save on the network throughput time (read or write).

22

© Informatica. Proprietary and Confidential.



Pushdown optimization is most efficient when you deal with large volumes of data on dedicated hardware, such as servers built to run only DB2, Oracle, and Teradata.

Note: Data transformation is faster using the Informatica secure agent Data Integration Service, rather than using the database engine. However, when you use pushdown optimization to push the transformation logic to the source or target database, you save on the network throughput time (read or write). This is crucial especially when you deal with large volumes of data.

Pushdown Optimization – Example

- **Scenario:**

- You have a mapping in which your source is a database resource.
- Your mapping includes a filter transformation to filter records based on some criteria.
- Your mapping also uses an expression transformation before loading the data to a target.

- **Solution:**

- If you select the source based pushdown optimization, then the data processing engine analyzes the logic of the mapping and decides the portion of the mapping that can be directly pushed to the source as a query.
- If there are simple expressions that can be sent to the database in the SELECT query, the SELECT query to the source includes both the filter and the expression.
- This allows the database to filter out the records, transform it according to the expression, and send the data to the Informatica Cloud engine.

23

© Informatica. Proprietary and Confidential.



Use pushdown optimization when you want to reduce the load of integration service. Using pushdown optimization on database resources improves the task performance.

Example: Consider that you have a mapping in which your source is a database resource. Your mapping includes a filter transformation to filter records based on some criteria. After the source data is filtered, you apply an expression transformation before loading the data to a target. If you select the source based pushdown optimization, then the data processing engine analyzes the logic of the mapping and decides the portion of the mapping that can be directly pushed to the source as a query.

In this case, if there are simple expressions that can be sent to the database in the SELECT query, the SELECT query to the source includes both the filter and the expression. This allows the database to filter out the records, transform it according to the expression, and send the data to the Informatica Cloud engine. This improves performance because the database engine only gets a subset of data and performs the required transformation. The processing can be even faster if the source and target are on the same database.

Topic

Secure Agent Groups



Secure Agent Groups Overview

- Use as a runtime environment to access data on-premise or in cloud repositories
- A secure agent within the group runs the task
- Prevents the activities of one department from impacting a different department
- All users in the organization can select the secure agent group as the runtime environment

25

© Informatica. Proprietary and Confidential.



You can use a secure agent group as the runtime environment to access data on-premise or in cloud repositories. When you select a secure agent group as the runtime environment for a connection or a task, a secure agent within the group runs the tasks.

You can create secure agent groups to prevent the activities of one department from impacting a different department. You can also create separate secure agent groups for test and production environments.

When you create a secure agent group, all users in the organization can select the secure agent group as the runtime environment.

Secure Agent Groups with Multiple Agents

Secure Agent Groups with Multiple Agents

By default, a secure agent is added to its own group

Use secure agent Cluster license to add multiple agents to one group

All agents within a group must be of the same type

File and directory structure of all secure agents within a group must be the same

26

© Informatica. Proprietary and Confidential.



When you install a secure agent, it is added to its own group by default. If you have the secure agent Cluster license, you can add multiple agents to one secure agent group. All agents within a group must be of the same type. This means that you must create a separate group for secure agents on Windows systems and a separate group for secure agents on Linux systems.

The file and directory structure of all secure agents within a group must be the same.

Benefits of Grouping Multiple Agents

Balance the distribution of tasks across machines

Group distributes tasks to available agents in a round-robin fashion

Improve scalability for connections and tasks

If the runtime environment is a secure agent group with multiple agents, the tasks can run if any secure agent in the group is up and running

27

© Informatica. Proprietary and Confidential.



You can add multiple agents to a group to balance the distribution of tasks across machines. When the runtime environment is a secure agent group with multiple agents, the group distributes tasks to the available agents in a round-robin fashion.

You can also add multiple agents to a group to improve scalability for connections and tasks. When you create a connection or task, you select the runtime environment to use. If the runtime environment is a secure agent group with multiple agents, the tasks can run if any secure agent in the group is up and running. You do not have to change the connection or task properties when you add or remove an agent, or if an agent in the group stops running.

Secure Agent Groups – Job Details

- View the job details to determine which secure agent ran the task
- To view job details:
 - Open Monitor Service, select Jobs, and click a job name

The first screenshot shows the 'My Services' dashboard with various integration components like Application Integration Console, ESB Gateway, Data Ingestion, Data Quality, Integration Hub, Administrator, DiscoveryIQ, and Monitor.

The second screenshot shows the 'Informatica Monitor' interface under 'All Jobs'. It lists several jobs, with 'XX_NormalizerAggregator_Task-2' highlighted in red.

The third screenshot is a 'Job Properties' window for 'XX_NormalizerAggregator_Task-2'. It displays details such as Task Name: XX_NormalizerAggregator_Task, Instance ID: 2, Task Type: Mapping Task, Started By: pkhow@informatica.com through API, Start Time: Aug 1, 2019 3:56:10 AM, End Time: Aug 1, 2019 3:56:25 AM, Duration: 15 seconds, Runtime Environment: CDI-JC-FIRSTNAME, and Secure Agent: CDI-JC-FIRSTNAME.

28

© Informatica. Proprietary and Confidential.



You can view the job details to determine which secure agent ran the task. To view these job details, open the Monitor service, select Jobs, and click a job name to determine which secure agent ran the job.

Shared Secure Agent Groups

- Administrator of a parent organization can share a secure agent group with the sub-organizations
- Secure agent group appears on the Runtime Environments page in all sub-organizations
- User in the sub-organization can select the shared secure agent group as the runtime environment
- Share a secure agent group to optimize the use of available secure agent resources
- You must have the Organization Hierarchy license to share a secure agent group

29

© Informatica. Proprietary and Confidential.



If you are the Administrator of a parent organization, you can share a secure agent group with the sub-organizations. When you share a secure agent group, the group appears on the Runtime Environments page in all sub-organizations. The sub-organizations can run tasks using the secure agents within the group.

When a user in the sub-organization creates a connection or a task, the user can select the shared secure agent group as the runtime environment.

You must share a secure agent group to optimize the use of available secure agent resources. To share a secure agent group, you must have the Organization Hierarchy license.

Manage Secure Agent Groups

- Create secure agent groups on the Runtime Environments page
- Share or stop sharing the secure agent group
- Rename or delete the secure agent group
- Add and remove secure agents
- Change group permissions

30

© Informatica. Proprietary and Confidential.



You can create secure agent groups on the Runtime Environments page. After you create a secure agent group, you can share or stop sharing the group, rename or delete the group, add and remove secure agents from the group, and change group permissions.

Topic

Data Transformation Manager Performance Properties



DTM Process Overview

- A DTM process is associated with the session run
- Creates and manages threads that carry out the session tasks
- Allocates process memory for the session and divides it into buffers
- Configure DTM session properties in mapping tasks

32

© Informatica. Proprietary and Confidential.



A DTM process is associated with the session run. The main purpose of the DTM process is to create and manage threads that carry out the session tasks. A DTM allocates process memory for the session and divides it into buffers. This is also known as buffer memory.

A DTM is an advanced session property that you can configure in mapping tasks.

DTM Buffer Size Configuration

- DTM buffer size specifies the amount of memory that is allocated to the task from the DTM process
- When you select the DTM buffer size advanced session property in a mapping task, you must specify the session property value as either 'Auto' or a numeric value
- When you select 'Auto' the task uses automatic memory settings
- You can also provide a numeric value for the session property

Advanced Session Properties		
Add	Session Property Name*	Session Property Value*
Remove	DTM Buffer Size	512MB

33

© Informatica. Proprietary and Confidential.



The DTM buffer size specifies the amount of memory that is allocated to the task from the DTM process. By default, a minimum of 12 MB is allocated to the buffer at run time.

When you select the DTM Buffer Size advanced session property in a mapping task, you must specify the session property value as either 'Auto' or a numeric value.

When you enter the session property value as 'Auto', the task uses automatic memory settings. When you use 'Auto', you must also configure the Maximum Memory Allowed for Auto Memory Attributes.

You can also provide a numeric value for the session property. For example, 512 KB or 512 MB.

DTM Buffer Size Configuration (continued)

- When a task contains large amounts of character data, increase the DTM buffer size to 24 MB
- When a task contains 'n' partitions, increase the DTM buffer size to at least 'n' times the value for the task with one partition
- When a source contains a large binary object with a precision larger than the allocated DTM buffer size, increase the DTM buffer size so that the task does not fail

Here are some scenarios where you may want to increase the DTM buffer size.

When a task contains large amounts of character data, you can increase the DTM buffer size to 24 MB.

When a task contains 'n' partitions, you can increase the DTM buffer size to at least 'n' times the value for the task with one partition.

When a source contains a large binary object with a precision larger than the allocated DTM buffer size, you can increase the DTM buffer size so that the task does not fail.

Module Summary

This module showed you how to:

- Describe partitions
- Explain types of partitions
- List partition rules and guidelines
- Discuss pushdown optimization
- List pushdown optimization types
- Discuss secure agent groups
- Discuss Data Transformation Manager (DTM) process and its configuration

35

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 18

Automating and Monitoring Tasks



Module Objectives

After completing this module, you will be able to:

- Define schedules
- Describe email notifications
- Discuss event monitoring

Topic

Automating Tasks



Schedules



Allows you to run tasks at a specific time or at regular intervals

Create a schedule from the Administrator service or from the Schedule step of the synchronization task wizard

Schedule Name

Start Time

Time Zone

Repeat Frequency

5

© Informatica. Proprietary and Confidential.



A schedule allows you to run tasks at a specific time or at regular intervals. You can create a schedule from the Administrator service or from the Schedule step of the synchronization task wizard.

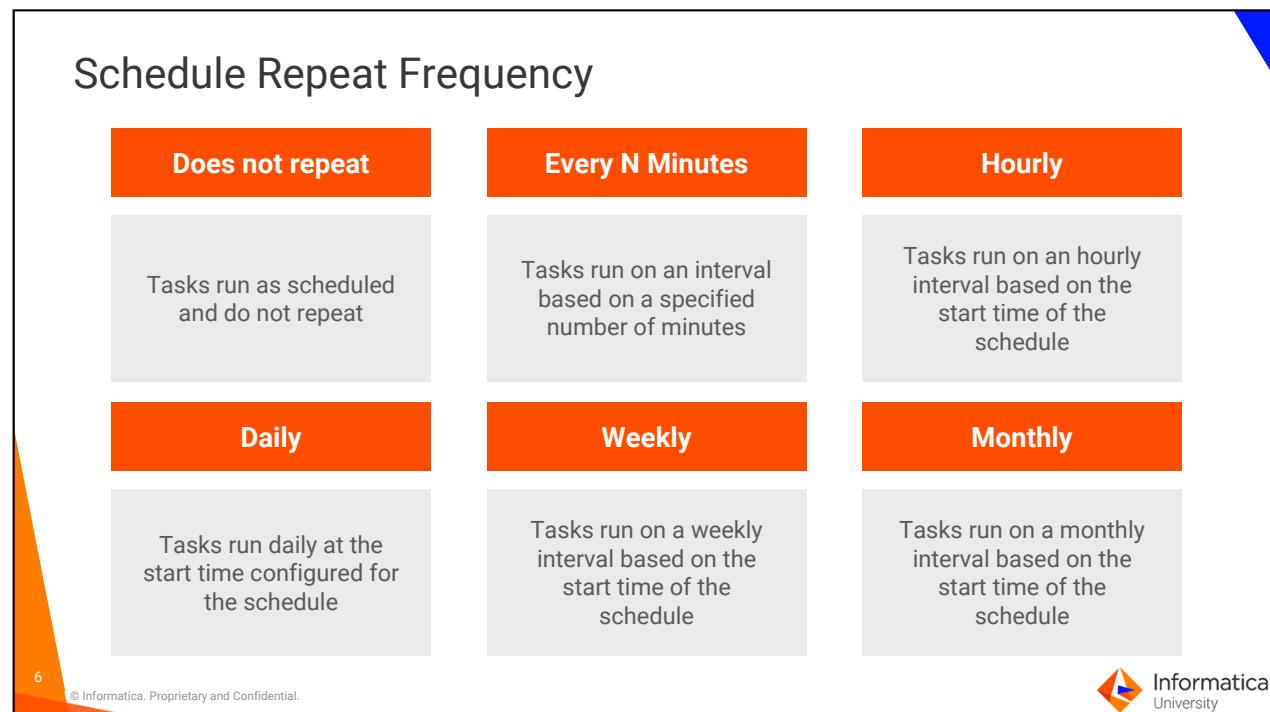
To create a schedule, you must specify the schedule information such as the schedule name, start time, time zone, and the repeat frequency for the schedule.

The schedule name specifies the name of the schedule.

The start time specifies the date and time for the schedule to start.

The time zone specifies the time zone for the schedule. The time zone can differ from the organization time zone or the user time zone.

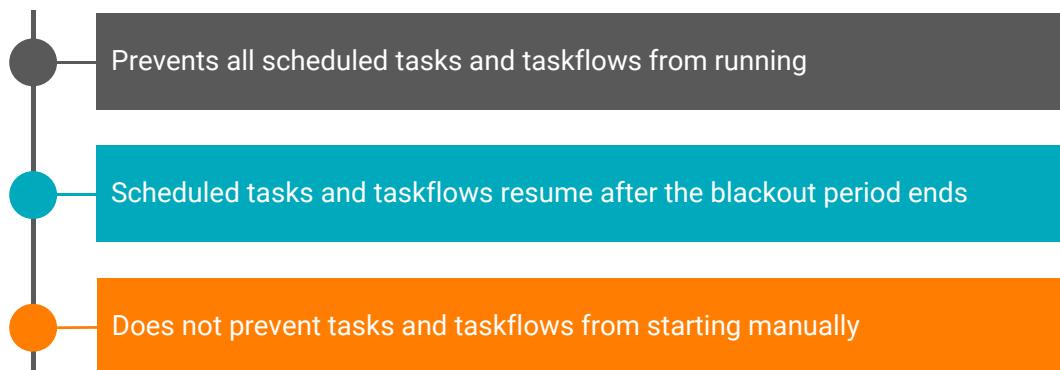
The repeat frequency specifies how often the tasks run.



When you create a schedule, you must choose one of the repeat frequency options such as does not repeat, every n minutes, hourly, daily, weekly, and monthly.

- Does not repeat: When you choose this option, the tasks run as scheduled and do not repeat.
- Every N minutes: This option allows you to run tasks on an interval, based on a specified number of minutes. You can set the value of the repeat frequency as 5, 10, 15, 20, 30, or 45 minutes.
- Hourly: When you select this option, the tasks run on an hourly interval, based on the start time of the schedule. You can set the value of the repeat frequency as 1, 2, 3, 4, 6, 8, or 12 hours.
- Daily: This option allows you to run tasks daily at the start time configured for the schedule. You can set the value of the repeat frequency as Every Day or Every Weekday.
- Weekly: You can choose this option to run tasks on a weekly interval, based on the start time of the schedule. You can set the value of the repeat frequency as one or more days of the week.
- Monthly: This option allows you to run tasks on a monthly interval, based on the start time of the schedule. You can set the value of the repeat frequency as the exact date of the month, between 1-28.

Schedule Blackout Period



7

© Informatica. Proprietary and Confidential.



What is a scheduled blackout period?

You can configure a black out period for the Org. A blackout period prevents all scheduled tasks and taskflows from running during a specified time period. After the blackout period ends, the scheduled tasks and taskflows run as per the defined schedule.

A blackout period does not prevent the tasks and taskflows from starting manually.

Topic

Monitoring Tasks



Email Notifications

- Allows you to monitor the status of tasks and taskflows via email messages
- You can configure notifications at the Org level or at the individual task and taskflow level

EXAMPLES

Error occurs and task and taskflow fails

Task and taskflow ran with a warning

Task and taskflow ran successfully

Email notifications allow you to monitor the status of the task and taskflow. You can configure email notifications at the Org level or at the individual task and taskflow level. When you configure email notifications at the Org level, the notifications are applicable to all tasks and taskflows in the Org. When you configure email notifications at the individual task and taskflow level, the notifications are applicable only to the individual task and taskflow.

Some examples where you can set up email notifications are:

- When an error occurs when the task and taskflow fails
- When a task and taskflow ran with a warning
- When a task and taskflow ran successfully

Email Notifications

Task Completed with Warnings

The Synchronization Task Demo_MultipleObject completed with warnings.

Organization ID:	iMZgmlu1DwSgxF9Np4Fq2T
Organization Name:	Informatica Demo
Start Time:	Jan 4, 2018 12:51:36 AM (PST)
End Time:	Jan 4, 2018 12:52:18 AM (PST)
Task Operation:	Insert
Source Connection:	SFDC_Demo
Source Object:	Price Book Entry, Price Book, Product
Target Connection:	SQLServer_Demo
Target Object:	Products
Result:	0 Success Rows, 34 Error Rows.

You can access your Informatica Cloud account at <https://dm-us.informaticacloud.com>

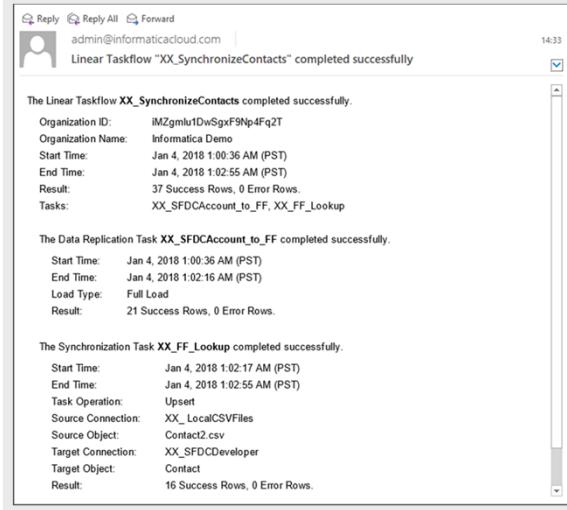
10 © Informatica. Proprietary and Confidential.

Informatica University

Here is an example of an email notification for a Synchronization task that completed with warnings. The notification provides information about the task, including the number of success and error rows.

Email Notifications

Taskflow Completed Successfully



The screenshot shows an email from admin@informaticacloud.com with the subject "Linear Taskflow "XX_SynchronizeContacts" completed successfully". The email body contains three sections of task completion details:

- The Linear Taskflow XX_SynchronizeContacts completed successfully.**

Organization ID:	IMZgmlu1DwSgxF9Np4Fq2T
Organization Name:	Informatica Demo
Start Time:	Jan 4, 2018 1:00:36 AM (PST)
End Time:	Jan 4, 2018 1:02:55 AM (PST)
Result:	37 Success Rows, 0 Error Rows.
Tasks:	XX_SFDCAccount_to_FF, XX_FF_Lookup
- The Data Replication Task XX_SFDCAccount_to_FF completed successfully.**

Start Time:	Jan 4, 2018 1:00:36 AM (PST)
End Time:	Jan 4, 2018 1:02:16 AM (PST)
Load Type:	Full Load
Result:	21 Success Rows, 0 Error Rows.
- The Synchronization Task XX_FF_Lookup completed successfully.**

Start Time:	Jan 4, 2018 1:02:17 AM (PST)
End Time:	Jan 4, 2018 1:02:55 AM (PST)
Task Operation:	Upsert
Source Connection:	XX_LocalCSVFiles
Source Object:	Contact2.csv
Target Connection:	XX_SFDCDeveloper
Target Object:	Contact
Result:	16 Success Rows, 0 Error Rows.

11 © Informatica. Proprietary and Confidential.

 Informatica University

Here is another example of an email notification for a taskflow that completed successfully. The notification provides information about the different tasks in the taskflow, and the number of success and error rows in each task.

Event Monitoring

- Use the asset and security logs to monitor events for the assets, licenses, users, and secure agents
- By default, the logs display events for the past 90 days
- To view the logs, you must be assigned a role that has the Audit Log – View privilege

12

© Informatica. Proprietary and Confidential.



You can monitor events for the assets, licenses, users, and secure agents in your organization through the asset and security logs. By default, the logs display events for the past 90 days. To change the length of time events appear in the logs, you must contact Informatica Global Customer Support. To view the logs, you must be assigned a role that has the **Audit Log View** privilege.

Event Monitoring – Asset Logs

- Displays events for assets
- Provides authentication events for users
- Displays information about events related to licenses
- To view the asset logs, in the Administrator Service, select **Logs**, and then select **Asset Logs**

The screenshot shows the Informatica Administrator Service interface. On the left is a sidebar with various navigation options: Organization, Licenses, SAML Setup, Users, User Groups, User Roles, Runtime Environments, Connections, Add-On Connectors, Schedules, Add-On Bundles, Swagger Files, and Logs. The 'Logs' option is highlighted with a blue bar at the bottom. On the right, the main content area has a title 'Asset Logs (1530)' with a dropdown arrow. Below it is a table with two columns: 'User Name' and 'Updated On'. The table lists several entries, each showing a user's email address and the date and time of the update.

User Name	Updated On
info.ilt.development@gmail.com	Jun 14, 2019, 2:08 AM
info.ilt.development@gmail.com	Jun 14, 2019, 2:08 AM
admin	Jun 13, 2019, 12:31 AM
info.ilt.development@gmail.com	Jun 12, 2019, 11:18 PM
info.ilt.development@gmail.com	Jun 11, 2019, 11:37 PM
info.ilt.development@gmail.com	Jun 11, 2019, 11:37 PM
info.ilt.development@gmail.com	Jun 11, 2019, 11:35 PM
info.ilt.development@gmail.com	Jun 11, 2019, 11:35 PM
info.ilt.development@gmail.com	Jun 11, 2019, 11:26 PM

13

© Informatica. Proprietary and Confidential.



The asset logs display events for assets such as when an asset was created, updated, copied, or deleted, and the name of the user who modified the asset.

You can view authentication events for users such as when a user in the organization logged in to IICS.

The asset logs also provide information about events related to licenses such as when a license was added, removed, or changed.

To view the asset logs, in the Administrator Service, select **Logs**, and then select **Asset Logs** at the top of the page.

Event Monitoring – Security Logs

- Displays events for secure agents and organizations
- To view the security logs, in the Administrator Service, select **Logs**, and then select **Security Logs**

The screenshot shows the Informatica Administrator Service interface. On the left is a sidebar with various navigation options: Organization, Licenses, SAML Setup, Users, User Groups, User Roles, Runtime Environments, Connections, Add-On Connectors, Schedules, Add-On Bundles, Swagger Files, and Logs. The 'Logs' option is highlighted with a yellow box. The main content area is titled 'Security Logs (8226)' and contains a table with two columns: 'User Name' and 'Updated On'. The table lists several entries, all of which are 'System built-in user' and were updated on 'Jun 15, 2019'. The last entry is 'Jun 15, 2019, 10:25 AM'.

User Name	Updated On
System built-in user	Jun 15, 2019, 2:02 PM
System built-in user	Jun 15, 2019, 11:28 AM
	Jun 15, 2019, 10:33 AM
	Jun 15, 2019, 10:31 AM
	Jun 15, 2019, 10:31 AM
	Jun 15, 2019, 10:26 AM
	Jun 15, 2019, 10:26 AM
	Jun 15, 2019, 10:26 AM
	Jun 15, 2019, 10:25 AM
	Jun 15, 2019, 10:25 AM

14

© Informatica. Proprietary and Confidential.



The security logs display events for secure agents and organizations such as when each agent was created or updated, when the organization information was updated, and the name of the user who modified the agent or the organization.

To view the security logs, in the Administrator Service, select **Logs**, and then select **Security Logs** at the top of the page.

Lab Activity

18-1 Creating a Schedule

In this lab, you will perform the following:

- Create a schedule

15

© Informatica. Proprietary and Confidential.



Module Summary

This module showed you how to:

- Define schedules
- Describe email notifications
- Discuss event monitoring



IICS: Cloud Data Integration Services

 Informatica
University

Module 19

Administration



Module Objectives

After completing this module, you will be able to:

- Discuss licenses in IICS
- Define user roles
- Explain types of user roles
- Describe users and user groups
- Explain permissions
- Discuss organization hierarchy
- Define sub-organization
- Explain asset import or export
- Define a bundle
- Manage bundles

3

© Informatica. Proprietary and Confidential.



Topic Licenses

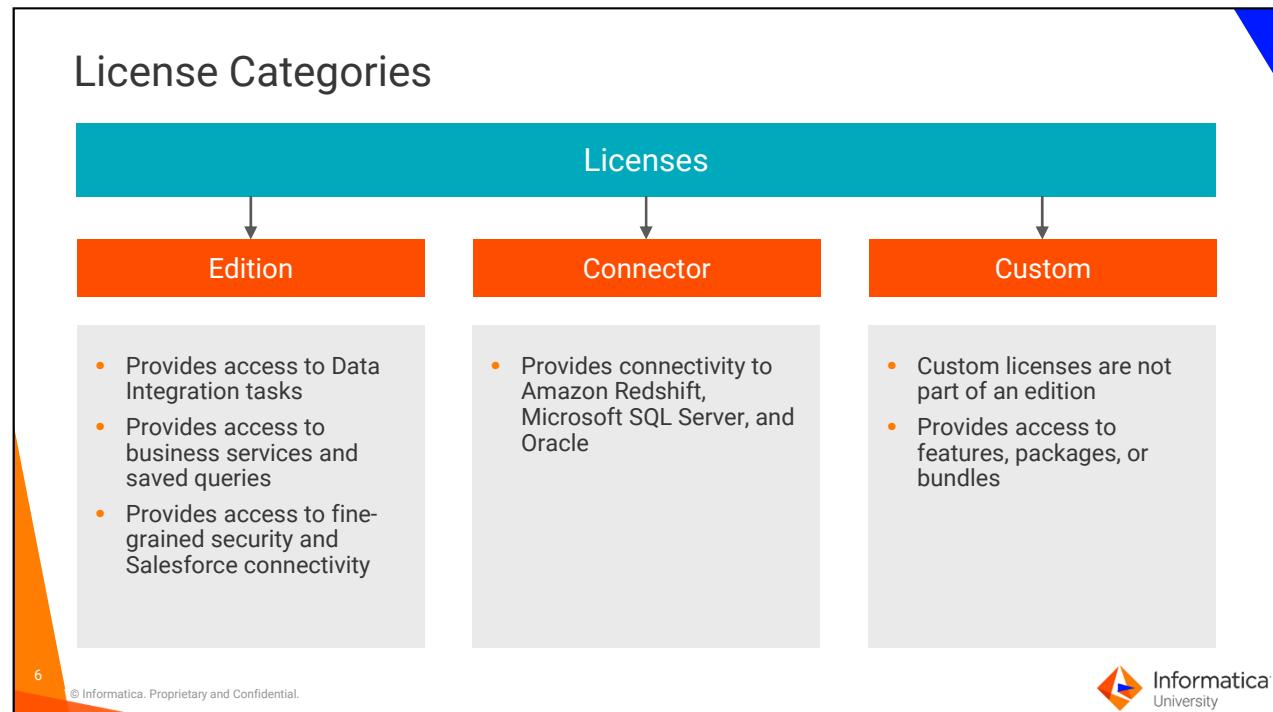


Licenses Overview

- Licenses determine the IICS subscription level for the organization and provide access to IICS features, connectors, and bundles
- The Administrator can review the licenses that are set up for the organization, verify license expiration dates, and check job limits and usage
- The Administrator can also manage sub-organization licenses

Licenses determine the IICS subscription level for the organization and provide access to IICS features, connectors, and bundles.

As an administrator, you can review the licenses set up for your organization, verify license expiration dates, and check job limits and usage. You can also manage sub-organization licenses and view job limits and usage for your sub-organizations.



Licenses are categorized as edition, connector, and custom licenses.

Edition licenses control the IICS features that you can use. They provide access to Data Integration tasks such as mapping tasks, replication tasks, and synchronization tasks. They also provide access to components such as business services and saved queries. Edition licenses also provide access to features such as fine-grained security and Salesforce connectivity.

Connector licenses provide connectivity to entities such as Amazon Redshift, Microsoft Sequel Server, and Oracle.

Custom licenses are licenses that are not part of an edition. They provide access to features, packages, or bundles. If your organization uses a custom license that provides access to a feature that is also included in an edition license, the terms of the custom license override the terms of the edition license.

License Types

Trial

- Use the edition free of charge for a 30-day period
- May provide limited access to the features, connectors, and packages

Subscription

- Use the licensed edition for the duration of the contract period
- Can renew the contract and continue to use the edition

Free Subscription

- Use the synchronization task free of charge
- May provide limited access to the features of the synchronization task

7

© Informatica. Proprietary and Confidential.



When you create an organization, IICS assigns the organization a license type for each licensed edition.

The types of licenses that IICS uses are trial, subscription, and free subscription.

Trial: You can use this edition free of charge for a 30-day period. At the end of the trial period, you can subscribe to the edition. A trial subscription may provide limited access to the features, connectors, and packages that are associated with the license.

Subscription: You can use the licensed edition for the duration of the contract period. Near the end of the contract period, IICS indicates that the contract is about to expire. You can renew the contract and continue to use the edition.

Free subscription: You can use the synchronization task free of charge. A free subscription may provide limited access to the features of the synchronization task.

Sub-organization Licenses

- A sub-organization represents different business environments within your organization
- Sub-organization has licenses maintained by the parent organization
- Sub-organization inherits all licenses from the parent organization, except for the Organization Hierarchy license and Bundle license
- The Administrator for the parent organization can disable, enable, and shorten the expiration dates for the inherited licenses
- The sub-organization administrator can view licenses but cannot change them

8

© Informatica. Proprietary and Confidential.



A sub-organization represents different business environments within your organization. For example, you can create separate sub-organizations for your development, testing, and production environments.

A sub-organization has licenses maintained by the parent organization. If a sub-organization requires a license that does not belong to the parent organization, you can contact Informatica Global Customer Support to obtain the license for the parent organization.

When you create a sub-organization, it inherits all the licenses from the parent organization, except for the Organization Hierarchy license and Bundle license. To use a bundle in the sub-organization, a user in the sub-organization must install the bundle.

The Administrator for the parent organization can disable, enable, and shorten the expiration dates for the inherited licenses. The sub-organization administrator can view licenses but cannot change them.

License Expiration

- When a license expires, you cannot access the features, connectors, or packages associated with the license
- Scheduled jobs associated with the license is also disabled
- If all licenses for the organization expire, you cannot log in to IICS
- Review the expiration date for licenses on the Licenses page in Administrator

Edition Name	License	Expiration Date
API and App Integration Trial	Trial	Nov 30, 2019
B2B Gateway Trial	Trial	Nov 30, 2019
Cloud Test Data Management Trial	Trial	Nov 30, 2019
Data Integration Advanced Trial	Trial	Nov 30, 2019
Data Masking Transformation Trial	Trial	Nov 30, 2019
Integration Hub Trial	Trial	Nov 30, 2019
Platform Advanced Trial	Trial	Nov 30, 2019

9

© Informatica. Proprietary and Confidential.



What happens when a license expires?

When a license expires, you cannot access the features, connectors, or packages associated with the license. Scheduled jobs associated with the license is also disabled. If all licenses for the organization expires, you cannot log in to IICS.

You can review the expiration date for licenses on the Licenses page in Administrator. To extend a license, you can contact Informatica Global Customer Support.

Topic

Administrator Service



Stop and Start Services that Run on a Secure Agent

- You can stop and start the microservices that run on a Secure Agent
- Stop and start Secure Agent services on the Agent details page in Administrator
- When you stop or start a Secure Agent service, other services that run on the agent are not impacted

Service Name	Enabled/Disabled	Status	Version	Last Update Time
File Integration Service	Enabled	Up and Running	7.3	Aug 16, 2019 4:27:43 PM
Data Integration Server	Enabled	Up and Running	55.0.3	Aug 16, 2019 4:29:03 PM
Mass Ingestion	Enabled	Up and Running	7.3	Aug 16, 2019 4:28:13 PM

11

© Informatica. Proprietary and Confidential.



You can stop and start the microservices that run on a Secure Agent to perform troubleshooting, to optimize resources on the agent machine, or when a service configuration changes. You can stop and start Secure Agent services on the Agent details page in Administrator. When you stop or start a Secure Agent service, other services that run on the agent are not impacted.

For example, if you encounter a problem with the Data Integration Server that runs on a Secure Agent, you can stop the service to perform troubleshooting. After you have finished troubleshooting, you can restart the service without affecting the other services that run on the agent.

Agent Blackout Periods

- You can configure blackout periods for a Secure Agent
- Blackout periods prevent Data Integration jobs from running during a certain time period
- Create an XML file that specifies the repeat frequency, start date, and end date for each blackout period

12

© Informatica. Proprietary and Confidential.



You can configure blackout periods for a Secure Agent. Blackout periods prevent Data Integration jobs from running on the agent during a certain time period.

To configure a blackout period on a Secure Agent, you create an XML file that specifies the repeat frequency, start date, and end date for each blackout period.

View Object Dependencies

- View object dependencies for Secure Agent groups and connections
- Administrator lists the connections and assets in each service that use the group as the runtime environment
- Administrator lists the runtime environments that the connection uses as well as the assets in each service that use the connection
- You can view, edit, and delete assets from the Dependencies page

Name	Type	Status
Analysis_mapping	Mapping	Valid
AT_Salesforce_to_FF	Synchronization Task	View Delete
Copy_Analysis_mapping	Mapping	Show Dependencies

13

© Informatica. Proprietary and Confidential.



You can view object dependencies for Secure Agent groups and connections.

When you view dependencies for a Secure Agent group, Administrator lists the connections and assets in each service that use the group as the runtime environment.

When you view dependencies for a connection, Administrator lists the runtime environments that the connection uses as well as the assets in each service that use the connection.

You can view, edit, and delete assets from the Dependencies page.

CLAIRE Recommendation Preferences

- Enable or disable CLAIRE recommendations for your organization
- CLAIRE recommendations allow in-product recommendations for mapping design

The screenshot shows the Informatica Administrator interface. On the left is a sidebar with various navigation options: Organization, Licenses, SAML Setup, Users, User Groups, User Roles, Runtime Environments, Connections, Add-On Connectors, Schedules, Add-On Bundles, and Swagger Files. The main panel is titled "CLAIRE Recommendation Preferences". It contains settings for the Data Integration Service, specifically for Jobs. Under "Jobs", the "Schedule Offset" is set to "30 seconds" and the "Time Zone" is set to "Pacific Daylight Time, Los Angeles". At the bottom of the preferences panel, there is a checkbox labeled "Enable CLAIRE recommendations" which is checked.

14

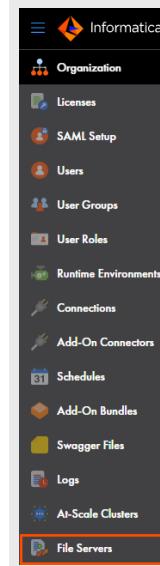
© Informatica. Proprietary and Confidential.



You can enable or disable CLAIRE recommendations for your organization. CLAIRE recommendations allow in-product recommendations for mapping design based on the analysis of metadata from your organization's assets, and assets from other IICS organizations. CLAIRE recommendations are enabled by default.

File Server Configuration

- You can configure file servers such as AS2 to run on each agent that uses the File Integration Service
- You can also configure partner users
- Configure file servers and partner users on the File Servers page in Administrator



 Informatica
University

15

You can configure file servers such as AS2 to run on each agent that uses the File Integration Service. You can also configure partner users so that they can connect to the servers to send files.

You can configure file servers and partner users on the **File Servers** page in Administrator.

Salesforce User Activation

- Activate the user account:
 - By using a verification code
 - By using Salesforce OAuth authentication

16

© Informatica. Proprietary and Confidential.



When you create a user account that uses Salesforce authentication, you can choose whether to activate the user account using a verification code or using Salesforce OAuth authentication.

Administration Overview

Use a combination of roles, user groups, and object-level permissions to secure objects

Role defines the general tasks that the user can perform

User group defines the objects that the user can work with, and the tasks the user can perform on objects

Object level permissions allows you to add or remove individual object instances from the user group domain

17

© Informatica. Proprietary and Confidential.



You can use a combination of roles, user groups, and object-level permissions to secure objects and data in your organization.

A role defines the general tasks that the user can perform within the organization. For example, users with Admin role can create and manage users.

A user group defines the objects that the user can work with and the tasks the user can perform on objects. For example, a user can only read and use connections in a task but cannot create new connections.

An object level permission allows you to add or remove individual object instances from the user group domain. For example, you can lock individual objects such as connections, tasks, or taskflows so that only users within an assigned user group can access the objects.

Topic

User Roles



What is a Role?

- Set of privileges that allows a user to perform tasks
- Roles determine the functionalities available to a user
- Assign at least one role to each user
- Can assign a system-defined role or a custom role

The screenshot shows the Informatica Administrator interface with the 'User Roles' page selected. The left sidebar lists various administrator functions: Organization, Licenses, SAML Setup, Users, User Groups, and User Roles (which is highlighted). Other listed items include Runtime Environments, Connections, Add-On Connectors, Schedules, Add-On Bundles, Swagger Files, and Logs. The main content area is titled 'User Roles' and displays a table with 10 entries. The table columns are 'Name', 'Last Updated', and 'Status'. The entries are: Admin (Mar 8, 2019, 4:20 PM, Enabled), Application Integration Business M... (Oct 13, 2018, 8:00 AM, Enabled), Application Integration Data Viewer (Oct 13, 2018, 8:00 AM, Enabled), Data Integration Data Previewer (Mar 8, 2019, 4:20 PM, Enabled), Data Integration Task Executor (Nov 27, 2018, 9:50 PM, Enabled), Deployer (Oct 13, 2018, 8:00 AM, Enabled), Designer (Mar 8, 2019, 4:20 PM, Enabled), Monitor (Nov 27, 2018, 9:50 PM, Enabled), Operator (Oct 13, 2018, 8:00 AM, Enabled), and Service Consumer (Mar 8, 2019, 4:20 PM, Enabled).

19

© Informatica. Proprietary and Confidential.



A role is a set of privileges that allows a user to perform tasks in the organization. Roles determine the functionalities available to a user. For example, to perform Administrative tasks, the user must have the Admin role. You must assign each user in the Org at least one role. While there is no technical limitation on assigning multiple roles to a single user, the best practice is to assign only one role to each user.

In IICS, you can assign a system-defined role or a custom role to a user.

To view the User Roles page, in the Administrator Service, select **User Roles**.

System-defined Roles

- System-defined roles are pre-defined roles that define access privileges for the services your organization uses
- You cannot edit or delete system-defined roles
- Roles available in IICS:

Admin	Designer	Monitor	Service Consumer
<ul style="list-style-type: none">• Have full access to all licensed services and can perform all tasks in the organization	<ul style="list-style-type: none">• Can create assets, tasks, connections, schedules, and runtime environments• Can monitor jobs	<ul style="list-style-type: none">• Can monitor Data Integration jobs	<ul style="list-style-type: none">• Can run tasks and taskflows

20

© Informatica. Proprietary and Confidential.



System-defined roles are pre-defined roles that define access privileges for the services your organization uses. The system-defined roles you assign to users and groups vary based on your organization's licenses. You cannot edit or delete system-defined roles.

Admin Role: Users with the Admin role have full access to all licensed services and can perform all tasks in the organization.

Designer Role: Users with the Designer role can create assets, tasks, connections, schedules, and runtime environments. They can also monitor jobs. They cannot perform administrative tasks for the organization.

Monitor Role: Users with a Monitor role can monitor Data Integration jobs.

Service Consumer Role: Users with the Service Consumer role can run tasks and taskflows. However, they cannot create or edit assets.

Custom Roles

Create custom roles based on the needs of the organization

You can edit and delete custom roles

To create custom roles, organization must have the Custom Roles license

21

© Informatica. Proprietary and Confidential.



You can create a custom role based on the business requirements of your organization. For example, you can create a custom administrator role that can configure roles, user groups, and access control, but cannot create, edit, or run Data Integration tasks.

You can edit and delete custom roles after you create them.

To create custom roles, your organization must have the “Custom Roles” license.

Role Details

- Role details page displays information about a role, including the asset and feature privileges
- For system-defined roles, you can view the role information and privileges
- For custom roles, you can view and change the role information and the assigned privileges
- Configure the following information on the role details page:
 - Role Name
 - Description
 - Services
 - Assets
 - Features

Asset Type	Create	Read	Update	Delete	Run	Set Permission
Azure Data Sync Task	<input checked="" type="checkbox"/>					
Business Service Definition	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Data Masking Task	<input checked="" type="checkbox"/>					
File Listener	<input checked="" type="checkbox"/>					

22

© Informatica. Proprietary and Confidential.



The role details page displays information about a role, including the asset and feature privileges associated with the role. For system-defined roles, you can only view the role information and privileges. For custom roles, you can view AND change the role information and the assigned asset and feature privileges.

To display the role details page, in the Administrator Service, select **User Roles**, and then click the role name.

- **Role Name:** This specifies the name you provide to the role.
- **Description:** This specifies a short description about the role.
- **Services:** This indicates the name of the service for which privileges are enabled or disabled. You can select a service to view the asset and feature privileges associated with it.
- **Assets:** This tab specifies the asset privileges for the selected service. The asset privileges control access to different types of assets.
- **Features:** This tab specifies the feature privileges for the selected service. The feature privileges are general privileges that control the ability to use the features of a service.

Topic

Users and User Groups



Users

- A user is an individual IICS account that allows secure access to an organization
- A user can perform tasks and access assets based on the roles assigned to the user
- While creating a new user, configure the following:
 - User Information
 - Login Settings
 - Assigned User Groups and Roles

Enabled	Group Name	Description	Enabled	Role Name	Description
<input type="checkbox"/>	Lab_Group		<input type="checkbox"/>	Admin	Role for performing administrative tasks for
					Application Integration Business M.
					Role used for business managers

24

© Informatica. Proprietary and Confidential.



A user is an individual IICS account that allows secure access to an organization. A user can perform tasks and access assets based on the roles that are assigned to the user. The Administrator can create and configure user accounts for the organization.

When you create a new user, you must configure certain properties such as the User Information, Login Settings, and Assigned User Groups and Roles.

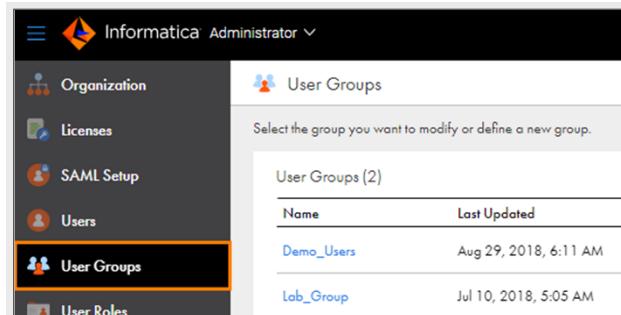
For user information, you must provide the user's first name, last name, job title, phone number, email, and description.

For Login Settings, you must provide the authentication method, user name, and maximum login attempts.

For the Assigned User Groups and Roles, you must assign at least one user group or role to each user.

User Groups

- Group of users in which all members can perform the same tasks and have the same access rights for different types of assets
- The Administrator can:
 - View and edit user group details
 - Create a group
 - Rename a group
 - Delete a group



The screenshot shows the Informatica Administrator interface. On the left is a sidebar with options: Organization, Licenses, SAML Setup, Users, User Groups (which is selected and highlighted in orange), and User Roles. The main panel title is "User Groups". It displays a message: "Select the group you want to modify or define a new group." Below this is a table titled "User Groups (2)".

Name	Last Updated
Demo_Users	Aug 29, 2018, 6:11 AM
Lab_Group	Jul 10, 2018, 5:05 AM

25

© Informatica. Proprietary and Confidential.



A user group is a group of users in which all members can perform the same tasks and have the same access rights for different types of assets. Members of a group can perform tasks and access assets based on the roles that the Administrator assigns to the group.

The Administrator can view and edit user group details, create a group, rename a group, and delete a group.

To view the User Groups page, in the Administrator Service, select **User Groups**.

Topic Permissions



Permissions

- Permissions determine the access rights that a user has for an object
- Permissions define which users and groups can read, update, delete, execute, and change permissions on the object

27

© Informatica. Proprietary and Confidential.



Permissions determine the access rights that a user has for a secure agent, secure agent group, connection, schedule, or an asset. Permissions add additional or custom security for an object. They define which users and groups can read, update, delete, execute, and change permissions on the object.

Permissions – Licenses and Privileges

- To configure permissions at the project level, your organization must have the **Set/Unset Security Permissions at Project Level license**
- If you want to configure permissions at the folder level, your organization must have the **Set/Unset Security Permissions at Folder Level license**
- To configure permissions for individual assets, your organization must have the **Fine Grained Security license**
- The role assigned to your user account or to a group in which you are a member must have the **Set Permission privilege** for the object type

28

© Informatica. Proprietary and Confidential.



To configure permissions at the project level for all assets in a project, your organization must have the **Set or Unset Security Permissions at Project Level license**.

If you want to configure permissions at the folder level for all assets in a folder, your organization must have the **Set or Unset Security Permissions at Folder Level license**.

To configure permissions for individual assets, your organization must have the **Fine Grained Security license**.

The role assigned to your user account or to a group in which you are a member must have the **Set Permission** privilege for the object type. For example, to configure permissions for a Secure Agent, you must be assigned a role that has the Set Permission privilege for Secure Agents.

Object-Level Permissions

- Navigate to the object or asset and set the appropriate permissions
- Permissions apply to the objects for which you configure them
- Permissions do not apply to copies of the object

The screenshot shows the Informatica Data Integration interface. On the left, there's a sidebar with options like 'New...', 'Home', 'Explore', 'Bundles', 'My Jobs', and 'My Import/Export Logs'. The main area is titled 'Explore' and shows a list of objects under 'Default' project. The list includes:

Name	Type	Updated On	Location	Description	Tags	Status
XX_FirstName_Employee	Synchronization	May 28, 2019, 6:04 AM	Default			Valid
XX_FirstName_Employees	Mapping	Jun 5, 2019, 3:52 AM	Default			
XX_FirstName_ErrorHandling	Mapping	Jun 3, 2019, 11:06 AM	Default			
XX_FirstName_HierarchyBuilder	Mapping	Jun 6, 2019, 12:32 AM	Default			
XX_FirstName_LinearTaskFlow	Linear Task	Jun 3, 2019, 10:45 PM	Default			
XX_FirstName_LookupOverride	Mapping	Jun 6, 2019, 2:58 AM	Default			
XX_FirstName_MappingParameters	Mapping	Jun 1, 2019, 5:46 AM	Default			
XX_FirstName_MappingParameters-Is...	Mapping	May 31, 2019, 3:35 AM	Default			
XX_FirstName_NormalizerAggregator	Mapping	May 30, 2019, 7:35 AM	Default			

A context menu is open over the first row ('XX_FirstName_Employee'). The menu items are: Properties..., View, Edit, Run, Download XML..., Rename..., Copy To..., Move To..., Export..., Permissions... (which is highlighted with a red box), and Delete.

29

© Informatica. Proprietary and Confidential.



To configure permissions for an object or asset in the Data Integration service, navigate to the object or asset and set the appropriate permissions.

You must note that, permissions apply to the objects for which you configure them but not to copies of the object. Therefore, when you copy or export an asset, the permissions are not copied or exported with the asset.

Permissions Types

Permission	Description
Read	Open and view objects
Create	Create objects
Update	Edit objects
Delete	Delete objects
Permission	Run objects
Run	Change the permissions that are assigned to objects

30

© Informatica. Proprietary and Confidential.



Rules and Guidelines for Permissions

Verify that you assign a role to the user or group with the appropriate privileges for the object type

To configure or edit a taskflow, you must have **Execute** permission for all tasks in the taskflow

To run a taskflow, you must have **Read** and **Execute** permissions on taskflows

To monitor or stop jobs, you must have **Execute** permission for the mapping, task, or taskflow

If you do not configure permission for an asset, then the asset has no permission restriction

Topic

Organization Hierarchy



What is an Organization Hierarchy?

- Hierarchy of related organizations
- Includes a parent organization and one or more sub-organizations
- To create an organization hierarchy, the parent organization must have the **Org hierarchy** license
- Administrator of the parent organization can create and manage organizations and organization hierarchy

33

© Informatica. Proprietary and Confidential.



An organization hierarchy is a hierarchy of related organizations. It includes a parent organization and one or more sub-organizations.

To create an organization hierarchy, the parent organization must have the Org hierarchy license. The Administrator of the parent organization can create and manage organizations and the organization hierarchy.

A user in one Org in the hierarchy cannot log into another Org in the hierarchy without a user account for the other Org.

Sub-Organization

- Administrator of the parent organization can create a sub-organization
- A sub-organization is related to the parent organization as part of the organization hierarchy
- A sub-organization inherits all licenses and subscriptions of the parent organization, except for the Org hierarchy license
- A sub-organization cannot act as a parent for any other organizations or be part of another organization hierarchy
- Configure sub-organization's security and create user accounts for the sub-organization

34

© Informatica. Proprietary and Confidential.



The Administrator of the parent organization can create a sub-organization. A sub-organization is an Informatica Cloud Org that is related to the parent organization as part of the organization hierarchy.

The sub-organization inherits all licenses and subscription options of the parent organization, except for Org hierarchy license. A sub-organization cannot act as a parent for any other organization or be part of another organization hierarchy.

An organization hierarchy can include a limited number of sub-organizations. If you want to increase the number of sub-organizations for the Org, you can contact Informatica Global Support.

When you create a sub-organization, you must also configure sub-organization's security and create user accounts for the sub-organization.

Linking Existing Organization as Sub-Organization

- To link an existing organization as a sub-organization, the following conditions must be met:

 You must have a user account in the existing organization

 The existing organization must not be a part of another organization hierarchy

 You must be the Administrator of the parent organization and the parent organization must have the Org hierarchy license

NOTE: After you link an existing organization as a sub-organization, the existing organization cannot act as a parent for other organizations.

35

© Informatica. Proprietary and Confidential.



There are certain conditions that must be fulfilled if you want to link an existing organization as a sub-organization.

- You must have a user account in the existing organization.
- The existing organization must not be a part of another organization hierarchy.
- You must be the Administrator of the parent organization and the parent organization must have the Org hierarchy license.

After you link an existing organization as a sub-organization, the existing organization cannot act as a parent for other organizations.

Importing/Exporting Assets – Methods

METHOD 1

- You can log in to the sub-organization and import or export assets from within the sub-organization

METHOD 2

- If you are the administrator of the parent organization, you can log in to the parent organization, switch to the sub-organization, and import or export assets

There are two ways you can import or export assets:

1. You can log in to the sub-organization and import or export assets from within the sub-organization.
2. If you are the administrator of the parent organization, you can log in to the parent organization, switch to the sub-organization, and import or export assets.

Importing/Exporting Assets

- You can import or export the following types of assets:
 - Tasks
 - Taskflows
 - Mapplets
 - Saved Queries
- IICS imports or exports all dependent assets
- IICS does not import or export schedule information

37

© Informatica. Proprietary and Confidential.



You can import or export Tasks, Taskflows, Mapplets, and Saved Queries.

When you import or export an asset, IICS imports or exports all dependent assets. For example, when you import a Task, IICS imports all connections in the Task.

IICS does not import or export the schedule information. So, you must re-assign schedules after the Task or Taskflow is imported or exported to the target Org.

Import/Export Rules

For security reasons, connections are imported or exported without passwords

When you import or export a connection, it is automatically assigned to the first secure agent in the target Org

You cannot import or export an asset that has the same name as an existing asset in the target Org

38

© Informatica. Proprietary and Confidential.



There are some important points to remember when you import or export assets.

For security reasons, connections are imported or exported without passwords. For Salesforce connections, IICS does not import or export the security tokens. So, you must update and verify all connections and Salesforce security tokens in the target Org.

When you import or export a connection, it is automatically assigned to the first secure agent in the target Org. If there are multiple secure agents in the target Org, you may have to re-assign the secure agent for the connection.

You cannot import or export an asset that has the same name as an existing asset in the target Org. As a best practice, you must assign version numbers to assets in the originating Org.

Topic

Bundle Management



Overview of Bundles

- A bundle is a set of related mappings, mapping tasks, and mapplets
- Data Integration users design, create, and publish the bundles
- Administrators manage the bundles
- To view the bundles for your organization, in Administrator, select 'Add-On Bundles'

The screenshot shows the Informatica Administrator interface. On the left is a sidebar with various management options: Organization, Licenses, SAML Setup, Users, User Groups, User Roles, Runtime Environments, Connections, Add-On Connectors, Schedules, and Add-On Bundles. The 'Add-On Bundles' option is selected and highlighted in orange. The main content area is titled 'Add-On Bundles' and displays three tabs: 'Installed Bundles' (selected), 'Copied Bundles', and 'Available Bundles'. Under 'Installed Bundles', there is a table with one entry:

Name	Description	Publisher	Version
EDI Gateway	Informatica Cloud B2B Gateway ...	Informatica	5.0

40

© Informatica. Proprietary and Confidential.



A bundle is a set of related mappings, mapping tasks, and mapplets that Data Integration users can use in Data Integration projects. The Data Integration users design, create, and publish these bundles. The Administrators manage the bundles.

To view the bundles that are installed or are available for your organization, in Administrator, select **Add-On Bundles**. The **Add-on Bundles** page displays information about installed bundles, copied bundles, and bundles that are available for installation or copying.

Bundle Management

INSTALL

- Install a public, private, or unlisted bundle that the bundle designer has configured to be used as a reference
- Data Integration users in your organization can use the assets in the bundle, but they cannot edit the assets

COPY

- Copy a public, private, or unlisted bundle that the bundle designer has configured for copying
- Data Integration users in your organization can edit the assets

UPGRADE

- If you installed a bundle and a newer version of the bundle is available, you can upgrade the bundle to the latest version

UNINSTALL

- If your organization no longer needs an installed bundle, you can uninstall it

41

© Informatica. Proprietary and Confidential.



If you are the administrator for an organization, you can install, copy, upgrade, and uninstall a bundle.

You can install a public, private, or unlisted bundle that the bundle designer has configured to be used as a reference. When you install a bundle, the bundle is installed into the Add-On Bundles project in Data Integration. Data Integration users in your organization can use the assets in the bundle, but they cannot edit the assets.

You can copy a public, private, or unlisted bundle that the bundle designer has configured for copying. When you copy a bundle, you select the Data Integration folder where you want to copy the bundle contents. You can copy a bundle multiple times and save the contents into different projects or folders. After you copy a bundle, Data Integration users in your organization can edit the assets.

If you installed a bundle and a newer version of the bundle is available, you can upgrade the bundle to the latest version. If your organization no longer needs an installed bundle, you can uninstall it.

Lab Activity

Appendix 1: Configure Administrative Settings for Your Informatica Cloud Org

In this lab, you will perform the following:

- Configure Administrative settings for the Org

Lab Activity

Appendix 2: Creating a Sub-Organization and Importing/Exporting Assets

In this lab, you will perform the following:

- Create a Sub-Org for testing environment
- Export an asset from an Org
- Import an asset to the Sub-Org

Demonstration

User Roles, User Groups, and Permissions

- View the video [Module19_UserRoles_UserGroups_and_Permissions.mp4](#)

Module Summary

This module showed you how to:

- Discuss licenses in IICS
- Define user roles
- Explain types of user roles
- Describe users and user groups
- Explain permissions
- Discuss organization hierarchy
- Define sub-organization
- Explain asset import or export
- Define a bundle
- Manage bundles

45

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica
University

Module 20

SAML Setup



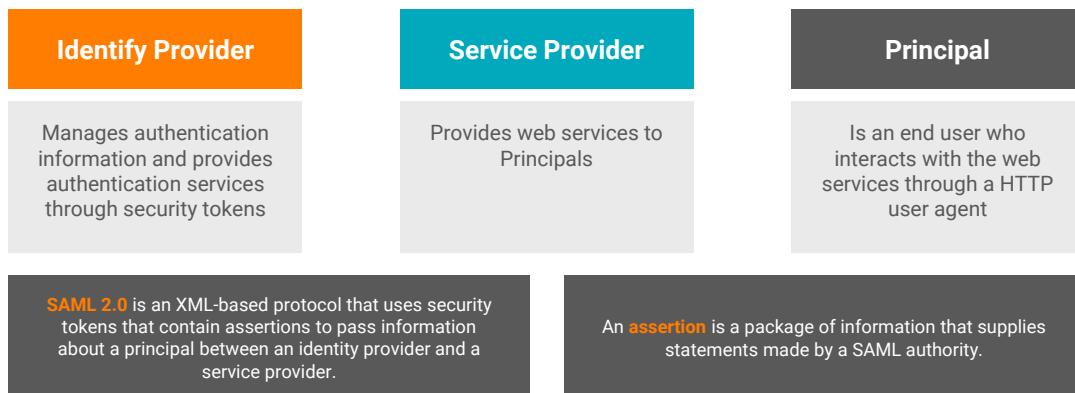
Module Objectives

After completing this module, you will be able to:

- Discuss SAML single sign-on
- List the single sign-on requirements
- Discuss single sign-on restrictions
- Explain SAML single sign-on configuration for IICS

SAML Single Sign-on

- Single sign-on allows users to access their organization without having to enter login information
- Single sign-on is based on the SAML 2.0 web browser single sign-on profile



4

© Informatica. Proprietary and Confidential.



You can enable single sign-on or SSO capability so that the users can access their organization without the need to enter login information.

Single sign-on to IICS is based on the Security Assertion Markup Language or SAML 2.0 web browser single sign-on profile. The SAML web browser single sign-on profile consists of an Identity provider, a Service provider, and a Principal.

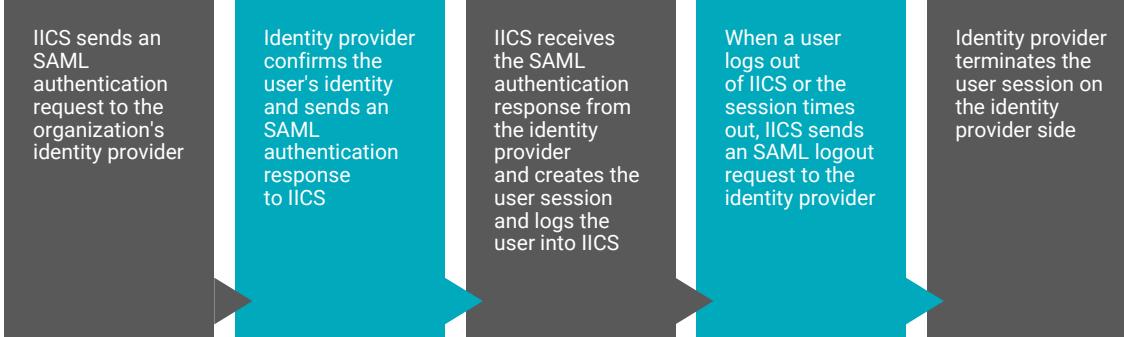
Identity provider is an entity that manages authentication information and provides authentication services through the use of security tokens.

Service provider is an entity that provides web services to Principals. For example, IICS is a service provider that hosts web applications.

Principal is an end user who interacts with the web services through a HTTP user agent.

SAML 2.0 is an XML-based protocol that uses security tokens that contain assertions to pass information about a principal between an identity provider and a service provider. An assertion is a package of information that supplies statements made by a SAML authority.

SAML Single Sign-on Process



5

© Informatica. Proprietary and Confidential.



When a user enters the IICS single sign-on URL in a browser:

- IICS sends an SAML authentication request to the organization's identity provider
- The identity provider confirms the user's identity and sends an SAML authentication response to IICS
- When IICS receives the SAML authentication response from the identity provider, it creates the user session and logs the user into IICS
- When a user logs out of IICS or the session times out, IICS sends an SAML logout request to the identity provider
- The identity provider then terminates the user session on the identity provider side

SAML Single Sign-on Requirements

- System must use a SAML 2.0-based identity provider
- Identity provider must be configured to use either the DSA-SHA1 or RSA-SHA1 algorithm to generate the signature
- IICS organization must have the SAML based Single Sign-On license
- You must be the Administrator of the IICS organization

6

© Informatica. Proprietary and Confidential.



To set up SAML single sign-on for an organization:

- The system must use an SAML 2.0-based identity provider. Some of the common identity providers include Microsoft Active Directory Federation Services, Okta, SSO Circle, and Open LDAP. The identity provider must be configured to use either the DSASHA1 or RSASHA1 algorithm to generate the signature.
- The IICS organization must have the SAML based Single Sign-On license.
- You must be the Administrator of the IICS organization.

Single Sign-on Restrictions

- You cannot use the single sign-on to register the secure agent in your organization
- To perform tasks that require a password, you must log in to IICS directly
- If your license with the identity provider expires, you cannot access IICS through single sign-on
- If the identity provider's service is down, you cannot log in to IICS through single sign-on
- If the identity provider certificate used for SAML single sign-on to IICS expires, you cannot access IICS through single sign-on
- If your organization uses trusted IP address ranges, you cannot log in to IICS from an IP address that is not within the trusted IP address ranges
- Connections are not authenticated when you use single sign-on

7

© Informatica. Proprietary and Confidential.



There are some restrictions that apply to SAML single sign-on access:

- You cannot use the single sign-on to register the secure agent in your organization. You must log in with an IICS user name and password to register the secure agent.
- When you access IICS through single sign-on, your password is not known to IICS. Therefore, to perform tasks that require a password, you must log in to IICS directly.
- If your license with the identity provider expires, you cannot access IICS through single sign-on.
- If the identity provider's service is down or IICS servers cannot reach it, you cannot log in to IICS through single sign-on.
- If the identity provider certificate used for SAML single sign-on to IICS expires, you cannot access IICS through single sign-on.
- If your organization uses trusted IP address ranges, you cannot log in to IICS from an IP address that is not within the trusted IP address ranges, and finally
- Connections are not authenticated when you use single sign-on.

SAML Single Sign-on Configuration for IICS

- IICS and your identity provider exchange configuration information when you set up single sign-on
- IICS requires identity provider metadata to send authentication requests to the identity provider
- Identity provider requires the service provider metadata from IICS to send authentication responses to IICS
- SAML and IICS attributes such as user roles must be mapped
- After you configure single sign-on, pass the IICS service provider metadata to your identity provider

IICS and your identity provider exchange configuration information when you set up single sign-on.

IICS requires the identity provider metadata to send authentication requests to the identity provider. The identity provider requires the service provider metadata from IICS to send authentication responses to IICS.

SAML and IICS attributes such as user roles must be mapped so that IICS can consume the data that is passed in authentication responses. After you configure single sign-on settings in IICS, you must pass the IICS service provider metadata to your identity provider.

SAML Single Sign-on Configuration for IICS (continued)

- To configure single sign-on for IICS, provide the following information:
 - Identity provider properties
 - Service provider properties
 - SAML attribute mapping properties
 - SAML role mapping properties

9

© Informatica. Proprietary and Confidential.



To configure SAML single sign-on for IICS, you must provide the Identity provider properties, Service provider properties, SAML attribute mapping properties, and SAML role mapping properties.

In the next few pages, you will see how to configure each of these properties on the **SAML Setup page** in IICS.

Identity Provider Properties

- You can use the identity provider XML file to populate some of the identity provider properties
- IICS parses and extracts most of the data from the XML file

10

© Informatica. Proprietary and Confidential.



If you have an identity provider XML file, you can use the file to populate some of the identity provider properties. IICS parses and extracts most of the data from the XML file. However, you may have to enter certain fields manually.

Identity Provider Properties (continued)

- **Use Identity Provider File** specifies the identity provider XML file
- **Disable auto provisioning of users** allows you to disable auto provisioning of SAML users
- **Issuer** specifies the entity ID of the identity provider
- **Single Sign-On Service URL** specifies the identity provider's HTTP-POST SAML binding URL for the SingleSignOnService
- **Single Logout Service URL** specifies the identity provider's HTTP-POST SAML binding URL for the SingleLogoutService
- **Signing Certificate** specifies the Base64-encoded PEM format identity provider certificate

Configure Single Sign-On (SSO) using Security Assertion Markup Language

SSO Configuration

Use Identity Provider File: Drop file here

Disable auto provision of users

Identity Provider Configuration

Issuer*:

Single Sign-On Service URL*:

Single Logout Service URL:

Signing Certificate*:

Use signing certificate for encryptions

Encryption Certificate:

Name Identifier Format:

Logout Service URL(SOAPBINDING):

Logout Page URL:

11

© Informatica. Proprietary and Confidential.



The **Use Identity Provider File** property specifies the identity provider XML file. If you have the identity provider XML file, you can click the **Choose File** button and select the file.

The **Disable auto provisioning of users** option allows you to disable auto provisioning of SAML users. So when a new SAML user logs in to IICS for the first time, the user will not be added to the organization in IICS.

The **Issuer** property specifies the entity ID of the identity provider, which is the unique identifier of the identity provider.

The **Single Sign-On Service URL** property specifies the identity provider's HTTP POST SAML binding URL for the Single Sign On Service. IICS sends login requests to this URL.

The **Single Logout Service URL** property specifies the identity provider's HTTP POST SAML binding URL for the Single Logout Service. IICS sends logout requests to this URL.

The **Signing Certificate** property specifies the Base 64-encoded Privacy Enhanced Mail format identity provider certificate that IICS uses to validate signed SAML messages from the identity provider.

Identity Provider Properties (continued)

- **Use signing certificate for encryption** uses the public key in your signing certificate to encrypt logout requests
- **Encryption Certificate** specifies the Base64-encoded PEM format identity provider certificate
- **Name Identifier Format** specifies the format of the name identifier
- **Logout Service URL (SOAP Binding)** specifies the identity provider's SAML SOAP binding URL for the single logout service
- **Logout Page URL** specifies the landing page to which a user is redirected after the user logs out of IICS

12

© Informatica. Proprietary and Confidential.



The **Use signing certificate for encryption** option uses the public key in your signing certificate to encrypt logout requests sent to your identity provider when a user logs out from IICS.

The **Encryption Certificate** property specifies the Base 64-encoded Privacy Enhanced Mail format identity provider certificate that IICS uses to encrypt SAML messages sent to the identity provider. You can specify this property if you don't enable use of the signing certificate for encryption property.

The **Name Identifier Format** property specifies the format of the name identifier in the authentication request that the identity provider returns to IICS.

The **Logout Service URL SOAP Binding** property specifies the identity provider's SAML SOAP binding URL for the single logout service. IICS sends logout requests to this URL.

The **Logout Page URL** property specifies the landing page to which a user is redirected after the user logs out of IICS.

Service Provider Properties

- **Informatica Cloud Platform SSO** displays the single sign-on URL for your organization
- **Clock Skew** specifies maximum time between time stamps in SAML response from the identity provider and the IICS clock
- If **Name Identifier value represents user's email address** is selected, IICS uses the name identifier as the email address
- If **Sign authentication requests** is selected, IICS signs authentication requests
- If **Sign logout requests sent using SOAP binding** is selected, IICS signs logout requests
- If **Encrypt name identifier in logout requests** is selected, IICS encrypts name identifier

Service Provider Settings	
Informatica Cloud Platform SSO:	
Clock Skew (seconds):	<input type="text" value="180"/>
<input checked="" type="checkbox"/> Name identifier value represents user's email address	
<input checked="" type="checkbox"/> Sign authentication requests	
<input checked="" type="checkbox"/> Sign logout requests sent using SOAP binding	
<input type="checkbox"/> Encrypt name identifier in logout requests	

13

© Informatica. Proprietary and Confidential.



The **Informatica Cloud Platform SSO** setting displays the single sign-on URL for your organization.

The **Clock Skew** setting specifies the maximum permitted time between the time stamps in the SAML response from the identity provider and the IICS clock.

If the **Name Identifier value represents user's email address** option is selected, then IICS uses the name identifier as the email address.

If the **Sign authentication requests** option is selected, then IICS signs authentication requests to the identity provider.

If the **Sign logout requests sent using SOAP binding** option is selected, then IICS signs logout requests sent to the identity provider.

If the **Encrypt name identifier in logout requests** option is selected, then IICS encrypts the name identifier in logout requests.

SAML Attribute Mapping Properties

- User login attributes such as name, email address, and user role are included in the authentication response from the identity provider to IICS

Property	Description
Use friendly SAML attribute names	If selected, IICS uses the human-readable form of the SAML attribute name
First Name	SAML attribute used to pass the user's first name
Last Name	SAML attribute used to pass the user's last name
Job Title	SAML attribute used to pass the user's job title
Email Addresses	SAML attribute used to pass the user's email addresses
Emails Delimiter	Delimiter to separate the email addresses if multiple email addresses are passed
Phone Number	SAML attribute used to pass the user's phone number
Time Zone	SAML attribute used to pass the user's time zone
User Roles	SAML attribute used to pass the user's assigned user roles
Roles Delimiter	Delimiter to separate the roles if multiple roles are passed

14

© Informatica. Proprietary and Confidential.



SAML Role Mapping Properties

- For each IICS role enter the equivalent SAML role
- Default Role** specifies the default role to use if the SAML authentication response does not include the SAML user role attributes
- Default Group** specifies the default user group for single sign-on users

SAML Role Mapping	
Admin:	role1, role2
Application Integration Business Manager:	role3, role4
Application Integration Data Viewer:	role5, role6
Data Integration Data Previewer:	role7, role8
Data Integration Task Executor:	role9, role10
Deployer:	role11, role12
Designer:	role13, role14
Monitor:	role15, role16
Operator:	role17, role18
Service Consumer:	role19, role20
Default Role: [*]	Designer
Default Group:	No Group Defined

15

© Informatica. Proprietary and Confidential.



The image shows the different IICS roles such as Admin, Application Integration Business Manager, Application Integration Data Viewer, and so on. For each of these roles, you must enter the equivalent SAML roles. If you want to enter multiple SAML role names for a single IICS role, then you must use a comma to separate the roles.

The **Default Role** property specifies the default role to use if the SAML authentication response does not include the SAML user role attributes.

The **Default Group** property specifies the default user group for single sign-on users.

Downloading the Service Provider Metadata

- To complete the SAML single sign-on setup, identity provider requires the SAML service provider metadata and IICS URL
- Download the service provider metadata file from the SAML Setup page



16

© Informatica. Proprietary and Confidential.



The identity provider requires the SAML service provider metadata and IICS URL to complete the SAML single sign-on setup process. You can download the service provider metadata file from the **SAML Setup** page.

Module Summary

This module showed you how to:

- Discuss SAML single sign-on
- List the single sign-on requirements
- Discuss single sign-on restrictions
- Explain SAML single sign-on configuration for IICS

17

© Informatica. Proprietary and Confidential.





IICS: Cloud Data Integration Services

 Informatica[®]
University

Module 21

Discovery IQ



Module Objectives

After completing this module, you will be able to:

- Discuss Informatica Discovery IQ
- Explain the features of Discovery IQ

Discovery IQ Overview

- Allows you to manage, monitor, and trouble-shoot your integration processes running in IICS
- Provides a comprehensive view of your product usage and consumption
- Provides contextual recommendations and best practices
- Supports analytics and log analysis for the Data Integration service and the Application Integration service

4

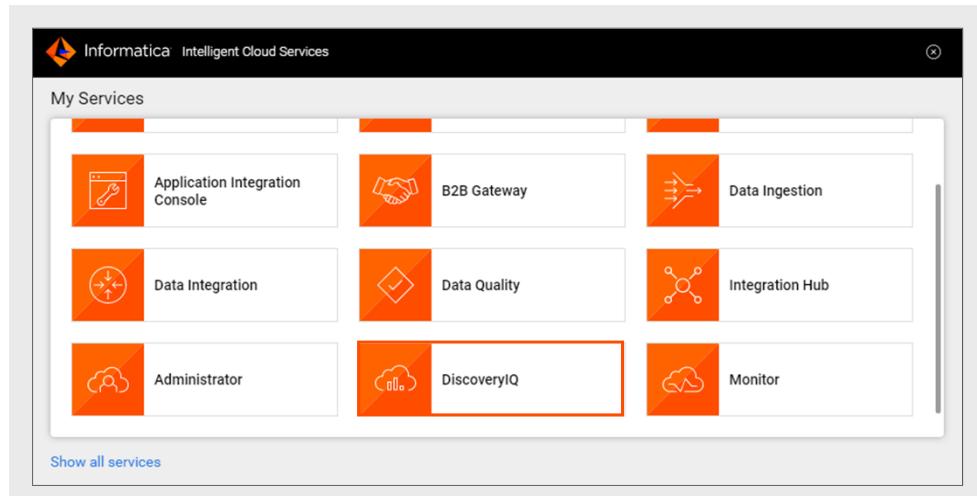
© Informatica. Proprietary and Confidential.



Informatica Discovery IQ is a cloud-based enterprise-grade solution that allows you to easily manage, monitor, and trouble-shoot your integration processes running in IICS. Discovery IQ provides a comprehensive view of your product usage and consumption. Its built-in intelligence provides contextual recommendations and best practices based on your product usage, product failure, and overall interaction history with Informatica.

Discovery IQ currently supports analytics and log analysis for the Cloud Data Integration service and the Cloud Application Integration service.

Accessing Discovery IQ



5

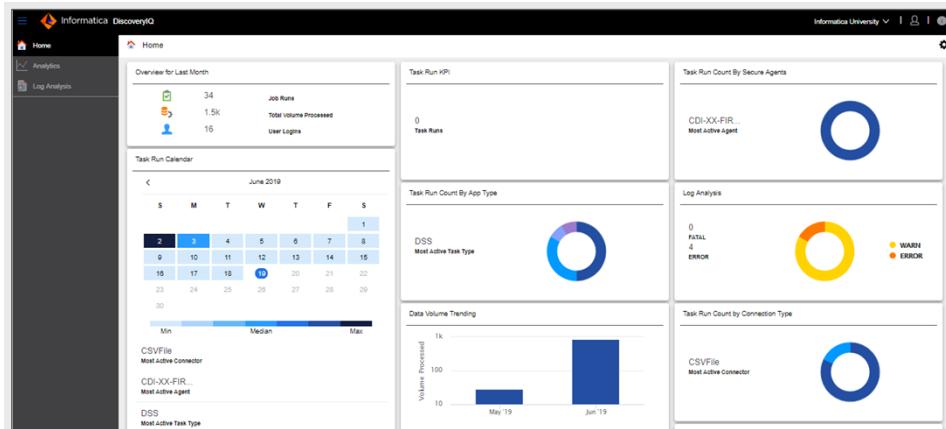
© Informatica. Proprietary and Confidential.



To access Discovery IQ, log in to your IICS account and then, from the My Services window, select Discovery IQ. Discovery IQ opens in a separate browser tab.

Discovery IQ Feature – Dashboard

- Provides a snapshot of metrics related to the logged in organization



6

The first screen or tab that you see after launching Discovery IQ is the home screen or the dashboard. The dashboard provides a snapshot of metrics related to the logged in organization.

The **Overview for Last Month** metric provides information about job runs, volumes processed, and user logins in IICS for the last one month.

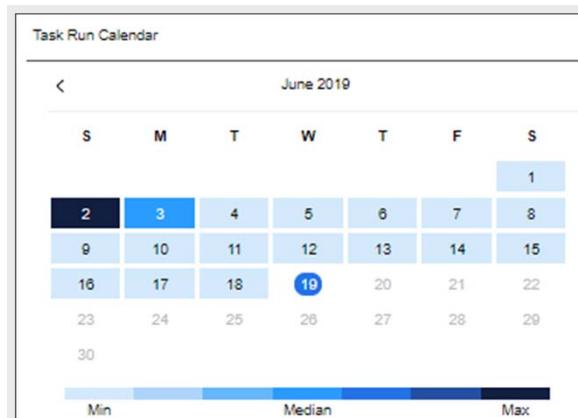
The **Task Run Calendar** metric shows the daily job run heat chart for the last 3 months along with most used connector, secure agent, and task type.

The dashboard also provides information about other metrics such as the Task Run KPI, Task Run Count By Secure Agents, Task Run Count By App Type, Log Analysis, and so on.

Relating Dashboard Information with Data Integration Service Example

Task

View the tasks run by task types on 3rd June



© Informatica. Proprietary and Confidential.



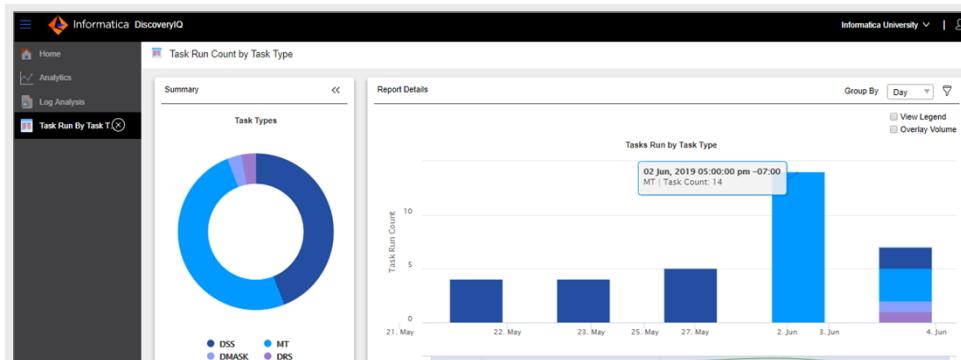
To better understand the information on the Discovery IQ dashboard, let's try to relate some information from this dashboard.

Assume that for a particular day of the month, you want to view the tasks run by task types. For example, you want to view the tasks run by task types on 3rd June. So on the dashboard, click 3rd June under the **Task Run Calendar** metric.

Relating Dashboard Information with Data Integration Service Example (continued)

Result

In the **Report Details** section, observe that between 2nd and 3rd June, the Mapping Task was run 14 times



© Informatica. Proprietary and Confidential.

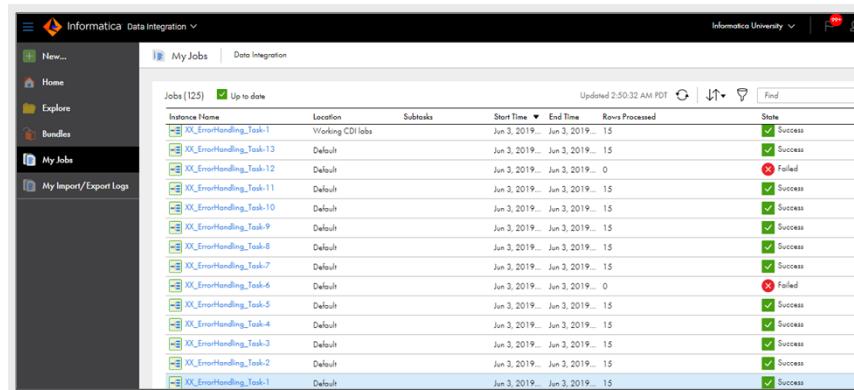
Informatica University

The analysis opens in a separate tab, named **Task Run Count by Task Type**. In the **Report Details** section, observe that between 2nd and 3rd June, the task M T, which represents a Mapping Task, was run 14 times.

Relating Dashboard Information with Data Integration Service Example (continued)

Verification

In the **My Jobs** page, verify that on 3rd June, the Mapping Task **XX_ErrorHandling** was run 14 times



The screenshot shows the Informatica Data Integration interface with the 'My Jobs' tab selected. A table lists 14 job instances, all of which are successful (green checkmarks) and completed on Jun 3, 2019, at 15:00. The tasks are named XX_ErrorHandling_Task-1 through XX_ErrorHandling_Task-14.

Instance Name	Location	Subtasks	Start Time	End Time	Rows Processed	State
XX_ErrorHandling_Task-1	Working CDI labs		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-13	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-12	Default		Jun 3, 2019...	Jun 3, 2019...	0	Failed
XX_ErrorHandling_Task-11	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-10	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-9	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-8	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-7	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-6	Default		Jun 3, 2019...	Jun 3, 2019...	0	Failed
XX_ErrorHandling_Task-5	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-4	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-3	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-2	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success
XX_ErrorHandling_Task-1	Default		Jun 3, 2019...	Jun 3, 2019...	15	Success

9

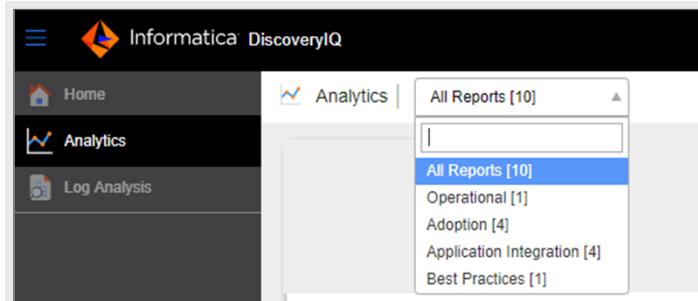
© Informatica. Proprietary and Confidential.



To check this task in IICS, navigate to the IICS Org and in the Data Integration service, click **My Jobs**. In the My Jobs page, you can verify that on 3rd June, the Mapping Task named **XX_ErrorHandling** was run 14 times.

Discovery IQ Feature – Analytics

- Provides Operational reports, Adoption reports, Application Integration reports, and also recommends Best Practices



10

© Informatica. Proprietary and Confidential.



The next tab in Discovery IQ is the Analytics tab.

The Analytics tab provides Operational reports, Adoption reports, Application Integration reports, and also recommends Best Practices.

Discovery IQ Feature – Analytics (continued)

- Operational reports like the 'Task Run KPI' provides a complete view of jobs run in the last one month

The screenshot shows the Informatica DiscoveryIQ web application. In the top navigation bar, the 'Analytics' tab is selected. Below the navigation, there's a search bar with 'Operational [1]' and a dropdown arrow. A callout box highlights the 'Task Run KPI' link. The main content area displays a summary of task runs: '1 Task Runs' and '1 WARNINGS'. A small circular icon with a yellow border and a dot inside is next to the 'WARNINGS' text.

11

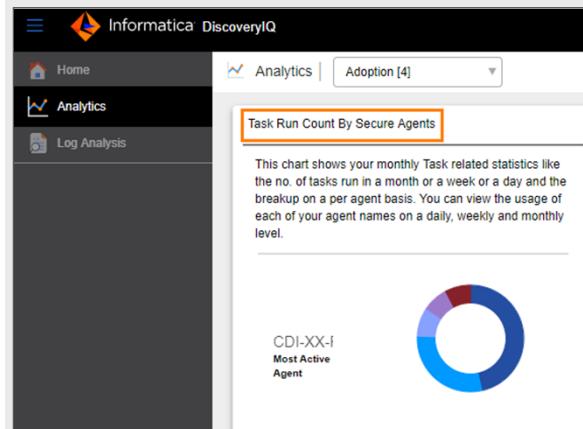
© Informatica. Proprietary and Confidential.



Operational reports like the **Task Run KPI** provides a complete view of jobs run in the last one month. You can drill down to a specific job to get detailed run-history of the job. The status and throughput of the job allows you to understand the performance of the job and isolate error trends.

Discovery IQ Feature – Analytics (continued)

- Adoption reports like the 'Task Run Count By Secure Agents' provides monthly task related statistics



12

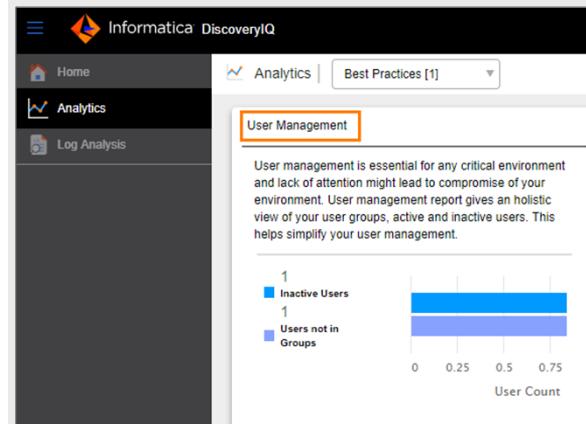
© Informatica. Proprietary and Confidential.



Adoption reports like the **Task Run Count By Secure Agents** provides monthly task related statistics such as the number of tasks run in a month or a week or a day, and the breakup on a per agent basis. You can view the usage of each of your agents on a daily, weekly, and monthly level.

Discovery IQ Feature – Analytics (continued)

- User Management Report suggests best practices for managing users and groups



13

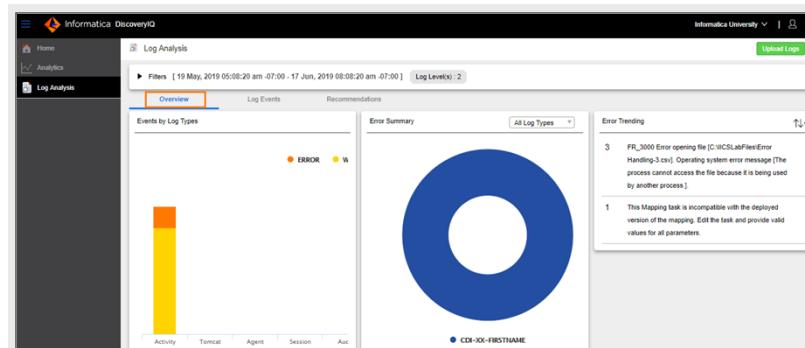
© Informatica. Proprietary and Confidential.



The User Management report provides information about the user groups, and the active and inactive users in your IICS Org. The User Management report helps you simplify the management of users groups and users in your Org.

Discovery IQ Feature – Log Analysis

- Provides log information for the jobs that are run in the IICS Org
- Overview sub-tab shows:
 - Log events distributed across log-types and secure agents
 - Top 10 frequently occurring errors



14

© Informatica. Proprietary and Confidential.



The Log Analysis tab provides the log information for the jobs that are run in the IICS Org. The Overview sub-tab shows log events distributed across log-types and secure agents. The Overview sub-tab also displays the top 10 frequently occurring errors.

Discovery IQ Feature – Log Analysis (continued)

- Log Events sub-tab provides tabular view of log events

The screenshot shows the Informatica DiscoveryIQ interface with the 'Log Analysis' sub-tab selected. The 'Log Events' sub-tab is highlighted with a blue border. The interface includes a search bar for 'Search Keyword' and a date range filter for 'Filters' (19 May, 2019 05:08:20 am -07:00 - 17 Jun, 2019 08:08:20 am -07:00). The main area displays a table titled 'Event Timeline' with columns for 'Time Stamp', 'Host', and 'Message'. The table contains four rows of log entries, each with a timestamp, host name (NICESS-LP051), and a detailed message about mapping tasks being incompatible with the deployed version.

Time Stamp	Host	Message
03 Jun, 2019 11:06:48 -07:00		This Mapping task is incompatible with the deployed version of the mapping. Edit the task and provide valid values for all ...
03 Jun, 2019 07:35:02 -07:00	NICESS-LP051	FR_3000 Error opening file [C:\IICSLabFiles\Error Handling-3.csv]. Operating system error message [The process cannot access the ...]
03 Jun, 2019 07:35:02 -07:00	NICESS-LP051	FR_3000 Error opening file [C:\IICSLabFiles\Error Handling-3.csv]. Operating system error message [The process cannot access the ...]
03 Jun, 2019 07:35:02 -07:00	NICESS-LP051	FR_3000 Error opening file [C:\IICSLabFiles\Error Handling-3.csv]. Operating system error message [The process cannot access the ...]

15

© Informatica. Proprietary and Confidential.



The Log Events sub-tab provides tabular view of log events with search and sort functionality.

Discovery IQ Feature – Log Analysis (continued)

- Recommendations sub-tab provides a summary of the error events
- It also provides recommendations for fixing the errors

The screenshot shows the Informatica DiscoveryIQ interface with the 'Log Analysis' tab selected. The 'Recommendations' sub-tab is highlighted with an orange border. On the left, there's a sidebar with 'Home', 'Analytics', and 'Log Analysis'. The main area has a 'Filters' section set to '19 May, 2019 05:08:20 am -07:00 - 17 Jun, 2019 08:08:20 am -07:00' and 'Log Level(s) : 2'. Below this is an 'Event Summarization' section with a message about a mapping task being incompatible with the deployed version. To the right is a 'Error Recommendations [10]' section containing two items: 'FAQ: In IICS, is there any RestAPI available to list all Assets (mappings/tasks) that uses certain connection/connection types?' and 'Converting Hierarchical Input To Flat Output Using Microsoft Azure Cosmos DB SQL API Connector'. At the bottom, there's a note about supported versions.

16

© Informatica. Proprietary and Confidential.



The Recommendations sub-tab provides a summary of the error events. It also provides recommendations for fixing the errors.

Module Summary

This module showed you how to:

- Discuss Informatica Discovery IQ
- Explain the features of Discovery IQ