



Informatica® Data Integration Hub
10.4.0

Operator Guide

© Copyright Informatica LLC 2013, 2022

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, the Informatica logo, Informatica Cloud, PowerCenter, PowerExchange, and Data Engineering Integration are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

See patents at <https://www.informatica.com/legal/patents.html>.

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

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2022-01-20

Table of Contents

Preface	12
Informatica Resources.	12
Informatica Network.	12
Informatica Knowledge Base.	12
Informatica Documentation.	12
Informatica Product Availability Matrices.	13
Informatica Velocity.	13
Informatica Marketplace.	13
Informatica Global Customer Support.	13
 Chapter 1: Introduction to Data Integration Hub.....	14
Data Integration Hub Overview.	14
Data Integration Hub Architecture.	17
Data Integration Hub Big Data.	18
Operation Console.	19
Changing the Operation Console Language.	19
Data Integration Hub Overview Diagram.	19
Data Integration Hub Topics.	20
Data Integration Hub Publications and Subscriptions.	21
Permissions and Privileges.	21
Operator User Role.	22
 Chapter 2: Catalog.....	23
Catalog Overview.	23
Topic Details.	23
Subscribing to a Topic in the Catalog.	25
 Chapter 3: Applications.....	26
Applications Overview.	26
Application Properties.	26
Managing Applications.	29
 Chapter 4: Topics.....	30
Topics Overview.	30
Publication Repository Types.	30
Topic Structure.	32
Topic Structure Updates.	32
Add Topic Tables.	34
Topic Table and Column Names.	36
Topic Table Relations.	37

Topic Data Retention.	37
Topic Permissions.	37
Managing Topics.	37

Chapter 5: Creating Topics..... 39

Creating Topics Overview.	39
Creating a Topic with a Relational Database Publication Repository.	39
Step 1. Access the New Relational Database Topic Wizard.	40
Step 2. Define Basic Topic Properties.	40
Step 3. Create Topic Data Structure.	40
Step 4. Define Topic Table Relations.	45
Step 5. Define Data Retention Period and Storage Location.	45
Step 6. Assign Topic Permissions.	46
Step 7. Review Topic Settings and Save the Topic.	46
Creating a Topic with a Hadoop Publication Repository.	46
Step 1. Access the New Big Data Topic Wizard.	46
Step 2. Define Basic Topic Properties.	47
Step 3. Create Topic Data Structure.	47
Step 4. Define Topic Table Relations.	52
Step 5. Define Data Retention Period.	52
Step 6. Assign Topic Permissions.	52
Step 7. Review Topic Settings and Save the Topic.	53
Creating a Topic with a File Store Publication Repository.	53
Step 1. Access the New File Store Topic Wizard.	53
Step 2. Define Basic Topic Properties.	53
Step 3. Create Topic Data Structure.	54
Step 4. Define Topic Table Relations.	54
Step 5. Define Data Retention Period and Storage Location.	55
Step 6. Assign Topic Permissions.	55
Step 7. Review Topic Settings and Save the Topic.	56
Creating a Topic with a Real-time Publication Repository.	56

Chapter 6: Topic Properties..... 57

Topic Properties Overview.	57
Topic General Properties.	58
Topic Structure Properties.	60
Add Tables from Database.	60
Add Tables from Enterprise Data Catalog.	61
Add a Table from Flat File.	61
Add Table from PowerCenter Workflow.	62
Add a Table from Metadata File Properties.	63
Create a Table.	63
Topic Table Relations Properties.	65

Topic Data Retention Properties for Relational Database and File Repositories.	65
Topic Data Retention Properties for a Hadoop Repository.	66
Topic Permissions Properties.	66
Topic Summary Properties.	66

Chapter 7: Publications..... 68

Publications Overview.	68
Publication Types.	68
Automatic Database Publication.	69
Automatic Flat File Publication.	70
Automatic Pass-through File Publication.	71
Automatic Data-driven Publication.	72
Custom Batch Publication.	72
Custom Multi-latency Publication.	73
Custom Cloud Publication.	73
Modular Cloud Publication.	74
Publication Process.	74
Publication Process with a Batch Workflow.	75
Publication Process with a Real-time Workflow.	75
Publication Process with a Data Integration Task.	76
Publication Process of a Data-driven Publication.	76
Publication Pre-process.	76
Publication Sources.	77
Relational Sources.	77
Flat File Sources.	77
Pass-through File Sources.	77
Publication Filters.	77
Publication Schedules.	78
Managing Publications.	79
Creating a Publication.	79
Searching for a Publication	79
Duplicating a Publication.	79
Editing a Publication.	80
Deleting a Publication.	80
Enabling and Disabling a Publication.	80
Running a Publication Manually.	80
Re-validating a Publication.	80

Chapter 8: Creating Publications..... 81

Creating Publications Overview.	81
Creating an Automatic Relational Database Publication.	81
Step 1. Access the Create Automatic Database Publication Wizard.	82
Step 2. Define Basic Publication Properties.	82

Step 3. Select Publication Source.	82
Step 4. View and Edit Publication Field Mapping.	83
Step 5. Define a Filter.	84
Step 6. Define Publication Schedule.	85
Step 7. Review Publication Settings and Save the Publication.	86
Creating an Automatic Flat File Publication.	86
Step 1. Access the Create Automatic File Publication Wizard.	87
Step 2. Define Basic Publication Properties.	87
Step 3. Select and Configure Publication Source.	87
Step 4. Define Joins.	89
Step 5. View and Edit Publication Field Mapping.	90
Step 6. Define a Filter.	91
Step 7. Define Publication Schedule.	91
Step 8. Review Publication Settings and Save the Publication.	92
Creating an Automatic Pass-through File Publication.	93
Step 1. Access the Create Automatic Pass-through File Publication Wizard.	93
Step 2. Define Basic Publication Properties.	93
Step 3. Configure Publication Source.	93
Step 4. Define Publication Schedule.	94
Step 5. Review Publication Settings and Save the Publication.	95
Creating an Automatic Data-driven Publication.	95
Step 1. Access the Create Automatic Data-driven Publication page.	95
Step 2. Define Publication Properties.	96
Step 3. Create a Request to Run the Publication.	96
Creating a Custom Batch Publication.	97
Step 1. Access the Create Custom Batch Publication Wizard.	97
Step 2. Define Basic Publication Properties.	97
Step 3. Select Publication Workflow.	98
Step 4. Define Publication Schedule.	98
Step 5. Review Publication Settings and Save the Publication.	99
Creating a Custom Multi-latency Publication.	99
Step 1. Access the Create Custom Multi-latency Publication Wizard.	99
Step 2. Define Basic Publication Properties.	100
Step 3. Select Publication Workflow.	100
Step 4. Define Publication Intervals.	100
Step 5. Review Publication Settings and Save the Publication.	100
Creating a Custom Cloud Publication.	100
Step 1. Access the Custom Cloud Publication Wizard.	101
Step 2. Define Basic Publication Properties.	101
Step 3. Select Publication Mapping.	101
Step 4. Define Publication Schedule.	101
Step 5. Review Publication Settings and Save the Publication.	102

Creating a Modular Cloud Publication.	102
1. Access the Create Modular Cloud Publication Wizard.	103
Step 2. Define Basic Publication Properties.	103
Step 3. Select Publication Mapping.	103
Step 4. Configure Publication Source.	103
Step 5. Configure Publication Target.	104
Step 6. Set Publication Parameters.	104
Step 7. Configure Publication Field Mapping.	104
Step 8. Define Publication Schedule.	105
Step 9. Review Publication Settings and Save the Publication.	105
Creating a Custom Pass-through Kafka Publication.	106
Step 1: Access the Create Custom Pass-Through Apache Kafka Publication Page.	106
Step 2. Define the Pass-through Kafka Publication Properties.	106
Chapter 9: Publication Properties.	107
Publication Properties Overview.	107
Publication General Properties.	107
Publication Processing Properties.	108
Publication Processing Properties for Data Integration Hub Workflows.	108
Publication Processing Properties for Informatica Intelligent Cloud Service Mappings.	109
Publication Source Properties.	109
Publication Source Properties for Relational Database.	109
Publication Source Properties for Flat File.	110
Publication Source Properties for Pass-through File.	113
Publication Target Properties.	114
Publication Join Properties.	114
Publication Field Mapping Properties.	114
Edit Table Mapping for Publications.	115
Edit Field Mapping for Publications.	116
Publication Filter Properties.	117
Basic Expression Properties.	117
Advanced Expression Properties.	118
Publication Schedule Properties.	118
Publication Summary Properties.	119
Properties of a Data-driven Publication.	120
Chapter 10: Subscriptions.	121
Subscriptions Overview.	121
Subscription Types.	122
Automatic Database Subscription	122
Automatic Flat File Subscription.	123
Automatic Pass-through File Subscription.	124
Automatic Data-driven Subscription.	125

Custom Batch Subscription	125
Custom Cloud Subscription	126
Modular Cloud Subscription.	127
Subscription Process.	127
Subscription Process with a Batch Workflow.	128
Subscription Process with a Data Integration Task.	128
Subscription Process of a Data-driven Subscription.	129
Post-Processing Mappings.	129
Unbound Subscription.	129
Subscription Targets	130
Relational Targets.	130
Flat File Targets.	130
Pass-through File Targets.	130
Subscription Filters.	131
Subscription Schedules.	131
Subscription Delivery.	132
Managing Subscriptions.	132
Creating a Subscription.	132
Searching for a Subscription	133
Duplicating a Subscription.	133
Editing a Subscription.	133
Deleting a Subscription.	133
Enabling and Disabling a Subscription.	134
Running a Subscription Manually.	134
Getting Previous Subscriptions.	134
Re-validating a Subscription.	134
Chapter 11: Creating Subscriptions.....	135
Creating Subscriptions Overview.	135
Creating an Automatic Relational Database Subscription.	135
Step 1. Access the Create Automatic Database Subscription Wizard.	136
Step 2. Define Basic Subscription Properties and Select a Topic.	136
Step 3. Select Post-process Workflow.	136
Step 4. Define Joins.	137
Step 5. Select Subscription Target.	137
Step 6. View and Edit Subscription Field Mapping.	137
Step 7. Define a Filter.	138
Step 8. Define Subscription Schedule.	139
Step 9. Define Delivery Options.	140
Step 10. Review Subscription Settings and Save the Subscription.	141
Creating an Automatic Flat File Subscription.	141
Step 1. Access the Create Automatic Flat File Subscription Wizard.	142
Step 2. Define Basic Subscription Properties and Select a Topic.	142

Step 3. Select Post-process Workflow.	142
Step 4. Define Joins.	142
Step 5. Select and Configure Subscription Target.	143
Step 6. View and Edit Subscription Field Mapping.	144
Step 7. Define a Filter.	145
Step 8. Define Subscription Schedule.	146
Step 9. Define Delivery Options.	147
Step 10. Review Subscription Settings and Save the Subscription.	147
Creating an Automatic Pass-through File Subscription.	147
Step 1. Access the Create Automatic Pass-through File Subscription Wizard.	148
Step 2. Define Basic Subscription Properties and Select a Topic.	148
Step 3. Select Post-process Workflow.	148
Step 4. Configure Subscription Target.	149
Step 5. Define a Filter.	149
Step 6. Define Subscription Schedule.	150
Step 7. Define Delivery Scope.	151
Step 8. Review Subscription Settings and Save the Subscription.	151
Creating an Automatic Data-driven Subscription.	151
Step 1. Access the Create Automatic Data-driven Subscription Page.	152
Step 2. Define Subscription Properties.	152
Step 3. Create a Request to Run the Subscription.	152
Creating a Custom Batch Subscription.	153
Step 1. Access the Create Custom Batch Subscription Wizard.	153
Step 2. Define Basic Subscription Properties and Select Topics.	153
Step 3. Select Subscription Workflows.	154
Step 4. Define Subscription Schedule.	154
Step 5. Review Subscription Settings and Save the Subscription.	155
Creating a Custom Cloud Subscription.	155
Step 1. Access the Custom Cloud Subscription Wizard.	155
Step 2. Define Basic Subscription Properties and Select Topics.	156
Step 3. Select Subscription Mappings.	156
Step 4. Define Subscription Schedule.	156
Step 5. Review Subscription Settings and Save the Subscription.	157
Creating a Modular Cloud Subscription.	157
Step 1. Access the Modular Cloud Subscription Wizard.	158
Step 2. Define Basic Subscription Properties and Select a Topic.	158
Step 3. Select Subscription Mapping.	158
Step 4. Configure Subscription Source.	158
Step 5. Configure Subscription Target.	158
Step 6. Set Subscription Parameters.	159
Step 7. Configure Subscription Field Mapping.	159
Step 6. Define Subscription Schedule.	159

Step 10. Review Subscription Settings and Save the Subscription.	160
Creating a Custom Pass-through Kafka Subscription.	160
Step1. Access the Create Custom Pass-through Kafka Subscription Page.	161
Step2. Define Basic Subscription Properties.	161

Chapter 12: Subscription Properties..... 162

Subscription Properties Overview.	162
Subscription General Properties.	162
Subscription Processing Properties.	164
Subscription Processing Properties for Automatic Mappings with Data Integration Hub Workflows.	164
Subscription Processing Properties for Automatic Mappings with Informatica Intelligent Cloud Service Mappings.	164
Subscription Processing Properties for Custom Mappings with Data Integration Hub Workflows.	165
Subscription Processing Properties for Custom Mappings with Informatica Intelligent Cloud Service Tasks.	165
Subscription Join Properties.	166
Subscription Source Properties.	166
Subscription Target Properties.	166
Subscription Target Properties for Relational Database.	166
Subscription Target Properties for Flat File.	167
Subscription Target Properties for Pass-through File.	168
Subscription Field Mapping Properties.	169
Edit Table Mapping for Subscriptions.	170
Edit Field Mapping for Subscriptions.	170
Subscription Filter Properties.	172
Basic Expression Properties.	172
Advanced Expression Properties.	172
Subscription Schedule Properties.	173
Subscription Delivery Properties.	173
Subscription Wizard Summary Page.	175
Properties of a Data-driven Subscription.	175

Chapter 13: Events and Event Monitoring..... 177

Events and Event Monitoring Overview.	177
Event Actions.	178
Managing Events on the Event List Page.	178
Basic Event Search Properties.	179
Performing a Basic Search for Events.	179
Advanced Event Search Properties.	180
Performing an Advanced Search for Events.	180
Publication and Subscription Event Types and Statuses.	181
Event Monitors.	182

Monitoring Rules.	182
Managing Monitoring Rules.	183
Creating a Monitoring Rule.	183
Monitoring Rule Properties.	185
Chapter 14: Dashboard and Reports.....	190
Dashboard and Reports Overview.	190
Dashboard Reports.	191
Errors by Application Report.	191
Events by Current Status Report.	193
Event Distribution Report.	194
Error Event Distribution Report.	195
Event Average Processing Time Report.	196
Dashboard Report CSV File Structure.	197
Managing the Dashboard.	199
Chapter 15: Glossary.....	200
Index.	203

Preface

See the *Data Integration Hub Operator Guide* to learn about the operational tasks that you can perform in the Data Integration Hub Operation Console. This guide includes information about creating and managing topics, publications, and subscriptions.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit <https://network.informatica.com>.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit <https://search.informatica.com>. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit <https://docs.informatica.com>.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at <http://velocity.informatica.com>. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:

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

To find online support resources on the Informatica Network, visit <https://network.informatica.com> and select the eSupport option.

CHAPTER 1

Introduction to Data Integration Hub

This chapter includes the following topics:

- [Data Integration Hub Overview, 14](#)
- [Data Integration Hub Architecture, 17](#)
- [Data Integration Hub Big Data, 18](#)
- [Operation Console, 19](#)
- [Data Integration Hub Overview Diagram, 19](#)
- [Data Integration Hub Topics, 20](#)
- [Data Integration Hub Publications and Subscriptions, 21](#)
- [Permissions and Privileges, 21](#)
- [Operator User Role, 22](#)

Data Integration Hub Overview

Data Integration Hub is an application integration solution that your organization can use to share and synchronize data between different applications in the organization.

To publish data to Data Integration Hub, first define the data set that you want to manage, for example, sales, customers, or orders. You define a data set by defining a topic. A topic defines the structure of the data that Data Integration Hub stores in the publication repository and the type of publication repository where data is stored. You can manage multiple topics that represent different data sets in Data Integration Hub. Applications publish data to topics and subscribe to data sets that are represented by topics.

Multiple applications can publish to the same topic, for example, different ordering applications can publish their orders to the same Orders topic. Multiple subscribers can consume the data from a topic. Different subscribing applications can consume the data in different formats and in different latencies based on a defined schedule.

Data Integration Hub stores the data that applications publish to topics in the Data Integration Hub publication repository. Data Integration Hub keeps the data in the publication repository until all subscribers consume the data and the retention period expires, and then deletes the data from the publication repository.

Applications can use PowerExchange® adapters and Informatica Intelligent Cloud Services® connectors to share data from different sources, such as database tables, files, or any sources that Informatica supports. Each application can be a publisher and a subscriber to different topics.

Publications publish to a specific topic. A publication defines the data source type and the location from where Data Integration Hub retrieves the data that the application publishes. Subscriptions subscribe to one or more topics. A subscription defines the data target type and the location in the subscribing application to where Data Integration Hub sends the published data.

When you create a publication or a subscription, you can choose to use either an automatic Data Integration Hub mapping or a custom Data Integration Hub mapping. Data Integration Hub creates automatic mappings based on the data structure that you define in the topic. Custom Data Integration Hub mappings are based on PowerCenter® workflows, Data Engineering Integration mappings, or Data Integration tasks that the developer creates and maintains for the publication or the subscription.

Data Integration Hub operator uses Enterprise Data Catalog to discover and leverage existing Data Integration Hub objects, and understand their lineage and impact on other entities in the enterprise.

Examples

You run a data center for a major retail chain. The main office has multiple applications. Some of the applications are located on-premises and some are located on the cloud. Each retail branch has a point-of-sale (POS) application and an inventory application. Your applications and branches require the following data:

Customer service applications

Require up-to-date customer order data.

Sales applications

Require up-to-date product sales data.

Marketing application

Requires a weekly deals report.

Accounting application

Requires a monthly deals report.

Branch applications

Require up-to-date inventory and pricing data.

Business Intelligence (BI) application

Requires a weekly report of sales and marketing data and of user interaction data from the corporate website, for the preceding 12 months.

With Data Integration Hub, you can address the following use-cases:

Share product catalog and prices.

You can share product price updates from the sales department with each branch, as follows:

1. Create a Products topic.
2. For the Product Information Management (PIM) application, define a publication that publishes product details and prices to the Products topic and set the schedule to publish the data daily.
3. For each branch application, define a subscription to the Products topic and set the subscription to consume the published data when it is available in Data Integration Hub.

Share daily sales details.

You can share the daily sales details that you receive from the stores with your central sales application and your customer service applications, as follows:

1. Create a Sales topic.

2. For each branch application, define a publication to the Sales topic, and set the schedule to publish daily.
3. For the sales application, define a subscription to the Sales topic, and set the schedule to consume the data when it is published.
4. For the customer service application, define a subscription to the Sales topic, and set the schedule to consume the data once a week.

Share deal details from Salesforce.

You can share deal details from a Salesforce cloud application with the marketing and accounting applications, as follows:

1. Create a Deals topic.
2. For the Salesforce application, define a cloud publication to the Deals topic, and set the schedule to publish weekly.
3. For the marketing application, define a subscription to the Deals topic, and set the schedule to consume the data once a week.
4. For the accounting application, define a subscription to the Deals topic, and set the schedule to consume the data once a month.

Share business intelligence data.

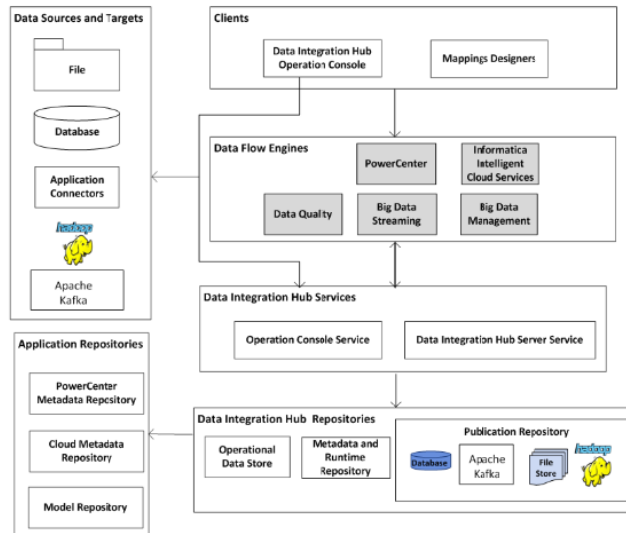
You can share sales data from Salesforce, marketing data from the marketing application, and user interaction data from the corporate website with the BI application as follows:

1. Create a Business Intelligence topic with a Hadoop publication repository and set the publication data retention period to 365 days.
2. For the Salesforce application, define a cloud publication to the Business Intelligence topic, and set the schedule to publish once a week.
3. For the marketing application, define a publication to the Business Intelligence topic, and set the schedule to publish once a week.
4. For the corporate website application, define a cloud publication to the Business Intelligence topic, and set the schedule to publish once a week.
5. For the BI application, define an aggregated subscription to the Business Intelligence topic, and set the schedule to consume the data once a week.

Data Integration Hub Architecture

The Data Integration Hub environment consists of user interface clients, data flow engines, Data Integration Hub services and repositories, and external metadata repositories.

The following image shows the Data Integration Hub components:



Data Integration Hub contains the following components:

Data Integration Hub Operation Console Web client

User interface to manage applications, topics, publications, and subscriptions, and to monitor publications, subscriptions, and events. Administrators also use the Operation Console to manage users and system settings. Developers use the Operation Console to manage Data Integration Hub workflows and connections.

Mappings designer clients

User interfaces to define sources and targets, build custom mappings, and create workflows and tasks. Use the mappings designers if you use custom mappings.

Data flow engines

Engines that retrieve data from publication sources and send the data to subscription targets. You can use different flow engines for different use cases. For example, use PowerCenter to publish and subscribe to on-premises applications, and use Informatica Intelligent Cloud Services to publish and subscribe to cloud applications.

Data Integration Hub Operation Console service

Service that processes actions that users perform on the Operation Console and creates the structure for published data sets in the publication repository.

Data Integration Hub Server service

Service that starts and monitors Data Integration Hub workflows for publications and subscriptions.

Data Integration Hub publication repository

Database that stores published data until the subscribers consume the data. After the data retention period ends, Data Integration Hub deletes the data from the publication repository.

Data Integration Hub metadata repository

Database that stores metadata for Data Integration Hub applications, topics, publications, subscriptions, and events.

Operational data store

A repository that contains aggregated information for reporting purposes. When you install the Data Integration Hub Dashboard and Reports component of Data Integration Hub, Data Integration Hub creates the operational data store repository based on the database connection details that you supply.

PowerCenter metadata repository

Database that stores metadata for PowerCenter mappings, workflows, and transformations.

Cloud metadata repository

Database that stores metadata for cloud mappings and tasks.

Model Repository Service

Database that stores metadata for Data Engineering Integration and Data Quality mappings and transformations.

Data sources and targets

Sources and targets that you use to publish and consume data. You can use the following types of sources and targets:

- Database. Tables and columns.
- File. Binary, text, or unstructured files.
- Application connectors. Connection objects for applications. Available when you use a custom mapping.
- Hadoop. Hadoop Distributed File System (HDFS) and Hive data warehouses.

Data Integration Hub Big Data

Publish and subscribe to high volumes of data, data streams, and data that you want to store for a long period of time with Data Integration Hub. For example, store business intelligence data that you need to review over time on the Data Integration Hub Hadoop publication repository, or publish from and subscribe to Hadoop Distributed File System (HDFS) and Hive data warehouses.

If you want to keep the published data in the Hadoop publication repository after the data is consumed by all subscribers, you can configure Data Integration Hub not to delete published data from the repository.

You can use both automatic mappings and custom mappings to publish and consume big data with Data Integration Hub. For custom mapping publications you can use Informatica Data Engineering Integration mappings and workflows and Informatica Data Engineering Streaming mappings. For custom mapping subscriptions you use Informatica Data Engineering Integration mappings and workflows.

Operation Console

Use the Operation Console user interface to manage applications, topics, publications, and subscriptions, and to monitor publications, subscriptions, and events. Administrators also use the Operation Console to manage users and system settings. Developers use the Operation Console to manage workflows and connections.

You can view the Operation Console in English or in Japanese. You can switch between the display languages.

The Operation Console contains two areas:

Navigator

Use the navigator to navigate between tasks that you can perform in the Operation Console. The navigator shows in the left pane of the Operation Console.

Current page

Main work area in which you perform the tasks that you select in the Navigator. The current page shows in the right pane of the Operation Console.

Changing the Operation Console Language

You can view the Operation Console in English or in Japanese. You can switch between the display languages.

1. In the browser from where you access Data Integration Hub, set the language to the required language.
2. The **Help** link opens the online help in English. To view the Japanese online help access the following URL:

```
http(s)://<host>:<port>/dih-help-ja
```

Where:

- <host> is the host name or the IP address of the Data Integration Hub server.
- <port> is the port number of the Data Integration Hub server.

For example:

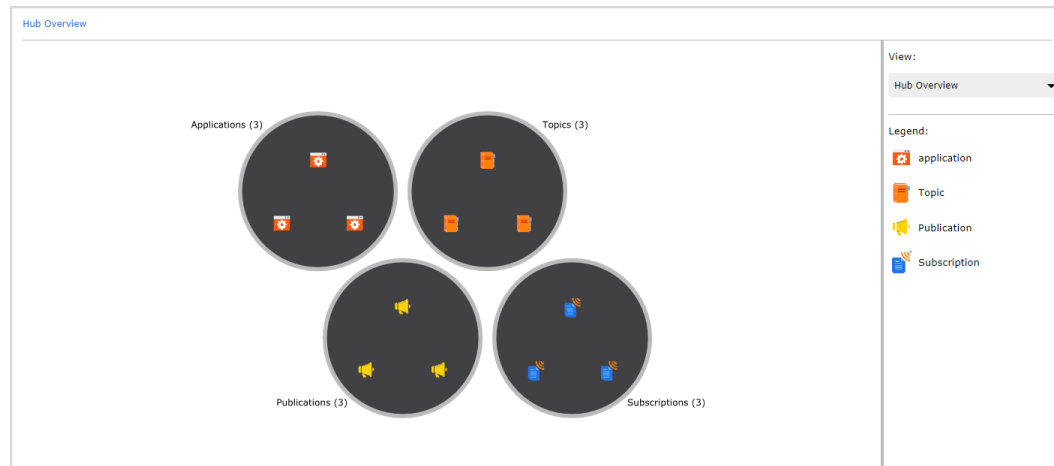
```
https://dih-releases:19443/dih-help-ja/
```

Data Integration Hub Overview Diagram

The Data Integration Hub Operation Console Home page displays the Hub Overview diagram when Data Integration Hub contains entities such as applications, topics, publications, or subscriptions.

The Hub Overview diagram provides a visual overview of the existing entities, grouped into categories.

The following image shows a sample Hub Overview diagram:



When you hover over an entity in the diagram, all related entities are highlighted. For example, when you hover over a topic, the applications and the publications that publish to the topic and the subscriptions that subscribe to the topic are highlighted. When you click an entity, a drill down view of the entity and its relations to other entities appears. For example, when you click a publication, the drill down view shows the publishing application, the topic to which the publication publishes data, and the subscriptions that subscribe to the topic. Click any entity in the drill down view. The edit window of the selected entity displays.

You can filter the Hub Overview diagram to the following views:

- Hub Overview. All entities including topics, related applications, publications, and subscriptions that are configured in Data Integration Hub.
- Process errors. Entities with current error events.
- Non-valid entities. Entities that are not valid.
- Topics with no publications. Topics with no associated publications.
- Topics with no subscriptions. Topics with no associated subscriptions.
- Most used topics. Three most used topics, based on the number of publications and subscriptions that use the topic.

When you filter the diagram, entities that are not relevant to the selected filter appear in view only mode.

To search topics in the Hub Overview diagram, click the **Find Topic** button and enter the name of the topic. Topics that match the search criteria and its related entities are displayed in Hub Overview diagram.

Data Integration Hub Topics

A Data Integration Hub topic is an entity that represents a data domain that is published and consumed in Data Integration Hub. A topic defines the canonical data structure and additional data definitions such as the data retention period.

For example, a Sales topic that represents sales data. Applications from all the stores in the organization publish sales data to the Sales topic. The accounting application subscribes to the Sales topic and consumes published sales data from all stores, or, if a filter is applied, from specific stores.

Before you define publications and subscriptions for the data that is published and consumed in Data Integration Hub, you need to define the canonical structure that will hold the data that is published to Data

Integration Hub in the Data Integration Hub publication repository. You define the canonical structure when you define the topic. You can define multiple topics that represent different source data sets.

Data Integration Hub Publications and Subscriptions

Publications and subscriptions are entities that define how applications publish data to Data Integration Hub and how applications consume data from Data Integration Hub. Publications publish data to a defined topic and subscriptions subscribe to topics.

Publications and subscriptions control the data flow and the schedule of data publication or data consumption. An application can be a publisher and a subscriber. Multiple applications can publish to the same topic. Multiple applications can consume data from the same topic.

You can use automatic, custom, and modular publications and subscriptions to publish data and to consume data. You can publish from and subscribe to different sources of data. Because the publishing process and the consuming process are completely decoupled, the publishing source and the consuming target do not have to be of the same data type. For example, you can publish data from a file and consume it into a database.

Automatic publications and subscriptions can publish from and subscribe to a relational database, a file, or a cloud application, or over a REST API.

Custom publications and subscriptions can publish from and subscribe to on-premises applications.

Modular publications and subscriptions can publish from and subscribe to cloud applications.

Permissions and Privileges

Permissions and privileges control access to Data Integration Hub objects and the actions that you can perform on the objects. The Data Integration Hub administrator assigns permissions and privileges to user groups to determine the objects that the users in the user groups can view or change.

For example, the administrator can assign the Sales category to all sales applications. The administrator can then assign permissions for the Sales category to specific user groups.

Permissions

Permissions control the objects and the data that you can access. For example, permissions can determine that you can access a particular topic, but cannot access another topic. Entities with no assigned permissions are accessible by all Data Integration Hub users.

The Data Integration Hub administrator creates categories and assigns categories to user groups to determine the users that can view or change the objects. You assign categories to topics and to applications to permit users to view or change the applications and the topics. Because publications and subscriptions are associated with applications and with topics, they inherit the permissions from both the parent application and from the associated topic. When you configure permissions for an application, only user groups with permissions to the application can access the associated subscriptions and publications. When you configure permissions for a topic, only user groups with permissions to the topic can access the associated subscriptions and publications. User permissions to publications and to subscriptions are determined either by the associated topic or by the associated application, whichever is the most restrictive.

Permissions also restrict access to monitors and to events that are related to the publications and the subscriptions.

Privileges

Privileges control the actions that you can perform on objects. For example, privileges can determine that you can view and edit publications, but cannot create or delete publications.

Operator User Role

The Data Integration Hub operator is responsible for managing operational entities in the Operation Console. The entities include applications, topics, publications, and subscriptions. The operator also monitors publication and subscription processing status with events and event monitoring rules.

The operator creates and modifies applications that need to publish or consume data. For each application, the operator creates and defines publications or subscriptions. The operator defined the data domains that Data Integration Hub manages by creating and modifying topics.

The operator configures and maintains source connections for publications and target connections for subscriptions. The operator can define connections when the operator creates a publication or a subscription, or manage connections directly in the Operation Console.

While publications and subscriptions are running, the operator monitors the status of the publication and subscription process with events and monitors, and in the Data Integration Hub Dashboard. If errors occur during the publication or the subscription process, the operator can perform actions on the events to troubleshoot issues.

The operator can create rules that monitor publication and subscription events, and perform actions on events that are in a defined status. For example, you can create a rule that disables publications whose events are in an Error status, a rule that invokes a PowerCenter workflow when a publication event is in an Error status, or a rule that sends an email to the Data Integration Hub administrator when a subscription event is in a Critical status.

CHAPTER 2

Catalog

This chapter includes the following topics:

- [Catalog Overview, 23](#)
- [Topic Details, 23](#)
- [Subscribing to a Topic in the Catalog, 25](#)

Catalog Overview

The catalog contains a list of all the topics that are defined in Data Integration Hub. In the catalog, you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

The catalog lists the topics in a grid format. You can drill down on each topic to view general details about the topic, the topic structure, a list of publishers to the topic, and a list of subscribers to the topic. You can perform a free text search to find topics in the catalog or search for topics based on topic tags.

You can subscribe to a specific topic from the catalog. When you subscribe to a topic from the catalog, the subscription wizard appears. You then create a subscription in the same way that you create a subscription directly from the application. After you subscribe to a topic from the catalog, the subscription appears in the subscribers list for the topic and in the subscription list for the application.

You can open topics from the catalog for editing if the associated publications and subscriptions are not running.

Topic Details

When you view a topic in the catalog, you can click the show details icon next to a topic name to view basic details, such as the name and type of the topic. You can also view the list of publishers and subscribers to the topic.

General Tab

The **General** tab displays basic information about the topic. The details appear in read-only mode and reflect settings that you define in the Topic wizard.

The following table describes the topic properties that appear on the **General** tab:

Property	Description
Topic Name	Name of the topic.
Type	Topic type, Delta or Full.
Description	Textual description of the topic.

Content Tab

The **Content** tab displays the tables and columns in the topic structure. You can click the columns to view details about the columns for each table, such as the name and data type.

The following table describes the topic properties that appear on the **Content** tab:

Property	Description
Table Name	Name of the table.
Columns	Link to the list of columns.

Publishers Tab

The **Publishers** tab displays the list of publishers to the topic. You can click the publisher names to view or to edit properties.

The following table describes the topic properties that appear on the **Publishers** tab:

Property	Description
Publication Name	Name of the publisher. You can click the name to open the Publication wizard and to view or change the publication properties. If the publication instance is running or if there is unconsumed data, the Publication wizard opens in read-only mode.
Publishing Application	Name of the application that publishes the data. You can click the name to open the application and to view or change the application properties.

Subscribers Tab

The **Subscribers** tab displays the list of subscribers to the topic. You can click the subscription names to view or to edit properties.

The following table describes the topic properties that appear on the **Subscribers** tab:

Property	Description
Subscription Name	Name of the subscription. You can click the name to open the Subscription wizard and to view or change the subscription properties. If the subscription instance is running or if there is unconsumed data, the Subscription wizard opens in read-only mode.
Subscribing Application	Name of the application that subscribes to the data. You can click the name to open the application and to view or change the application properties.

Subscribing to a Topic in the Catalog

Use the Navigator to view the topic catalog and to subscribe to one or more topics.

1. In the Navigator, click **Hub Management > Catalog**.

The **Catalog** page appears. You can perform a free text search to find topics in the catalog or search for topics based on topic tags.

2. To view topic details, click the **Details** icon next to the topic that you want to view.

The **Topic Details** pane displays basic information about the topic and a list of subscribers.

3. To subscribe to a topic, perform the following steps to create a subscription:

- a. Click the **Subscribe** icon next to the topic to which you want to subscribe.

The **Create Subscription** dialog box appears.

- b. Select the following entities and then click **OK**:

- Application Name. Subscribing application.
- Subscription Type. Type of subscription to use when subscribing to the topic.

The **Create Subscription** dialog box closes. The subscription wizard for the type of subscription that you selected appears.

- c. Follow the steps in the subscription wizard.

The subscription is added to the subscribers list of the topic.

CHAPTER 3

Applications

This chapter includes the following topics:

- [Applications Overview, 26](#)
- [Application Properties, 26](#)
- [Managing Applications, 29](#)

Applications Overview

An application represents an entity in your organization that needs to share data with other applications in your organization, such as sales applications or customer service applications. In Data Integration Hub, an application is a container for publications and subscriptions.

Applications can reside either inside or outside the Data Integration Hub network, and can be located on-premises or on the cloud. To share file-based data between applications that reside outside the Data Integration Hub network, you can use the SSH File Transfer Protocol (SFTP).

An application can publish data to a defined topic and can subscribe to data from a topic. For example, a sales application can publish sales reports and subscribe to inventory updates from an operations application. When you add a publication to an application, you define the schedule according to which topic data will be published from the application. You also define the schedule according to which topic data will be retrieved from the application and published to the Data Integration Hub publication repository. When you add a subscription to an application, you define the topic to which the application subscribes and the schedule and scope of data that the application consumes from the topic. The topic defines the structure of the data that the associated publications and subscriptions publish and consume.

You can use permissions to restrict access to applications to specific users. Only users with permissions to the application can define publications and subscriptions for this application.

Application Properties

Applications properties include general information, contacts, and permissions. The properties appear on tabs for the application.

General Tab

Use the **General** tab to enter basic information about the application.

The **General** tab includes the following properties:

Property	Description
Application Name	Name of the application. The name can contain up to 64 characters and can contain special characters.
Description	Description of the application. The description can contain up to 255 characters.

Publications Tab

Use the **Publications** tab to create, run, edit, and delete publications. A publication defines the source from where data is published to a defined topic, and the publication frequency. The publication list displays the list of all defined publications for the specific application. You can use the run button to run a publication manually. Data Integration Hub then triggers the publication process.

The **Publications** tab includes the following properties:

Property	Description
Publication Name	Name of the publication.
Mode	Mode of the publication: enabled or disabled. You can disable and enable the publication by selecting the required mode. For more information, see “Managing Publications” on page 79 .
Description	Description of the publication.
Topics	Names of the topics to which the publication publishes data. If the structure of the topic was changed since the publication was last run, and the topic structure does not invalidate the publication, an Updated indication appears next to the topic name. For example, when a table was added to the topic structure. If an Updated indication appears next to the topic name, run the publication wizard to see the change.
Status	Publication status icon. If the publication is not valid, a warning icon appears.

Subscriptions Tab

Use the **Subscriptions** tab to create, run, edit, and delete subscriptions. A subscription defines the set of data that the application consumes from a defined topic, and the frequency at which the application consumes the data. The subscription list displays the list of all defined subscriptions for the specific application.

You can use the following options to run subscriptions manually and to generate subscription events for undelivered data:

- **Run Subscription.** Run the subscription. Enabled for subscriptions that have pending published data. Not applicable for subscriptions with the **When published data is ready** scheduling option.
- **Get Previous Publications.** Get data that was published before the subscriber subscribed to the topic and therefore was not sent to the subscriber. Data Integration Hub will generate subscription events for undelivered data that was published during the date range that you define and which is still available in the publication repository. The subscriptions will run according to the subscription schedule and the delivery preferences that are defined for each subscription.

The **Subscription** tab includes the following properties:

Property	Description
Subscription Name	Name of the subscription.
Mode	Mode of the subscription: enabled or disabled. You can disable and enable the subscription by selecting the required mode. For more information, see “Managing Subscriptions” on page 132
Description	Description of the subscription.
Topics	Name of the topic from which the subscription consumes data. If the structure of the topic was changed since the subscription was last run, and the topic structure does not invalidate the subscription, an Updated indication appears next to the topic name. For example, when a table was added to the topic structure. If an Updated indication appears next to the topic name, run the subscription wizard to see the change.
Status	Subscription status icon. If the subscription is not valid, a warning icon appears.

Contacts Tab

Use the **Contacts** tab to enter information about contact people for the application.

The **Contacts** tab includes the following properties:

Property	Description
Name	Name of the contact person. The name can contain up to 60 characters.
Title	Title of the contact person, such as manager or coordinator. The title can contain up to 60 characters.
Business Phone	Telephone number for the contact person. The number can contain up to 20 characters.

Permissions Tab

Use the **Permissions** tab to assign and remove categories that restrict access to the application. If you do not assign categories to an application, all Data Integration Hub users can access the application.

The **Permissions** tab includes the following application properties:

Property	Description
Available Categories	List of categories that you can assign to the application.
Selected Categories	List of assigned categories for the application.

Managing Applications

Use the Navigator to create, edit, and delete applications. You also manage publications and subscriptions for the application. You can only manage applications for which you have editing privileges.

1. In the Navigator, click **Hub Management > Applications**.

The **Applications** page appears.

2. Choose to create, edit, or delete an application.

- To create an application, click **New Application** and enter the application properties. You can add publications and subscriptions after you save the application.

Note: You must save the application before you add a publication or a subscription.

- To edit an application, click the **Edit** icon next to the application that you want to edit and change the application properties.
- To delete an application, click the **Delete** icon next to the application that you want to delete and confirm the deletion. You cannot delete an application with publications or subscriptions. To delete an application with publications or subscriptions, first delete the publications or subscriptions.

CHAPTER 4

Topics

This chapter includes the following topics:

- [Topics Overview, 30](#)
- [Publication Repository Types, 30](#)
- [Topic Structure, 32](#)
- [Topic Table Relations, 37](#)
- [Topic Data Retention, 37](#)
- [Topic Permissions, 37](#)
- [Managing Topics, 37](#)

Topics Overview

A topic is an entity that represents a data domain that is published and consumed in Data Integration Hub. A topic defines the data structure and additional data definitions, such as the data retention period. Multiple applications can publish to the same topic. An application can subscribe to multiple topics.

For example, you can create a Sales topic that represents sales data. Applications from all the stores in the organization publish sales data to the Sales topic. The accounting application subscribes to the Sales topic and consumes published sales data from all stores, or, if a filter is applied, from specific stores.

Publication Repository Types

When you create a topic, you choose the type of publication repository in which Data Integration Hub manages and stores published data for the topic.

Data Integration Hub can store topic data on the following types of publication repository:

- **Relational database.** Choose this type of repository to store published data in a relational database structure that represents the structure in which you want to keep the data. For example, data that is published from a relational database or from files. A relational database publication repository usually stores the data for a short intermediate period after the data is consumed by all subscribers. Data Integration Hub supports the following databases on which to store relational database topic data: Oracle, Microsoft SQL Server.

- **Big Data.** Choose this type of repository if you publish high volumes of data that you want to store for a long period of time or if you do not want Data Integration Hub to delete published data after the data is consumed. The availability of the Hadoop repository depends on whether or not the Hadoop component is installed on your system.

To publish and subscribe to a Hadoop-based repository with custom publications and subscriptions, you must use workflows that are based on a Data Engineering Integration mapping and workflow. When you create a custom publication, if one of the topics that you select for the publication is a Hadoop-based topic, only workflows that are based on a Data Engineering Integration mapping or workflow are listed for selection as the publication mapping.

When you create a compound subscription, that subscribes to multiple topics, all topics that you select must be Hadoop-based, and only workflows that are based on a Data Engineering Integration mapping or workflow are listed for selection as the subscription mapping. You also enable the mandatory option for topics in compound subscription to prioritize a few topics in the compound subscription over other topics.

Data Integration Hub triggers a processing of subscription after the publication event for all topics are completed. If the wait time of the publication event is complete and Data Integration Hub has not published all mandatory topics, an error event is generated during run-time.

Before you use a Hadoop-based publication repository for publications and subscriptions, consider the following restrictions:

- You cannot assign a pre-process to a custom publication that publishes to a Hadoop-based repository.
- You cannot configure a custom publication that publishes files to a Hadoop-based repository to run immediately when the files are ready to be published.
- You cannot use a Hadoop repository to publish and subscribe to pass-through files and Hadoop Distributed File System (HDFS) files.
- **File Store.** Choose this type of repository to publish files that you want to keep as-is without loading the data into a relational database. For example, if you publish PDF or .zip files into a file repository, Data Integration Hub delivers the files without processing them.
- **Real-time.** Choose this type of repository to monitor real-time Apache Kafka data streaming. Apache Kafka is a distributed streaming platform that can publish and subscribe to stream of records, store and process streams of record.
In order to track the Kafka flows, you must configure the Apache Kafka server URL in the System property of the Data Integration Hub.

You must then create a topic with the publication repository type of Real-time and create an application to define the publisher and subscriber. Also, create a workflow that maps to the Apache Kafka server. The publication and subscription of the Data Integration hub associated with the source and target of the Kafka server.

The Data Integration Hub records streaming of data in the Apache Kafka server at regular intervals. The Data Integration Hub operator configures the interval at which Data Integration Hub must record the data streaming value in the topic. The Events List stores the log of events. The Processing Information tab in the Events List, stores the Offset and LogEndOffset values that define the difference between data values at intervals in each partition. For more information about events, see [“Managing Events on the Event List Page” on page 178](#).

Note: When you create a compound subscription, that is, a subscription that consumes data sets from multiple topics with a single batch workflow, all topics must be of the same type. Data Integration Hub operator enables the mandatory option for topics in compound subscription to prioritize a few topics in the compound subscription over other topics.

Data Integration Hub triggers a processing of subscription after the publication event for all topics are completed. If the wait time of the publication event is complete and Data Integration Hub has not published all mandatory topics, an error event is generated during run-time.

Topic Structure

When you create the structure of a topic, you define the data structure on the Data Integration Hub publication repository to where the publications that are associated with the topic publish data, and from where subscribers to the topic consume the data. The topic structure can consist of multiple tables.

When you create a topic, Data Integration Hub generates the tables in the publication repository where it retains the data that is published for the topic. Data Integration Hub uses the data structure for the publications and subscriptions that are associated with the topic.

Topic Structure Updates

When you edit the structure of a topic with associated publications or subscriptions, it might affect the associated publications and subscriptions. Topic structure changes might also impact the data in the publication repository and sometimes cause data loss.

Based on the nature of the update, you might have to edit the associated publications and subscriptions to align with the updated topic structure. The following table describes the effects of topic structure updates on data in the publication repository and on the associated publications and subscriptions, and the resulting optional or required changes to the associated publications and subscriptions.

Topic Structure Update	Effect on Data in Publication Repository ¹	Effect on Associated Publications and Subscriptions	Optional/Required Changes to Associated Publications and Subscriptions
Add table	Table added	None	Optional: To publish and to consume the additional table, edit the mapping to include the additional table. If you do not update the mapping, data in the table will not be published to the publication repository and subscribers will not receive it.
Delete table	Table deleted, including data that was published to the table	Become not valid	Remove references to the table from the mapping of publications and from the mapping and the filter of subscriptions.
Add column	Column added	None	Optional: To publish and to consume the additional column, edit the mapping to include the additional column. If you do not update the mapping, data in the column will not be published to the publication repository and subscribers will not receive it.
Delete column	Column deleted, including data that was published to the column	Become not valid	Remove references to the column from the mapping of publications and from the mapping and the filter of subscriptions.
Rename column	Column name updated	Automatic subscriptions with filters and joins become not valid	Run the subscription wizard for all associated subscriptions.

Topic Structure Update	Effect on Data in Publication Repository ¹	Effect on Associated Publications and Subscriptions	Optional/Required Changes to Associated Publications and Subscriptions
Change column physical name	Physical column name updated	Become not valid	To publish and to consume the column with a new name, edit the mapping to change the name. If you do not update the mapping, publishers will not publish and subscribers will not receive the data in the updated column.
Change column data type	Column deleted, including data that was published to the column, and another column created with new data type	Become not valid	Remove references to the changed column from the mapping of publications and from the mapping and the filter of subscriptions. Optional: To publish or to consume the column that is created with the new data type, edit the mapping to include the new column. If you do not update the mapping, data in the new column will not be published to the publication repository and subscribers will not receive it.
Increase column precision, scale unchanged	Column updated	Become not valid	Run the publication or the subscription wizard for all associated publications and subscriptions.
Increase column precision, increase scale by a lower value than the precision increase or by the same value as the precision increase	Column updated	Become not valid	Run the publication or the subscription wizard for all associated publications and subscriptions.
Any other precision or scale updates	Column deleted, including data that was published to the column, and another column created with updated precision or scale	Become not valid	Remove references to the changed column from the mapping of publications and from the mapping and the filter of subscriptions. Optional: To publish or to consume the column that is created with the new precision or the new scale, edit the mapping to include the new column. If you do not update the mapping, data in the new column will not be published to the publication repository and subscribers will not receive it.

¹. Deleting columns in the publication repository might take a long time, based on the number of rows in the table.

Add Topic Tables

For topics that store the topic data on a relational database or on a Hadoop repository, you can use one or more of the following methods to add tables to the topic:

- Add tables from a database. Use this method if the structure of the data domain that the topic represents exists in a relational database in your organization. For example, if the structure exists in one of the applications that publish to the topic.
- Add tables from Enterprise Data Catalog. Use this method when the structure of the data domain that the topic represents exists in Enterprise Data Catalog. Select tables from an Enterprise Data Catalog asset that the Data Integration Hub topic must use.
- Add tables from a flat file. Use this method when the structure of the data domain that the topic represents exists in a flat file. You can only use delimited files to add a topic table.
- Add tables from a PowerCenter workflow. Use this method if your organization has a PowerCenter workflow that extracts data from the source applications and the targets of the workflow represent the structure in which you want to keep the data. The developer creates the workflow in PowerCenter Workflow Manager. Data Integration Hub generates the topic structure based on the targets in the workflow.
- Add tables from a metadata file. Use this method when the structure of the table in the data domain that the topic represents exists in a JSON, XLS, XLSX, or XML file.
- Create a table. Use this method when the structure of the data domain that the topic represents does not exist in a database in your organization, in a PowerCenter workflow, or in a file, and you want to define the structure manually in Data Integration Hub.

You can use more than one method to add tables to a single topic. For example, add two tables from a database and one table from a PowerCenter workflow definition file.

For topics that store the topic data on a file repository, the only method that you can use to add a topic table is by creating the table. You need to define a table for each expected file. The table stores metadata only.

Add Topic Table Rules and Guidelines

When you add topic tables, consider the following rules and guidelines:

- A topic table name and a topic physical table name can contain ASCII alphanumeric characters and underscores and cannot start with a numeric character.
- A topic table name and a topic physical table name must be unique in the Data Integration Hub publication repository.

Using Metadata Files to Create Topic Tables

You can load a metadata file to Data Integration Hub and create a topic table based on the structure of the file. You can use JSON, XLS, XLSX, and XML metadata files to create topic tables.

When you use a metadata file to create a topic table, you can define table attributes in the file before you load it to Data Integration Hub. For example, define column data type and precision, or define a column as a filter accelerator.

All the following fields are mandatory for XLS/XLSX files, some of the fields are optional for JSON and XML files:

columnName

Mandatory. Name of the table column. The name must begin with an alphabetic character or underscore and can contain only alphanumeric characters or underscores. The name cannot include server-reserved

keywords. For a list of server-reserved keywords, see the Microsoft SQL Server or Oracle documentation. *columnName* can contain up to 80 characters. Any special characters and spaces in *columnName* are replaced with an underscore ("_").

physicalColumnName

Mandatory. Special characters and spaces in *physicalColumnName* are replaced with an underscore ("_"). *physicalColumnName* for Oracle and Hadoop can contain up to 30 characters. For SQL, it can contain up to 80 characters.

description

Optional. The description can contain up to 1024 characters. If a description exceeds the limit, Data Integration Hub truncates the value before saving it.

filterAccelerator

Optional. The column is used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions. By default, the value is false.

When you use filter accelerators, consider the following guidelines:

- Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
- Filter accelerators have no impact on subscriptions that do not use filters.
- Any change in the table resets the value of filterAccelerator.

datatype

Optional. Data type of the field. By default, the value is string.

The file can contain fields of the following data types:

- string
- decimal
- double
- int32
- int64
- date_time
- text

precision

Optional. Applies to data types that support precision. The default precision value depends on the data type of the field:

- String: 255
- Decimal: 15
- Text: 50000

scale

Optional. Applies to data types that support data scaling. The default scale value depends on the data type of the field:

- Decimal: 0
- All other data types: empty

Example table in a JSON file

The following example shows the table format in a JSON file:

```
[
  {
    "columnName": "id", "physicalColumnName": "id1", "description": "name_id",
    "filterAccelerator": "false", "dataType": "int32"},
  {
    "columnName": "name", "physicalColumnName": "name1", "description": "name_desc",
    "filterAccelerator": "false", "dataType": "string", "precision": 100}
]
```

Example table in an XML file

The following example shows the table format in an XML file:

```
<table>
  <column>
    <columnName>id</columnName>
    <physicalColumnName>id</physicalColumnName>
    <description>id_desc</description>
    <dataType>int32</dataType>
    <filterAccelerator>true</filterAccelerator>
  </column>
</table>
```

Example table in an XLS or XLSX file

The following table shows a sample table structure in an XLS or XLSX file:

columnName	physicalColumnName	description	filterAccelerator	dataType	precision	scale
id	id	desc_id	TRUE	int32		
name	name	desc_name	FALSE	String	255	
salary	salary	desc_salary	FALSE	decimal	15	2

Topic Table and Column Names

When Data Integration Hub creates the topic tables in the publication repository, it might change the table names and the column names that you assign when you create the topic.

Data Integration Hub applies the following rules when it determines the table and column names:

- Table and column names must be from one through 80 characters long.
- Data Integration Hub replaces non alphanumeric characters with '_' .
- When a table name or a column name starts with numeric characters, Data Integration Hub adds a 'T_' prefix to the table name or a 'F_' prefix to the column name.
- When a table name or a column name starts with an underscore, Data Integration Hub adds a 'T' prefix to the table name or a 'F' prefix to the column name.

Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic .

For example, the topic contains a Customer table and a Business Unit table. A customer can belong to more than one business unit. You create a subscription that subscribes to both tables in the topic, and add a filter to the Business Unit table to consume the data for a single business unit. Using the relationship information between the tables, the filter is implicitly applied to the customers table, and the subscriber consumes only the data for customers that belong to the business unit that interests you.

You define the relations between topic tables by assigning primary and foreign keys to the relevant tables. In our example, you define a column in the Business Unit as a primary key, and one or more columns in the Customer table as foreign keys.

Topic Data Retention

The data retention period of the topic defines how long Data Integration Hub retains the data in the publication repository after the data is consumed. When you use a Hadoop publication repository, you can configure Data Integration Hub not to delete published data from the Hadoop repository.

Topic Permissions

Topic permissions control who can access the topic.

The Data Integration Hub administrator creates categories and assigns categories to user groups to determine the users that can view or change topics. You assign categories to a topic to permit users to view or change the topic. Because publications and subscriptions are associated with topics, they inherit the permissions from the associated topic.

When you configure permissions for a topic, only user groups with permissions to the topic can access the associated subscriptions and publications.

Managing Topics

Use the Navigator to manage topics.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Choose to create, duplicate, edit, or delete a topic.
 - To create a topic, click **New Topic**, select the type of topic to create, and then click **Create**. Follow the Topic wizard.

- To duplicate a topic, click the Duplicate icon next to the topic. This creates a new topic with a different name and identical properties. Rename the topic in the **Confirmation** pop-up window, and click **Save**.

Note: If you don't rename the duplicate topic, Data Integration Hub saves it with the default name format, *<topic name>_Copy*.

- To edit a topic, click the Edit icon next to the topic. For more information about how you edit topic tables, see [“Topic Structure Properties” on page 60](#).
- To delete a topic, click the Delete icon next to the topic.

CHAPTER 5

Creating Topics

This chapter includes the following topics:

- [Creating Topics Overview, 39](#)
- [Creating a Topic with a Relational Database Publication Repository, 39](#)
- [Creating a Topic with a Hadoop Publication Repository, 46](#)
- [Creating a Topic with a File Store Publication Repository, 53](#)
- [Creating a Topic with a Real-time Publication Repository, 56](#)

Creating Topics Overview

In the Data Integration Hub Operation Console, use the topic wizard to create topics.

When you create a topic, you choose the topic type and the type of repository on which to store data for the topic, define the data structure and the data retention period, select a data storage location, and assign topic permissions.

Creating a Topic with a Relational Database Publication Repository

To create a topic with a relational database publication repository, perform the following tasks:

1. Access the **New Topic** wizard.
2. Define basic topic properties.
3. Create the topic data structure.
4. Optionally, define topic table relations.
5. Define the data retention period and the storage location.
6. Optionally, assign topic permissions.
7. Review the topic settings and save the topic.

Step 1. Access the New Relational Database Topic Wizard

Access the **New Relational Database Topic** wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **Relational Database**, and then click **Create**.
The **New Relational Database Topic** wizard appears.

Step 2. Define Basic Topic Properties

Define topic properties in the **General** page of the **New Topic** wizard.

1. Enter the topic name.
2. Optionally, enter a description of the topic.
3. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
4. Choose the topic type.
 - **Delta**. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data source includes delta indicators.
 - **Full**. The topic instance contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.
5. Under **Publication Repository** select **Relational database**.
Note: If you select the option **Prevent new publications and new subscriptions to this topic** you cannot create publications and subscriptions that publish to and subscribe from the topic.
6. Click **Next**.
The **Structure** page appears.

Step 3. Create Topic Data Structure

Create and preview the data structure of the topic in the **Structure** page of the **New Topic** wizard.

You can use one or more of the following methods to add tables to the topic:

Add tables from a database

Use this method if the structure of the data domain that the topic represents exists in a database in your organization. For example, if the structure exists in one of the applications that publish to the topic.

Add tables from Enterprise Data Catalog

Use this method when the structure of the data domain that the topic represents exists in Enterprise Data Catalog. Select tables from an Enterprise Data Catalog asset that the Data Integration Hub topic must use.

Add a table from a flat file

Use this method when the structure of the data domain that the topic represents exists in a flat file. You can only use delimited files to add a topic table with this method.

Add tables from a PowerCenter workflow

Use this method if your organization has a PowerCenter workflow that extracts data from the source applications, and the targets of the workflow represent the structure in which you want to keep the data.

The developer creates the workflow in PowerCenter Workflow Manager. Data Integration Hub generates the topic structure based on the targets in the workflow.

Add tables from a Metadata file

Use this method when the structure of a table in the data domain that the topic represents exists in a JSON, XML, XLS, or XLSX file.

Create a table

Use this method when the structure of the data domain that the topic represents does not exist in a database in your organization, in a PowerCenter workflow, or in a file, and you want to define the structure manually in Data Integration Hub.

Note: You must add at least one table to the topic structure.

1. Click **Add Tables** and then select the method by which to add the tables to the structure.
Note: Data Integration Hub adds the columns DIH_PUBLICATION_INSTANCE_DATE, DIH_PUBLICATION_INSTANCE_ID, and DIH__UPDATE_STRATEGY_FLAG to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.
2. Click **Next**.
The **Table Relations** page appears.

Adding Topic Tables from a Database

Add tables to the topic structure from a database that exists in your organization.

Perform the following steps to add topic tables from a database:

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Database**.
The **Add Tables from Database** page appears.
2. Click the browse button next to **Database Connection**.
The **Select Metadata Connection** dialog box appears.
3. Choose a connection from the list of connections, or click **New Connection** to create a connection.
Note: To use a Teradata connection you must create the connection before you create the topic. You cannot create a new Teradata connection here.
Data Integration Hub uses the metadata access connection details to access the database schema information.
4. In the **Select Metadata Connection** dialog box click **OK**.
The **Search Results** area shows tables in the database that use the default schema.
5. Optionally, search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Topic Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.
 - You can search for a substring.

6. In the **Select Topic Tables** area, select tables in the database to include in the topic structure and then click **Add Tables**.

The **Add Tables from Database** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding Tables from Enterprise Data Catalog

You add tables from Enterprise Data Catalog when the structure of the data domain that the topic represents exists in Enterprise Data Catalog.

Perform the following steps in the **Add Tables from Enterprise Data Catalog** window to add tables from Enterprise Data Catalog:

1. Search and select a resource in the **Enterprise Data Catalog Resource** field and click **OK**.
Data Integration Hub retrieves all the objects of the asset from Enterprise Data Catalog and displays them as options in the Select Topic Tables pane.
2. **Select Topic Tables**. Select tables to include in the topic structure.
3. Click **OK**.

Note: To use tables from Enterprise Data Catalog, verify that the Data Integration Hub administrator has configured the Enterprise Data Catalog connectivity with Data Integration Hub. For more information about configuring the Enterprise Data Catalog connection, see the *Data Integration Hub Administrator Guide*.

Adding a Topic Table from a Flat File

Add a table to the topic structure from a sample flat file. You can only use delimited files to add a topic table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Flat File**.

The **Add Table from File** wizard appears, showing the **Definition** page.

2. Click the browse button next to **Sample file**.

The **Sample File** dialog box appears.

3. Click **Browse**, select the file, and then click **Upload**.

The **Sample File** dialog box closes.

4. In the **File Format** area, define the following parameters:

Code page

Character encoding used in the file.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Default text length

Optional. Length of the text fields in the table.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

The **Preview** area shows the columns in the table.

5. Click **Next**.

The **Structure** page of the **Add Table from File** wizard appears.

6. Review the table structure and adjust it if required. The following list describes the elements of the **Structure** page of the **Add Table from File** wizard:

Table Name

Name of the table. By default, Data Integration Hub derives the table name from the name of the file that you select in the **Definition** page. You can edit the default name.

The table name can contain only ASCII alphanumeric characters and underscores and cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Column table

Use the add, up, down, and remove buttons to add, order, and remove table columns. The table must contain at least one column.

Each column must contain the following information:

- Column Name. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
If you select the option **Import column names** in the **Definition** page, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1**, **Field2**, **Fieldn**.
- Data Type. Select from the list of available data types. By default, Data Integration Hub reads the data as string.
- Precision. Enabled only for data types that support precision.
- Scale. Enabled only for data types that support data scaling.

Preview

This area shows the columns in the table. If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file.

7. Click **Finish**.

The **Add Table from File** wizard closes. The **Structure** page of the **New Topic** wizard shows the topic table and a preview of the topic structure.

Adding Topic Tables from a PowerCenter Workflow

Add tables to the topic structure from a PowerCenter workflow.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From PowerCenter**.
The **Add Tables from PowerCenter Workflow** page appears.
2. Choose one of the following options to add tables to the topic structure:
 - Choose **Select a workflow from the PowerCenter repository** to select a workflow from the PowerCenter repository. Either enter a string in the **Folder name** text box and then click **Search**, or click **Show All**, and then select a workflow from the list.

- Choose **Select a workflow definition file (.xml)** to select a workflow definition file. Browse to select the file and then click **Upload**.
3. Click **Save**.
The **Add Tables from PowerCenter Workflow** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding Topic Tables from a Metadata File

You can load a metadata file to Data Integration Hub and create a topic table that is based on the structure of the file. Data Integration Hub supports JSON, XML, XLS, and XLXS file formats.

1. In the **Structure** page of the New Topic wizard, click **Add Tables** and then select **From Metadata File**.
The **Add table from Metadata File** page appears.
2. Browse to select the file and click **Upload**.
3. Enter the table name, and click **Save**.
The **Structure** page shows a list of topic tables and a preview of the topic structure.

Creating a Topic Table

Add a table to the topic structure by creating the table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **Create**.
The **Create Table** page appears.
2. Enter the logical table name in the **Table Name** field.
3. If required, change the physical table name.
4. Optionally, enter a description of the table.
5. To apply delta detection to the table click **Apply Delta Detection**.
Note: Data Integration Hub applies delta detection to tables for which a primary key is defined. You assign primary keys to tables in the **Table Relations** tab of the topic wizard.
6. Use the add, up, and down buttons to add and order table columns. You must add at least one column to the table.

For each column, enter the following information:

- Column Name. Logical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- Physical Name. Physical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- Data Type. Select from the list of available data types.
- Precision. Enabled only for data types that support precision.
- Scale. Enabled only for data types that support data scaling.
- Description. Optional description of the column.

Note: Data Integration Hub adds the columns DIH_PUBLICATION_INSTANCE_DATE, DIH_PUBLICATION_INSTANCE_ID, and DIH__UPDATE_STRATEGY_FLAG to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

7. Click **OK**.

The **Create Table** page closes. The **Structure** page shows the topic table and a preview of the topic structure.

Step 4. Define Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic in the **Table Relations** page of the **New Topic** wizard. You can also define filter accelerators for topics that you plan to use for unbound subscriptions.

Tip: In the **Show** list, you can select to show all topic tables or to show a specific table.

1. In the **Filter Accelerator** column, define a column or multiple columns as filter accelerators. A filter accelerator indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
 - Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
 - Filter accelerators have no impact on subscriptions that do not use filters.
2. In the **Key** column, define a column from a topic table for which to define relations as a primary key.
 3. Define primary keys for as many tables as required, one primary key for each table.
 4. For each table for which you defined a primary key, define one or more foreign keys. A foreign key must relate to a primary key from a different table.
 5. Click **Next**.

The **Data Retention** page appears.

Step 5. Define Data Retention Period and Storage Location

Define the data retention period and the data storage location in the **Data Retention** page of the **New Topic** wizard.

1. Enter the number of days that Data Integration Hub retains the data in the Data Integration Hub publication database after the data is consumed in the **Publication data retention period** field.

2. Click **Advanced**.

The **Data Storage Location** area appears.

3. Choose the location in the database where Data Integration Hub stores data that is published to the topic.
 - Default. Data Integration Hub stores the data based on the default storage configuration. For example, the default File Group.
 - Custom. Browse to select an available data storage group in the publication repository database. For example, browse the available file groups.
4. Click **Next**.

The **Permissions** page appears.

Step 6. Assign Topic Permissions

Control access to the topic in the Operation Console in the **Permissions** page of the **New Topic** wizard. If you do not assign permissions, the topic is accessible by all Data Integration Hub users.

1. Select the category to which you want to assign permission to the topic under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

2. Repeat step [1](#) to assign additional categories.
3. Click **Next**.

The **Summary** page appears.

Step 7. Review Topic Settings and Save the Topic

Review the topic settings and save the topic in the **Summary** page of the **New Topic** wizard.

1. Review the topic settings.
2. Click **Finish**.

The **New Topic** wizard closes. The **Topics** page shows the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The topic is listed in the catalog. In the catalog you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

Creating a Topic with a Hadoop Publication Repository

To create a topic with a Hadoop publication repository, perform the following tasks:

1. Access the **New Topic** wizard.
2. Define basic topic properties.
3. Create the topic data structure.
4. Optionally, define topic table relations.
5. Define the data retention period and the storage location.
6. Optionally, assign topic permissions.
7. Review the topic settings and save the topic.

Step 1. Access the New Big Data Topic Wizard

Access the **New Big Data Topic** wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **Big Data**, and then click **Create**.

The **New Big Data Topic** wizard appears.

Step 2. Define Basic Topic Properties

Define topic properties in the **General** page of the **New Topic** wizard.

1. Enter the topic name.
2. Optionally, enter a description of the topic.
3. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
4. Choose the topic type.
 - Delta. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data source includes delta indicators.
 - Full. The topic instance contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.
5. Under **Publication Repository** select **Hadoop**.

Note: If you select the option **Prevent new publications and new subscriptions to this topic** you cannot create publications and subscriptions that publish to and subscribe from the topic.
6. Click **Next**.

The **Structure** page appears.

Step 3. Create Topic Data Structure

Create and preview the data structure of the topic in the **Structure** page of the **New Topic** wizard.

You can use one or more of the following methods to add tables to the topic:

Add tables from a database

Use this method if the structure of the data domain that the topic represents exists in a database in your organization. For example, if the structure exists in one of the applications that publish to the topic.

Add tables from Enterprise Data Catalog

Use this method when the structure of the data domain that the topic represents exists in Enterprise Data Catalog. Select tables from an Enterprise Data Catalog asset that the Data Integration Hub topic must use.

Add a table from a flat file

Use this method when the structure of the data domain that the topic represents exists in a flat file. You can only use delimited files to add a topic table with this method.

Add tables from a PowerCenter workflow

Use this method if your organization has a PowerCenter workflow that extracts data from the source applications, and the targets of the workflow represent the structure in which you want to keep the data. The developer creates the workflow in PowerCenter Workflow Manager. Data Integration Hub generates the topic structure based on the targets in the workflow.

Add tables from a Metadata file

Use this method when the structure of a table in the data domain that the topic represents exists in a JSON, XML, XLS, or XLSX file.

Create a table

Use this method when the structure of the data domain that the topic represents does not exist in a database in your organization, in a PowerCenter workflow, or in a file, and you want to define the structure manually in Data Integration Hub.

Note: You must add at least one table to the topic structure.

1. Click **Add Tables** and then select the method by which to add the tables to the structure.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

2. Click **Next**.

The **Table Relations** page appears.

Adding Topic Tables from a Database

Add tables to the topic structure from a database that exists in your organization.

Perform the following steps to add topic tables from a database:

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Database**.

The **Add Tables from Database** page appears.

2. Click the browse button next to **Database Connection**.

The **Select Metadata Connection** dialog box appears.

3. Choose a connection from the list of connections, or click **New Connection** to create a connection.

Note: To use a Teradata connection you must create the connection before you create the topic. You cannot create a new Teradata connection here.

Data Integration Hub uses the metadata access connection details to access the database schema information.

4. In the **Select Metadata Connection** dialog box click **OK**.

The **Search Results** area shows tables in the database that use the default schema.

5. Optionally, search for tables in the database. Note the following guidelines:

- To search for tables by table name, enter a string in the **Find Topic Table** text box and then click **Search**.
- To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
- To clear the search results and show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
- The search is not case sensitive.
- You can search for a substring.

6. In the **Select Topic Tables** area, select tables in the database to include in the topic structure and then click **Add Tables**.

The **Add Tables from Database** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding a Topic Table from a Flat File

Add a table to the topic structure from a sample flat file. You can only use delimited files to add a topic table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From Flat File**.

The **Add Table from File** wizard appears, showing the **Definition** page.

2. Click the browse button next to **Sample file**.
The **Sample File** dialog box appears.
3. Click **Browse**, select the file, and then click **Upload**.
The **Sample File** dialog box closes.
4. In the **File Format** area, define the following parameters:

Code page

Character encoding used in the file.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table.
Enter the number of the row that serves as the file's header row in **From row**.

Default text length

Optional. Length of the text fields in the table.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

The **Preview** area shows the columns in the table.

5. Click **Next**.
The **Structure** page of the **Add Table from File** wizard appears.
6. Review the table structure and adjust it if required. The following list describes the elements of the **Structure** page of the **Add Table from File** wizard:

Table Name

Name of the table. By default, Data Integration Hub derives the table name from the name of the file that you select in the **Definition** page. You can edit the default name.

The table name can contain only ASCII alphanumeric characters and underscores and cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Column table

Use the add, up, down, and remove buttons to add, order, and remove table columns. The table must contain at least one column.

Each column must contain the following information:

- **Column Name**. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
If you select the option **Import column names** in the **Definition** page, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1**, **Field2**, **Fieldn**.

- **Data Type.** Select from the list of available data types. By default, Data Integration Hub reads the data as string.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.

Preview

This area shows the columns in the table. If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file.

7. Click **Finish**.

The **Add Table from File** wizard closes. The **Structure** page of the **New Topic** wizard shows the topic table and a preview of the topic structure.

Adding Tables from Enterprise Data Catalog

You add tables from Enterprise Data Catalog when the structure of the data domain that the topic represents exists in Enterprise Data Catalog.

Perform the following steps in the **Add Tables from Enterprise Data Catalog** window to add tables from Enterprise Data Catalog:

1. Search and select a resource in the **Enterprise Data Catalog Resource** field and click **OK**.
Data Integration Hub retrieves all the objects of the asset from Enterprise Data Catalog and displays them as options in the Select Topic Tables pane.
2. **Select Topic Tables.** Select tables to include in the topic structure.
3. Click **OK**.

Note: To use tables from Enterprise Data Catalog, verify that the Data Integration Hub administrator has configured the Enterprise Data Catalog connectivity with Data Integration Hub. For more information about configuring the Enterprise Data Catalog connection, see the *Data Integration Hub Administrator Guide*.

Adding Topic Tables from a PowerCenter Workflow

Add tables to the topic structure from a PowerCenter workflow.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **From PowerCenter**.
The **Add Tables from PowerCenter Workflow** page appears.
2. Choose one of the following options to add tables to the topic structure:
 - Choose **Select a workflow from the PowerCenter repository** to select a workflow from the PowerCenter repository. Either enter a string in the **Folder name** text box and then click **Search**, or click **Show All**, and then select a workflow from the list.
 - Choose **Select a workflow definition file (.xml)** to select a workflow definition file. Browse to select the file and then click **Upload**.
3. Click **Save**.

The **Add Tables from PowerCenter Workflow** page closes. The **Structure** page shows a list of topic tables and a preview of the topic structure.

Adding Topic Tables from a Metadata File

You can load a metadata file to Data Integration Hub and create a topic table that is based on the structure of the file. Data Integration Hub supports JSON, XML, XLS, and XLSX file formats.

1. In the **Structure** page of the New Topic wizard, click **Add Tables** and then select **From Metadata File**.

The **Add table from Metadata File** page appears.

2. Browse to select the file and click **Upload**.
3. Enter the table name, and click **Save**.

The **Structure** page shows a list of topic tables and a preview of the topic structure.

Creating a Topic Table

Add a table to the topic structure by creating the table.

1. In the **Structure** page of the **New Topic** wizard, click **Add Tables** and then select **Create**.

The **Create Table** page appears.

2. Enter the logical table name in the **Table Name** field.
3. If required, change the physical table name.
4. Optionally, enter a description of the table.
5. Use the add, up, and down buttons to add and order table columns. You must add at least one column to the table.

For each column, enter the following information:

- Column Name. Logical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- Physical Name. Physical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- Data Type. Select from the list of available data types.
- Precision. Enabled only for data types that support precision.
- Scale. Enabled only for data types that support data scaling.
- Description. Optional description of the column.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

6. Click **OK**.

The **Create Table** page closes. The **Structure** page shows the topic table and a preview of the topic structure.

Step 4. Define Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic in the **Table Relations** page of the **New Topic** wizard. You can also define filter accelerators for topics that you plan to use for unbound subscriptions.

Tip: In the **Show** list, you can select to show all topic tables or to show a specific table.

1. In the **Filter Accelerator** column, define a column or multiple columns as filter accelerators. A filter accelerator indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
 - Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
 - Filter accelerators have no impact on subscriptions that do not use filters.
2. In the **Key** column, define a column from a topic table for which to define relations as a primary key.
 3. Define primary keys for as many tables as required, one primary key for each table.
 4. For each table for which you defined a primary key, define one or more foreign keys. A foreign key must relate to a primary key from a different table.
 5. Click **Next**.

The **Data Retention** page appears.

Step 5. Define Data Retention Period

Define the data retention period in the **Data Retention** page of the **New Topic** wizard.

1. Choose one of the following options:

Retain published data for ... days after it is consumed

Enter the number of days that Data Integration Hub retains the data in the Data Integration Hub publication repository after the data is consumed. When the retention period ends, Data Integration Hub deletes the data from the repository.

Do not delete published data for this topic

Data Integration Hub does not delete the data from the publication repository.

2. Click **Next**.

The **Permissions** page appears.

Step 6. Assign Topic Permissions

Control access to the topic in the Operation Console in the **Permissions** page of the **New Topic** wizard. If you do not assign permissions, the topic is accessible by all Data Integration Hub users.

1. Select the category to which you want to assign permission to the topic under **Available Categories** and click the right arrow.

The category appears under **Selected Categories**.

2. Repeat step [1](#) to assign additional categories.
3. Click **Next**.
The **Summary** page appears.

Step 7. Review Topic Settings and Save the Topic

Review the topic settings and save the topic in the **Summary** page of the **New Topic** wizard.

1. Review the topic settings.
2. Click **Finish**.

The **New Topic** wizard closes. The **Topics** page shows the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The topic is listed in the catalog. In the catalog you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

Creating a Topic with a File Store Publication Repository

To create a topic with a file store publication repository, perform the following tasks:

1. Access the **New Topic** wizard.
2. Define basic topic properties.
3. Create the topic data structure.
4. Optionally, define topic table relations.
5. Define the data retention period and the storage location.
6. Optionally, assign topic permissions.
7. Review the topic settings and save the topic.

Step 1. Access the New File Store Topic Wizard

Access the **New File Store Topic** wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **File Store**, and then click **Create**.
The **New File Store Topic** wizard appears.

Step 2. Define Basic Topic Properties

Define topic properties in the **General** page of the **New Topic** wizard.

1. Enter the topic name.
2. Optionally, enter a description of the topic.

3. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
4. Choose the topic type.
 - Delta. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data source includes delta indicators.
 - Full. The topic instance contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.
5. Under **Publication Repository** select **File repository**.

Note: If you select the option **Prevent new publications and new subscriptions to this topic** you cannot create publications and subscriptions that publish to and subscribe from the topic.
6. Click **Next**.

The **Structure** page appears.

Step 3. Create Topic Data Structure

Create and preview the data structure of the topic in the **Structure** page of the **New Topic** wizard.

1. Click **Add Tables**.

The **Create Table** page appears. The column table fields are populated with the default Data Integration Hub values, and you can only edit the column precision fields.

Note: Data Integration Hub adds the columns DIH_PUBLICATION_INSTANCE_DATE, DIH_PUBLICATION_INSTANCE_ID, and DIH__UPDATE_STRATEGY_FLAG to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.
2. Enter the logical table name in the **Table Name** field.
3. If required, change the physical table name.
4. Optionally, enter a description of the table.
5. If required, edit the column precision fields.
6. Click **OK**.

The **Create Table** page closes. The **Structure** page shows the topic table and a preview of the topic structure.
7. Click **Next**.

The **Table Relations** page appears.

Step 4. Define Topic Table Relations

You can define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic in the **Table Relations** page of the **New Topic** wizard. You can also define filter accelerators for topics that you plan to use for unbound subscriptions.

Tip: In the **Show** list, you can select to show all topic tables or to show a specific table.

1. In the **Filter Accelerator** column, define a column or multiple columns as filter accelerators. A filter accelerator indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
 - Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
 - Filter accelerators have no impact on subscriptions that do not use filters.
2. In the **Key** column, define a column from a topic table for which to define relations as a primary key.
 3. Define primary keys for as many tables as required, one primary key for each table.
 4. For each table for which you defined a primary key, define one or more foreign keys. A foreign key must relate to a primary key from a different table.
 5. Click **Next**.
- The **Data Retention** page appears.

Step 5. Define Data Retention Period and Storage Location

Define the data retention period and the data storage location in the **Data Retention** page of the **New Topic** wizard.

1. Enter the number of days that Data Integration Hub retains the data in the Data Integration Hub publication database after the data is consumed in the **Publication data retention period** field.
2. Click **Advanced**.
The **Data Storage Location** area appears.
3. Choose the location in the database where Data Integration Hub stores data that is published to the topic.
 - Default. Data Integration Hub stores the data based on the default storage configuration. For example, the default File Group.
 - Custom. Browse to select an available data storage group in the publication repository database. For example, browse the available file groups.
4. Click **Next**.
The **Permissions** page appears.

Step 6. Assign Topic Permissions

Control access to the topic in the Operation Console in the **Permissions** page of the **New Topic** wizard. If you do not assign permissions, the topic is accessible by all Data Integration Hub users.

1. Select the category to which you want to assign permission to the topic under **Available Categories** and click the right arrow.
The category appears under **Selected Categories**.
2. Repeat step [1](#) to assign additional categories.
3. Click **Next**.
The **Summary** page appears.

Step 7. Review Topic Settings and Save the Topic

Review the topic settings and save the topic in the **Summary** page of the **New Topic** wizard.

1. Review the topic settings.
2. Click **Finish**.

The **New Topic** wizard closes. The **Topics** page shows the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The topic is listed in the catalog. In the catalog you can view all the data sets that are available in Data Integration Hub and verify that the data that the target applications require exists in Data Integration Hub.

Creating a Topic with a Real-time Publication Repository

This topic describes how to create a topic that fetches consumer metrics from the Apache Kafka messaging platform.

Task Prerequisites

Before you start this task, obtain the details of the topic structure that you want to create.

To create a new topic with the Apache Kafka repository, perform the following tasks:

1. In the Navigator, click **Hub Management > Topics**.
The **Topics** page appears.
2. Click **New Topic**, select **Real-Time**, and then click **Create**.
The **Create Real-timeTopic** wizard appears.
3. Enter the topic name.
4. Optionally, enter a description of the topic.
5. Optionally, assign a tag to the topic so that you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.
6. Enter the number of partitions.
The maximum number of partitions that the topic can have is 999.
7. Select the intervals in which the topic must fetch the consumer metrics from the Apache Kafka server.
The default value is 30 seconds.
8. Click **Permissions** to provide permissions of a category to the topic.
 - a. Select a category from the **Available Categories** and click the right arrow.
The category that you select is displayed in the list of **Selected Categories**. The topic now has permissions that are defined for the selected category.
9. Click **Save**.

The **Create Real-time Topic** wizard closes. The **Topics** page displays the topic you created. You can create publications that publish data to the topic and subscriptions that consume data from the topic. The **Catalog** page lists the new topic.

CHAPTER 6

Topic Properties

This chapter includes the following topics:

- [Topic Properties Overview, 57](#)
- [Topic General Properties, 58](#)
- [Topic Structure Properties, 60](#)
- [Topic Table Relations Properties, 65](#)
- [Topic Data Retention Properties for Relational Database and File Repositories, 65](#)
- [Topic Data Retention Properties for a Hadoop Repository, 66](#)
- [Topic Permissions Properties, 66](#)
- [Topic Summary Properties, 66](#)

Topic Properties Overview

Use the topic wizard to create and edit topics to which applications publish data and from which applications consume data.

The topic wizard contains the following pages:

General page

Define basic topic properties and select the type of repository on which Data Integration Hub stores published data for the topic.

Structure page

Create and preview the structure of the data that the topic captures. When you save the topic, Data Integration Hub creates the structure on the Data Integration Hub publication repository.

Table Relations page

Define relations between topic tables, to filter the data that automatic database and flat file subscriptions consume from the topic.

Data Retention page

Define how long Data Integration Hub retains the data in the publication repository and select a data storage location.

Permissions page

Control access to the topic in the Operation Console by selecting categories to grant users permissions to access the topic. Only users with data access permissions for the selected categories can access the

topic. If you do not select any category, all Data Integration Hub users can access the topic in the Operation Console.

Summary page

Review topic properties and save the topic.

Topic General Properties

The **General** page of the topic wizard includes the following properties:

Topic Name

Name of the topic. The name is not case-sensitive and can contain up to 48 characters. The name can contain only alphanumeric characters and underscores.

You cannot change the topic name after you create the topic.

Description

Optional description of the topic.

Tag

Topic tag. When you add a tag to a topic, you can search for the topic on the **Topics** page and the **Catalog** page based on the tag.

Topic Type

Type of the topic. Topic type depends on the type of data that applications publish to the topic and has an impact on the delivery options to the subscribers to the topic.

Choose one of the following options:

- Delta. The topic instance contains only the latest data changes. If you choose this topic type, verify that the data sources include delta indicators.
- Full. The topic contains all of the data changes that occurred after the last publication. Choose this topic type if you associate publications with custom mapping with the topic.

Publication Repository

Type of repository on which Data Integration Hub stores published data for the topic.

Choose one of the following options:

- Relational database. Choose this type of repository to store published data in a relational database structure that represents the structure in which you want to keep the data. For example, data that is published from a relational database or from files. A relational database publication repository usually stores the data for a short intermediate period after the data is consumed by all subscribers. Data Integration Hub supports the following databases on which to store relational database topic data: Oracle, Microsoft SQL Server.
- Big Data. Choose this type of repository if you publish high volumes of data that you want to store for a long period of time or if you do not want Data Integration Hub to delete published data after the data is consumed. The availability of the Hadoop repository depends on whether or not the Hadoop component is installed on your system.
To publish and subscribe to a Hadoop-based repository with custom publications and subscriptions, you must use workflows that are based on a Data Engineering Integration mapping and workflow. When you create a custom publication, if one of the topics that you select for the publication is a

Hadoop-based topic, only workflows that are based on a Data Engineering Integration mapping or workflow are listed for selection as the publication mapping.

When you create a compound subscription, that subscribes to multiple topics, all topics that you select must be Hadoop-based, and only workflows that are based on a Data Engineering Integration mapping or workflow are listed for selection as the subscription mapping. You also enable the mandatory option for topics in compound subscription to prioritize a few topics in the compound subscription over other topics.

Data Integration Hub triggers a processing of subscription after the publication event for all topics are completed. If the wait time of the publication event is complete and Data Integration Hub has not published all mandatory topics, an error event is generated during run-time.

Before you use a Hadoop-based publication repository for publications and subscriptions, consider the following restrictions:

- You cannot assign a pre-process to a custom publication that publishes to a Hadoop-based repository.
- You cannot configure a custom publication that publishes files to a Hadoop-based repository to run immediately when the files are ready to be published.
- You cannot use a Hadoop repository to publish and subscribe to pass-through files and Hadoop Distributed File System (HDFS) files.
- File Store. Choose this type of repository to publish files that you want to keep as-is without loading the data into a relational database. For example, if you publish PDF or .zip files into a file repository, Data Integration Hub delivers the files without processing them.
- Real-time. Choose this type of repository to monitor real-time Apache Kafka data streaming. Apache Kafka is a distributed streaming platform that can publish and subscribe to stream of records, store and process streams of record. In order to track the Apache Kafka flows, you must configure the Apache Kafka server URL in the System property of the Data Integration Hub. You must then create a topic with the publication repository type of Real-time and create an application to define the publisher and subscriber. Also, create a workflow that maps to the Apache Kafka server. The publication and subscription of the Data Integration hub associated with the source and target of the Kafka server.

The Data Integration Hub records streaming of data in the Apache Kafka server at regular intervals. The Data Integration Hub operator configures the interval at which Data Integration Hub must record the data streaming value in the topic. The Events List stores the log of events. The Processing Information tab in the Events List, stores the Offset and LogEndOffset values that define the difference between data values at intervals in each partition.

Prevent new publications and new subscriptions to this topic

Prevent new publications from publishing to the topic and prevent new subscriptions from subscribing to the topic. For example, when you plan to delete the topic. The topic is not available for selection when creating publications and subscriptions.

Existing publications can publish data to the topic and existing subscriptions can consume data from the topic.

Note: You can use the Data Integration Hub import and export utility to import new publications and subscriptions to the topic.

Topic Structure Properties

The **Structure** page of the topic wizard includes the following properties:

Topic Tables

Details of the tables that are part of the topic structure.

To add topic tables, click **Add Tables**. For topics that use a relational database or a Hadoop publication repository, select the method by which to add the tables to the structure. You must add one or more tables.

Structure Preview

Shows details of the tables you add to the topic structure. You can show all tables or choose a table for which to show the details.

Add Tables from Database

For topics that use a relational database or a Hadoop publication repository, define the structure of the topic by selecting tables from a relational database.

To add tables from a database, you must select or define the metadata connection to the database.

Database Table Properties

When you define the structure of the topic by adding tables from a database, you select or create the database metadata connection and then select the tables to include in the topic structure.

The **Add Tables from Database** page contains the following properties:

Database Connection

Name of the data access connection. You can browse to choose from the list of connections, or create a connection, in the **Select Metadata Connection** dialog box.

Find Topic Table

Optionally, search for tables in the database. Note the following guidelines:

- To search for tables by table name, enter a string in the **Find Topic Table** text box and then click **Search**.
- To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
- To clear the search results and show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
- The search is not case sensitive.
- You can search for a substring.

Select Topic Tables

Select tables in the database to include in the topic structure.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

Add Tables from Enterprise Data Catalog

You add tables from Enterprise Data Catalog when the structure of the data domain that the topic represents exists in Enterprise Data Catalog.

The **Add Tables from Enterprise Data Catalog** page contains the following properties:

Enterprise Data Catalog Resource

Name of the data access connection. You can browse to choose from the list of resources.

Select Topic Tables

Select tables to include in the topic structure.

Note: To use tables from Enterprise Data Catalog, verify that the Data Integration Hub administrator has configured the Enterprise Data Catalog connectivity with Data Integration Hub. For more information about configuring the Enterprise Data Catalog connection, see the *Data Integration Hub Administrator Guide*.

Add a Table from Flat File

For topics that use a relational database or a Hadoop publication repository, define the structure of the topic by extracting the table structure from a flat file. Data Integration Hub creates a single table from each file.

To extract a table from a flat file, you must select a file and define the file format. You then review the table structure and adjust it if required.

You can only use delimited files to add a topic table.

File Table Definition Properties

Use the **Definition** page of the **Add Table from Flat File** wizard to select the file that contains the structure of the data domain that the topic represents, to configure the file format, and to preview the file.

The **Definition** page of the **Add Table from Flat File** wizard includes the following elements:

File

Location and name of the file that contains the structure of the data domain that the topic represents. You can browse to select the file.

Code page

Character encoding used in the file.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Default text length

Optional. Length of the text fields in the table.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

Preview

This area shows the columns in the table. If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file.

File Table Structure Properties

Use the **Structure** page of the **Add Table from Flat File** wizard to review the table structure and to adjust it if required.

The **Structure** page of the **Add Table from Flat File** wizard includes the following elements:

Table Name

Name of the table. By default, Data Integration Hub derives the table name from the name of the file that you select in the **Definition** page. You can edit the default name.

The table name can contain only ASCII alphanumeric characters and underscores and cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Column table

Use the add, up, down, and remove buttons to add, order, and remove table columns. The table must contain at least one column.

Each column must contain the following information:

- **Column Name.** Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
If you select the option **Import column names** in the **Definition** page, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1**, **Field2**, **Fieldn**.
- **Data Type.** Select from the list of available data types. By default, Data Integration Hub reads the data as string.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.

Preview

This area shows the columns in the table. If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to the table. When you finish adding the table, the columns show in the **Structure Preview** area of the **Structure** tab of the **Topic** wizard.

Add Table from PowerCenter Workflow

For topics that use a relational database or a Hadoop publication repository, define the structure of the topic by extracting tables from a PowerCenter workflow.

To extract tables from a PowerCenter workflow, you must select either a PowerCenter workflow or a workflow definition file that the developer exported from PowerCenter. Data Integration Hub generates the topic structure based on the targets in the workflow.

When Data Integration Hub creates the topic tables in the publication repository, it might change the table names and the column names that you assign when you create the topic. If you want to use the same

workflow for a publication with custom mappings, you must first verify that the target tables in the workflow are identical to the topic target tables that Data Integration Hub generated on the Data Integration Hub publication repository. If required, update the tables in the workflow. For more information, see [“Topic Table and Column Names” on page 36](#).

Workflow Table Properties

When you add tables from a PowerCenter workflow that defines the tables to include in the topic structure, you either select a workflow from the PowerCenter repository, or select a workflow definition file.

The **Add Tables from PowerCenter Workflow** page contains the following properties:

Use one of the following methods to add a table based on a PowerCenter workflow

- Choose **Select a workflow from the PowerCenter repository** to select a workflow. To show the workflows that you can select, either enter a string in the **Folder name** text box and then click **Search**, or click **Show All**. Select a workflow from the list, and then click **Save**.
- Choose **Select a workflow definition file (.xml)** to select a workflow definition file. Browse to select a workflow definition file, then click **Upload**, and then click **Save**.

Note: Data Integration Hub adds the columns DIH_PUBLICATION_INSTANCE_DATE, DIH_PUBLICATION_INSTANCE_ID, and DIH_UPDATE_STRATEGY_FLAG to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

Add a Table from Metadata File Properties

Use the **Add Table from Metadata File** page select the file that contains the structure of the data domain that the topic represents and name the table to create.

The **Add Table from Metadata File** page contains the following parameters:

File

Location and name of the file that contains the structure of the data domain that the topic represents. You can browse to select the file.

Table Name

Logical name of the table. Can contain only ASCII alphanumeric characters and underscores. Cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Create a Table

Define the structure of the topic by creating a table.

To create a table, you must define the table name. For topics that use a relational database or a Hadoop publication repository, define the structure of the table by adding and configuring table columns. For topics that use a file publication repository, Data Integration Hub defines the structure of the table that holds the metadata of the file and you cannot add columns to the table or remove columns from the table. You can only edit the precision of the columns that Data Integration Hub creates.

Create Table Properties

When you define the structure of the topic by creating a table, you name the table. For topics that use a relational database or a Hadoop publication repository, you also configure the table columns.

The **Create Table** page contains the following properties:

Table Name

Logical name of the table. Can contain only ASCII alphanumeric characters and underscores. Cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Physical Table Name

Physical name of the table. Can contain only ASCII alphanumeric characters and underscores. Cannot start with a numeric character.

The name must be unique in the Data Integration Hub publication repository.

Table Description

Optional description of the table.

Apply Delta Detection

Data Integration Hub detects the delta between the current publication event and the previous publication event. Data Integration Hub applies the delta that it detects in publications to subscriptions that consume the data from a relational database topic with the delivery option **Insert new rows and update changed rows**.

Note: Data Integration Hub applies delta detection to tables for which a primary key is defined. You assign primary keys to tables in the **Table Relations** tab of the topic wizard.

Column table

For topics that use a relational database or a Hadoop publication repository, use the add, up, and down buttons to add and order table columns. You must add at least one column to the table. For each column, enter the following information:

- **Column Name.** Logical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- **Physical Name.** Physical name of the column. Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.
- **Data Type.** Select from the list of available data types.
- **Precision.** Enabled only for data types that support precision.
- **Scale.** Enabled only for data types that support data scaling.
- **Description.** Optional description of the column.

For topics that use a file publication repository, Data Integration Hub defines the tables and columns that hold the file metadata and you can only edit the precision field.

Note: Data Integration Hub adds the columns `DIH_PUBLICATION_INSTANCE_DATE`, `DIH_PUBLICATION_INSTANCE_ID`, and `DIH__UPDATE_STRATEGY_FLAG` to each topic table. When you finish creating the table, the columns show in the **Structure Preview** area of the **Structure** page of the topic wizard.

Topic Table Relations Properties

The **Table Relations** page of the topic wizard includes the following properties:

Table Name

Name of the table.

Column table

Name of the column.

Filter Accelerator

Indicates that the column will be used in subscription queries and requires performance-related handling by Data Integration Hub. Use this indicator with topics that you plan to use for unbound subscriptions.

When you use filter accelerators, consider the impact that performance-related handling has on system performance:

- Filter accelerators speed up the handling of subscriptions that use both filters and PowerCenter pushdown optimization. By default, the PowerCenter pushdown optimization option is selected for unbound subscriptions only.
- Filter accelerators slow down the writing of publication data to the Data Integration Hub publication repository.
- Filter accelerators have no impact on subscriptions that do not use filters.

Key

Assign primary and foreign keys to define relations between topic tables, to filter the data that automatic database subscriptions consume from the topic.

- Define columns as primary keys for as many tables as required, one primary key for each table.
- For each table for which you defined a primary key, define one or more columns as foreign keys. A foreign key must be related to a primary key in a different table.

Topic Data Retention Properties for Relational Database and File Repositories

The **Data Retention** page of the topic wizard includes the following properties for relational database and file repositories:

Publication data retention period

Determines how long Data Integration Hub retains published data in the publication repository after the data is consumed.

For each publication instance, the retention period starts after all subscribers have either successfully consumed the published data or discarded the data. That is, after all the events that are associated with the publication instance are either in a Complete or in a Discarded event status.

Note: The data retention period applies to the published data in the publication repository. Data Integration Hub does not delete the source data from the document store.

Advanced

Data Storage Location. Location in the database where Data Integration Hub stores the topic. Choose one of the following options:

- Default Location. Data Integration Hub stores the topic based on the default storage configuration. For example, in the default File Group.
- Custom Location. Browse to select an available data storage group in the publication repository database. For example, browse the available file groups.

Topic Data Retention Properties for a Hadoop Repository

The **Data Retention** page of the topic wizard includes the following properties for a Hadoop repository:

Retain published data for ... days after it is consumed

Determines how long Data Integration Hub retains published data in the publication repository after the data is consumed.

For each publication instance, the retention period starts after all subscribers have either successfully consumed the published data or discarded the data. That is, after all the events that are associated with the publication instance are either in a Complete or in a Discarded event status.

Note: The data retention period applies to the published data in the publication repository. Data Integration Hub does not delete the source data from the document store.

Do not delete published data for this topic

Data Integration Hub does not delete published data from the publication repository for the topic.

Topic Permissions Properties

The **Permissions** page of the topic wizard includes the following properties:

Available Categories

List of categories that you can assign to the topic.

Selected Categories

List of assigned categories for the topic.

Topic Summary Properties

The **Summary** page of the topic wizard includes the following properties:

Topic Name

Name of the topic.

Type

Topic type: Delta or Full.

Data Retention Period

Indicates how long Data Integration Hub retains the data in the Data Integration Hub publication repository after all subscribers consume it.

CHAPTER 7

Publications

This chapter includes the following topics:

- [Publications Overview, 68](#)
- [Publication Types, 68](#)
- [Publication Process, 74](#)
- [Publication Pre-process, 76](#)
- [Publication Sources, 77](#)
- [Publication Filters, 77](#)
- [Publication Schedules, 78](#)
- [Managing Publications, 79](#)

Publications Overview

Publications are entities that define how applications publish data to Data Integration Hub. Publication definitions can include the type, format, and schedule of data publication. Publications publish data to topics. Multiple publications can publish to the same topic. The topic defines the structure to which the data is published.

You create and edit publications in the Data Integration Hub Operation Console. During the configuration you choose the topic to which to publish the data. The steps that you take when you configure a publication depend on the publication type.

Publication Types

You can use an automatic publication, a custom publication, or a modular publication to publish data. The publication type you choose is based on the organization requirements and the complexity of the business logic that you want to apply when publishing the data.

Automatic Publications

With automatic publications you can publish data from a relational database source or a flat file source, or over a REST API. If required, Data Integration Hub generates the automatic mapping based on details that you provide when you create the publication.

You can use the following automatic publication types to publish data with Data Integration Hub:

- Automatic database publication. A publication with an automatic mapping and a relational database source.
- Automatic file publication. A publication with an automatic mapping and a flat file source.
- Automatic pass-through file publication. A publication with an automatic mapping and a pass-through file source.
- Automatic data-driven publication. A publication that transmits data over a REST API when the data is published, not based on a defined schedule. Data Integration Hub keeps the data as-is.

Custom Publications

With custom publications you can publish from any type of source that PowerCenter, Data Engineering Integration, Data Engineering Streaming, Data Quality, and Informatica Intelligent Cloud Services support.

You can use the following custom publication types to publish data with Data Integration Hub:

- Custom batch publication. A publication that uses a Data Integration Hub batch workflow. A batch workflow can be based on a PowerCenter batch workflow, a Data Engineering Integration mapping or workflow, or a Data Quality mapping or workflow.
- Custom multi-latency publication. A publication that uses a Data Integration Hub real-time workflow or a Data Integration Hub Data Engineering Streaming workflow. A Data Integration Hub real-time workflow is based on a PowerCenter real-time workflow. A Data Integration Hub Data Engineering Streaming workflow is based on a Data Engineering Streaming mapping.
- Custom cloud publication. A publication that uses an Informatica Intelligent Cloud Services task.

Modular Publication

A publication with an Informatica Intelligent Cloud Services mapping and a cloud source. With a modular cloud publication you can publish data from any cloud source that Informatica Intelligent Cloud Services supports. Data Integration Hub generates the publication mapping based on details that you provide when you create the publication, including the parameters values that you enter into the template that you select for publication processing.

Automatic Database Publication

Use this type of publication to publish data from a relational database source if you want Data Integration Hub to create the PowerCenter workflow for the publication.

When you create the publication, you select a relational database connection. Data Integration Hub explores the source connection, searches for the tables and fields that match the names of tables and fields that are defined in the topic structure, and performs one of the following actions:

- If all or some of the tables that are defined in the topic structure exist in the source connection, Data Integration Hub maps the source tables and fields to topic tables and fields, based on a name match. You can edit the mapped tables and fields and manually map them to other tables and fields in the topic.
- If none of the tables that are defined in the topic structure exist in the source connection, Data Integration Hub displays a warning and generates the mapping based on the topic tables. You can exclude fields or tables from the mapping, or manually map the tables to other tables in the topic.
- If the connection does not contain metadata access, Data Integration Hub cannot explore the source connection. Data Integration Hub displays a warning and generates the mapping based on the topic tables. You can exclude tables and fields from the mapping. You cannot edit the mapping.

Data Integration Hub generates the workflow based on the mapping. The workflow runs during the publication process. The publisher publishes only tables and fields that are mapped from the source to the topic. If the tables that are defined in the mapping do not exist in the source connection, you must add the required tables to the source before you run the publication. Otherwise, the publishing workflow fails during run time.

If you want to run a pre-process on the publication, you might be required to set pre-processing settings. Before you start creating the publication, get the required parameter settings.

The following pages appear in the publication wizard for this type of publication:

- General. Define basic publication properties.
- Processing. If you want to run a pre-process on the publication, select a publication pre-processing workflow. If the pre-process includes parameters, the parameters appear in this page. The content of the page depends on the workflow parameters that the developer defines. The developer imports the workflow to Data Integration Hub with the parameter definitions. The developer also determines the layout of this page. If required, set the values of the parameters.
- Source. Select the data access connection to the source from which Data Integration Hub reads the data. You can combine data from multiple source tables into a single topic table by creating a join or multiple joins. A join combines columns from two or more tables in the database. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the source database or in the topic.
- Field Mapping. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default. You can manually map source tables and fields to topic tables and fields.
- Filter. Define the data that the publication published by setting filter conditions on table columns.
- Schedule. Define the method and the frequency of data publishing. You can select to run the publication manually or by an external trigger, or you can select to publish the data according to a schedule that you define.
- Summary. Review the publication settings and save the publication.

Automatic Flat File Publication

Use this type of publication to publish data from a flat file source if you want Data Integration Hub to create the PowerCenter workflow for the publication.

For automatic publications with a flat file source, define the location of the source files, and then configure flat file sources for the tables in the topic that is associated with the publication. If you are publishing from a Hadoop Distributed File System (HDFS), select the HDFS connection. If you use the file transfer protocol to move the files, select the file transfer connection.

You must associate at least one topic table with a source file. Data Integration Hub maps the source fields to topic fields, based on a name match. You can edit the mapping and manually map source fields to topic fields.

Data Integration Hub creates the workflow with a mapping, based on the topic tables that are associated with a file, and runs the workflow during the publication process. Topic tables that are not associated with a file are not mapped.

Data Integration Hub deletes the files after it reads them. When you use file transfer, you can select whether or not Data Integration Hub deletes the files from the remote server after it reads them.

Note: When you publish files from HDFS, Data Integration Hub does not delete the files after it reads them. If required, you must delete the files yourself.

The following pages appear in the publication wizard for this type of publication:

If you want to run a pre-process on the publication, you might be required to set pre-processing settings. Before you start creating the publication, get the required parameter settings.

- **General.** Define basic publication properties.
- **Processing.** If you want to run a pre-process on the publication, select a publication pre-processing workflow. If the pre-process includes parameters, the parameters appear in this page. The content of the page depends on the workflow parameters that the developer defines. The developer imports the workflow to Data Integration Hub with the parameter definitions. The developer also determines the layout of this page. If required, set the values of the parameters.
- **Source.** Choose the source type. For an HDFS source or when you use file transfer, select the connection to the source from which Data Integration Hub reads the files. Enter the location of the file or files from where Data Integration Hub reads the data, and configure the format and the structure of flat file sources for the topic tables. You must configure a file source for at least one topic table.
- **Join.** You can pull data from multiple source files into a single topic table by creating joins. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the topic.
- **Field Mapping.** Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default. You can manually map source fields to topic fields.
- **Filter.** Define the data that the publication publishes by setting filter conditions on source columns.
- **Schedule.** Define the method and the frequency of data publishing. You can select to publish the data immediately after the published files are ready, to run the publication manually or by an external trigger, or to publish the data according to a schedule that you define.
If you select to publish the data immediately after the published files are ready, define the maximal period of time that Data Integration Hub waits before it publishes available files. When the maximal period of time ends, Data Integration Hub runs the publication even if not all the files are ready to be published.
- **Summary.** Review the publication settings and save the publication.

Automatic Pass-through File Publication

Use this type of publication to publish data from a file source that you want to keep as-is without loading the data into a relational database if you want Data Integration Hub to create the PowerCenter workflow for the publication. For example, to publish pass-through files such as PDF or .zip files. Data Integration Hub published the files as-is, without processing them.

For automatic publications with a pass-through file source, define the location of the source files, and then define the expected file names or the expected file name patterns for the tables in the topic that is associated with the publication. If you use the file transfer protocol to move the files, select the file transfer connection.

You must associate at least one topic table with a source file. Data Integration Hub creates the workflow with a mapping, based on the topic tables that are associated with a file, and runs the workflow during the publication process. Topic tables that are not associated with a file are not mapped.

Data Integration Hub deletes the files after it reads them. When you use file transfer, you can select whether or not Data Integration Hub deletes the files from the remote server after it reads them.

If you want to run a pre-process on the publication, you might be required to set pre-processing settings. Before you start creating the publication, get the required parameter settings.

The following pages appear in the publication wizard for this type of publication:

- **General.** Define basic publication properties.
- **Processing.** If you want to run a pre-process on the publication, select a publication pre-processing workflow. If the pre-process includes parameters, the parameters appear in this page. The content of the

page depends on the workflow parameters that the developer defines. The developer imports the workflow to Data Integration Hub with the parameter definitions. The developer also determines the layout of this page. If required, set the values of the parameters.

- **Source.** Enter the location and name of the file or files that Data Integration Hub publishes. You must configure a file source for at least one topic table. When you use file transfer, select the connection to the source from which Data Integration Hub reads the files.
- **Schedule.** Define the method and the frequency of data publishing. You can select to publish the data immediately after the published files are ready, to run the publication manually or by an external trigger, or to publish the data according to a schedule that you define.
If you select to publish the data immediately after the published files are ready, define the maximal period of time that Data Integration Hub waits before it publishes available files. When the maximal period of time ends, Data Integration Hub runs the publication even if not all the files are ready to be published.
- **Summary.** Review the publication settings and save the publication.

Automatic Data-driven Publication

Use this type of publication to publish data that you want to keep as-is over a REST API, so that Data Integration Hub runs the publication as data arrives without a predefined schedule. Data Integration Hub publishes the data as-is, without processing it.

For automatic data-driven publications, you configure publication properties and select the topic into which to publish the data. You then create a POST request to run the publication.

You configure a data-driven publication in the **Create Automatic Data-driven Publication** page.

Note: When you use an automatic data-driven publication, you cannot publish topics to a Hadoop repository.

Custom Batch Publication

Use this type of publication to publish data when you want to use a Data Integration Hub batch workflow that the developer created in the Operation Console.

A Data Integration Hub workflow is based on a PowerCenter batch workflow, a Data Engineering Integration mapping or workflow, or a Data Quality mapping or workflow. Batch workflows run according to a trigger and not continuously.

Note: To use file transfer to publish files from remote servers you must run a Data Integration Hub batch workflow. You cannot run a Data Integration Hub real-time workflow for file transfer.

If the workflow includes parameters, you might be required to set publication or publication pre-processing parameters. Before you start creating the publication, get the required parameter values.

The following pages appear in the publication wizard for this type of publication:

- **General.** Define basic publication properties.
- **Processing.** Select the publication mapping. If you use file transfer, select the connection to the source from which Data Integration Hub reads the files. If running a pre-process on the publication is enabled and you want to run a pre-process, select a publication pre-processing workflow.
The content of this page depends on the PowerCenter workflow, Data Engineering Integration mapping or workflow, or Data Quality mapping or workflow parameters that the developer defines. The developer imports the workflow or the mapping to Data Integration Hub with the parameter definitions. The developer also determines the layout of the page. If required, set the values of the parameters.

- **Schedule.** Define the method and the frequency of data publishing. You can select to run the publication manually or by an external trigger, or you can select to publish the data according to a schedule that you define.
- **Summary.** Review the publication settings and save the publication.

Custom Multi-latency Publication

Use a custom multi-latency publication when you want to publish data at a low frequency.

To create a custom multi-latency publication, you can use the following workflows that the developer creates in the Data Integration Hub Operation Console:

- Data Integration Hub real-time workflow
- Data Integration Hub Data Engineering Streaming workflow

A Data Integration Hub real-time workflow is based on a PowerCenter real-time workflow. A real-time workflow continuously processes data from real-time sources. They read data from real-time sources and write the data to Data Integration Hub. The real-time sources can be Apache Kafka streaming applications, web-service providers, or Java Message Service (JMS) queues.

A Data Integration Hub Data Engineering Streaming workflow is based on a Data Engineering Streaming mapping.

The following pages appear in the publication wizard for this type of publication:

- **General.** Define basic publication properties.
- **Processing.** Select the publication workflow. To run a pre process on the publication, select a publication pre processing workflow.
- **Schedule.** Define the time interval according to which Data Integration Hub groups the published data and publishes the data.
- **Summary.** Review the publication settings and save the publication.

Custom Cloud Publication

Use this type of publication to publish data from a cloud source with an Informatica Intelligent Cloud Services task that the developer created in Data Integration.

Data Integration Hub uses a Cloud REST API to trigger the task, and the Data Integration Hub cloud connector writes the published data to the Data Integration Hub publication repository.

The following pages appear in the publication wizard for this type of publication:

- **General.** Define basic publication properties.
- **Processing.** Select an Informatica Cloud task.
- **Schedule.** Define the method and the frequency of data publishing. You can select to run the publication manually or by an external trigger, or you can select to publish the data according to a schedule that you define.
- **Summary.** Review the publication settings and save the publication.

Modular Cloud Publication

Use this type of publication to publish data from a cloud source with an Informatica Intelligent Cloud Services mapping.

For modular publications with a cloud source, the developer creates the mapping in Data Integration. You select the mapping when you create the publication in Data Integration Hub.

If the mappings includes parameterized transformations, you enter values for the parameters in the publication wizard.

If the target connection is parameterized in the mapping, you select the Data Integration Hub publication repository as the publication target in the publication wizard.

The following pages appear in the publication wizard for this type of publication:

- General. Define basic publication properties.
- Processing. Select the publication mapping.
- Source. If the publication source is not configured by the mapping, configure the publication source from which Data Integration Hub reads the data.
- Target. If the publication target is not configured by the mapping, configure the target on the Data Integration Hub publication repository to which Data Integration Hub writes the data.
- Input Parameters. If the mapping includes parameterized transformations, set parameter values.
- Field Mapping. If required, map source table fields to topic table fields.
- Schedule. Define the method and the frequency of data publishing. You can select to run the publication manually or by an external trigger, or you can select to publish the data according to a schedule that you define.
- Summary. Review the publication settings and save the publication.

Publication Process

The publication process includes retrieving the data from the publisher, running any associated mappers, such as a mapping or a task, and writing the data to the relevant topic in the Data Integration Hub publication repository. After the publication process ends, subscribers can consume the published data from the publication repository.

The publication process depends on the publication type.

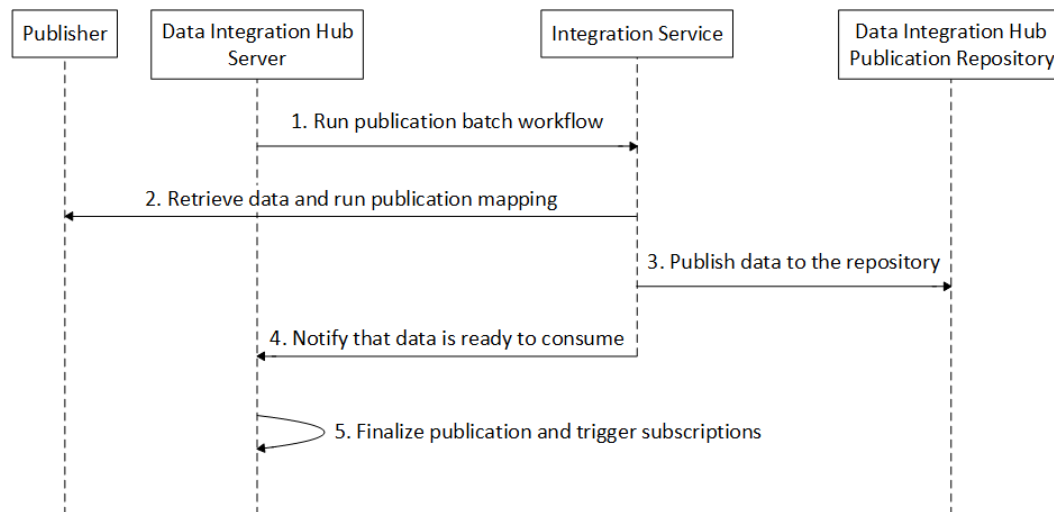
- Automatic publications can run a Data Integration Hub workflow that is based on a PowerCenter batch workflow or run over a REST API.
- Custom publications can either run a Data Integration Hub workflow that is based on a PowerCenter batch workflow, PowerCenter real-time workflow, Data Engineering Integration mapping or workflow, Data Engineering Streaming mapping, or Data Quality mapping or workflow, or run an Informatica Intelligent Cloud Services task.
- Modular publications run an Informatica Intelligent Cloud Services mapping.

Publication Process with a Batch Workflow

The publication process for publications that run a Data Integration Hub batch workflow includes the following stages:

1. When the publisher is ready to publish the data, the Data Integration Hub server runs the publication batch workflow and sends a request to the relevant Integration Service, either the PowerCenter Integration Service or the Data Integration Service.
2. The Integration Service extracts the data from the publisher and runs the automatic or custom mapping on the data.
3. The Integration Service writes the data to the Data Integration Hub publication repository.
4. The Integration Service notifies the Data Integration Hub server that the published data is ready for subscribers.
5. The Data Integration Hub server changes the status of the publication event to complete and triggers subscription processing.

The following image shows the main stages of the publication process for publications that run a batch workflow:



Publication Process with a Real-time Workflow

The publication process for publications that run a Data Integration Hub real-time workflow includes the following stages:

1. The developer runs the real-time workflow. The workflow writes the data to the relevant tables in the Data Integration Hub publication repository.
2. The Data Integration Hub server triggers a scheduled process and checks for new data in the relevant tables in the Data Integration Hub publication repository.
3. If new data is found, Data Integration Hub updates the publication ID and the publication date of the data to indicate that the data is ready for consumption and creates a publication event in the Data Integration Hub repository.
4. The Data Integration Hub server changes the status of the publication event to complete and triggers subscription processing.

Publication Process with a Data Integration Task

The publication process for publications that run a Data Integration task includes the following stages:

1. When the publication is triggered, either according to schedule or by an external API, the Data Integration Hub server triggers the Data Integration task that is defined for the publication through an Informatica Intelligent Cloud Services REST API.
2. The publication process uses the Data Integration Hub cloud connector to write the data to Data Integration Hub.
3. The Data Integration Hub server changes the status of the publication event to complete and triggers subscription processing.

Publication Process of a Data-driven Publication

The publication process for data-driven publications includes the following stages:

1. After you create a data-driven publication, you create a POST request to run the publication.
2. When you post the request, Data Integration Hub transfers published data from the request directly to the Data Integration Hub publication repository, to the topic that you define in the publication.
3. Data Integration Hub creates a Data-driven Publication event, based on the event grouping that is defined for the publication:
 - If the grouping time is set to zero, that is, no grouping is defined for the publication, Data Integration Hub creates an event each time data is published to the publication repository.
 - If you define a grouping time, Data Integration Hub creates an event at the end of each grouping period that contains publications. For example, if you configure the publication to group publications every ten seconds, Data Integration Hub creates an event every ten seconds, providing that data was published to the publication repository during the 10-second period.

Publication Pre-process

For publications with a mapping that is based on a PowerCenter workflow, you can use a publication pre-process to prepare the data before you publish it.

When a preparation process has to run on the data that you want to publish before the source application publishes the data to Data Integration Hub, Data Integration Hub can trigger a preparation process before it runs the publication, so that the data that it publishes is always up-to-date.

When you configure a publication to run a publication pre-process, Data Integration Hub triggers the pre-process before it runs the publication. After the pre-process runs, Data Integration Hub runs the publication and publishes the pre-processed data.

The developer creates a PowerCenter workflow that runs the pre-process. You select the pre-process workflow and define pre-processing parameters in the Publication wizard. You can select a pre-process for both publications that use automatic mapping and publications that use custom mapping.

Publication Sources

Automatic publications can publish from a relational database source, from a flat file source, from a pass-through file source, or over a REST API.

Custom publications can publish from any type of source that PowerCenter, Data Engineering Integration, Data Engineering Streaming, Data Quality, or Informatica Intelligent Cloud Services support.

Modular publications can publish from any type of source that Informatica Intelligent Cloud Services supports.

Relational Sources

For publications that use an automatic mapping with a relational source, Data Integration Hub supports Microsoft SQL Server, Oracle, IBM DB2, PostgreSQL and Teradata relational databases as publication sources.

Flat File Sources

For publications that use an automatic mapping with a flat file source, Data Integration Hub supports only delimited, flat files as publication sources.

You publish flat files from file systems, including Hadoop Distributed File Systems (HDFS). When you publish files from HDFS, Data Integration Hub does not delete the files after it reads them. If required, you must delete the files yourself.

You can use file transfer to publish flat files from applications that reside outside the Data Integration Hub network. You cannot use file transfer to publish files from HDFS. When you use file transfer, you can select whether or not Data Integration Hub deletes the files from the remote server after it reads them.

Pass-through File Sources

For publications that use an automatic mapping with a pass-through file source, Data Integration Hub publishes the files without processing them and stores them on a file publication repository. For example, unstructured files such as PDF or .zip files.

You can use file transfer to publish flat files from applications that reside outside the Data Integration Hub network. When you use file transfer, you can select whether or not Data Integration Hub deletes the files from the remote server after it reads them.

You cannot publish pass-through files from Hadoop Distributed File Systems (HDFS).

Publication Filters

To define the data that an automatic relational database or flat file publication publishes, add a filter to the publication. When you add a filter to a publication, you can define filter conditions to table rows.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2016.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub publishes only the rows that meet all of the conditions.

Publication Schedules

The publication schedule defines the frequency of the publication for publications that are triggered by Data Integration Hub.

When you define a schedule for a publication, you can select to publish the data manually or by an external trigger, at defined intervals, or after a specific publication or subscription run completes. For publications with a flat file source or with a pass-through file source, you can also select to run the publication immediately after the published files are ready.

Note: If you use file transfer, you can run the publication immediately after the published files are ready only if Data Integration Hub deletes the files after it reads them. This ensures that files are not read multiple times.

For file publications that are published manually, by an external trigger, or at defined intervals, and that publish multiple files, all the files must be present in the source location when the publication starts.

For publications with a defined schedule, the publication starts when one of the following conditions is true:

- The scheduled start time arrives.
- You run the publication manually.
- A publication pre-process starts the publication process.
- The developer starts the publication from a command-line API or from a REST API.
- The developer starts the publication from PowerCenter with the DX_StartPublication transformation.
- The running of the publication or subscription after which the current publication is configured to run is complete.

For automatic file publications, you can define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the publication in case of failure, and the retry interval. The policy does not apply to publications that you run manually.

Data-driven publications and multi-latency publications are not triggered by Data Integration Hub. For publications of those types you define the time intervals according to which Data Integration Hub groups the published data.

Managing Publications

Use the Navigator to create, search, duplicate, re-validate, edit, or delete publications. In the Navigator, you can also enable and disable publications, and run publications manually.

Creating a Publication

Use the Navigator to create publications. Ensure that the following conditions are met before you create publication:

- The application from which you want to publish data must exist. You can either use an existing application, or create and save a new application.
- The topic to which you want to publish data must exist. You can either use an existing topic, or create and save a new topic.
- For a custom publication, the Data Integration Hub workflow task must exist.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the application from which you want to publish data.
The **Application Details** page appears.
3. Click the **Publications** tab.
The publication list appears.
4. Click **New**, select the type of publication to create, and then click **Create**.
The create publication wizard or page for the selected publication type opens.
5. Enter the required details in each tab, and click **Next**.
6. Click **Finish**.

Searching for a Publication

To search for a publication in the **Applications** page, type the name of the publication in the **Find** text box and click **Search**. Publications that match the search text that you entered appear. Perform the following tasks to search for a publication:

- Click **Reload Default Search Results** to clear search results and display the list of publications.
- Click **Refresh Search Results** to refresh search results.

Duplicating a Publication

Click the **Duplicate** icon next to the Publication to create a publication with a different name and identical properties. Perform the following tasks to duplicate a publication:

1. On the **Publications** tab, click the **Duplicate** icon next to the publication.
The **Confirmation** widow appears.
2. Rename the duplicate publication as required, and click **Save**.
If you don't rename the duplicate publication, Data Integration Hub saves it with the default name format, `<name>_Copy`.

Editing a Publication

When you open a publication with running subscriptions, the details appear in read-only mode. If you open a publication with scheduled subscribers that did not start running, you can choose to open the publication in read-only mode or discard unconsumed data and edit the publication. Perform the following tasks to edit a publication:

1. On the **Publications** tab, click the **Edit** icon next to the publication name.
The **Edit** page appears.
2. Change the required properties in the publication wizard or page and click **Save**.

Deleting a Publication

To delete a publication with associated subscriptions, you must delete the subscriptions before you delete the publication. You cannot delete a publication with associated subscribers. Perform the following tasks to delete a publication:

1. On the **Publications** tab, click the **Delete** icon next to the publication name.
A confirmation message appears.
2. Click **Ok**.

Enabling and Disabling a Publication

On the **Publications** tab, perform the following tasks to enable or disable publications:

- To disable a publication, select **Disabled** from the **Mode** column.
- To enable a disabled publication, select **Enabled** from the **Mode** column.
- To disable all publications, select **Actions > Disable All**.
- To enable all publications, select **Actions > Enable All**.

Running a Publication Manually

Perform the following tasks to run a publication manually:

1. On the **Publications** tab, click the **Run** icon next to the publication.
A confirmation message appears.
2. Click **Yes**.
After you click **Finish**, Data Integration Hub refreshes the publication properties and the publication status icon shows that the publication is valid.

Re-validating a Publication

When the publication status icon shows that the publication is not valid, re-validate a publication by opening the publication wizard in edit mode, following the wizard steps, and verifying the settings. If the publication became not valid due to a change in the topic that does not impact the publication, you might not have to change any setting. Completing the wizard re-generates the publication mapping. Verify that all associated connections, entities, and mappings are valid before you re-validate the publication.

CHAPTER 8

Creating Publications

This chapter includes the following topics:

- [Creating Publications Overview, 81](#)
- [Creating an Automatic Relational Database Publication, 81](#)
- [Creating an Automatic Flat File Publication, 86](#)
- [Creating an Automatic Pass-through File Publication, 93](#)
- [Creating an Automatic Data-driven Publication, 95](#)
- [Creating a Custom Batch Publication, 97](#)
- [Creating a Custom Multi-latency Publication, 99](#)
- [Creating a Custom Cloud Publication, 100](#)
- [Creating a Modular Cloud Publication, 102](#)
- [Creating a Custom Pass-through Kafka Publication, 106](#)

Creating Publications Overview

To create a publication, in the Data Integration Hub Operation Console, use the create publication wizard or page that is applicable to the type of publication that you want to create.

Each type of publication wizard or page contains the properties that are relevant to the publication type.

Creating an Automatic Relational Database Publication

To create an automatic publication with a relational database source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication source. Optionally, define joins for the publication.
4. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default.
5. Optionally, define a filter for the publication.

6. Define the publication schedule.
7. Review the publication settings and save the publication.

Step 1. Access the Create Automatic Database Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Relational Database**, and then click **Create**.
The **Create Automatic Database Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.
5. Click **Next**.
The **Source** page appears.

Step 3. Select Publication Source

Select the publication source from which Data Integration Hub reads data in the **Source** page of the publication wizard.

1. Select a connection from the **Database Connection** list.
The **Define new joins by joining existing tables** section appears.
2. Optionally, create joins for the publication. A join combines columns from two or more tables in the database. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the source database or in the topic.

To create joins, perform the following actions:

- a. By default, the **Source** page shows the default schema only and the default schema is selected for both the left and right tables. To select non default schemas, clear the option **Show default schema only** and then select schemas for the left right tables.
Note: When you select non default schemas, do not use two source tables with the same name from different schemas to define the join. The source table names that you use to create the join must be unique across the schemas.
 - b. Select the tables to join from the **Select Left Table** and **Select Right Table** lists.
 - c. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a topic table.
 - d. Select the column to join from the left table and then select the column to join from the right table. To add more columns to the join click the plus sign and then select the columns to join.
 - e. Click **Create Join**.
The new join appears on the **Source** page.
 - f. Repeat step [2](#) to create as many joins as required.
3. Click **Next**.
The **Field Mapping** page appears.

Step 4. View and Edit Publication Field Mapping

View and edit the mapping of source tables and fields to topic tables and fields on the **Field Mapping** page of the publication wizard. If you created joins in "Step 3. Select Publication Source", configure field mapping of joins and fields to topic tables and fields.

1. Perform the following actions to map a join to a topic table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Click the row of the join that you want to map and then click **Map Source Table**.
The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the join you mapped in the **Mapped Source Table** list.
 - c. Go to Step [3](#) to map the fields in the join to fields in the topic table.
2. Perform the following actions to view and edit the mapping of a source table to a topic table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Source Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and to show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.

- You can search for a substring.

The **Search Results** section displays the source tables that match the search string and the schema that each table uses. The name of the topic table for which you are editing the mapping shows in the row of the source table that is mapped to the topic table, in the **Mapped To** column.

- To map a different source table to the topic table, click the row of the source table that you want to map and then click **Map Source Table**.

The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the revised mapping.

- Perform the following actions to configure, view, and edit the mapping of fields in a source table or in a join to fields in a topic table:

- Click the edit field mapping icon.

The **Edit Field Mapping** dialog box appears.

- To search for fields that are used in the topic table and in the source table, enter a string for the field name in **Find Field**, and then click **Search**. The search is not case sensitive. You can search for a substring.

Choose **Display unmapped fields only** to show only unmapped fields in the search results.

Source and topic fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.

- To map a source table field to a topic table field, drag the field from the **Source Table** section to the **Topic Table** section.
- To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
- To clear the mapping of a field, in the **Actions** column, click the clear icon.
- To view field details, rest on the details icon to the left of a field name.
- To revert to the default Data Integration Hub field mapping, click **Auto Map**.
- To clear all field mappings, click **Clear All**.
- Click **OK** to map the fields.

The **Edit Field Mapping** dialog box closes.

- To revert to the default Data Integration Hub table and field mapping, click **Auto Map** in the **Field Mapping** page.

- To clear all table and field mapping, click **Clear All** in the **Field Mapping** page.

Note: The publication must contain at least one mapped source table.

- In the **Field Mapping** page, click **Next**.

The **Filter** page appears.

Step 5. Define a Filter

To define the data that an automatic relational database or flat file publication publishes, add a filter to the publication. When you add a filter to a publication, you can define filter conditions to table rows.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2016.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub publishes only the rows that meet all of the conditions.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator.
 - d. Click the plus icon to the right of the expression line.The condition shows in the condition list.
5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.The condition shows in the condition list.
6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 6. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 7. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating an Automatic Flat File Publication

To create an automatic publication with a flat file source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select and configure the publication source.
4. Optionally, define joins for the publication.
5. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default.
6. Optionally, define a filter for the publication.
7. Define the publication schedule.
8. Review the publication settings and save the publication.

Step 1. Access the Create Automatic File Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Flat File**, and then click **Create**.
The **Create Automatic File Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.
5. Click **Next**.
The **Source** page appears.

Step 3. Select and Configure Publication Source

Select the type and the location of the publication source, from where Data Integration Hub picks up the files, in the **Source** page of the publication wizard.

1. Select the source type from the **Source Type** list.
If you select the HDFS source type, the **HDFS Connection** field appears on the page.
2. If you select the HDFS source type, select a connection from the **HDFS Connection** list.
3. To use file transfer to publish files from remote servers, select **Use File Transfer**.
The **Connection** and **Delete Files from Source After Pickup** fields appear on the page.
4. If you selected to use file transfer, configure the following file transfer properties:
Connection
Select a connection from the list.
Delete Files from Source After Pickup
Data Integration Hub deletes the files after it reads them. You must have permissions on the remote server to delete the files.
5. Enter the location of the files that contain the source data in the **Directory** field. If you use file transfer, you cannot use patterns to define the directory path. You can, however, use patterns to define the file name.
6. Click **Configure**.

The **Configure Flat File Source** dialog box appears.

7. In the **Configure Flat File Source** dialog box configure the following properties and then click **OK**:

Logical Name

Description name of the source. The name identifies the source when you configure field mapping for the publication.

File Name

Name of the file that Data Integration Hub picks up and publishes to the publication repository. The file name can include asterisk wildcards and variables. Enter variables in the following format: `($pattern)`. For example: `Input_($PublicationName)_($sequence).in`.

Configure source based on

Select one of the following options:

- **Topic Table**. Select the topic table that represents the structure of the flat file.
- **Sample file**. A sample file that represents the structure of the flat file. Browse to select and upload the file.

File Format

Expand the **File Format** area to view and edit the format of the file. Changing the file format might affect the column structure.

The **File Format** area can include the following properties:

Code page

Character encoding used in the file.

Number of initial rows to skip

Number of rows at the beginning of the file that Data Integration Hub ignores when it reads the file. Data Integration Hub published only the subsequent rows.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

Thousands Separator

Optional. Symbol used in the file as a thousands separator.

Decimal Separator

Symbol used in the file as a decimal separator.

Datetime Format

Date and time format used in the file. Select a predefined format or select **Other** to define a custom format. For information about the supported datetime formats, see the *PowerCenter Transformation Language Reference*.

Note: The datetime format can contain up to 50 characters.

Column Structure

The column structure depends on the file format settings that you configure in the **File Format** area. The table must contain at least one column.

If you provide a sample file, Data Integration Hub reads the file according to the file format settings and presents the columns that it detects in the sample file. Use the **Add Column** button, the up and down arrows, and the remove icon to add, order, and remove table columns.

Click the edit icon to the right of a column to edit column parameters.

Each column must contain the following parameters:

Column Name

Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.

If you select the option **Import column names**, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1**, **Field2**, **Fieldn**.

Data Type

Select from the list of available data types. By default, Data Integration Hub reads the data as string.

Precision

Enabled only for data types that support precision.

Scale

Enabled only for data types that support data scaling.

Sample File Preview

If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file. The **Preview** area presents the data in the structure that Data Integration Hub applies when it distributes the data into topic fields in the topic table, according to the **File Format** parameters.

The **Configure Flat File Source** dialog box closes.

8. Repeat steps [6](#) through [7](#) for each table topic that appears on the **Source** page.

9. In the **Source** page click **Next**.

The **Join** page appears.

Step 4. Define Joins

Pull data from multiple source files into a single topic table by creating a join. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the topic.

1. Optionally, create joins for the publication. To create a join, perform the following actions:
 - a. Select the files to join from the **Select Left File** and **Select Right File** lists.
 - b. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a topic table. The new join appears on the **Join** page.

- c. Select the column to join from the left file and then select the column to join from the right file. To add more columns to the join click the plus sign and then select the columns to join.
 - d. Click **Create Join**.
2. Repeat step [1](#) to create as many joins as required.
3. Click **Next**.
The **Field Mapping** page appears.

Step 5. View and Edit Publication Field Mapping

View the mapping of source tables and fields to topic tables and fields and edit field mapping on the **Field Mapping** page of the publication wizard. If you created joins in "Step 4. Define Joins", configure field mapping of joins and fields to topic tables and fields.

1. Perform the following actions to map a source table or a join to a topic table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Click the row of the table or join that you want to map and then click **Map Source Table**.
The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the table you mapped in the **Mapped Source Logical Name** list.
2. Perform the following actions to configure, view, and edit the mapping of fields in a source table or in a join to fields in a topic table:
 - a. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears. Data Integration Hub maps the fields automatically, according to name match.
 - b. To search for fields that are used in the topic table and the source table, enter a string for the field name in **Find Field**, and then click **Search**. The search is not case sensitive. You can search for a substring.
Choose **Display unmapped fields only** to show only unmapped fields in the search results.
Source and topic fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.
 - c. To map a source table field to a topic table field, drag the field from the **Source Table** section to the **Topic Table** section.
 - d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
 - e. To clear the mapping of a field, in the **Actions** column, click the clear icon.
 - f. To view field details, rest on the details icon to the left of a field name.
 - g. To revert to the default Data Integration Hub field mapping, click **Auto Map**.
3. Click **OK** to apply the mapping.
The **Edit Field Mapping** dialog box closes.
4. To clear all table and field mapping, click **Clear All** in the **Field Mapping** page.
Note: The publication must contain at least one mapped source table.
5. In the **Field Mapping** page, click **Next**.
The **Filter** page appears.

Step 6. Define a Filter

To define the data that an automatic relational database or flat file publication publishes, add a filter to the publication. When you add a filter to a publication, you can define filter conditions to table rows.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2016.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub publishes only the rows that meet all of the conditions.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
 - d. Click the plus icon to the right of the expression line.The condition shows in the condition list.
5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.The condition shows in the condition list.
6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 7. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

When the file is ready to be published

Runs the publication after the published files are ready, the next time it polls the remote sources. Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the publication in case of failure, and the retry interval. The policy does not apply to publications that you run manually.
3. Click **Next**.

The **Summary** page appears.

Step 8. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating an Automatic Pass-through File Publication

To create an automatic publication with a pass-through file source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Configure the publication source.
4. Define the publication schedule.
5. Review the publication settings and save the publication.

Step 1. Access the Create Automatic Pass-through File Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Pass-through File**, and then click **Create**.
The **Create Automatic Pass-through File Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.
5. Click **Next**.
The **Source** page appears.

Step 3. Configure Publication Source

Configure the source from which Data Integration Hub publishes files in the **Source** page of the publication wizard.

1. To use file transfer to publish files from remote servers, select **Use File Transfer**.
The **Connection** and **Delete Files from Source After Pickup** fields appear on the page.
2. If you selected to use file transfer, configure the following file transfer properties:

Connection

Select a connection from the list.

Delete Files from Source After Pickup

Data Integration Hub deletes the files after it reads them. You must have permissions on the remote server to delete the files.

3. Enter the location of the file or files that Data Integration Hub publishes in the **Directory** field. If you use file transfer, you cannot use patterns to define the directory path. You can, however, use patterns to define the file name.
4. Click the edit icon next to a table topic and enter the name of the source file in the **File Name** column. The file name can include asterisk wildcards and variables. Enter variables in the following format: `($pattern)`. For example: `Input_($PublicationName)_($sequence).in`.
5. Repeat step 4 for each table topic that appears on the **Source** page.
6. Click **Next**.

The **Schedule** page appears.

Step 4. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

When the file is ready to be published

Runs the publication after the published files are ready, the next time it polls the remote sources. Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the publication in case of failure, and the retry interval. The policy does not apply to publications that you run manually.
3. Click **Next**.

The **Summary** page appears.

Step 5. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating an Automatic Data-driven Publication

To create an automatic data-driven publication, perform the following tasks:

1. Access the create publication page.
2. Configure publication properties and select the topic into which to publish the data.
After you configure the publication properties, you can copy the following URLs from the create publication page:
 - REST URL of the API. Use this URL to publish the data.
 - REST URL of the API Swagger file with the Swagger structure for the topic into which the publication publishes data. Use the structure in the publication request.You use the URLs when you create the request that runs the publication.
3. Create a request to run the publication.

Step 1. Access the Create Automatic Data-driven Publication page

Access the create publication page in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Automatic > Data-driven**, and then click **Create**.
The **Create Automatic Data-driven Publication** page appears.

Step 2. Define Publication Properties

Define publication properties in the create publication page.

1. Enter the publication name.

The following URLs are automatically updated:

- REST URL of the API. Use this URL to publish the data.
- REST URL of the API Swagger file with the Swagger structure for the topic into which the publication publishes data. Use the structure in the publication request.

You use the URLs when you create the request that runs the publication.

2. Optionally, enter a description of the publication.

3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Optionally, define the time interval according to which Data Integration Hub groups the published data.

- If no grouping is defined, that is, **Group publications every ... seconds** is set to zero, Data Integration Hub does not group the published data. Data Integration Hub writes the data to the publication repository as it is being published.
- If you define a grouping time interval, Data Integration Hub groups the published data and writes it to the publication repository based on the time interval. For example, if you define **Group publications every 10 seconds**, Data Integration Hub will group the published data at the end of each 10-second period.

5. Click **Save**.

Step 3. Create a Request to Run the Publication

To create a request to run the publication, you have to copy the URL of the REST API and the URL of the API Swagger file from the publication that you created in [“Step 2. Define Publication Properties” on page 96](#).

- Create a POST request with the following details:

- Request URL: REST API URL. For example:

```
http://hostname:18080/dih-console/api/v1/publication/MyPublication/data
```

Where `MyPublication` is the publication name.

- Request body: Use the structure definition in the Swagger structure for the topic into which the publication publishes data to prepare the body of the request. Access the REST URL of the API Swagger file to view and copy the structure, for example:

```
http://hostname:18080/dih-console/api/v1/publication/MyPublication/data?swagger
```

Where `MyPublication` is the publication name.

The following is an example of request payload for a topic that contains two topic tables, each containing two columns:

```
{
  "Employee": [
    {
      "EmployeeName": "John Smith"
    },
    {
      "EmployeeId": "AA18"
    }
  ]
}
```



```

    ],
    "Department": [
      {
        "DepartmentName": "Computer Science"
      },
      {
        "DepartmentId": "Dep13"
      }
    ]
  }
}

```

Creating a Custom Batch Publication

To create a custom publication with a batch workflow, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication workflow. If you use file transfer, select the connection to the source from which Data Integration Hub reads the files and choose whether Data Integration Hub deletes files from the source after pickup.
4. If you want to run a pre-process on the publication, select the pre-processing mapping.
5. If the workflow includes parameters, set the publication parameters.
6. Define the publication schedule.
7. Review the publication settings and save the publication.

Step 1. Access the Create Custom Batch Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Custom > Batch**, and then click **Create**.
The **Create Custom Batch Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.

Step 3. Select Publication Workflow

Select a Data Integration Hub publication workflow in the **Processing** page of the publication wizard. If required, set parameter values.

1. Select a workflow from the **Custom Mapping** list. Only publication workflows that are based on a batch workflow are available for selection.
If the workflow supports file transfer, the **Use File Transfer** option is enabled.
If the workflow contains parameters, the parameters are shown in the **Publication Parameters** area.
2. To use file transfer, select the **Use File Transfer** option and then perform the following actions:
 - a. Select the connection to the source from which Data Integration Hub reads the files in the **Connection** field.
 - b. To prevent Data Integration Hub from deleting files from the remote server after reading the files, clear the option **Delete Files from Source After Pickup**.
3. If you want to run a pre-process on the publication, select a publication pre-processing workflow from the **Pre-Processing Mapping** list.
4. If the workflow includes parameters, set the values of the parameters in the **Publication Parameters** area.
5. Click **Next**.
The **Schedule** page appears.

Step 4. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 5. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Custom Multi-latency Publication

To create a multi-latency publication that is based on a PowerCenter real-time workflow or a Data Engineering Streaming real-time mapping, perform the following tasks:

Note: You can only use a Data Engineering Steam real-time mapping if Data Integration Hub is installed on a UNIX operating system and the Data Integration Hub Data Engineering Integration / Informatica Data Quality component is installed on your system.

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication workflow or mapping.
4. Define the publication intervals.
5. Review the publication settings and save the publication.

Step 1. Access the Create Custom Multi-latency Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Custom > Multi-latency**, and then click **Create**.
The **Create Custom Multi-latency Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Click **Next**.

The **Processing** page appears.

Step 3. Select Publication Workflow

Select the publication workflow in the **Processing** page of the publication wizard.

1. Select a workflow from the **Custom Mapping** list. Only real-time publication and Data Engineering Streaming workflows are available for selection.
2. If required, enter workflow parameters. Click **Next**.

The **Schedule** page appears.

Step 4. Define Publication Intervals

Define the time interval according to which Data Integration Hub groups the published data and publishes the data in the **Schedule** page of the publication wizard.

1. Select the publication interval in **Publish every**. You can set a value of 10 seconds up to 59 minutes and 50 second, with intervals of 10 seconds.
2. Click **Next**.

The **Summary** page appears.

Step 5. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Custom Cloud Publication

To create a custom publication with an Informatica Intelligent Cloud Services task, perform the following tasks:

1. Access the create publication wizard.

2. Define basic publication properties.
3. Select the Informatica Intelligent Cloud Services task that runs the publication mapping.
4. Define the publication schedule.
5. Review the publication settings and save the publication.

Step 1. Access the Custom Cloud Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management** > **Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Custom** > **Cloud**, and then click **Create**.
The **Create Custom Cloud Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.

Step 3. Select Publication Mapping

Select an Informatica Cloud task that defines the publication mapping in the **Processing** page of the publication wizard.

1. Select a task from the **Informatica Cloud Task** list.
2. Click **Next**.
The **Schedule** page appears.

Step 4. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.

- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 5. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Modular Cloud Publication

To create a modular publication with a cloud source, perform the following tasks:

1. Access the create publication wizard.
2. Define basic publication properties.
3. Select the publication mapping.
4. If required, configure the publication source.
5. If required, configure the publication target.
6. If the mapping includes parameterized transformations, set the parameter values.

7. If required, configure field mapping.
8. Define the publication schedule.
9. Review the publication settings and save the publication.

1. Access the Create Modular Cloud Publication Wizard

Access the create publication wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the **Publications** tab. Click **New**, then select **Modular > Cloud**, and then click **Create**.
The **Create Modular Cloud Publication** wizard appears.

Step 2. Define Basic Publication Properties

Define publication properties in the **General** page of the publication wizard.

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.

Step 3. Select Publication Mapping

Select an Informatica Cloud publication mapping in the **Processing** page of the publication wizard.

1. Select a mapping from the **Cloud Mapping** list.
2. Click **Next**.
The **Source** page appears.

Step 4. Configure Publication Source

If the publication source is not configured by the mapping, configure the source in the **Source** page of the publication wizard.

1. Configure source settings as applicable.
2. Click **Next**.
The **Target** page appears.

Step 5. Configure Publication Target

If the publication target is not configured by the mapping, configure the target in the **Target** page of the publication wizard. The target of the publication is the Data Integration Hub publication repository.

1. Configure target settings as applicable.
2. Click **Next**.

The **Input Parameters** page appears.

Step 6. Set Publication Parameters

If the mapping contains parameters, the parameters show in the **Input Parameters** page of the publication wizard. Set parameter values as applicable.

1. Click the edit icon next to the parameter for which to define a value.
2. In the **Edit Parameter** dialog box, define the parameter value in the **Expression** area. Click a field in the **Fields** area to add it to the expression.

Click **OK**.

The parameter value shows in the **Input Parameters** page.

3. Repeat steps [1](#) through [2](#) to set the required parameter values.
4. Click **Next**.

The **Field Mapping** page appears.

Step 7. Configure Publication Field Mapping

If field mapping is not configured by the mapping, map source table fields to topic table fields on the **Field Mapping** page of the publication wizard.

1. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears.
2. To select to view all the fields, mapped fields, or unmapped fields in the topic table and the source table, select the relevant option from the **Show** lists.
3. To search for fields that are used in the topic table and the source table, enter a string for the field name in **Search Fields**, and then press Enter. The search is not case sensitive. You can search for a substring.
4. To map a source table field to a topic table field, drag the field from the **Default** section to the topic table section.
5. To add an expression to a field, in the **Mapped Field/Expression** column, click the expression icon to open the **Field Expression** dialog box, and then select fields and functions to add to the field. To validate the expression, click **Validate**.
6. Click **OK** to apply the mapping.
The **Edit Field Mapping** dialog box closes.
7. In the **Field Mapping** page, click **Next**.
The **Schedule** page appears.

Step 8. Define Publication Schedule

Define the method and the frequency of data publishing in the **Schedule** page of the publication wizard.

1. Select the method and the frequency of data publishing.

When the file is ready to be published

Runs the publication after the published files are ready, the next time it polls the remote sources. Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 9. Review Publication Settings and Save the Publication

Review the publication settings and save the publication in the **Summary** page of the publication wizard.

1. Review the publication settings.
2. Click **Finish**.

The publication wizard closes. The **Publications** tab of the **Edit Application** page shows the publication you created. The publication is listed in the Publications catalog. If you configured a publication schedule, the publication publishes data according to the defined schedule.

Creating a Custom Pass-through Kafka Publication

This topic describes how to create a publication for an Apache Kafka topic.

Perform the following steps to create a publication for a Kafka topic.

1. Access the create publication page.
2. Define the basic publication properties.
3. Save the publication.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The publication mapping and the task are configured in Informatica Intelligent Cloud Services.
- You have created an application.
- You have created a topic.

Step 1: Access the Create Custom Pass-Through Apache Kafka Publication Page

This topic describes how to create a publication for a Apache Kafka topic.

In order to create an Apache Kafka publication, you must configure the Apache Kafka server broker URL in the System Properties and then create a topic for Apache Kafka.

Perform the following steps to create an Apache Kafka publication:

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Select the Publications tab and click **New Publication > Custom > Pass-through Kafka > Create**.

The Create Publication page appears.

Step 2. Define the Pass-through Kafka Publication Properties

Enter a short description of the task here (optional).

1. Enter the publication name.
2. Optionally, enter a description of the publication.
3. Select a topic from the **Topic** list.

The **Topic Structure** area shows the structure of the topic to which the application publishes the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.

4. Select the topic that you have created to monitor the Apache Kafka streaming.
5. Click **Save**.

You can associate a maximum of one real-time topic to a publication. If you have created more than one publication, the publication you have created recently is active.

The publication is created and saved in the **Publication** tab.

CHAPTER 9

Publication Properties

This chapter includes the following topics:

- [Publication Properties Overview, 107](#)
- [Publication General Properties, 107](#)
- [Publication Processing Properties, 108](#)
- [Publication Source Properties, 109](#)
- [Publication Target Properties, 114](#)
- [Publication Join Properties, 114](#)
- [Publication Field Mapping Properties, 114](#)
- [Publication Filter Properties, 117](#)
- [Publication Schedule Properties, 118](#)
- [Publication Summary Properties, 119](#)
- [Properties of a Data-driven Publication, 120](#)

Publication Properties Overview

Use the publication wizard or page to create or edit a publication. The wizard or page you use depends on the type of publication that you create or edit.

Each type of publication wizard or page contains the properties that are relevant to the publication type.

Publication General Properties

The **General** page can include the following properties:

Publication Name

Name of the publication. The name is not case-sensitive and can contain up to 64 characters. The name must begin with an alphabetic character or an underscore, and can contain only alphanumeric characters and underscores.

You cannot change the publication name after you create the publication.

Description

Textual description of the publication.

Mode

Mode of the publication. Choose one of the following options:

- **Enabled.** The publication runs according to the schedule that you define in the **Schedule** page. You can also run the publication from within the Operation Console.
- **Disabled.** The publication does not run according to schedule or by an external API. You can only run the publication from within the Operation Console.

Topic

Topic to which the application publishes the data. Browse to choose a topic from the list of topics. For pass-through file publications, only topics that store published data on a document store publication repository are available for selection. For all other publication types, both topics that store published data on a relational database and topics that store published data on a Hadoop repository are available for selection.

Topic Structure

Shows the structure of the topic to which the application publishes the data. You can view all the tables in the topic, or choose the tables to show.

Tables with Delta Detection shows the names of topic tables for which delta detection is enabled.

Publication Processing Properties

Processing properties depend on the type of publication mapping and the type of data flow engine that the publication uses:

- [“Publication Processing Properties for Data Integration Hub Workflows” on page 108](#)
- [“Publication Processing Properties for Informatica Intelligent Cloud Service Mappings” on page 109](#)

Publication Processing Properties for Data Integration Hub Workflows

The **Processing** page can include the following properties:

Custom Mapping

Select a workflow for the publication from the list. Only publication workflows that are applicable to the type of publication that you create are available for selection. For example, if you create a batch workflow publication, only based workflows are listed.

Use File Transfer

Use file transfer to publish files that reside on remote applications.

Connection

Applicable for file transfer. Name of the connection from where the publication workflow reads the files to be published.

Delete Files from Source After Pickup

Prevent Data Integration Hub from deleting files from the remote server after reading the files.

Run publication pre-processing

Run a pre-processing workflow before the publication process runs. Select the workflow from the list.

Pre-Processing Mapping

Select a publication pre-processing workflow from the list.

Publication Parameters

The content of the **Publication Parameters** area depends on the parameters of the PowerCenter workflow, Data Engineering Integration mapping or workflow, or Data Quality mapping or workflow. The developer imports the workflow or mapping to Data Integration Hub with the parameter definitions. The developer also determines the layout of the area with the Forms Designer.

If required, set the parameter values.

Pre-processing Parameters

The content of the **Pre-processing Parameters** area depends on the parameters of the publication pre-processing workflow. The developer imports the PowerCenter workflow to Data Integration Hub with the parameter definitions. The developer also determines the layout of the area with the Forms Designer.

If required, set the parameter values.

Publication Processing Properties for Informatica Intelligent Cloud Service Mappings

The **Processing** page includes the following properties:

Cloud Mapping

Select an Informatica Intelligent Cloud Service mapping for the publication from the list.

Publication Source Properties

Source properties depend on the type of publication source from which Data Integration Hub reads data in an automatic mapping:

- [“Publication Source Properties for Relational Database” on page 109](#)
- [“Publication Source Properties for Flat File” on page 110](#)
- [“Publication Source Properties for Pass-through File” on page 113](#)

For modular cloud publications, the source properties depend on the settings of the Informatica Intelligent Cloud Services mapping that you select for the publication.

Publication Source Properties for Relational Database

When you use an automatic mapping with a relational database source, the **Source** page of the publication wizard can include the following properties:

Database Connection

Data access connection from where the publication workflow reads the data to be published.

Use pushdown optimization

Fine tune the performance of data retrieval from the source database.

Define new joins by joining existing tables

Combine data from multiple source tables into a single table by creating a join or multiple joins. A join combines columns from two or more tables in the database. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the source database or in the topic.

Publication Source Properties for Flat File

When you use an automatic mapping with a flat file source, the **Source** page of the publication wizard can include the following properties:

Source Type

Type of publication source, **Flat File** or **HDFS**.

Use File Transfer

Applicable for a flat file source. Use file transfer to publish files from remote servers that reside outside the Data Integration Hub network. Data Integration Hub supports the FTP, SFTP, and FTPS protocols for file transfer.

Connection

Applicable for a flat file source. Name of the flat file connection from where the publication workflow reads the files to be published.

HDFS Connection

Applicable for an HDFS source. Name of the HDFS connection from where the publication workflow reads the files to be published.

Delete Files from Source After Pickup

Applicable for a flat file source that uses file transfer. Data Integration Hub deletes the files after it reads them. You must have permissions on the remote server to delete the files.

Directory

Location of the files that contain the source data. You can only use delimited files as publication sources. If you use file transfer, you cannot use patterns to define the directory path. You can, however, use patterns to define the file name.

The publication workflow reads the source files that you define from this directory. If the files that you define are not found in the this directory when the publication workflow runs, Data Integration Hub generates an error event.

Configure flat file source based on a sample file or on a topic table.

Click **Configure** to configure the source files for the publication.

Preview area

Click the edit icon next to a topic table to open the **Configure Flat File Source** dialog box and edit source settings for the table. Click the delete icon to delete the table.

Configure Flat File Source Properties

The following image shows a sample **Configure Flat File Source** dialog box:

Configure Flat File Source

Logical name* ⓘ

File Name*

Configure source based on:

☒ Topic Table ▼

☐ Sample file

► File Format

▼ Column Structure (0 columns)

The column structure depends on the file format settings.

No.	Column Name	Data Type	Precision	Scale
-----	-------------	-----------	-----------	-------

► Sample File Preview

The **Configure Flat File Source** dialog box can include the following properties:

Logical Name

Description name of the source. The name identifies the source when you configure field mapping for the publication.

File Name

Name of the file that Data Integration Hub picks up and publishes to the publication repository. The file name can include asterisk wildcards and variables. Enter variables in the following format: `($pattern)`. For example: `Input_($PublicationName)_($sequence).in`.

Configure source based on

Select one of the following options:

- Topic Table. Select the topic table that represents the structure of the flat file.

- **Sample file.** A sample file that represents the structure of the flat file. Browse to select and upload the file.

File Format

Expand the **File Format** area to view and edit the format of the file. Changing the file format might affect the column structure.

The **File Format** area can include the following properties:

Code page

Character encoding used in the file.

Number of initial rows to skip

Number of rows at the beginning of the file that Data Integration Hub ignores when it reads the file. Data Integration Hub publishes only the subsequent rows.

Import column names

Optional. Select **Yes** to use the column names in the file as the default column headers in the table. Enter the number of the row that serves as the file's header row in **From row**.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** to define a custom delimiter. For information about the supported column delimiters, see the sections about importing delimited flat files and about updating delimited file properties in the *PowerCenter Designer Guide*.

Text qualifier

Optional. Symbols used in the file to enclose a string.

Thousands Separator

Optional. Symbol used in the file as a thousands separator.

Decimal Separator

Symbol used in the file as a decimal separator.

Datetime Format

Date and time format used in the file. Select a predefined format or select **Other** to define a custom format. For information about the supported datetime formats, see the *PowerCenter Transformation Language Reference*.

Note: The datetime format can contain up to 50 characters.

Column Structure

The column structure depends on the file format settings that you configure in the **File Format** area. The table must contain at least one column.

If you provide a sample file, Data Integration Hub reads the file according to the file format settings and presents the columns that it detects in the sample file. Use the **Add Column** button, the up and down arrows, and the remove icon to add, order, and remove table columns.

Click the edit icon to the right of a column to edit column parameters.

Each column must contain the following parameters:

Column Name

Must begin with an alphabetic character or an underscore and can contain only alphanumeric characters and underscores.

If you select the option **Import column names**, Data Integration Hub populates the column names with the strings of the defined row. If you do not select the option **Import column names**, Data Integration Hub assigns default names to the columns. For example, **Field1**, **Field2**, **Fieldn**.

Data Type

Select from the list of available data types. By default, Data Integration Hub reads the data as string.

Precision

Enabled only for data types that support precision.

Scale

Enabled only for data types that support data scaling.

Sample File Preview

If you select a sample file that represents the structure of the flat file, this area shows the data in the sample file. The **Preview** area presents the data in the structure that Data Integration Hub applies when it distributes the data into topic fields in the topic table, according to the **File Format** parameters.

Publication Source Properties for Pass-through File

When you use an automatic mapping with a pass-through file source, the **Source** page of the publication wizard can include the following properties:

Use File Transfer

Use file transfer to publish files from remote servers that reside outside the Data Integration Hub network. Data Integration Hub supports the FTP, SFTP, and FTPS protocols for file transfer.

Connection

Applicable when you use file transfer. Name of the flat file connection from where the publication workflow reads the files to be published.

Delete Files from Source After Pickup

Applicable when you use file transfer. Data Integration Hub deletes the files after it reads them. You must have permissions on the remote server to delete the files.

Directory

Location of the source files. If you use file transfer, you cannot use patterns to define the directory path. You can, however, use patterns to define the file name.

The publication workflow reads the publication files that you define from this directory. If the files that you define are not found in the this directory when the publication workflow runs, Data Integration Hub generates an error event.

Source Definition area

Click the edit icon next to a topic table to enter the name of the source file for the table. The file name can include asterisk wildcards and variables. Enter variables in the following format: `($pattern)`. For example: `Input_($PublicationName)_($sequence).in`.

You must enter a source file name for at least one topic table.

Publication Target Properties

For modular cloud publications, the target properties depend on the settings of the Informatica Intelligent Cloud Services mapping that you select for the publication.

Publication Join Properties

The **Join** page can include the following properties:

Define new joins by joining existing files

Pull data from multiple source files into a single topic table by creating a join. You can create multiple joins, and you can combine data from joins into new joins. Joins are virtual entities, and are not created in the topic.

Publication Field Mapping Properties

The **Field Mapping** page shows the tables in the topic that is associated with the publication, and the mapping between topic tables and source tables.

For relational database publications, if you added joins to the publication, the **Field Mapping** page shows the joins in the new tables list. For flat file publications, if you added joins to the publication, by default, the **Field Mapping** page shows only topic tables. The **Field Mapping** page shows a join only after you configure table mapping for the join.

The **Field Mapping** page can include the following properties:

Field Mapping

Shows the total number of tables in the topic that is associated with the publication and the number of tables that are mapped.

Auto Map

Data Integration Hub maps the tables and fields according to table and field names. If a source table name or field name is identical to a name in the topic, Data Integration Hub maps between the tables or the fields. Data Integration Hub does not map tables and fields that do not match, and you can manually map them.

Clear All

Clears table and field mapping.

Note: The publication must contain at least one mapped source table.

Mapping table

Lists all the tables in the topic that is associated with the publication and the source tables that are mapped to the topic tables. Use the mapping table to manually map source tables to topic tables, to map source fields to topic fields, and to clear the mapping.

Each line in the mapping table contains the following elements:

- Mapped Source Table or Mapped Source Logical Name. Name of the source or join and the number of fields in the table.
- Topic Table. Name of the topic table to which the source table is mapped and the number of fields in the topic table.
- Mapped Fields. Number of fields that are mapped between the source table and the topic table.
- Edit Table Mapping. Click the edit icon to edit the table mapping. For more information, see [“Edit Table Mapping for Publications” on page 115](#).
- Edit Field Mapping. Click the edit icon to edit the field mapping. A source table must be mapped to a topic table before you can edit the field mapping of the table. For more information, see [“Edit Field Mapping for Publications” on page 116](#).
- Clear Mapping. Clear the mapping between the source table and the topic table.

Note: For publications with a relational database source, if the source does not contain any of the topic tables or if the selected connection does not contain metadata access details, Data Integration Hub generates the mapping based on the topic tables. The table mapping and field mapping editing options are disabled. To change the default mapping, click **Clear All**, and then edit the table and field mapping.

Edit Table Mapping for Publications

For publications with an automatic mapping and a relational database or a flat file source, manually map source tables and join tables to topic tables.

When you click the edit table mapping icon next to a table, the **Edit Table Mapping** dialog box opens.

Edit Table Mapping Properties for Relational Database Publications

To map a join table to a topic table, select the table from the join table list and map it to the topic table. To map a source table to a topic table, find a source table and map it to the topic table.

The **Edit Table Mapping** dialog box for relational database publications can include the following properties:

Select a new table to map to the topic table

Select a join table to map to the topic table.

Find Source Table

Search for source tables in the database. Note the following guidelines:

- To search for tables by table name, enter a string in the **Find Source Table** text box and then click **Search**.
- To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
- To clear the search results and show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
- The search is not case sensitive.
- You can search for a substring.

Search Results

Lists the source tables that match your search and shows which source table is mapped to the topic table for which you edit the mapping. To map a different source table to the topic table, select a table from the **Source Table** list.

Map Source Table

Maps between the table that you select and the topic table, and closes the **Edit Table Mapping** dialog box.

Edit Table Mapping Properties for Flat File Publications

The **Edit Table Mapping** dialog box for flat file publications can include the following properties:

Select a source to map to the topic table

Select a table to map to the topic table.

Map Source Table

Maps between the table that you select and the topic table, and closes the **Edit Table Mapping** dialog box.

Edit Field Mapping for Publications

For publications with an automatic mapping and a relational database, flat file, or cloud source, manually map fields in source tables to fields in topic tables.

The **Field Mapping** page of the publication wizard lists the tables in the topic that is associated with the publication, and the mapping between source tables and topic tables. When you click the edit field mapping icon next to a table, the **Edit Field Mapping** dialog box opens.

Note: For publications with a relational database or a flat file source, a source table must be mapped to a topic table before you can edit the field mapping for the table.

Edit Field Mapping Properties for Publications

To map fields in a table, select the source table field and the topic table field from the corresponding columns of the **Edit Field Mapping** dialog box and map them to each other.

The following list describes the elements of the **Edit Field Mapping** dialog box.

Find Field

Searches fields by name in both the topic table and the source table. Enter the string for which to search, and then click **Search**. The search is not case sensitive. You can search for a substring.

After you enter a string in the **Find Field** text box, you can click **X** to clear the search and show all fields.

Choose **Display unmapped fields only** to show only unmapped fields in the search results.

Auto Map

Data Integration Hub maps the fields according to field names. If a name in the source is identical to a name in the topic, Data Integration Hub maps between the fields. Data Integration Hub does not map fields that do not match, and you can manually map them.

Clear All

Clears the field mapping. Data Integration Hub does not publish unmapped fields.

Note: The table must contain at least one mapped field.

Source Table

Lists the fields in the source table that match your search. Mapped fields are denoted by a green checkmark. Select the field that you want to map.

Rest on the details icon to the left of a field name to view the details of the field.

Note: If the source table contains the Data Integration Hub system fields DIH_PUBLICATION_INSTANCE_ID, DIH_PUBLICATION_INSTANCE_DATE, and DIH_UPDATE_STRATEGY_FLAG, do not map any of these fields to the topic table.

Topic Table

Lists the fields in the topic table that match your search, and shows the mapping between the listed topic fields and source fields. Mapped fields are denoted by a green checkmark. Select the field that you want to map.

You can perform the following actions on a topic field:

- Add an expression to the field. In the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
- Clear the mapping of the field. In the **Actions** column, click the clear icon.
- View field details. Rest on the details icon to the left of a field name.

Publication Filter Properties

The **Filter** page can include the following properties:

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2016.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub publishes only the rows that meet all of the conditions.

Basic Expression Properties

When you define a filter condition with a basic expression for a publication, the **Filter** page shows the following properties:

Select Table or Select Source

Select the table or the source file to which the filter applies.

Select Column

Select the column to filter.

Select Operator

Select the filter operator. The available operators depend on the type of content in the column.

Enter Value

When this field is enabled, select or enter a value for the operator.

Advanced Expression Properties

When you define a filter condition with an advanced expression for a publication, the **Filter** page shows the following properties:

Select Table or Select Source

Select the table or the source file to which the filter applies.

Expression text area

Enter a valid PowerCenter expression that defines the filtering criteria. The expression is limited to the selected table. The expression can contain up to 1024 characters.

Publication Schedule Properties

The **Schedule** page can include the following properties:

When the file is ready to be published

Runs the publication immediately after the published files are ready.

Define the maximal period of time that Data Integration Hub waits for the files to be available in the directory that you defined in the **Source** page in **Wait for all data to be published ... hours**. When the maximal period of time ends, Data Integration Hub discards the file events of the publication and changes the status of the publication event to Error.

Note: When you use file transfer to publish the files, after the published files are ready, Data Integration Hub runs the publication the next time it polls the remote sources.

Manually or by an external trigger

No schedule. You can use the following methods to run the publication:

- Run manually. Click the Run arrow on the **Publications** page.
- Run by an API. Call a command-line API or a REST API that starts the publication.
- Run by a PowerCenter workflow. Start a publication transformation.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

By schedule

Runs the publication according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the publication in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the publication in intervals of up to 24 hours. You select the number of hours from the list. The publication runs at the beginning of the hour. For example, if you enter 2, the publication runs at 00:00, 02:00, and at consecutive two-hour intervals.
- Daily. Runs the publication at the same hour every day.
- Weekly. Runs the publication every week on one or more days at the same hour.
- Monthly. Runs the publication every month on a specific date or a specific day at the same hour.

Define the publication intervals in the **Repeat running** area.

For file publications that use this scheduling option and that publish multiple files, all the files must be present in the source location when the publication starts.

After the following publication or subscription run completes

Runs the publication after the run of the publication or subscription that you select here completes.

Publish every

For publications with real-time workflows, Data Integration Hub groups the published data and publishes the data at the intervals that you define here. You can set a value of 10 seconds up to 59 minutes and 50 second, with intervals of 10 seconds.

You can click the Run arrow on the **Publications** page to run the publication manually.

Retry policy

Defines a retry policy for automatic file publications. A retry policy defines the number of times Data Integration Hub retries to run the publication in case of failure, and the retry interval. The policy does not apply to publications that you run manually.

Publication Summary Properties

The **Summary** page can include the following properties:

Element	Description
Publication Name	Name of the publication.
Topic Name	The topic to which the application publishes the data.
Topic Type	Type of topic to which the application publishes the data.
Mapping	Mapping type. If you use a custom mapping, the field displays the name of the publication workflow.
Informatica Cloud Task	The name of the Informatica Cloud task associated with the publication.
Pre-Processing Mapping	Indicates whether publication pre-processing is enabled. If pre-processing is enabled, the field displays the name of the pre-processing workflow.
Source Type	Type of data source.
Source	Names of the source data connection.
Schedule	Time interval for publishing the data.
Status	Publication status.

Properties of a Data-driven Publication

The **Create Data-driven Publication** page includes the following properties:

Publication Name

Name of the publication. The name is not case-sensitive and can contain up to 64 characters. The name must begin with an alphabetic character or an underscore, and can contain only alphanumeric characters and underscores.

You cannot change the publication name after you create the publication.

Description

Textual description of the publication.

Mode

Mode of the publication. Choose one of the following options:

- Enabled. The publication runs when you post the publication request.
- Disabled. The publication does not run when you post the publication request.

Topic

Topic to which the application publishes the data. Browse to choose a topic from the list of topics.

Topic Structure

Shows the structure of the topic to which the application publishes the data. You can view all the tables in the topic, or choose the tables to show.

Group publications every ... seconds

Data Integration Hub groups the published data and publishes the data at the intervals that you define here.

- If no grouping is defined, that is, **Group publications every ... seconds** is set to zero, Data Integration Hub does not group the published data. Data Integration Hub writes the data to the publication repository as it is being published.
- If you define a grouping time interval, Data Integration Hub groups the published data and writes it to the publication repository based on the time interval. For example, if you define **Group publications every 10 seconds**, Data Integration Hub will group the published data at the end of each 10-second period.

Publish Data through the URL

REST URL of the API. Use this URL to publish the data.

Click **Copy** to copy the URL to your clipboard.

The Publish Data REST API returns the Swagger structure for the topic into which the publication publishes data

REST URL of the API Swagger file with the Swagger structure for the topic into which the publication publishes data. Use the structure in the publication request.

Click **Copy** to copy the URL to your clipboard.

CHAPTER 10

Subscriptions

This chapter includes the following topics:

- [Subscriptions Overview, 121](#)
- [Subscription Types, 122](#)
- [Subscription Process, 127](#)
- [Post-Processing Mappings, 129](#)
- [Unbound Subscription, 129](#)
- [Subscription Targets , 130](#)
- [Subscription Filters, 131](#)
- [Subscription Schedules, 131](#)
- [Subscription Delivery, 132](#)
- [Managing Subscriptions, 132](#)

Subscriptions Overview

Subscriptions are entities that define how applications consume data from Data Integration Hub.

Subscriptions subscribe to topics. A subscription can subscribe to multiple topics. Multiple subscriptions can consume data from the same topic.

When you create a subscription, choose the topic or topics to which the application subscribes. If applicable, define the schedule and the delivery scope of the data to consume and the delivery behavior for the published data, for example, to aggregate all data sets to a single data set, or to consume the latest published data set.

In the Data Integration Hub Operation Console, you use the subscription wizard or page to create and edit subscriptions. During the configuration you choose the topic to which to subscribe. The steps that you take when you configure a subscription depend on the subscription type and on the type of mapping that the subscription uses.

Subscription Types

You can use an automatic subscription, a custom subscription, or a modular subscription to publish data. The subscription type you choose is based on the organization requirements and the complexity of the business logic that you want to apply when publishing the data.

Automatic Subscriptions

With automatic subscriptions you can consume data into a relational database target or a flat file target, or over a REST API. If required, Informatica Intelligent Cloud Services generates the automatic mapping based on details that you provide when you create the subscription.

You can use the following automatic subscription types to subscribe to data with Data Integration Hub:

- Automatic database subscription. A subscription with an automatic mapping and a relational database target.
- Automatic file subscription. A subscription with an automatic mapping and a flat file target.
- Automatic pass-through file subscription. A subscription with an automatic mapping and a pass-through file target.
- Automatic data-driven subscription. A subscription that consumes data over a REST API when the data is ready, not based on a defined schedule.

Custom Subscriptions

With custom subscriptions you can consume data into any type of target that PowerCenter, Data Engineering Integration, Data Engineering Streaming, Data Quality, and Informatica Intelligent Cloud Services support.

You can use the following custom subscription types to subscribe to data with Data Integration Hub:

- Custom batch subscription. A subscription that uses a Data Integration Hub batch workflow. A batch workflow can be based on a PowerCenter batch workflow, a Data Engineering Integration mapping or workflow, or a Data Quality mapping or workflow.
- Custom cloud subscription. A subscription with an Informatica Intelligent Cloud Services task and a cloud target.

Modular Subscription

A subscription with an Informatica Intelligent Cloud Services mapping and a cloud target. With a modular cloud subscription you can consume data into any cloud target that Informatica Intelligent Cloud Services supports. Data Integration Hub generates the subscription mapping based on details that you provide when you create the subscription, including the parameters values that you enter into the template that you select for subscription processing.

Automatic Database Subscription

Use this type of subscription to consume data into a relational database target if you want Data Integration Hub to create the PowerCenter workflow for the subscription.

For subscriptions with a relational database target, select the relational database connection. Data Integration Hub explores the target connection, searches for the tables and fields that match the names of tables and fields that are defined in the topic structure, and performs one of the following actions:

- If all or some of the tables that are defined in the topic structure exist in the target connection, Data Integration Hub maps the topic tables and fields to target tables and fields, based on a name match. You can edit the mapped tables and fields and manually map them to other tables and fields in the target.

- If none of the tables that are defined in the topic structure exist in the target connection, Data Integration Hub displays a warning and generates the mapping based on the topic tables. You can exclude fields or tables from the mapping, or manually map the tables to other tables in the target.
- If the connection does not contain metadata access, Data Integration Hub cannot explore the target connection. Data Integration Hub displays a warning and generates the mapping based on the topic tables. You can exclude tables and fields from the mapping. You cannot edit the mapping.

If the subscribing application requires only a subset of the data that is published to the topic, you can define which tables and columns to consume. You can also filter the rows to consume.

Data Integration Hub generates the workflow based on the mapping. The workflow runs during the subscription process. The subscriber subscribes only to tables and fields that are mapped from the topic to the target. If the tables that are defined in the mapping do not exist in the target connection, you must add the required tables to the target before you run the subscription. Otherwise, the subscription workflow fails during run time.

The following pages appear in the subscription wizard for this type of subscription:

- General. Define basic subscription properties.
- Processing. If you want to run a post-process on the subscription, select a subscription post-processing workflow. If the post-process includes parameters, the parameters appear in this page and are optional. If required, set the values of the parameters.
- Join. You can pull data from multiple topic tables into a single target table by creating joins. You can create multiple joins, and you can combine data from joins into new joins.
- Target. Select the data access connection to the target to which Data Integration Hub writes the data.
- Field Mapping. Review the subscription field mapping, and if required, edit the mapping that Data Integration Hub generates by default. You can manually map topic tables and fields to target tables and fields.
- Filter. Define the data that the subscription consumes by setting filter conditions on table columns.
- Schedule. Define the method and the frequency of data consumption. You can select to consume the published data immediately after the data is ready, to run the subscription manually or by an external trigger, or to consume the data according to a schedule that you define.
- Delivery. Define the delivery options of the data to consume. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format: all available subscriptions, all available subscriptions - aggregated, or the latest subscription only. For all subscriptions, choose how Data Integration Hub handles data that exists in the target application: append the new data to the existing data in the target, insert new rows, or overwrite the existing data in the target. For subscriptions where the type of topic to which the subscription subscribes is Full, if you select to append the new data or to insert new rows, you can select to delete rows that don't exist in the topic from the target.
- Summary. Review the subscription settings and save the subscription.

Automatic Flat File Subscription

Use this type of subscription to consume data into a flat file target if you want Data Integration Hub to create the PowerCenter workflow for the subscription.

For subscriptions with a flat file target, define the location of the target files and the expected file format. If you are delivering the files to a Hadoop Distributed File System (HDFS), select the HDFS connection. If you use the file transfer protocol to move the files, select the file transfer connection.

Data Integration Hub creates the mapping based on the topic tables that are associated with the file, and based on the fields in the topic structure. Topic tables that are not associated with a file are not mapped. You

can edit the mapping of fields and exclude fields and tables. For each table that exists in the mapping, the subscription process generates a file. The file contains only the fields that are mapped from the topic to the target.

The following pages appear in the subscription wizard for this type of subscription:

- General. Define basic subscription properties.
- Processing. If you want to run a post-process on the subscription, select a subscription post-processing workflow. If the post-process includes parameters, the parameters appear in this page and are optional. If required, set the values of the parameters.
- Join. You can pull data from multiple topic tables into a single target table by creating joins. You can create multiple joins, and you can combine data from joins into new joins.
- Target. Choose the target type. For an HDFS target or when you use file transfer, select the connection to the target to which Data Integration Hub delivers the data. Enter the location of the file or files into which Data Integration Hub writes the data, and configure flat file targets for the topic tables. You must configure a file target for at least one topic table.
- Field Mapping. Review the subscription field mapping, and if required, edit the mapping that Data Integration Hub generates by default. You can manually map topic tables and fields to target tables and fields.
- Filter. Define the data that the subscription consumes by setting filter conditions on table columns.
- Schedule. Define the method and the frequency of data consumption. You can select to consume the published data immediately after the data is ready, to run the subscription manually or by an external trigger, or to consume the data according to a schedule that you define.
- Delivery. Define the delivery options of the data to consume. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format: all available subscriptions, all available subscriptions - aggregated, or the latest subscription only. For subscriptions that do not use file transfer, choose how Data Integration Hub handles data that exists in the target application: append the new data to the existing data in the target, insert new rows, or overwrite the existing data in the target. For subscriptions where the type of topic to which the subscription subscribes is Full, if you select to append the new data or to insert new rows, you can select to delete rows that don't exist in the topic from the target.
- Summary. Review the subscription settings and save the subscription.

Automatic Pass-through File Subscription

Use this type of subscription to consume files that were published to topics that use a file publication repository and that provide files as-is, without loading them to a relational database, if you want Data Integration Hub to create the PowerCenter workflow for the subscription.

You cannot deliver the files with a pass-through file target to a Hadoop Distributed File System (HDFS).

For subscriptions with a pass-through file target, define the location of the target files, and then define the expected file names or the expected file name patterns for the tables in the topic that is associated with the subscription. If you use the file transfer protocol to move the files, select the file transfer connection.

Data Integration Hub creates the mapping based on the topic tables that are associated with the file. Topic tables that are not associated with a file are not mapped. You can edit the mapping of fields and exclude fields and tables. For each table that exists in the mapping, the subscription process generates a file. The file contains only the fields that are mapped from the topic to the target.

The following pages appear in the subscription wizard for this type of subscription:

- General. Define basic subscription properties.

- **Processing.** If you want to run a post-process on the subscription, select a subscription post-processing workflow. If the post-process includes parameters, the parameters appear in this page and are optional. If required, set the values of the parameters.
- **Target.** Choose the target type. If you use file transfer, select the connection to the target to which Data Integration Hub delivers the data. Enter the location of the file or files that Data Integration Hub delivers, and name the file or files that Data Integration Hub writes to the target. You must name a file target for at least one topic table.
- **Filter.** Define the data that the subscription consumes by setting filter conditions on the file metadata.
- **Schedule.** Define the method and the frequency of data consumption. You can select to consume the published data immediately after the data is ready, to run the subscription manually or by an external trigger, or to consume the data according to a schedule that you define.
- **Delivery.** Define the delivery options of the data to consume. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format: all available publications or the latest publication.
- **Summary.** Review the subscription settings and save the subscription.

Automatic Data-driven Subscription

Use this type of subscription to subscribe to data over a REST API, so that Data Integration Hub runs the subscription as data is published to the publication repository without a predefined schedule.

For automatic data-driven subscriptions, you configure subscription properties and select the topic from which to consume the data. You then create a POST request to run the subscription.

You configure a data-driven subscription in the **Create Automatic Data-driven Subscription** page.

Note: When you use an automatic data-driven subscription, you cannot subscribe from topics in a Hadoop repository.

Custom Batch Subscription

Use this type of subscription to deliver data when you want to use a Data Integration Hub batch workflow that the developer created in the Data Integration Hub Operation Console. Batch workflows run according to a trigger and not continuously.

A custom mapping uses a PowerCenter workflow, a Data Engineering Integration mapping or workflow, or a Data Quality mapping or workflow to consume data. Workflows and mappings can perform complex transformations on the data.

When a custom mapping uses a PowerCenter workflow, the developer creates a workflow in PowerCenter Designer. When a custom mapping uses a Data Engineering Integration mapping or workflow or a Data Quality mapping or workflow, the developer creates a mapping in Informatica Developer. The developer then imports the workflow or the mapping to a Data Integration Hub workflow. You select the Data Integration Hub workflow when you create the custom subscription in Data Integration Hub.

The Data Integration Hub workflow includes information about the target data structure and the database connection.

To create a custom batch subscription, select a Data Integration Hub workflow that reads from topic tables and writes to the target. Workflow sources must include at least one of the topic tables, and must not include any source table that is not defined in the topic.

If a PowerCenter workflow, a Data Engineering Integration mapping or workflow, or a Data Quality mapping or workflow includes parameters, you can enter values for the parameters in the Subscription wizard. The developer imports the workflows and mappings to Data Integration Hub with the parameter definitions.

When you create a custom batch subscription, you can create a compound subscription, where the subscription consumes data sets from multiple topics of the same type. The subscription process starts after all publications from all mandatory topics publish data. You can specify the maximum time to wait for all publications to finish publishing, from the time the first publication is ready to consume.

The following pages appear in the subscription wizard for this type of subscription:

- General. Define basic subscription properties.
- Processing. Select the subscription mapping. If you use file transfer, select the connection to the target to which Data Integration Hub writes the files. If you want to run a post-process on the subscription, select a subscription post-processing workflow.
The content of this page depends on the PowerCenter workflow, Data Engineering Integration mapping or workflow, or Data Quality mapping or workflow parameters that the developer defines. The developer imports the workflow or the mapping to Data Integration Hub and determines the layout of the page. Post-processing parameters are optional. If required, set the values of the parameters.
- Schedule. Define the method and the frequency of data consumption. You can select to consume the published data immediately after the data is ready, to run the subscription manually or by an external trigger, or to consume the data according to a schedule that you define.
- Summary. Review the subscription settings and save the subscription.

Custom Cloud Subscription

Use this type of subscription to deliver data to a cloud application and you want to use a cloud task that the developer created in Informatica Intelligent Cloud Services.

A custom cloud mapping uses an Informatica Intelligent Cloud Services task to consume data. Tasks can perform complex transformations on the data.

The developer creates the task in Informatica Intelligent Cloud Services. The task includes information about the target data structure and the database connection. You select the task when you create the subscription in Data Integration Hub. Data Integration Hub triggers the task when the publication is ready for subscribers and uses the Data Integration Hub cloud connector to read the data from Data Integration Hub.

To create a custom cloud subscription, select an Informatica Intelligent Cloud Services task that reads from topic tables and writes to the target. Task sources must include at least one of the topic tables, and must not include any source table that is not defined in the topic.

When you create a custom cloud subscription, you can create a compound subscription, where the subscription consumes data sets from multiple topics of the same type. The subscription process starts after all publications from all mandatory topics publish data. You can specify the maximum time to wait for all publications to finish publishing, from the time the first publication is ready to consume.

The following pages appear in the subscription wizard for this type of subscription:

- General. Define basic subscription properties.
- Processing. Select an Informatica Intelligent Cloud Services task.
- Schedule. Define the method and the frequency of data consumption. You can select to consume the published data immediately after the data is ready, to run the subscription manually or by an external trigger, or to consume the data according to a schedule that you define.
- Summary. Review the subscription settings and save the subscription.

Modular Cloud Subscription

Use this type of subscription to consume data into a cloud target with an Informatica Intelligent Cloud Services mapping.

For modular subscriptions with a cloud target, the developer creates the mapping in Data Integration. You select the mapping when you create the subscription in Data Integration Hub.

If the mappings includes parameterized transformations, you enter values for the parameters in the subscription wizard.

If the source connection is parameterized in the mapping, you select the Data Integration Hub publication repository as the subscription source in the subscription wizard.

The following pages appear in the subscription wizard for this type of subscription:

- General. Define basic subscription properties.
- Processing. Select the subscription mapping.
- Source. If the subscription source is not configured by the mapping, configure the publication source on the Data Integration Hub publication repository from which Data Integration Hub reads the data.
- Target. If the subscription target is not configured by the mapping, configure the target to which Data Integration Hub writes the data.
- Input Parameters. If the mapping includes parameterized transformations, set parameter values.
- Field Mapping. If required, map topic table fields to target table fields.
- Schedule. Define the method and the frequency of data consumption. You can select to consume the published data immediately after the data is ready, to run the subscription manually or by an external trigger, or to consume the data according to a schedule that you define.
- Summary. Review the subscription settings and save the subscription.

Subscription Process

The subscription process includes retrieving the required data from the Data Integration Hub subscription repository, running any associated mappers, such as a mapping or a task, and writing the data to one or more subscriber targets. Data Integration Hub keeps the data in the subscription repository until the retention period of the topic expires.

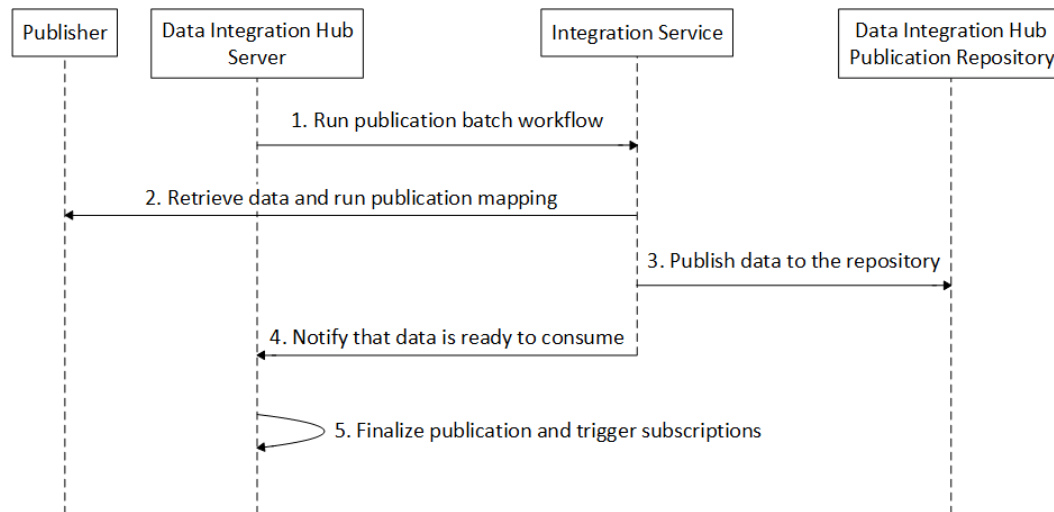
- Automatic subscriptions can run a Data Integration Hub workflow that is based on a PowerCenter batch workflow or run over a REST API.
- Custom subscriptions can either run a Data Integration Hub workflow that is based on a PowerCenter batch workflow, Data Engineering Integration mapping or workflow, Data Engineering Streaming mapping, or Data Quality mapping or workflow, or run an Informatica Intelligent Cloud Services task.
- Modular subscriptions run an Informatica Intelligent Cloud Services mapping.

Subscription Process with a Batch Workflow

The subscription process for subscriptions that run a Data Integration Hub batch workflow includes the following stages:

1. When the publication is ready for subscribers, the Data Integration Hub server runs the subscription batch workflow and sends a request to the relevant Integration Service, either the PowerCenter Integration Service or the Data Integration Service.
2. The Integration Service extracts the data from the Data Integration Hub publication repository, and runs the automatic or custom mapping on the data.
3. The Integration Service sends the required data to the subscriber.
4. The Integration Service notifies the Data Integration Hub server after the subscriber consumed the published data that they require.
5. The Data Integration Hub server changes the status of the subscription event to complete.

The following image shows the main stages of the subscription process for each subscription:



Subscription Process with a Data Integration Task

The subscription process for subscriptions that run a Data Integration task includes the following stages:

1. When the publication is ready for subscribers, the Data Integration Hub server triggers the Data Integration task that is defined for the subscription through an Informatica Intelligent Cloud Services REST API.
2. The subscription process uses the Data Integration Hub cloud connector to read data from Data Integration Hub.
3. The Data Integration task reads the data from Data Integration Hub and then writes the data to the cloud application.
4. The Data Integration Hub server changes the status of the subscription event to complete.

Subscription Process of a Data-driven Subscription

The subscription process for data-driven subscriptions includes the following stages:

1. When you configure the properties of a data-driven subscription, you enter the URL to where Data Integration Hub sends notifications when data is ready to consume from the Data Integration Hub publication repository, from the topic that you define in the subscription.
2. You create a POST request to run the subscription and fetch the data from the Data Integration Hub publication repository, from the topic that you define in the subscription.
3. When Data Integration Hub sends notifications that data is ready to be consumed from the topic, you post the request to run the subscription and to fetch the data.

Post-Processing Mappings

For on-premises subscriptions you can use a subscription post-process to implement business logic after data consumption.

When you configure a subscription to run a post-process, Data Integration Hub triggers the post-process after it runs the subscription.

The developer creates a PowerCenter workflow that runs the post-process. You select the post-process workflow and define post-processing parameters in the subscription wizard. You can select a post-process for both subscriptions that use automatic mapping and subscriptions that use custom mapping.

Unbound Subscription

You can define a subscription as an unbound subscription.

While regular subscriptions retrieve published data according to publication instances, unbound subscriptions ignore publication instances and retrieve the data according to the defined filter, regardless of when or in what batch the data was published.

Use this type of subscription when you want to receive the data according to a business-related logic and not according to publication batches.

When you define the topic from which the unbound subscription consumes data, you can define one or more filter accelerators. Filter accelerators are table columns that define how Data Integration Hub retrieves data based on business-related logic. Data Integration Hub indexes filter accelerator columns to accelerate data retrieval.

For example, an application publishes orders to Data Integration Hub when the orders are sent to the ordering system. The accounting system must receive the orders according to the business-date range, which is not necessarily according to the batches in which the orders are published. To deliver the orders according to the business-date range, perform the following steps:

- In the topic to which the application publishes the orders, set the business-date column as a filter accelerator.
- Create an unbound subscription that subscribes to the topic that publishes the orders and define a filter that is based on the business-date column.

The subscription consumes the orders according to the relevant business date, regardless of when and how they are published.

Subscription Targets

Automatic subscriptions can consume data into a relational database target, into a flat file target, or into a pass-through file target, or over a REST API.

Custom subscriptions can consume data into any type of source that PowerCenter, Data Engineering Integration, and Data Quality support.

Modular subscriptions can consume data into any type of source that Informatica Intelligent Cloud Services supports.

Relational Targets

For subscriptions that use an automatic mapping and a relational database target, Data Integration Hub supports Microsoft SQL Server, Oracle, IBM DB2, PostgreSQL, and Teradata relational databases as subscription targets.

Data Integration Hub can create target tables if they do not exist in the target database. If you choose to create target tables when you create or edit a subscription, Data Integration Hub creates the target tables based on table mapping. Data Integration Hub does not create unmapped tables.

Flat File Targets

For subscriptions that use an automatic mapping and a flat file target, Data Integration Hub supports only delimited files as subscription targets.

You can deliver flat files to file systems, including Hadoop Distributed File Systems (HDFS). You can use file transfer to deliver flat files to applications that reside outside the Data Integration Hub network. You cannot use file transfer to deliver files to HDFS.

Pass-through File Targets

For subscriptions that use an automatic mapping and a pass-through file target, Data Integration Hub delivers the files without processing them. For example, unstructured files such as PDF or .zip files.

You can use file transfer to deliver files to applications that reside outside the Data Integration Hub network.

When you select a pass-through file target, you can select only topics that use a file publication repository. Data Integration Hub delivers the files as-is.

You cannot deliver pass-through files to Hadoop Distributed File Systems (HDFS).

Subscription Filters

To define the data that an automatic mapping subscription consumes, add a filter to the subscription. When you add a filter to a subscription, you can define filter conditions to table rows.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

For automatic database and flat file subscriptions, if data relations are defined for the topic to which the subscription subscribes, filters are applied to the tables for which relations are defined, and the subscriber consumes only the data that is defined by the data relations.

Subscription Schedules

The subscription schedule defines the frequency of the subscription for subscriptions that are triggered by Data Integration Hub.

You can consume published data when it is published, manually, by an external trigger, at defined intervals, or after a specific publication or subscription run completes. If you create a compound subscription, you can only choose to consume data when it is published, manually, or by an external trigger.

Consumption of data by the subscription starts when one of the following conditions exist:

- The subscription schedule is set to consumes data immediately after the publisher publishes the data to Data Integration Hub.
- The scheduled start time arrives.
- The developer starts the subscription from a command line API or from a REST API.
- You manually run a subscription that is in a Delayed status.
- You manually get previous publications.
- The running of the publication or subscription after which the current subscription is configured to run is complete.

For automatic subscriptions that are triggered by Data Integration Hub, you can define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.

Data-driven subscriptions are not triggered by Data Integration Hub. For subscriptions of this type you define the time intervals according to which subscriptions groups and consumes data.

Subscription Delivery

The subscription delivery defines the data that the subscription consumes and the update strategy for data that exists in the target, and provides advanced performance tuning options. Subscription delivery options are applicable for automatic mapping subscriptions.

Data to Consume

The subscription can consume all the available publications as multiple data sets or as a single data set, or consume the latest publication.

Unbound subscriptions consume the data according to the defined filter, regardless of when or in what batch the data was published.

Update Data in Target

Select whether to append the new data to existing data in the target, insert new rows and update the data that exists in the target, or overwrite the existing data in the target.

When you select to insert new rows and update the data that exists in the target, you define the update strategy for rows that exist in the target application, either update all existing rows or update only changed rows. Data Integration Hub applies the update strategy that you define when the following conditions exist:

- No joins are defined for the topic table.
- Primary keys are applied to the topic table.

If joins are defined for the topic table, Data Integration Hub detects the update strategy based on the joins. If no joins are defined for the topic table and primary keys are applied to the topic table, Data Integration Hub applies the update strategy that you define in the subscription delivery options.

If the type of topic to which the subscription subscribes is Full or if the subscription is an unbound subscription, and you choose to insert new rows and with an update changed rows strategy, you can select to delete rows that don't exist in the topic from the target.

Advanced Performance Tuning

Fine tune the performance of data retrieval from the Data Integration Hub publication repository for subscriptions that use an automatic mapping. Performance tuning is important when you use unbound subscriptions.

Managing Subscriptions

Use the Navigator to create, search, duplicate, edit, or delete subscriptions. In the Navigator you can also disable and enable subscriptions, and run subscriptions manually.

Creating a Subscription

Use the Navigator to create subscriptions:

The following conditions must exist before you create a subscription:

- The application from which you want to consume the published data must exist. You can either use an existing application, or create and save a new application.

- The topic to which you want to subscribe must exist.
 - For a custom subscription, the Data Integration Hub workflow or the Informatica Intelligent Cloud Services task must exist.
1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
 2. Click the application from which you want to consume the published data.
The **Application Details** page appears.
 3. Click the **Subscriptions** tab.
The subscriptions list appears.
 4. Click **New**, select the type of subscription to create, and then click **Create**.
The create subscription wizard or page for the selected subscription type opens.
 5. Enter the required details in each tab, and click **Next**.
 6. Click **Finish**.

Searching for a Subscription

Perform the following tasks to search for a subscription:

To search a subscription, type the name of the subscription in the **Find** text box and click **Search**. Subscriptions that match the search text that you entered appear.

- Click **Reload Default Search Results** to clear search results and display the list of subscriptions.
- Click **Refresh Search Results** to refresh search results.

Duplicating a Subscription

Click the **Duplicate** icon next to the Subscription to create a subscription with a different name and identical properties. Perform the following tasks to duplicate a subscription:

1. On the **Subscriptions** tab, click the **Duplicate** icon next to the subscription name.
The **Confirmation** window appears.
2. Rename the duplicate subscription as required, and click **Save**.

Editing a Subscription

When you open a subscription while it is running, the details appear in read-only mode. Perform the following tasks to edit a subscription:

1. On the **Subscriptions** tab, click the **Edit** icon next to the subscription name.
The **Edit** page appears.
2. Change the required properties in the subscription wizard or page, and click **Save**.

Deleting a Subscription

Perform the following tasks to delete a subscription:

1. On the **Subscriptions** tab, click the **Delete** icon next to the subscription name and confirm the deletion.
You cannot delete a subscription while it is running.

2. Click **Ok**.

Enabling and Disabling a Subscription

On the **Subscriptions** tab, perform the following tasks to enable or disable a subscription:

- To disable a subscription select **Disabled** from the **Mode** column .
- To enable a disabled subscription, select **Enabled** from the **Mode** column.
- To disable all subscriptions, select **Actions > Disable All**.
- To enable all subscriptions, select **Actions > Enable All**.

A disabled subscription does not run according to schedule or by an external API. You can only run a disabled subscription from within the Operation Console.

Running a Subscription Manually

Perform the following tasks to manually run subscriptions:

1. On the **Subscriptions** tab, click the **Run Subscription** icon next to the subscription that you want to run.
A confirmation message appears.
2. Click **Yes**.

Getting Previous Subscriptions

To get previous subscriptions, click the **Get Previous Publications** icon next to the subscription for which you want to get the data. In the **Get Previous Publications** dialog box, define the date range for which to get subscriptions and click **Run**. Data Integration Hub delivers undelivered data that was published during the date range that you define and which is still available in the publication repository.

Re-validating a Subscription

When the subscription status icon shows that the subscription is not valid, open the Subscription wizard in edit mode and follow the wizard steps. Ensure to verify the settings. If the subscription shows that it is not valid due to topic changes that do not impact the subscription, you need not change any setting. Data Integration Hub re-generates the subscription mapping and refreshes the subscription properties. The subscription status icon shows that the subscription is valid. Verify that all associated topics, connections, entities, and mappings are valid before you re-validate the subscription.

CHAPTER 11

Creating Subscriptions

This chapter includes the following topics:

- [Creating Subscriptions Overview, 135](#)
- [Creating an Automatic Relational Database Subscription, 135](#)
- [Creating an Automatic Flat File Subscription, 141](#)
- [Creating an Automatic Pass-through File Subscription, 147](#)
- [Creating an Automatic Data-driven Subscription, 151](#)
- [Creating a Custom Batch Subscription, 153](#)
- [Creating a Custom Cloud Subscription, 155](#)
- [Creating a Modular Cloud Subscription, 157](#)
- [Creating a Custom Pass-through Kafka Subscription, 160](#)

Creating Subscriptions Overview

To create a subscription, in the Data Integration Hub Operation Console, use the subscription wizard or page that is applicable to the type of subscription that you want to create.

Each type of subscription wizard or page contains the properties that are relevant to the subscription type.

Creating an Automatic Relational Database Subscription

To create a subscription with an automatic mapping and a relational database target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. If you want to run a post-process on the subscription, select the post-process workflow.
4. Optionally, define joins for the subscription.
5. Select the subscription target.

6. Review the publication field mapping, and if required, edit the mapping that Data Integration Hub generates by default.
7. Define a filter for the subscription.
8. Define the subscription schedule.
9. Define the delivery scope and the delivery format for the subscription.
10. Review the subscription settings and save the subscription.

Step 1. Access the Create Automatic Database Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Relational database**, and then click **Create**.
The **Create Automatic Database Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Select whether to apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.
5. Click **Next**.
The **Processing** page appears.

Step 3. Select Post-process Workflow

If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a post-process workflow from the **Pos-process Mapping** list and determine whether or not to run the workflow if the subscription fails.
2. If the workflow includes parameters, the parameters are shown in the **Post Process Parameters** area. If required, set the values of the parameters.
3. Click **Next**.
The **Join** page appears.

Step 4. Define Joins

Pull data from multiple topic tables into a single table by creating a join.

1. Optionally, create joins for the subscription. To create a join, perform the following actions:
 - a. Select the topic tables to join from the **Select Left Table** and **Select Right Table** lists.
 - b. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a target table. The new join appears on the **Join** page.
 - c. Select the column to join from the left table and then select the column to join from the right table. To add more columns to the join click the plus sign and then select the columns to join.
 - d. Click **Create Join**.
2. Repeat step [1](#) to create as many joins as required.
3. Click **Next**.
The **Target** page appears.

Step 5. Select Subscription Target

Select the type of the target to which Data Integration Hub writes data and choose the connection to the target in the **Target** page of the subscription wizard.

1. Select the data access connection and click **Next**.
The **Field Mapping** page appears.
2. Optionally, select **Create target tables if target tables do not exist**. If the target tables do not exist in the target database, Data Integration Hub creates them based on table mapping. Data Integration Hub does not create unmapped tables.

Step 6. View and Edit Subscription Field Mapping

View and edit the mapping of topic tables and fields to target tables and fields on the **Field Mapping** page of the subscription wizard. If you created joins in "Step 4. Define Joins", configure field mapping of joins to target tables and fields.

1. Perform the following actions to view and edit the mapping of a topic table or a join to a target table:
 - a. Click the edit table mapping icon.
The **Edit Table Mapping** dialog box appears.
 - b. Search for tables in the database. Note the following guidelines:
 - To search for tables by table name, enter a string in the **Find Target Table** text box and then click **Search**.
 - To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
 - To clear the search results and to show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
 - The search is not case sensitive.
 - You can search for a substring.

The **Search Results** section displays the target tables that match the search string and the schema that each table uses. The name of the topic table for which you are editing the mapping shows in row of the target table to which the topic table is mapped, in the **Mapped To** column.

- c. To map the topic table to a different target table, click the row of the target table to which you want to map the topic table and then click **Map Target Table**.

The **Edit Table Mapping** dialog box closes. The **Field Mapping** page shows the revised mapping.

2. Perform the following actions on the **Field Mapping** page to view and to map fields in a topic table to fields in a target table:

- a. Click the edit field mapping icon.

The **Edit Field Mapping** dialog box appears.

- b. To search for fields that are used in the topic table or in the target table, enter a string for the field name in the **Find Field** text box, and then click **Search**.

Choose **Display unmapped fields only** to show only unmapped fields in the search results.

Topic and target fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.

- c. To map a topic table field to a target table field, drag the field from the **Topic Table** section to the **Target Table** section.
- d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
- e. To clear the mapping of a field, in the **Actions** column, click the clear icon.
- f. To view field details, rest on the details icon to the left of a field name.
- g. To revert to the default Data Integration Hub table and field mapping, click **Auto Map**.
- h. Click **OK** to map the fields.

The **Edit Field Mapping** dialog box closes.

3. To revert to the default Data Integration Hub table and field mapping, click **Auto Map** in the **Field Mapping** page.
4. To clear all table and field mappings, click **Clear All** in the **Field Mapping** page.
Note: The subscription must contain at least one mapped topic table.
5. In the **Field Mapping** page, click **Next**.

The **Filter** page appears.

Step 7. Define a Filter

Define the data that the subscription consumes by setting filter conditions on table columns in the **Filter** page of the subscription wizard.

If data relations are defined for the topic to which the subscription subscribes, filters are applied to the tables for which the relations are defined, and the subscriber consumes only the data that is defined by the data relations.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
 - d. Click the plus icon to the right of the expression line.The condition shows in the condition list.
5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.The condition shows in the condition list.
6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 8. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.
3. Click **Next**.
The **Delivery** page appears.

Step 9. Define Delivery Options

Define the delivery options of the data to consume in the **Delivery** page of the subscription wizard.

1. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format. For subscriptions that run immediately after the published data is ready, go to step [2](#).
 - All available publications. Processes each published data set with a separate subscription mapping.
 - All available publications - aggregated. Groups and processes all published data sets and delivers a single data set.
 - Latest publication only. Delivers only the latest published data set.
2. Choose how Data Integration Hub handles data that exists in the target application.
 - Append the new data to the existing data in the target. Data Integration Hub adds rows to the table.
 - Insert new rows and update remaining rows. Data Integration Hub inserts new rows and then updates rows that exist in the target. Applicable if the database has a primary key.
 - Insert new rows and update changed rows. Data Integration Hub inserts new rows and then updates changed rows that exist in the target. Updating only changed rows slows down system performance. Applicable if the database has a primary key.

- Delete rows that don't exist in the topic from the target. Applicable if the type of topic to which the subscription subscribes is Full or if the subscription is an unbound subscription, and you choose to append new data or to insert new rows.
- Overwrite existing data in the target. Data Integration Hub truncates the target tables before it inserts new rows.

Note: If the new subscription subscribes to a topic where data exists and you want to select the option **Insert new rows and update remaining rows** or the option **Insert new rows and update changed rows**, perform the following steps:

1. Select the delivery option **Overwrite existing data in the target**, complete the subscription wizard, and save the subscription.
 2. Access the application to which the subscription subscribes, open the **Subscriptions** tab, and click the Get Previous Publications icon next to the subscription.
 3. In the Get Previous Publications dialing box, select **Include consumed publications** and then click **Run**.
 4. Edit the subscription to apply the required delivery option.
3. Click **Next**.
The **Summary** page appears.

Step 10. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating an Automatic Flat File Subscription

To create a subscription with an automatic mapping and a flat file target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. If you want to run a post-process on the subscription, select the post-process workflow.
4. Optionally, define joins for the subscription.
5. Select a target type and define the structure of the target file .
6. Define a filter for the subscription.
7. Define the subscription schedule.
8. Define the delivery scope and the delivery format for the subscription.
9. Review the subscription settings and save the subscription.

Step 1. Access the Create Automatic Flat File Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Flat File**, and then click **Create**.
The **Create Automatic Flat File Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Select whether to apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.
5. Click **Next**.
The **Processing** page appears.

Step 3. Select Post-process Workflow

If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a post-process workflow from the **Pos-process Mapping** list and determine whether or not to run the workflow if the subscription fails.
2. If the workflow includes parameters, the parameters are shown in the **Post Process Parameters** area. If required, set the values of the parameters.
3. Click **Next**.
The **Join** page appears.

Step 4. Define Joins

Pull data from multiple topic tables into a single table by creating a join.

1. Optionally, create joins for the subscription. To create a join, perform the following actions:
 - a. Select the topic tables to join from the **Select Left Table** and **Select Right Table** lists.
 - b. Select the join type from the **Select Join Type** list and then enter the join name in the **Join Name** field. The name that you assign to the join shows in the **Field Mapping** page, where you can map the join to a target table. The new join appears on the **Join** page.
 - c. Select the column to join from the left table and then select the column to join from the right table. To add more columns to the join click the plus sign and then select the columns to join.

- d. Click **Create Join**.
2. Repeat step [1](#) to create as many joins as required.
3. Click **Next**.

The **Target** page appears.

Step 5. Select and Configure Subscription Target

Select the type of the target to which Data Integration Hub writes data and define the location and the structure of the target file in the **Target** page of the subscription wizard.

1. Select the target type from the **Target Type** list.
If you select the HDFS target type, the **HDFS Connection** field appears on the page.
2. If you select the HDFS target type, select a connection from the **HDFS Connection** list.
3. To use file transfer to deliver files to remote servers, select **Use File Transfer**.
The **Connection** field appears on the page.
4. If you selected to use file transfer, select a connection from the **Connection** list.
5. Define the location and the structure of the target file:

Directory

Location where Data Integration Hub creates the target file or files. Data Integration Hub generates a file for each table in the topic.

Target Filename Pattern

Pattern of the names of the target files. For more information, see [“Target File Name Pattern for Flat Files” on page 144](#).

Code Page

Character encoding used in the file.

Use column names as the first line

Use the first line in the source data as the table header.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** and then define a custom delimiter. For information about the supported column delimiters, see the sections about delimited files in the *PowerCenter Designer Guide*.

Text Qualifier

Optional. Symbols used in the file to enclose a string.

Thousands Separator

Optional. Symbol used in the file as a thousands separator.

Decimal Separator

Symbol used in the file as a decimal separator.

Datetime Format

Date and time format used in the file. Select a predefined format or select **Other** to define a custom format. For information about the supported datetime formats, see the chapter about dates in the *PowerCenter Transformation Language Reference*.

Note: The datetime format can contain up to 50 characters.

6. Click **Next**.

The **Field Mapping** page appears.

Target File Name Pattern for Flat Files

When you define a flat file target, the pattern of the file name, which you define in the **Target Filename Pattern** field, can contain the following variables:

Variable	Description
(\$Table_Name)	Name of the table in the topic that contains the data. Use this variable in topics that include more than one table to create an output file for each table column.
(\$Time_Stamp)	Date and time when the subscription runs, in the following format: yyyy-mm-dd-hh24_mi_ss
(\$Publication_Instance_Date)	Date and time when the publication instance is created, in the following format: yyyy-mm-dd-hh24_mi_ss Use this variable to differentiate between multiple publication instances and to sort publications by date.

For example, the file name pattern `file_($Table_Name)_($Publication_Instance_Date).txt` creates the following file name:

```
file_ORDERS_2015-03-27-12_43_26.txt
```

Step 6. View and Edit Subscription Field Mapping

View and edit the mapping of topic tables and fields to target tables and fields on the **Field Mapping** page of the subscription wizard. If you created joins in "Step 4. Define Joins", configure field mapping of joins to target tables and fields.

1. Perform the following actions on the **Field Mapping** page to view and to map fields in a topic table to fields in a target table:
 - a. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears.
 - b. To search for fields that are used in the topic table or in the target table, enter a string for the field name in the **Find Field** text box, and then click **Search**.
Choose **Display unmapped fields only** to show only unmapped fields in the search results.
Topic and target fields that match the search string appear. A green checkmark next to a field name indicates that the field is mapped.
 - c. To map a topic table field to a target table field, drag the field from the **Topic Table** section to the **Target Table** section.
 - d. To add an expression to a field, in the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
 - e. To clear the mapping of a field, in the **Actions** column, click the clear icon.

- f. To view field details, rest on the details icon to the left of a field name.
- g. To revert to the default Data Integration Hub table and field mapping, click **Auto Map**.
- h. Click **OK** to map the fields.

The **Edit Field Mapping** dialog box closes.

2. To revert to the default Data Integration Hub table and field mapping, click **Auto Map** in the **Field Mapping** page.
3. To clear all table and field mappings, click **Clear All** in the **Field Mapping** page.

Note: The subscription must contain at least one mapped topic table.

4. In the **Field Mapping** page, click **Next**.

The **Filter** page appears.

Step 7. Define a Filter

Define the data that the subscription consumes by setting filter conditions on table columns in the **Filter** page of the subscription wizard.

If data relations are defined for the topic to which the subscription subscribes, filters are applied to the tables for which the relations are defined, and the subscriber consumes only the data that is defined by the data relations.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
 - d. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.

The condition shows in the condition list.

6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 8. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.
3. Click **Next**.

The **Delivery** page appears.

Step 9. Define Delivery Options

Define the delivery options of the data to consume in the **Delivery** page of the subscription wizard.

1. For subscriptions that run either manually or by an external trigger or that run by schedule, choose the data delivery scope and delivery format. For subscriptions that run immediately after the published data is ready, go to step [2](#).
 - All available publications. Processes each published data set with a separate subscription mapping.
 - All available publications - aggregated. Groups and processes all published data sets and delivers a single data set.
 - Latest publication only. Delivers the latest published data set.
2. Choose how Data Integration Hub handles data that exists in the target application. The selection is disabled for subscriptions that subscribe to a topic with a Delta publication type and are configured to consume all available publications, and for subscriptions that use file transfer.
 - Append the new data to the existing data in the target. Data Integration Hub appends records to the file.
 - Overwrite existing data in the target. Data Integration Hub overwrites the file.
3. Click **Next**.
The **Summary** page appears.

Step 10. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.
The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating an Automatic Pass-through File Subscription

To create a subscription with an automatic mapping and a pass-through file target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. If you want to run a post-process on the subscription, select the post-process workflow.
4. Configure the subscription target.
5. Define a filter for the subscription.
6. Define the subscription schedule.
7. Define the delivery scope of the subscription.
8. Review the subscription settings and save the subscription.

Task Prerequisites

Before you start this task, verify the following prerequisites:

- The subscription mapping and the task are configured in Informatica Intelligent Cloud Services.
- You have created an application.
- You have created a topic.

Step 1. Access the Create Automatic Pass-through File Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Pass-through File**, and then click **Create**.
The **Create Automatic Pass-through File Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Click **Next**.
The **Processing** page appears.

Step 3. Select Post-process Workflow

If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a post-process workflow from the **Pos-process Mapping** list and determine whether or not to run the workflow if the subscription fails.
2. If the workflow includes parameters, the parameters are shown in the **Post Process Parameters** area. If required, set the values of the parameters.
3. Click **Next**.
The **Target** page appears.

Step 4. Configure Subscription Target

Configure the target to which Data Integration Hub delivers files in the **Target** page of the subscription wizard.

1. To use file transfer to deliver files to remote servers, select **Use File Transfer**.
The **Connection** field appears on the page.
2. If you selected to use file transfer, select a connection from the **Connection** list.
3. Enter the location of the file or files that Data Integration Hub delivers in the **Directory** field.
4. Click the edit icon next to a topic table to assign a target file to the table. By default, Data Integration Hub assigns the name of the topic table as the file name, in the following format:
<topic_table_name>_(\$sequence). For example, if the name of the topic table is `orderId`, the default name of the file that is assigned to the table is `orderId_($sequence)`. You can edit the <topic_table_name> section of file name.

The pattern of the file name can contain the following variables:

Variable	Description
(\$sequence)	Required. Use this variable to differentiate between multiple publication instances.
(\$Time_Stamp)	Optional. Date and time when the subscription runs, in the following format: yyyy-mm-dd-hh24_mi_ss

For example:

The file name pattern `file_($Time_Stamp)_($sequence).txt` creates the following file name:

`file_2015-03-27-12_43_26_1260.txt`

You must assign a target file for at least one topic table.

5. Repeat step [4](#) for each table topic that appears on the **Target** page.
6. Click **Next**.
The **Filter** page appears.

Step 5. Define a Filter

Define the data that the subscription consumes by setting filter conditions on the file metadata.

You can use basic expressions and advanced expressions to define filter conditions. You can use more than one method to add filters to a table row.

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an `Orders` table, add a condition that the date in the `ShippedDate` column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an `Orders` table, add the following expression:

`(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))`

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

1. Choose the type of expression that you want to create.
2. Select the table to which to apply the filter from the **Select Table** list.
3. Enter the condition parameters according to the method that you selected:
 - Basic expression: go to step [4](#).
 - Advanced expression: go to step [5](#).
4. To create a basic expression, perform the following steps:
 - a. Select the column to filter from the **Select Column** list.
 - b. Select the filter operator from the **Select Operator** list. The available operators depend on the type of content in the column.
 - c. When **Select Value** is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.
 - d. Click the plus icon to the right of the expression line.The condition shows in the condition list.
5. To create an advanced expression, perform the following steps:
 - a. Enter a valid PowerCenter expression in the text field. Alternatively, prepare a basic expression, select **Advanced Expression**, and then, in the expression text area, complete the expression. The expression is limited to the selected table. The expression can contain up to 1024 characters.
 - b. Click the plus icon to the right of the expression line.The condition shows in the condition list.
6. Repeat steps [1](#) through [5](#) to add the required conditions.
7. Click **Next**.

The **Schedule** page appears.

Step 6. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.

- **Daily.** Runs the subscription at the same hour every day.
- **Weekly.** Runs the subscription every week on one or more days at the same hour.
- **Monthly.** Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Optionally, define a retry policy. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.
3. Click **Next**.
The **Delivery** page appears.

Step 7. Define Delivery Scope

Define the delivery scope of the data to consume in the **Delivery** page of the subscription wizard.

1. Choose the data delivery scope:
 - All available publications. Processes each published data set with a separate subscription mapping.
 - Latest publication only. Delivers only the latest published data set.
2. Click **Next**.
The **Summary** page appears.

Step 8. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.
The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating an Automatic Data-driven Subscription

To create an automatic data-driven subscription, perform the following tasks:

1. Access the create subscription page.
2. Configure subscription properties and select the topic from which to consume the data.
After you configure the subscription properties, you can copy the following URLs from the create subscription page:
 - REST URL of the API. Use this URL to subscribe to the data.
 - REST URL of the Swagger structure for the topic from which the subscription consumes data.

You use the URLs when you create the request that runs the subscription.

3. Create a request to run the subscription.

Step 1. Access the Create Automatic Data-driven Subscription Page

Access the create subscription page in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Automatic > Data-driven**, and then click **Create**.
The **Create Automatic Data-driven Subscription** page appears.

Step 2. Define Subscription Properties

Define subscription properties and select a topic in the create subscription page.

1. Enter the subscription name.
The following URLs are automatically updated:
 - REST URL of the API. Use this URL to subscribe to the data.
 - REST URL of the Swagger structure for the topic from which the subscription consumes data.You use the URLs when you create the request that runs the subscription.
2. Optionally, enter a description of the subscription.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Enter the URL to where Data Integration Hub sends notifications when data is ready to consume.
5. Click **Save**.

Step 3. Create a Request to Run the Subscription

To create a request to run the subscription, you have to copy the URL of the REST API and the URL of the API Swagger file from the subscription that you created in [“Step 2. Define Subscription Properties” on page 152](#).

- Create a POST request with the following details:
- Request URL: REST API URL. For example:

```
http://hostname:18080/dih-console/api/v1/subscription/MySubscription/data
```


Where `MySubscription` is the subscription name.
 - Request body:

```
{"batchSize":<records_max_number> } :
```


Where `records_max_number` is the maximal number of records from a table that can be put on the HTTP response stream before flushing it.

For example:

```
{ "batchSize": 5 } :
```

If `records_max_number` is left empty, Data Integration Hub applies the default batch size of 500 records with a limit of 5000 records.

Creating a Custom Batch Subscription

To create a subscription with a custom mapping and a batch workflow, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. Select the subscription workflow. If you use file transfer, select the connection to the target to which Data Integration Hub writes the files. If you want to run a post-process on the subscription, select the post-process workflow.
4. If the workflow includes parameters, set subscription parameter values.
5. Define the subscription schedule.
6. Review the subscription settings and save the subscription.

Step 1. Access the Create Custom Batch Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Custom > Batch**, and then click **Create**.
The **Create Custom Batch Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select Topics

Define subscription properties and select topics in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select the topic to which you want to subscribe and then click **Add Topic**.
The topic appears in the topic table.
4. Optionally, subscribe to additional topics and create a compound subscription.

In compound subscriptions that consume data from multiple topics, topic information includes a **Mandatory** option that determines whether the topic is mandatory or not. Data Integration Hub waits for all mandatory topics to be available for consumption before it runs the subscription. By default, all topics in a compound subscription are mandatory. If a topic is not mandatory for the subscription, clear the **Mandatory** option for the topic.

5. If the subscription is a compound subscription, in the **Wait for all mandatory topics to be available for consumption for ... minutes** field, specify the number of minutes to wait from the time the first topic is ready to consume until all mandatory topics are ready to consume.
6. Click **Next**.
The **Processing** page appears.

Step 3. Select Subscription Workflows

Select a Data Integration Hub subscription workflow in the **Processing** page of the subscription wizard. If you want to run a post-process on the subscription, select the post-process workflow. If required, set parameter values.

1. Select a subscription workflow from the **Custom Mapping** list. Only subscription workflows that are based on a batch workflow are available for selection.
If the workflow supports file transfer, the **Use File Transfer** option is enabled.
2. To use file transfer, select the **Use File Transfer** option and then select the connection to the target to which Data Integration Hub writes the files in the **Connection** field.
3. To run a post-process on the subscription, select a post-process workflow from the **Pos-process Mapping** list and determine whether or not to run the workflow if the subscription fails.
4. If the workflows include parameters, the parameters are shown on the page. Set the values of the parameters in the **Subscription Parameters** area and the **Post Process Parameters** area, as applicable.
5. Click **Next**.
The **Schedule** page appears.

Step 4. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard. If you create a compound subscription, where the subscription consumes data sets from multiple topics, you can only choose to consume data when it is published.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Note: This option is not applicable to subscriptions that subscribe to a Hadoop-based topic.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.

- **Daily.** Runs the subscription at the same hour every day.
- **Weekly.** Runs the subscription every week on one or more days at the same hour.
- **Monthly.** Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 5. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating a Custom Cloud Subscription

To create a subscription with a custom mapping with an Informatica Intelligent Cloud Services task, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. Select the Informatica Intelligent Cloud Services task that runs the subscription mapping.
4. Define the subscription schedule.
5. Review the subscription settings and save the subscription.

Step 1. Access the Custom Cloud Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Custom > Cloud**, and then click **Create**.
The **Create Custom Cloud Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select Topics

Define subscription properties and select topics in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select the topic to which you want to subscribe and then click **Add Topic**.

The topic appears in the topic table.

4. Optionally, subscribe to additional topics and create a compound subscription.

In compound subscriptions that consume data from multiple topics, topic information includes a **Mandatory** option that determines whether the topic is mandatory or not. Data Integration Hub waits for all mandatory topics to be available for consumption before it runs the subscription. By default, all topics in a compound subscription are mandatory. If a topic is not mandatory for the subscription, clear the **Mandatory** option for the topic.

5. If the subscription is a compound subscription, in the **Wait for all mandatory topics to be available for consumption for ... minutes** field, specify the number of minutes to wait from the time the first topic is ready to consume until all mandatory topics are ready to consume.
6. Click **Next**.

The **Processing** page appears.

Step 3. Select Subscription Mappings

Select an Informatica Cloud task that defines the subscription mapping in the **Processing** page of the subscription wizard. If you want to run a post-process on the subscription, select the post-process workflow.

1. Select a task from the **Informatica Cloud Task** list.
2. Click **Next**.

The **Schedule** page appears.

Step 4. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard. If you create a compound subscription, where the subscription consumes data sets from multiple topics, you can only choose to consume data when it is published.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Note: This option is not applicable to subscriptions that subscribe to a Hadoop-based topic.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 5. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating a Modular Cloud Subscription

To create a subscription with a modular mapping and a cloud target, perform the following tasks:

1. Access the create subscription wizard.
2. Define basic subscription properties and select the topic to which you want to subscribe.
3. Select the subscription mapping.
4. If required, configure the subscription source.
5. If required, configure the subscription target.
6. If the mapping includes parameterized transformations, set the parameter values.
7. If required, configure field mapping.
8. Define the subscription schedule.
9. Review the subscription settings and save the subscription.

Step 1. Access the Modular Cloud Subscription Wizard

Access the create subscription wizard in the Data Integration Hub Operation Console.

1. In the Navigator, click **Hub Management** > **Applications**.
The **Applications** page appears.
2. Click the name of the application to which you want to deliver content.
The **Edit Application** page appears.
3. Select the **Subscriptions** tab. Click **New**, then select **Modular** > **Cloud**, and then click **Create**.
The **Create Modular Cloud Subscription** wizard appears.

Step 2. Define Basic Subscription Properties and Select a Topic

Define subscription properties and select a topic in the **General** page of the subscription wizard.

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select a topic from the **Topic** list.
The **Topic Structure** area shows the structure of the topic from which Data Integration Hub delivers the data. You can view the structure of all the tables in the topic, or choose a table for which to show the structure.
4. Select whether to apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.
5. Click **Next**.
The **Processing** page appears.

Step 3. Select Subscription Mapping

Select an Informatica Cloud subscription mapping in the **Processing** page of the subscription wizard.

1. Select a mapping from the **Cloud Mapping** list.
2. Click **Next**.
The **Source** page appears.

Step 4. Configure Subscription Source

If the subscription source is not configured by the mapping, configure the source in the **Source** page of the subscription wizard. The source of the subscription is the Data Integration Hub publication repository.

1. Configure source settings as applicable.
2. Click **Next**.
The **Target** page appears.

Step 5. Configure Subscription Target

If the subscription target is not configured by the mapping, configure the target in the **Target** page of the subscription wizard.

1. Configure target settings as applicable.

2. Click **Next**.

The **Input Parameters** page appears.

Step 6. Set Subscription Parameters

If the mapping contains parameters, the parameters show in the **Input Parameters** page of the subscription wizard. Set parameter values as applicable.

1. Click the edit icon next to the parameter for which to define a value.
2. In the **Edit Parameter** dialog box, define the parameter value in the **Expression** area. Click a field in the **Fields** area to add it to the expression.

Click **OK**.

The parameter value shows in the **Input Parameters** page.

3. Repeat steps [1](#) through [2](#) to set the required parameter values.
4. Click **Next**.

The **Field Mapping** page appears.

Step 7. Configure Subscription Field Mapping

If field mapping is not configured by the mapping, map topic table fields to target table fields on the **Field Mapping** page of the subscription wizard.

1. Click the edit field mapping icon.
The **Edit Field Mapping** dialog box appears.
2. To select to view all the fields, mapped fields, or unmapped fields in the topic table and the target table, select the relevant option from the **Show** lists.
3. To search for fields that are used in the topic table and the target table, enter a string for the field name in **Search Fields**, and then press Enter. The search is not case sensitive. You can search for a substring.
4. To map a topic table field to a target table field, drag the field from the **Default** section to the target table section.
5. To add an expression to a field, in the **Mapped Field/Expression** column, click the expression icon to open the **Field Expression** dialog box, and then select fields and functions to add to the field. To validate the expression, click **Validate**.
6. Click **OK** to apply the mapping.
The **Edit Field Mapping** dialog box closes.
7. In the **Field Mapping** page, click **Next**.

The **Schedule** page appears.

Step 6. Define Subscription Schedule

Define the method and the frequency of the subscription in the **Schedule** page of the subscription wizard.

1. Select the method and the frequency of the subscription.

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

2. Click **Next**.

The **Summary** page appears.

Step 10. Review Subscription Settings and Save the Subscription

Review the subscription settings and save the subscription in the **Summary** page of the subscription wizard.

1. Review the subscription settings.
2. Click **Finish**.

The subscription wizard closes. The **Subscription** tab of the **Edit Application** page shows the subscription you created. If you configured a subscription schedule, the subscription consumes data according to the defined schedule.

Creating a Custom Pass-through Kafka Subscription

This topic describes how to create a custom pass-through Kafka subscription.

In order to create a subscription to consume data streams into Apache Kafka, you must create a topic that monitors the Apache Kafka streaming.

1. Access the create subscription page.
2. Define the basic subscription properties.
3. Save the subscription.

Step1. Access the Create Custom Pass-through Kafka Subscription Page

Perform the following tasks to create a custom pass-through Kafka subscription:

1. In the Navigator, click **Hub Management > Applications**.
The **Applications** page appears.
2. Click the name of the application from which you want to publish content.
The **Edit Application** page appears.
3. Click the **Subscriptions** tab and click **New** and then select **Custom > Pass-through Kafka > Create**.
The custom pass-through Kafka subscription wizard appears.

Step2. Define Basic Subscription Properties

Perform the following tasks to define the basic subscription properties:

1. Enter the subscription name.
2. Optionally, enter a description of the subscription.
3. Select the topic to which you want to subscribe and then click **Add Topic**.
4. Click **Save**.
The subscription is saved in the Subscriptions tab.

CHAPTER 12

Subscription Properties

This chapter includes the following topics:

- [Subscription Properties Overview, 162](#)
- [Subscription General Properties, 162](#)
- [Subscription Processing Properties, 164](#)
- [Subscription Join Properties, 166](#)
- [Subscription Source Properties, 166](#)
- [Subscription Target Properties, 166](#)
- [Subscription Field Mapping Properties, 169](#)
- [Subscription Filter Properties, 172](#)
- [Subscription Schedule Properties, 173](#)
- [Subscription Delivery Properties, 173](#)
- [Subscription Wizard Summary Page, 175](#)
- [Properties of a Data-driven Subscription, 175](#)

Subscription Properties Overview

Use the Subscription wizard or page to create or edit a subscription. The wizard or page you use depends on the type of subscription that you create or edit.

Each type of Subscription wizard or page contains the properties that are relevant to the subscription type.

Subscription General Properties

The **General** page of the subscription wizard can include the following properties:

Subscription Name

Name of the subscription. The name is not case-sensitive and can contain up to 64 characters. The name can contain only alphanumeric characters and underscores.

You cannot change the subscription name after you create the subscription.

Description

Textual description of the subscription.

Mode

Mode of the subscription. Choose one of the following options:

- **Enabled.** The subscription runs according to the schedule that you define in the **Schedule** page. You can also run the subscription from within the Operation Console.
- **Disabled.** The subscription does not run according to schedule or by an external API. You can only run the subscription from within the Operation Console.

Unbound subscription

An unbound subscription is not restricted to specific publication instances. It subscribes to all the data that a publication publishes and consumes the data based on the defined filter, regardless of when or in what batch the data or the files were published.

Select Topic

Select the topic to which to subscribe.

For pass-through file subscriptions, only topics that store published data on a file publication repository are available for selection. For all other subscription types, both topics that store published data on a relational database and topics that store published data on a Hadoop repository are available for selection.

For custom batch and cloud subscriptions, where you can select to subscribe to multiple topics, click **Add Topic** to add the topic to the topic list. You must add at least one topic to the list.

Topic Structure

Details of the topic or the topics to which the subscription subscribes. The information that appears for each topic depends on the type of subscription that you create or edit and on the topic definition and appears in read-only mode.

In compound subscriptions that consume data from multiple topics, topic information includes a **Mandatory** option that determines whether the topic is mandatory or not. Data Integration Hub waits for all mandatory topics to be available for consumption before it runs the subscription. By default, all topics in a compound subscription are mandatory. If a topic is not mandatory for the subscription, clear the **Mandatory** option for the topic.

Wait for all mandatory topics to be available for consumption for ... minutes

In compound subscriptions, the maximum time to wait from the time that the first topic is ready to consume until all mandatory topics are ready to consume. If all the mandatory topics are ready within the defined wait time, Data Integration Hub runs the subscription as soon as the last topic is ready, regardless of the remaining wait time. If not all the mandatory topics are ready to consume within the defined wait time, Data Integration Hub doesn't run the subscription and creates an error event.

Apply PK/FK

Apply primary and foreign keys to topic table columns where topic table relations are defined. If joins are manually defined for the subscription, the keys are not applied.

Subscription Processing Properties

Processing properties depend on the type of subscription mapping and the type of data flow engine that the subscription uses:

- [“Subscription Processing Properties for Automatic Mappings with Data Integration Hub Workflows” on page 164](#)
- [“Subscription Processing Properties for Automatic Mappings with Informatica Intelligent Cloud Service Mappings” on page 164](#)
- [“Subscription Processing Properties for Custom Mappings with Data Integration Hub Workflows” on page 165](#)
- [“Subscription Processing Properties for Custom Mappings with Informatica Intelligent Cloud Service Tasks” on page 165](#)

Subscription Processing Properties for Automatic Mappings with Data Integration Hub Workflows

The **Processing** page of the subscription wizard can include the following properties for automatic subscriptions that use a Data Integration Hub workflow:

Post-process Mapping

Select a workflow for the post-process from the list. Only subscription post-processing workflows are available for selection.

Run workflow even if subscription fails

Determines whether or not the post-process mapping runs if the subscription fails.

Post Process Parameters

The content of the **Post Process Parameters** area depends on the parameters of the PowerCenter workflow. The developer imports the workflows to Data Integration Hub with the parameter definitions. The developer also determines the layout of the area with the Forms Designer. If required, set the parameter values.

Subscription Processing Properties for Automatic Mappings with Informatica Intelligent Cloud Service Mappings

The **Processing** page of the subscription wizard includes the following properties for custom subscriptions that use an Informatica Intelligent Cloud Service task:

Cloud Mapping

Select an Informatica Intelligent Cloud Service mapping for the subscription from the list.

Subscription Processing Properties for Custom Mappings with Data Integration Hub Workflows

The **Processing** page of the subscription wizard can include the following properties for custom subscriptions that use a Data Integration Hub workflow:

Custom Mapping

Select a workflow for the subscription from the list. Only subscription workflows that are applicable to the type of subscription that you create are available for selection. For example, if you create a batch workflow subscription, only batch workflows are listed.

Use File Transfer

Use file transfer to consume files into remote applications.

Connection

Applicable for file transfer. Select the connection to where the subscription workflow writes the files that the subscription consumes.

Post-process Mapping

Select a workflow for the post-process from the list. Only subscription post-processing workflows are available for selection.

Run workflow even if subscription fails

Determines whether or not the post-process mapping runs if the subscription fails.

Subscription Parameters

The content of the **Subscription Parameters** area depends on the parameters of the PowerCenter workflow, Data Engineering Integration mapping or workflow, or Data Quality mapping or workflow. The developer imports the workflows and mappings to Data Integration Hub with the parameter definitions. The developer also determines the layout of the area with the Forms Designer. If required, set the parameter values.

Post Process Parameters

The content of the **Post Process Parameters** area depends on the parameters of the PowerCenter workflow. The developer imports the workflows to Data Integration Hub with the parameter definitions. The developer also determines the layout of the area with the Forms Designer. If required, set the parameter values.

Subscription Processing Properties for Custom Mappings with Informatica Intelligent Cloud Service Tasks

The **Processing** page of the subscription wizard includes the following properties for custom subscriptions that use an Informatica Intelligent Cloud Service task:

Informatica Intelligent Cloud Service Task

Select an Informatica Cloud Task for the subscription from the list.

Post-process Mapping

Select a workflow for the post-process from the list. Only subscription post-processing workflows are available for selection.

Run workflow even if subscription fails

Determines whether or not the post-process mapping runs if the subscription fails.

Post Process Parameters

The content of the **Post Process Parameters** area depends on the parameters of the PowerCenter workflow. The developer imports the workflows to Data Integration Hub with the parameter definitions. The developer also determines the layout of the area with the Forms Designer. If required, set the parameter values.

Subscription Join Properties

The **Join** page of the subscription wizard can include the following properties:

Define new joins by joining existing topic tables

Pull data from multiple topic tables into a single table by creating a join.

Subscription Source Properties

For modular cloud subscriptions, the source properties depend on the settings of the Informatica Intelligent Cloud Services mapping that you select for the subscription.

Subscription Target Properties

Target properties depend on the type of target to which Data Integration Hub writes data in an automatic mapping:

- [“Subscription Target Properties for Relational Database” on page 166](#)
- [“Subscription Target Properties for Flat File” on page 167](#)
- [“Subscription Target Properties for Pass-through File” on page 168](#)

For modular cloud subscriptions, the target properties depend on the settings of the Informatica Intelligent Cloud Services mapping that you select for the subscription.

Subscription Target Properties for Relational Database

The **Target** page of the subscription wizard can include the following properties when you use an automatic mapping with a relational database source:

Database Connection

Choose the data access connection to where the subscription workflow writes the data.

Subscription Target Properties for Flat File

The **Target** page of the subscription wizard can include the following properties when you use a flat file target:

Target Type

Type of subscription target.

Use File Transfer

Applicable for a flat file target. Use file transfer to deliver files to remote servers that reside outside the Data Integration Hub network. Data Integration Hub supports the FTP, SFTP, and FTPS protocols for file transfer.

Connection

Applicable for a flat file target. Name of the file transfer connection to where the subscription workflow writes the files that the subscription consumes.

HDFS Connection

Applicable for an HDFS target. Select the HDFS connection to where the subscription workflow writes the files that the subscription consumes.

Directory

Location where Data Integration Hub creates the target file or files. Data Integration Hub generates a file for each table in the topic.

Target Filename Pattern

Pattern of the names of the target files. For more information, see [“Target File Name Pattern for Flat Files” on page 144](#).

Code Page

Character encoding used in the file.

Use column names as the first line

Use the first line in the source data as the table header.

Delimiter

Delimiter used in the file to separate between columns. Select a predefined delimiter or select **Custom** and then define a custom delimiter. For information about the supported column delimiters, see the sections about delimited files in the *PowerCenter Designer Guide*.

Text Qualifier

Optional. Symbols used in the file to enclose a string.

Thousands Separator

Optional. Symbol used in the file as a thousands separator.

Decimal Separator

Symbol used in the file as a decimal separator.

Datetime Format

Date and time format used in the file. Select a predefined format or select **Other** to define a custom format. For information about the supported datetime formats, see the chapter about dates in the *PowerCenter Transformation Language Reference*.

Note: The datetime format can contain up to 50 characters.

Target File Name Pattern for Flat Files

When you define a flat file target, the pattern of the file name, which you define in the **Target Filename Pattern** field, can contain the following variables:

Variable	Description
(\$Table_Name)	Name of the table in the topic that contains the data. Use this variable in topics that include more than one table to create an output file for each table column.
(\$Time_Stamp)	Date and time when the subscription runs, in the following format: yyyy-mm-dd-hh24_mi_ss
(\$Publication_Instance_Date)	Date and time when the publication instance is created, in the following format: yyyy-mm-dd-hh24_mi_ss Use this variable to differentiate between multiple publication instances and to sort publications by date.

For example, the file name pattern `file_($Table_Name)_($Publication_Instance_Date).txt` creates the following file name:

```
file_ORDERS_2015-03-27-12_43_26.txt
```

Subscription Target Properties for Pass-through File

The **Target** page of the subscription wizard can include the following properties when you use a pass-through file target:

Use File Transfer

Use file transfer to deliver files to remote servers that reside outside the Data Integration Hub network. Data Integration Hub supports the FTP, SFTP, and FTPS protocols for file transfer.

Connection

Applicable when you use file transfer. Name of the file transfer connection to where the subscription workflow writes the files that the subscription consumes.

Directory

Location where Data Integration Hub writes the target file.

Target Definition area

Click the edit icon next to a topic table to assign a target file to the table. By default, Data Integration Hub assigns the name of the topic table as the file name, in the following format:

`<topic_table_name>_($sequence)`. For example, if the name of the topic table is `orderId`, the default name of the file that is assigned to the table is `orderId_($sequence)`. You can edit the `<topic_table_name>` section of file name.

The pattern of the file name can contain additional variables. For more information, see [“Target File Name Pattern for Pass-through Files” on page 169](#).

You must assign a target file for at least one topic table.

Target File Name Pattern for Pass-through Files

When you define a pass-through file target, the pattern of the file name, which you define in the **File Name** field, can contain the following variables:

Variable	Description
(\$sequence)	Required. Use this variable to differentiate between multiple publication instances and to sort publications by order.
(\$Time_Stamp)	Optional. Date and time when the subscription runs, in the following format: yyyy-mm-dd-hh24_mi_ss

For example:

The file name pattern `file_($Time_Stamp)_($sequence).txt` creates the following file name:

`file_2015-03-27-12_43_26_1260.txt`

Subscription Field Mapping Properties

The **Field Mapping** page of the subscription wizard can include the following properties:

Field Mapping

Shows the total number of tables in the topic, which is associated with the subscription, and the number of tables that are mapped.

Auto Map

Data Integration Hub maps the tables and fields according to table and field names. If a table name or a field name in the target is identical to a name in the topic, Data Integration Hub maps between the tables or the fields. Data Integration Hub does not map tables and fields that do not match, and you can manually map them.

Clear All

Clears table and field mapping.

Note: The subscription must contain at least one mapped target table.

Mapping table

Lists all tables in the topic, which is associated with the subscription, and the target tables to which the topic tables are mapped. Use the mapping table to manually map topic tables to target tables, to map topic fields to target fields, and to clear the mappings.

Each line in the mapping table contains the following elements:

- Topic Table. Name of the topic table and the number of fields in the topic table.
- Mapped Target Table. Name of the target table and the number of fields in the target table.
- Mapped Fields. Number of fields that are mapped between the topic table and the target table.
- Edit Table Mapping. Click the edit icon to edit the table mapping. For more information, see the section "Edit Table Mapping for Subscriptions".
This option is disabled for subscriptions with a flat file source.

- **Edit Field Mapping.** Click the edit icon to edit the field mapping. A topic table must be mapped to a target table before you can edit the field mapping of the table. For more information, see the section "Edit Field Mapping for Subscriptions".
- **Clear Mapping.** Clear the mapping between the topic table and the target table.

Note: If the target does not contain any of the topic tables or if the selected connection does not contain metadata access details, Data Integration Hub generates the subscription mapping based on the topic tables, and the table mapping and field mapping editing options are disabled. To change the default mapping, click **Clear All**, and then edit the table and field mapping.

Edit Table Mapping for Subscriptions

For subscriptions with an automatic mapping and a relational database target, manually map topic tables to target tables.

The **Field Mapping** page of the subscription wizard shows the tables in the topic that is associated with the subscription, and the mapping between topic tables and target tables. When you click the edit table mapping icon next to a table, the **Edit Table Mapping** dialog box opens.

Edit Table Mapping Properties for Subscriptions

To map a topic table to a target table, find a target table and map it to the topic table that you are editing.

The following list describes the properties of the **Edit Table Mapping** dialog box:

Find Target Table

Search for tables in the database. Note the following guidelines:

- To search for tables by table name, enter a string in the **Find Target Table** text box and then click **Search**.
- To search for tables that use a schema other than the default schema, clear the option **Show default schema only**, enter a string in the **Schema** text box, and then click **Search**.
- To clear the search results and to show only tables that use the default schema, select the option **Show default schema only** and then click **Show All**.
- The search is not case sensitive.
- You can search for a substring.

Search Results

Lists the target tables that match your search and shows any mapping between the listed target tables and topic tables. Select a table to map from the list.

Map Target Table

Maps between the table that you select in **Search Results** and the topic table for which you edit the mapping.

Edit Field Mapping for Subscriptions

For subscriptions with an automatic mapping, manually map fields in topic tables to fields in target tables.

The **Field Mapping** page of the subscription wizard shows the tables in the topic that is associated with the subscription, and the mapping between topic tables and target tables. When you click the edit field mapping icon next to a table, the **Edit Field Mapping** dialog box opens.

Note: For subscriptions with a relational database or a flat file target, a topic table must be mapped to a target table before you can edit the field mapping for the table.

The published data includes the publication ID and the publication time stamp, in the `DIH_PUBLICATION_INSTANCE_ID` and `DIH_PUBLICATION_INSTANCE_DATE` fields. The resolution of the time stamp is seconds. By default, those fields are not mapped. If you map these fields, verify that the target structure includes these fields. The target application can use these fields to differentiate between multiple publication instances and to sort publications by date.

Note: Subscribers that consume the publication time stamp into an Oracle database: By default, Oracle clients display the event date and do not display the time. To display the complete time stamp, configure the client display options.

Edit Field Mapping Properties for Subscriptions

To map fields in a table, select the target table field and the topic table field from the corresponding columns of the **Edit Field Mapping** dialog box and map them to each other.

The following list describes the elements of the **Edit Field Mapping** dialog box.

Find Field

Searches fields by name in both the topic table and the target table. Enter the string for which to search, and then click **Search**. The search is not case sensitive. You can search for a substring.

After you enter a string in the **Find Field** text box, you can click **X** to clear the search and show all fields.

Choose **Display unmapped fields only** to show only unmapped fields in the search results.

Auto Map

Data Integration Hub maps the fields according to field names. If a name in the target is identical to a name in the topic, Data Integration Hub maps between the fields. Data Integration Hub does not map fields that do not match, and you can manually map them.

Clear All

Clears the field mapping. In subscriptions with a relational database target, fields that are not mapped will not be written to the subscribing database. In subscriptions with a flat file target, the target file will not include unmapped fields.

Note: The table must contain at least one mapped field.

Topic Table

Lists the fields in the topic table that match your search. Mapped fields are denoted by a green checkmark. Select the field that you want to map.

Rest on the details icon to the left of a field name to view the details of the field.

Target Table

Lists the fields in the target table that match your search, and shows the mapping between the listed target fields and topic fields. Mapped fields are denoted by a green checkmark. Select the field that you want to map.

You can perform the following actions on a target field:

- Add an expression to the field. In the **Actions** column, click the expression icon to open the **Expression Editor**, and then select fields and functions to add to the field.
- Clear the mapping of the field. In the **Actions** column, click the clear icon.
- View field details. Rest on the details icon to the left of a field name.

Subscription Filter Properties

The **Filter** page of the subscription wizard can include the following properties:

Basic Expression

Use this method if you want to apply a condition to a table row and the operators and values in the basic condition builder meet your requirements. For example, for an Orders table, add a condition that the date in the ShippedDate column is greater than April 1, 2015.

Advanced Expression

Add filter conditions to tables with PowerCenter expressions. For example, for an Orders table, add the following expression:

```
(ShipCountry='USA') and ((ShipCity='New York') or (ShipCity='Los Angeles'))
```

Condition list

Select the table for which you want to define a condition and define the expression. The properties that show in the list depend on the type of expression that you select.

The filters use an AND logic. If multiple filters exist for a table, Data Integration Hub writes only the rows that meet all of the conditions to the target.

Basic Expression Properties

When you define a filter condition with a basic expression for a subscription, the **Filter** page shows the following properties:

Select Table

Select the table to which to apply the filter.

Select Column

Select the column to filter.

Note: You denote columns as filter accelerators at the topic level. Data Integration Hub indexes the columns that are denoted as filter accelerators in the Data Integration Hub publication repository and uses them to accelerate data retrieval.

Select Operator

Select the filter operator. The available operators depend on the type of content in the column.

Enter Value

When this field is enabled, select or enter a value for the operator. Values of string operators can contain up to 90 digits. Values of numeric operators can contain up to 15 digits.

Advanced Expression Properties

When you define a filter condition with an advanced expression for a subscription, the **Filter** page shows the following properties:

Select Table

Select the table to which the filter applies.

Expression text area

Enter a valid PowerCenter expression that defines the filtering criteria. The expression is limited to the selected table. The expression can contain up to 1024 characters.

Subscription Schedule Properties

The **Schedule** page of the subscription wizard can include the following properties:

When published data is ready

Runs the subscription immediately after the published data is ready.

Manually or by an external trigger

No schedule. You can use the following methods to run the subscription:

- Run manually. Click the Run arrow on the **Subscriptions** page.
- Run by an API. Call a command-line API or a REST API that starts the subscription.

By schedule

Runs the subscription according to the defined schedule. Select one of the following options:

- Every n minutes. Runs the subscription in intervals of up to 60 minutes. You select the number of minutes from the list.
- Hourly. Runs the subscription in intervals of up to 24 hours. You select the number of hours from the list.
- Daily. Runs the subscription at the same hour every day.
- Weekly. Runs the subscription every week on one or more days at the same hour.
- Monthly. Runs the subscription every month on a specific date or a specific day at the same hour.

Define the delivery intervals in the **Repeat running** area.

After the following publication or subscription run completes

Runs the subscription after the run of the publication or subscription that you select here completes.

Retry policy

Defines the retry policy for automatic subscriptions that are triggered by Data Integration Hub. A retry policy defines the number of times Data Integration Hub retries to run the subscription in case of failure, and the retry interval. The policy does not apply to subscriptions that you run manually.

Subscription Delivery Properties

The **Delivery** page of the Subscription wizard can include the following properties:

Data to Consume

Delivery scope and format of the data to consume. Choose one of the following options:

- All available publications. Processes each published data set with a separate subscription mapping.
- All available publications - aggregated. Groups and processes all published data sets and delivers a single data set. When you use an automatic mapping, the subscription sorts the data according to the publication date and time of the publication instances.
- Latest publication only. Delivers only the latest published data set.

This section is disabled for unbound subscriptions.

Update Data in Target

Handling of data that exists in the target application. Choose one of the following options:

- Append the new data to the existing data in the target. If the target type is Relational Database, Data Integration Hub adds rows to the table. If the target type is Flat File, Data Integration Hub appends records to the file.
- Insert new rows and update remaining rows. Data Integration Hub inserts new rows and then updates rows that exist in the target.
- Insert new rows and update changed rows. Data Integration Hub inserts new rows and then updates changed rows that exist in the target. Updating only changed rows slows down system performance. Applicable for subscriptions that consume data from topic tables where delta detection is enabled.
- Delete rows that don't exist in the topic from the target. Applicable if the type of topic to which the subscription subscribes is Full or if the subscription is an unbound subscription, and you choose to append new data or to insert new rows.
- Overwrite existing data in the target. If the target type is Relational Database, Data Integration Hub truncates the target tables before it inserts new rows. If the target type is Flat File, Data Integration Hub overwrites the file.

Advanced Performance Tuning

Use the options in this section to fine tune the performance of data retrieval from the Data Integration Hub publication repository for subscriptions that use an automatic mapping. Performance tuning is important when you use unbound subscriptions.

- Use pushdown optimization. Run querying and filtering on the Data Integration Hub publication repository instead of using a mapping with a PowerCenter Filter transformation.

Note:

- Pushdown optimization offloads processing jobs from the PowerCenter repository to the publication repository. Because the publication repository is critical to system health, use pushdown optimization with caution.
- In subscribers that subscribe to a topic that is managed on a Hadoop publication repository, basic filters are pushed to the Hadoop publication repository. Advanced filter expressions are not pushed to Hadoop and are processed in PowerCenter.

For more information, see the chapter "Pushdown Optimization" in the *PowerCenter Advanced Workflow Guide*.

- Use PowerCenter partitions with ... partitions. Enter the number of partitions in the text field. It is recommended that the number of partitions equals the number of CPUs on the computer that runs the PowerCenter Integration Service.

The number of partitions you enter determines how many PowerCenter processes handle the request in parallel. While parallel processing is faster, it consumes more resources from the PowerCenter Integration Service. For more information, see the chapter "Understanding Pipeline Partitioning" in the *PowerCenter Advanced Workflow Guide*.

- Sort by publications order. Sorts the data that the subscription consumes according to the order of batches by which the data is published. By default, Data Integration Hub does not sort the data.

Subscription Wizard Summary Page

The **Summary** page of the subscription wizard can include the following properties:

Subscription Name

Name of the subscription.

Topics

Name of the topic or topics to which the subscription subscribes.

Mapping

Mapping type. If you use a custom mapping, the field displays the workflow name.

Schedule

Time interval for consuming published data.

Data to Consume

Scope of the published data to consume.

Append or Overwrite Existing Data

Indicates whether to delete all data from the target application before consuming the new data.

Wait for all data to be published

Indicates whether a compound subscription can run only after all publications finish publishing data.

Properties of a Data-driven Subscription

The **Create Data-driven Subscription** page includes the following properties:

Subscription Name

Name of the subscription. The name is not case-sensitive and can contain up to 64 characters. The name can contain only alphanumeric characters and underscores.

You cannot change the subscription name after you create the subscription.

Description

Textual description of the subscription.

Mode

Mode of the subscription. Choose one of the following options:

- Enabled. The subscription runs when you post the subscription request.
- Disabled. The subscription does not run when you post the subscription request.

Topic

Topic from which the application consumes data. Browse to choose a topic from the list of topics.

Topic Structure

Shows the structure of the topic from which the application consumes data. You can view all the tables in the topic, or choose the tables to show.

Send notification when data is ready to consume

URL to where Data Integration Hub sends notifications when data is ready to consume.

Consume Data through a REST API

REST URL of the API. Use this URL to subscribe to the data.

Click **Copy** to copy the URL to your clipboard.

The Consume Data REST API returns the Swagger structure for the topic from which the application consumes data

REST URL of the Swagger structure for the topic from which the subscription consumes data.

Click **Copy** to copy the URL to your clipboard.

CHAPTER 13

Events and Event Monitoring

This chapter includes the following topics:

- [Events and Event Monitoring Overview, 177](#)
- [Event Actions, 178](#)
- [Basic Event Search Properties, 179](#)
- [Advanced Event Search Properties, 180](#)
- [Publication and Subscription Event Types and Statuses, 181](#)
- [Event Monitors, 182](#)

Events and Event Monitoring Overview

An event is a representation of a publication or a subscription instance, at a particular stage of processing. The Data Integration Hub server generates events as it processes publications and subscriptions, and it changes the status of the events as they go through the transformation process. When an application triggers a publication, if the triggered publication has a pre-process, the publication event also tracks the pre-process. When a subscription triggers a post-process, the subscription event also tracks the post-process.

When an application that runs a publication pre-process publishes data or files to the publication repository, the Data Integration Hub server assigns an event to the publication as follows:

- If the pre-process passes the Publication event ID to the publication process, the publication uses the same event, and the Data Integration Hub server does not generate an additional Publication event for the publication process.
- If the pre-process does not pass the event ID to the publication process, the Data Integration Hub server generates another Publication event for the publication process.

If a file publication publishes more than one file, Data Integration Hub creates a File event for each file that it picks up. Data Integration Hub creates a Publication event after all the files are picked up.

The Publication event is the root event and the parent event for all of the subscription events that the Data Integration Hub server generates during processing. After the published data is ready for subscribers, the Data Integration Hub server generates a Subscription child event for each subscriber that needs to consume the published data. The Publication event contains aggregated status information for all Subscription child events.

You monitor and perform actions on events in the Operation Console. You can view all events on the Event List page. You can perform a basic or advanced search for events, and drill down to a specific event to view event details.

By default, the Event List page displays root events: Publication, File, Aggregated Subscription, and Compound Subscription. After a publication is ready for subscribers, you can drill down to the associated Subscription child events of the publication.

You can create rules that monitor publication and subscription events, and perform actions on events that are in a defined status. For example, you can create rules to perform the following tasks:

- Disable publications that have events with an Error status.
- Invoke a PowerCenter workflow when a publication event is in an Error status.
- Send an email to the Data Integration Hub administrator when a subscription event is in a Critical status.

Event Actions

You can perform actions on events from the **Event List** page based on the event status, state, or other requirements in your organization.

Note: You cannot perform any actions on File events.

The following table describes actions that you can perform on events:

Action	Description
Reprocess	Applicable to subscription events in a final state. Reprocesses the subscription of the selected event. When you reprocess a Subscription event, a new Subscription child event is created with a Reprocessed status. The Subscription event is a child event of the Publication event that Data Integration Hub generated for the publication that the subscription consumes. When you reprocess a Subscription event that is associated with a Compound Subscription event, other Subscription events that are associated with the Compound Subscription are not reprocessed.
Change event status	Changes the status of the selected events in the Change Event Status pop-up window.
Discard	Discards all of the selected events. The status of the selected events is changed to Discarded, but the events are not deleted.

Managing Events on the Event List Page

In the **Event List** page, you can view a list of events, an overview summary of events, and event details. You can also drill to parent and child events and perform actions on events.

1. In the Navigator, click **Monitoring and Tracking > Event List**.
2. To display events that were created within a specific time range, select a value from the **Time frame** list.
3. To change the default time frame, select a new value and click **Set as default time frame**.

Note: On Windows operating systems, the default time frame applies to the current browser for the current Windows account. If you use multiple browsers for the same Windows account, you need to define a default time frame for each browser.

4. To view a summary of all of the events that were created within the time frame according to the event type and status, click **Overview**.

The overview displays the total number of events for each event type and status. Each total number contains a link that you can click to display the list of events for that event type or status.

5. To view details for a specific event, click the event ID for the event that you want to view.

The **Details** section displays information about the event.

6. To drill to parent or child events for which you have viewing privileges, click the **Drill Up** or **Drill Down** icon next to the event that you want to view.

The **Details** section displays information about the child or parent event.

7. To perform an action on one or more events, select the events and click **Actions**.

The **Actions** list displays the actions that you can perform on the selected events.

Basic Event Search Properties

You can perform a basic search for events based on the event type or status. You can also filter the search results according to a specific time frame.

The following table describes basic event properties for which you can perform a basic search:

Property	Description
Event ID	Unique identifier for the event.
Application	Name of the related application.
Publication/Subscription	Name of the related publication or subscription.
Topic	Name of the related topic.
Event Type	Type of the event.
Event Status	Status of the event. This list contains two properties: <ul style="list-style-type: none">- State. Indicates whether the event is still processing or reached a final status.- Status. Current progress of the event, regardless of whether the event finished processing.

Performing a Basic Search for Events

You can perform a basic search for events. The search results appear in a list or in an overview summary according to the event type and status. You cannot perform a basic search in child events.

1. In the Navigator, click **Monitoring and Tracking > Event List**.
2. If the Advanced Search pane appears, click **Basic search**.
3. In the **Find** field, enter the search string.
4. Click **Go** to start the search.

The Event List page displays the list of events that match the search criteria.

Advanced Event Search Properties

In addition to the basic search criteria, you can perform an advanced search for specific event properties.

The following table describes event properties for which you can perform an advanced search:

Property	Description
Account	Account associated with the event.
Event Attributes	Attributes of the event. Click Browse and select up to 3 attributes and operators in the Event Attributes Search Criteria window.
Event ID	ID of the event.
Event Status	Status of the event.
Event Type	The type of event.
Operation	Select whether all search properties apply together or whether any of them may apply to the search.
Parent Event ID	Unique identifier of the parent events for any events that match the search criteria.
Partner	Partner associated with the event.
Subject	Subject associated with the event.
Show child events	Indicates whether to perform the search in child events. Cleared by default.
Time frame	Time frame in which the event occurred.
Topic	Name of the related topic.

Performing an Advanced Search for Events

In addition to the basic search criteria, you can perform an advanced search for specific event properties. You can also search for parent and child events.

1. In the Navigator, click **Monitoring and Tracking > Event List**.
2. If the Advanced Search pane is hidden, click **Advanced search**.
3. Select a search operator:
 - **And**. Search for all of the fields that contain values.
 - **Or**. Search for any of the fields that contain values.
4. To search in child events, select the **Show child events** check box.
5. Enter values in the event properties that you want to search and click **Go**.

The **Event List** page displays the search results. If you selected to show child or parent events, the **Drill Up** or **Drill Down** columns display links to parent or child events for which you have viewing privileges, and the event ID appears when you hover the mouse over the icon for the event that you want to view.

Publication and Subscription Event Types and Statuses

Data Integration Hub assigns a default set of event types and event statuses to publication and subscription events when the following conditions exist:

- A publication with a publication pre-process runs.
- A publication or a subscription with an automatic mapping runs.
- A publication or subscription with a custom mapping that is associated with a Data Integration task runs.
- A subscription post-process runs.
- A data-driven publication runs.

Note: It is recommended that publications and subscriptions with a custom mapping that is associated with a PowerCenter workflow use the same event statuses and types as those that Data Integration Hub assigns to the automatic mappings. The developer assigns event statuses and types for custom mapping in the PowerCenter workflow.

Default Event Types

Data Integration Hub assigns the following event types:

- **Publication.** Assigned to a publication process that is not initiated by a data-driven publication. Acts as the parent event for all Subscription events and for File events of publications that publish multiple files. For File events of publications that publish a single file, the event log includes a link to the file on the Data Integration Hub document store.
- **Data-driven Publication.** Assigned to a publication process of a data-driven publication.
- **File Event.** Assigned to the publication of a file in publications that publish multiple files. The event log includes a link to the file on the Data Integration Hub document store.
- **Subscription.** Assigned to a subscription process. Acts as a child event for a publication event. For events of subscriptions that consume pass-through files and do not use file transfer, the event log includes a link to the file on the Data Integration Hub document store. For events of subscriptions that use file transfer to consume files, the event log includes a link to the file transfer list.
- **Aggregated Subscription.** Assigned to a subscription process that consumes multiple data sets from the same topic with a single subscription mapping. The event contains references to all Subscription events that were created when the associated topic finished publishing each data set. The Subscription events inherit their status from the Aggregated Subscription event.
- **Compound Subscription.** Assigned to a subscription process that consumes data sets from multiple topics with a single subscription mapping. The event contains references to all Subscription events that Data Integration Hub creates when each topic publication finished publishing the data set.

Default Event Statuses

For publications, Data Integration Hub assigns the following event statuses:

- **Pre-processing.** Indicates that the publication pre-processing instance is running.
- **Processing.** Indicates that the publication instance is running.
- **Delayed.** Relevant for File events, for publications that publish multiple files. Indicates that the file that is related to the event is ready but that not all the files that the publication publishes are ready. When all the files that the publication publishes are ready, Data Integration Hub creates a Publication event, and the related File events inherit the status of the Publication event. You cannot run delayed File events.

- **Completed.** Indicates that the publication instance finished running and that the data is ready for subscribers.
- **Error.** Indicates that the publication instance encountered errors and did not finish running.

Each Publication event also shows the consumption status of the child Subscription events. The status reflects the overall consumption and changes after all Subscription events changed status. For example, the consumption status changes to complete after all subscribers finished consuming the published data.

For subscriptions, Data Integration Hub assigns the following event statuses:

- **Delayed.** Indicates that the published data is ready but that the subscribing application did not start consuming the data. Relevant for subscriptions with a defined schedule and for subscriptions that are run manually or by an external trigger. You can run delayed subscription events from the **Subscriptions** tab of the **Application Details** page.
- **Processing.** Indicates that the subscription instance is running.
- **Completed.** Indicates that the subscription instance finished running and that the subscribing application consumed all published data.
- **Post-processing.** Indicates that the subscription post-processing instance is running.
- **Error.** Indicates that the subscription instance encountered errors and did not finish running.

Event Monitors

You can create event monitors that track publications and subscriptions based on their event status, and instigate actions when an event is in a defined status.

You create monitoring rules that define which entities to monitor, what are the event statuses for which to take action, and what actions Data Integration Hub takes when an event reaches the defined status.

You can create rules that monitor publication and subscription events, and perform actions on events that are in a defined status. For example, you can create rules to perform the following tasks:

- Disable publications that have events with an Error status.
- Invoke a PowerCenter workflow when a publication event is in an Error status.
- Send an email to the Data Integration Hub administrator when a subscription event is in a Critical status.

Monitoring Rules

A monitoring rule defines which entities to monitor, the event statuses that trigger actions, and the actions to take when an event is in a defined status.

When you create a monitoring rule, you define the following elements:

- **Entity or entities to which the rule applies.** A rule can apply to a single publication, to multiple publications, or to all current and future publications, or to a single subscription, to multiple subscriptions, or to all current and future subscriptions.
- **Event status or statuses to which the rule applies.** Data Integration Hub applies the rule only to events that are in a final state.
- **Rule action or actions.** You can select one or more of the following actions:
 - **Send email notification.** You define the user or users to which Data Integration Hub sends an email notification when the rule conditions are true.

- Disable publications and subscriptions that are in the status or statuses to which the rule applies.
- Invoke a monitoring workflow. For example, when a publication fails, invoke a workflow that deletes the data from the publishing source. The developer creates a PowerCenter workflow and imports it to Data Integration Hub. You select the workflow when you create the rule.

Managing Monitoring Rules

Use the Navigator to create, edit, or delete monitoring rules. In the Navigator you can also disable and enable monitoring rules.

1. In the Navigator, click **Monitoring and Tracking > Monitors**.
The **Monitors** page appears.
2. Choose the action that you want to perform:
 - To create a rule, click **New Monitoring Rule** and then follow the **Create Monitoring Rule** wizard. For more details, see [“Creating a Monitoring Rule” on page 183](#).
 - To edit a rule, click the Edit button next to the rule that you want to edit.
 - To delete a rule, click the Delete button next to the rule that you want to delete.
 - To disable a rule, select **Disabled** from the **Mode** list. A disabled rule does not run and does not perform the defined actions.
 - To enable a disabled rule, select **Enabled** from the **Mode** list.

Creating a Monitoring Rule

1. In the Navigator, click **Monitoring and Tracking > Monitors**.
The **Monitors** page appears.
2. Click **New Monitoring Rule**.
The **Create Monitoring Rule** wizard appears.
3. Enter the rule name. Optionally, enter a description of the rule. Click **Next**.
The **Select Entity** page appears.
4. In **Event Type**, select the type of entity to which the rule applies, publication or subscription, and then select the entity or entities to which the apply the rule. You must apply the rule to at least one publication or one subscription.
 - To apply the rule to all publications or to all subscriptions, including current publications or subscriptions and publications or subscriptions that are added to Data Integration Hub after you create the rule, select **Apply to all publications** or **Apply to all subscriptions**.
 - To apply the rule to all current publications or to all current subscriptions, select the check box to the left of the **Publication Name** or the **Subscription Name** title.
 - To select a single publication or a single subscription, select the check box to the left of the publication name or the subscription name.
 - To select multiple publications or multiple subscriptions, select multiple check boxes to the left of the publication names or the subscription names.
 - To apply the rule for system events, select **System**.

Click **Next**.

The **Select Status** page appears.

5. From the **Event Statuses** list, select a status to which to apply the rule and then click the right-arrow. You can apply monitoring rules to statuses in a Final state only. Add as many statuses as required to the **Selected Statuses** list.

The statuses that you select are removed from the **Event Statuses** list and show in the **Selected Statuses** list.

6. Click **Next**.

The **Select Actions** page appears.

7. Select one or more of the following rule actions:

Send email notification

Send email notifications when a publication or a subscription is in one of the statuses that you selected in the **Select Status** page. You can send notifications to all the users in a Data Integration Hub user group, to specific users in a Data Integration Hub user group, and to email addresses that you specify.

To send email notifications, perform the following steps for each user group, user, or email address:

1. Expand the option and then click **Add**.
2. To send notifications to all users in a user group, select the group from the **User Group** list, and then select **All** from the **User Name** list.
3. To send notifications to a specific user in a user group, select the group from the **User Group** list, and then select the user from the **User Name** list. If you select a user for which an email address is not defined in Data Integration Hub, Data Integration Hub does not send notifications to that user until the administrator updates the email address of the user in the system. A warning message shows when you click **Next** in the **Select Actions** page.
4. To send notifications to a specific email address, select **None** from the **User Group** list, then select **None** from the **User Name** list, and then enter the email address in the **Email** field.

Data Integration Hub sends email notifications to the recipients that you define here when any of the publications or subscriptions that you selected in the **Select Entity** page are in any of the statuses that you selected in the **Select Status** page.

Disable publications and subscriptions that are in <selected_statuses> statuses

Expand the option and then select **Disable publications and subscriptions**.

Data Integration Hub disables the publications or subscriptions that you selected in the **Select Entity** page when they are in any of the statuses that you selected in the **Select Status** page. A disabled publication or subscription does not run according to schedule or by an external API. You can only run a disabled publication or subscription from within the Operation Console.

Note: This option is disabled if you have selected an event type of System.

Tip: To enable a disabled publication or a disabled subscription, access the corresponding **Publication** or **Subscription** wizard.

Invoke monitoring workflow

Select a Data Integration Hub workflow that Data Integration Hub runs when a publication or a subscription is in any of the statuses that you selected in the **Select Status** page. The developer creates a workflow in PowerCenter and then imports the PowerCenter workflow to Data Integration Hub.

Click **Next**.

The **Summary** page appears.

8. Review the rule settings and then click **Finish**.

The **Create Monitoring Rule** wizard closes. The rule is added to the **Monitors** page.

Monitoring Rule Properties

Use the **Create Monitoring Rule** wizard to create and edit event monitoring rules.

The **Create Monitoring Rule** wizard contains the following pages:

General page

Define basic rule properties.

Select Entity page

Select the entity or the entities to which the rule applies.

Select Status page

Select the event status or statuses to which the rule applies.

Select Actions page

Select the rule action or actions.

Summary page

Review rule settings and save the rule.

Monitoring Rule General Properties

When you create a monitoring rule, configure the rule name, description , and mode in the **General** page of the **Create Monitoring Rule** wizard.

The following image shows a sample **General** page:

The screenshot shows the 'Create Monitoring Rule' wizard with the 'General' page selected. The wizard has five steps: 1. General, 2. Select Entity, 3. Select Status, 4. Select Actions, and 5. Summary. The 'General' page contains three fields: 'Rule Name*' (a text input field), 'Description' (a text area), and 'Mode' (a dropdown menu set to 'Enabled' with an information icon). At the bottom right, there are three buttons: 'Previous', 'Next', and 'Cancel'.

The **General** page includes the following properties:

Rule Name

Name of the rule. The name is not case-sensitive and can contain up to 80 characters. The name must begin with an alphabetic character or an underscore, and can contain only alphanumeric characters and underscores.

Description

Textual description of the rule.

Mode

Mode of the rule. Choose one of the following options:

- Enabled. The rule runs and performs the defined actions.
- Disabled. The rule does not run and does not perform the defined actions.

Monitoring Rule Select Entity Properties

When you create a monitoring rule, select the publications or subscriptions that you want to monitor in the **Select Entity** page of the **Create Monitoring Rule** wizard.

The following image shows a sample **Select Entity** page:

Create Monitoring Rule

1 General | **2 Select Entity** | 3 Select Status | 4 Select Actions | 5 Summary

Select the publication or subscription that you want to monitor.

Entity Type* Publication ▼

Publications (3 | 0 selected)

☐ Apply to all publications ⓘ

Find...

<input type="checkbox"/>	Publication Name	Description
<input type="checkbox"/>	amir_publication	
<input type="checkbox"/>	FF_publication	
<input type="checkbox"/>	relational_multiple	

Previous Next Cancel

The **Select Entity** page includes the following properties:

Entity Type

Type of entity to which the rule applies, publication or subscription.

When you select the entity type, the page shows the properties that are applicable to the type of entity that you select in the **Publications** or **Subscriptions** area, accordingly.

Apply to all publications or Apply to all subscriptions

Applies the rule to all publications or to all subscriptions, including current publications or subscriptions and publications or subscriptions that are added to Data Integration Hub after you create the rule.

Find

Search for publications or for subscriptions by name or by description.

Publication table or Subscription table

Lists all the publications or subscriptions, or, if you enter a string in the **Find** text box, lists the publications or the subscriptions that match your search.

Monitoring Rule Select Status Properties

When you create a monitoring rule, select the event statuses that you want to monitor in the **Select Status** page of the **Create Monitoring Rule** wizard.

The following image shows a sample **Select Status** page:

The screenshot shows the 'Create Monitoring Rule' wizard with five tabs: General, Select Entity, Select Status (active), Select Actions, and Summary. Below the tabs, a message states: 'Select event statuses that you want to monitor. You must select at least one status from the list below.' The main area is divided into two columns: 'Event Statuses (7)' and 'Selected Statuses (0)'. The 'Event Statuses' column contains a list of statuses with corresponding icons: Complete (green checkmark), Critical (red exclamation mark), Discarded (blue document with X), Error (red X), Rejected (red circle with slash), Reprocessed (blue circular arrow), and Warning (yellow triangle). The 'Selected Statuses' column is empty. Between the columns are two arrow buttons for moving items. At the bottom right, there are 'Previous', 'Next', and 'Cancel' buttons.

The **Select Status** page includes the following properties:

Event Statuses

Event statuses to that you can apply the rule. You can apply monitoring rules to statuses in a Final state only.

Selected Statuses

Event statuses to which the rule applies.

Monitoring Rule Select Actions Properties

When you create a monitoring rule, select the rule action or actions in the **Select Actions** page of the **Create Monitoring Rule** wizard.

The following image shows a sample **Select Actions** page:

The screenshot shows the 'Create Monitoring Rule' wizard with the 'Select Actions' step highlighted. The wizard has five steps: 1 General, 2 Select Entity, 3 Select Status, 4 Select Actions, and 5 Summary. Below the steps, a message states 'You must select at least one action.' There are three expandable sections: 'Send email notification (0 users)' with an 'Add' button, 'Disable publications and subscriptions that are in Critical statuses' with a checkbox and an information icon, and 'Invoke monitoring workflow' with a search bar. At the bottom are 'Previous', 'Next', and 'Cancel' buttons.

User Group	User Name	Email	Delete
------------	-----------	-------	--------

The **Select Actions** page includes the following properties:

Send email notification

Send email notifications when a publication or a subscription is in any of the statuses that you selected in the **Select Status** page. You can send notifications to all the users in a Data Integration Hub user group, to specific users in a Data Integration Hub user group, and to email addresses that you specify. You define the email recipients in the user list by defining one or more of the following properties:

- **User Group.** A Data Integration Hub user group. From the **User Name** list, select either **All** to send notifications to all group users or select a specific user to send notifications to that user only. If you select a user for which an email address is not defined in Data Integration Hub, Data Integration Hub does not send notifications to that user until the administrator updates the email address of the user in the system. A warning message shows when you click **Next** in the **Select Actions** step.
- **User Name.** A user in a Data Integration Hub user group. You must first select the group in the **User Group** list.
- **Email.** An email addresses to which to sent notifications. You must first select **None** from the **User Group** list and from the **User Name** list

Disable publications and subscriptions that are in <selected_statuses> statuses

Data Integration Hub disables the publications or the subscriptions that you selected in the **Select Entity** page when they are in the status or statuses that you selected in the **Select Status** page. A disabled publication or subscription does not run according to schedule or by an external API. You can only run a disabled publication or subscription from within the Operation Console.

Invoke monitoring workflow

Select a Data Integration Hub workflow that Data Integration Hub runs when a publication or a subscription is in any of the statuses that you selected in the **Select Status** page. The developer creates a workflow in PowerCenter and then imports the PowerCenter workflow to Data Integration Hub.

Monitoring Rule Summary Properties

When you create a monitoring rule, review rule settings and save the rule in the **Summary** page of the **Create Monitoring Rule** wizard.

The following image shows a sample **Summary** page:

The screenshot shows the 'Create Monitoring Rule' wizard with five steps: 1 General, 2 Select Entity, 3 Select Status, 4 Select Actions, and 5 Summary (highlighted). The Summary page displays the following properties:

Rule Name	Disable_Pubs_on_Error
Description	
Rule	If the events of publications: All Are in status: Error Then: Disable related publications and subscriptions

At the bottom right, there are three buttons: Previous, Finish, and Cancel.

The **Summary** page includes the following properties:

Rule Name

Name of the rule.

Description

Textual description of the rule.

Rule

Rule conditions and actions.

CHAPTER 14

Dashboard and Reports

This chapter includes the following topics:

- [Dashboard and Reports Overview, 190](#)
- [Dashboard Reports, 191](#)
- [Dashboard Report CSV File Structure, 197](#)
- [Managing the Dashboard, 199](#)

Dashboard and Reports Overview

The Dashboard is a collection of panels that display personalized visual reports about information that Data Integration Hub processes. Each report appears in a panel that you view in the **Dashboard** page of the Operation Console.

Use the Dashboard to view summary information about Data Integration Hub events, such as the number of publication events or the number of errors by application. You can drill down to each report to view event details or manage daily tasks that require further investigation. You can focus on areas of interest and analyze statistical information about the processed data. You can export the report to a comma-separated-values (CSV) file for further analysis.

Data that appears in the charts is filtered by global Dashboard filters, for example, by a time frame or by an application. The data is then filtered by permissions. You can view data pertaining to applications to which you have access. All charts are visible to users with dashboard privileges.

In the Dashboard, all the Data Integration Hub reports appear in the Events tab. You can organize the display of the charts, and add tabs to display selected reports only. Each report displays event information from the operational data store or the run-time Data Integration Hub repository. You can apply global filters for all of the panels in all of the tabs, including the time frame and the application for which to display data. In each report, you can select whether to view events for both publications and subscription, events for publications only, or events for subscriptions only. You can drill down to view the events from each panel in the **Event List** page.

Dashboard Reports

The Dashboard contains panels in which you can view different reports based on information that Data Integration Hub collects from the operational data store or the run-time Data Integration Hub repository.

The following list describes the Dashboard reports:

Errors by Application

Up to 10 or up to 20 applications with the highest number of error events created during the selected time frame. You can select to show all errors, including current events, that is, events that are stored in the Data Integration Hub repository, and events that are stored in the operational data store, or to show only current, unresolved errors.

Use the report to analyze application activity or identify potential bottlenecks.

Events by Current Status

Up to 10 or up to 20 event statuses with the highest number of events that are in that status during the selected time frame.

Use the report to identify potential bottlenecks or other issues that might require further attention.

Event Distribution

Total number of publications and subscriptions that were completed during the selected time frame, per hour or per day. The report displays events that reached their final status, including both successful and error events.

Use the report to identify event creation peaks and valleys, and identify time periods that might require further attention.

Error Event Distribution

Total number of publications and subscriptions that reached their final state with an Error status during the selected time frame, per hour or per day.

Use the report to identify error event creation peaks, and identify time periods that might require further attention.

Event Average Processing Time

Visual indication of the distribution of publication and subscription processing time during the selected time frame. Error events are not calculated.

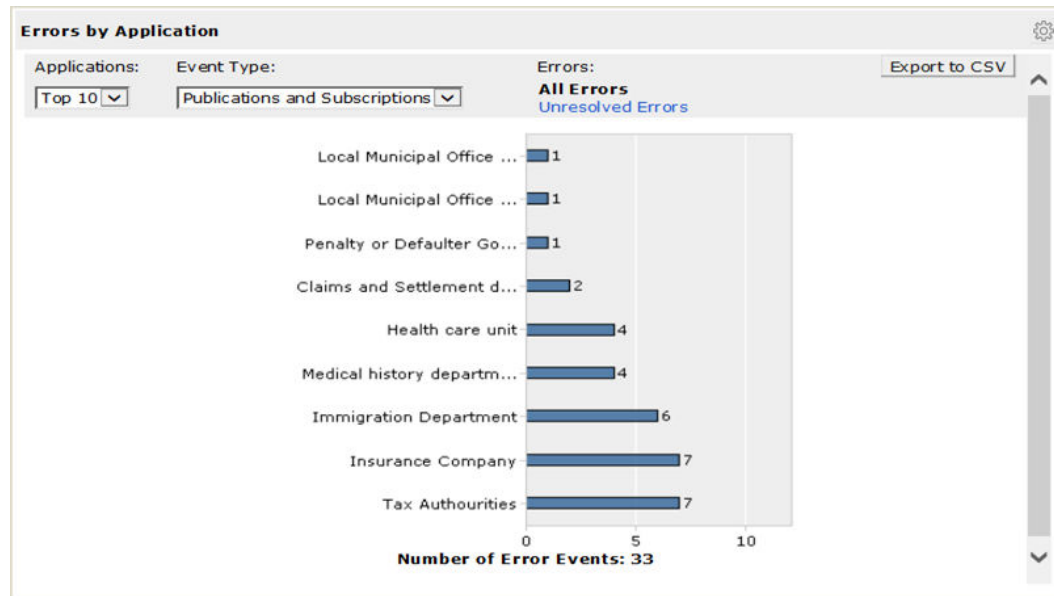
Use the report to identify potential bottlenecks or other issues that might require further attention.

Errors by Application Report

The Errors by Application report displays up to 10 or up to 20 applications with the highest number of error events created during the selected time frame, for the selected event type. Use the report to analyze application activity or identify potential bottlenecks.

You can click each bar on the graph to display the error events or unresolved error events for the application in the **Event List** page.

The following image shows the Errors by Application report:



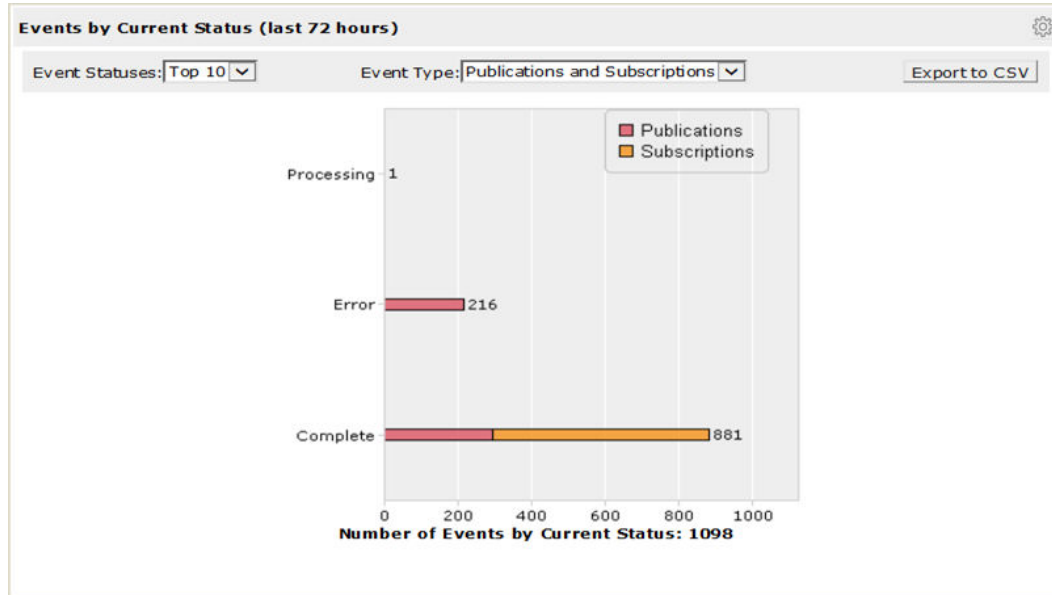
The following table describes the Errors by Application report elements:

Element	Description
Applications	Number of applications for which to display the number of error events for the selected time frame. You can choose one of the following options: <ul style="list-style-type: none"> - Top 10. Up to 10 applications with the highest number of error events created. - Top 20. Up to 20 applications with the highest number of error events created.
Event Type	Type of events to display. You can choose one of the following options: <ul style="list-style-type: none"> - Publications and Subscriptions. - Publications Only. - Subscriptions Only.
Errors	Type of errors to display. You can choose one of the following options: <ul style="list-style-type: none"> - All. Displays all events that reached an error state for the applications during the selected time frame, including current events, that is, events that are stored in the Data Integration Hub repository, and events that are stored in the operational data store. You cannot change the status of events that are stored in the operational data store. - Unresolved. Displays current error events for the applications. That is, events that are stored in the Data Integration Hub repository. The events appear based on the time frame that you select in the unresolved error events filter. You can change the status of current events in the Event List page.
Application Name	Names of the applications with the highest number of error events for the selected time frame. Appears on the Y-axis of the panel.
Events	Number of error events for the applications during the selected time frame. Appears on the X-axis of the panel.
Number of Error Events	Total number of error events for the applications during the selected time frame.
Export to CSV	Saves the data in the report as a CSV file.

Events by Current Status Report

The Events by Current Status report displays up to 10 or up to 20 event statuses with the highest number of events created during the selected time frame. Use the report to identify potential bottlenecks or other issues that might require further attention. You can click each bar on the graph to display the events for the event status in the **Event List** page.

The following image shows the Events by Current Status report:



The following table describes the Events by Current Status report elements:

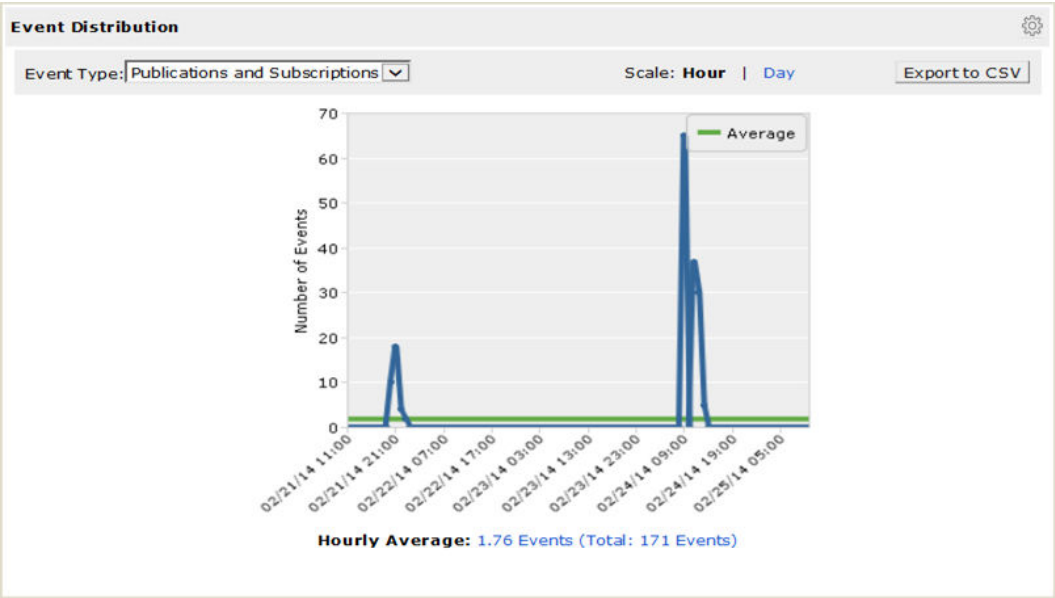
Element	Description
Event Statuses	Number of event statuses for which to display the number of events for the selected time frame. You can choose one of the following options: <ul style="list-style-type: none">- Top 10. Up to 10 event statuses with the highest number of events created.- Top 20. Up to 20 event statuses with the highest number of events created.
Event Type	Type of events to display. You can choose one of the following options: <ul style="list-style-type: none">- Publications and Subscriptions.- Publications Only.- Subscriptions Only.
Event Status	Names of the event statuses with the highest number of events for the selected time frame. Appears on the Y-axis of the panel.
Events	Number of events for the event statuses during the selected time frame. Appears on the X-axis of the panel.
Number of Events by Current Status	Total number of events for the event statuses during the selected time frame.
Export to CSV	Saves the data in the report as a CSV file.

Event Distribution Report

The Event Distribution report displays the total number of publications and subscriptions that were completed during the selected time frame, per hour or per day. The report displays events that reached their final status, including both successful and error events. Use the report to identify event creation peaks and valleys, and identify time periods that might require further attention.

The report displays all event types and statuses that reached a final state. You can click any point on the graph to view the events in the **Event List** page.

The following image shows the Event Distribution report:



The following table describes the Event Distribution report elements:

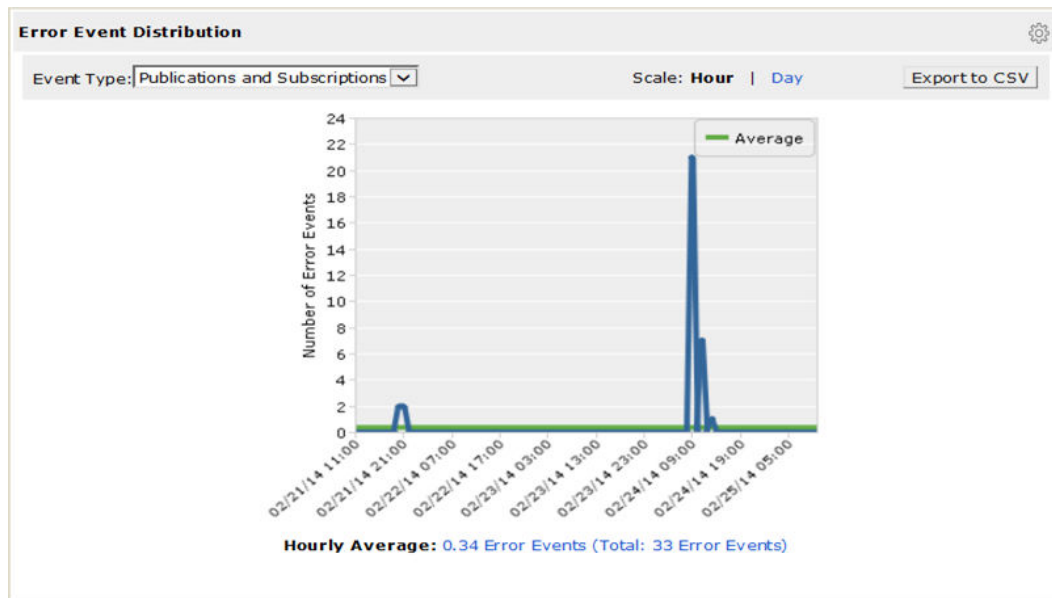
Element	Description
Event Type	Type of events to display. You can choose one of the following options: <ul style="list-style-type: none">- Publications and Subscriptions.- Publications Only.- Subscriptions Only.
Scale	Granularity of the calculations. You can choose one of the following options: <ul style="list-style-type: none">- Hour. Displays the number of events at hourly intervals as points on the graph.- Day. Displays the number of events at daily intervals as points on the graph. If you select a time frame filter longer than seven days, you can only view event distribution on a day scale.
Number of Events	Number of events that Data Integration Hub generated during the selected time frame. Appears on the Y-axis of the panel.
Time Distribution	Date or time intervals during which Data Integration Hub generated events. Appears on the X-axis of the panel.
Average Line	Average, over the selected time frame, of the average number of final events that Data Integration Hub generated each day or each hour and the total number of events for the selected time frame.

Element	Description
Daily/Hourly Average	Total count of final events within the entire selected time frame. You can click the link to display the events in the Event List page.
Export as CSV	Saves the data in the report as a CSV file.

Error Event Distribution Report

The Error Event Distribution report displays the total number of publications and subscriptions that reached their final state with an Error status during the selected time frame, per hour or per day. Use the report to identify error event creation peaks, and identify time periods that might require further attention.

The following image shows the Error Event Distribution report:



The following table describes the Error Event Distribution report elements:

Element	Description
Event Type	Type of events to display. You can choose one of the following options: <ul style="list-style-type: none"> - Publications and Subscriptions. - Publications Only. - Subscriptions Only.
Scale	Granularity of the calculations. You can choose one of the following options: <ul style="list-style-type: none"> - Hour. Displays the number of error events at hourly intervals as points on the graph. - Day. Displays the number of error events at daily intervals as points on the graph. If you select a time frame filter longer than seven days, you can only view error event distribution on a day scale.
Number of Error Events	Number of error events that Data Integration Hub generated during the selected time frame. Appears on the Y-axis of the panel.

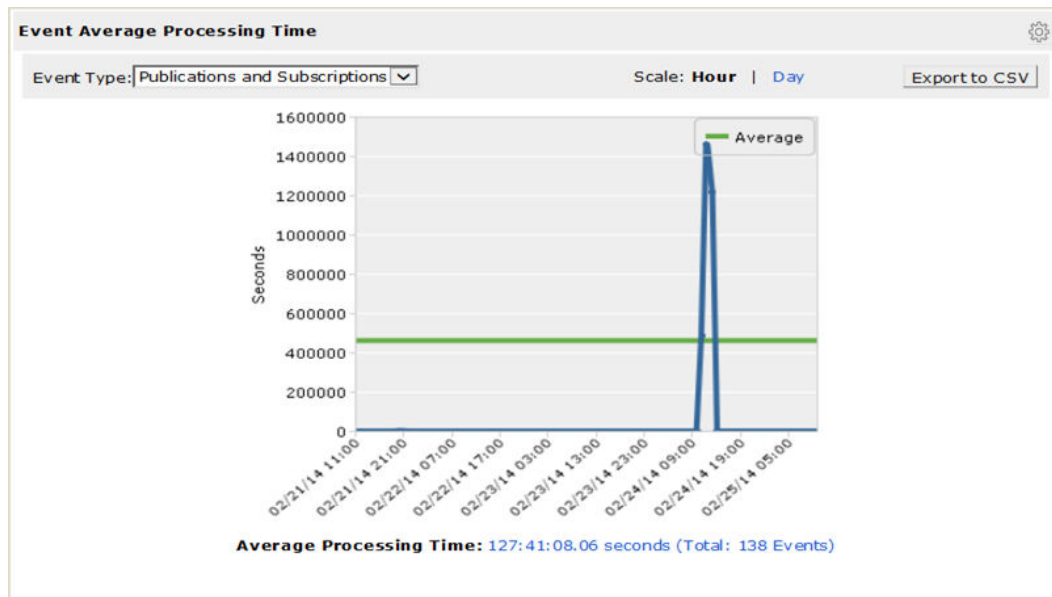
Element	Description
Time Distribution	Date or time intervals during which Data Integration Hub generated error events. Appears on the X-axis of the panel.
Average Line	Average, over the selected time frame, of the average number of final error events that Data Integration Hub generated each day or each hour and the total number of error events for the selected time frame.
Daily/Hourly Average	Overall error event average and the total number of events for the selected time frame. You can click the link to display the error events in the Event List page.
Export as CSV	Saves the data in the report as a CSV file.

Event Average Processing Time Report

The Event Average Processing Time report provides a visual indication of the average processing time of publications and subscriptions in Data Integration Hub during the selected time frame. Publications and subscriptions that are in Error status are not calculated. Use the report to identify deviations or other issues that might require further attention. You can click any point on the graph to display the events in the **Event List** page.

The Dashboard calculates the average processing time for time intervals during which Data Integration Hub processed events.

The following image shows the Event Average Processing Time report:



The following table describes the Event Average Processing Time report elements:

Element	Description
Event Type	Type of events to display. You can choose one of the following options: <ul style="list-style-type: none">- Publications and Subscriptions.- Publications Only.- Subscriptions Only.
Scale	Granularity of the calculations. You can choose one of the following options: <ul style="list-style-type: none">- Hour. Displays the average processing time of events at hourly intervals as points on the graph.- Day. Displays the average processing time of events at daily intervals as points on the graph. If you select a time frame filter longer than seven days, you can only view events on a day scale.
Seconds	Average number of seconds from when Data Integration Hub generated the events until first time that the events finished processing during the selected time frame. Appears on the Y-axis of the panel.
Time Distribution	Date or time intervals during which Data Integration Hub generated events. Appears on the X-axis of the panel.
Average Line	Average processing time for the entire time frame.
Average Processing Time	Average processing time for events that Data Integration Hub generated each day or hour and the total number of events for the selected time frame. You can click the link to display the events in the Event List page.
Export as CSV	Saves the data in the report as a CSV file.

Dashboard Report CSV File Structure

You can export the Dashboard reports to a comma-separated-values (CSV) file. The file contains all data in the report in a single column, separated by semicolons. The file does not contain subtotals or formatting information.

The following table describes the CSV elements for the Errors by Application report:

Element	Description
APPLICATION_NAME	Name of the application.
APPLICATION_ID	Identifier for the application in Data Integration Hub.
ERROR_COUNT	Number of error events that Data Integration Hub generated for the application during the selected time frame.
Other file elements are for internal use and can be ignored.	

The following table describes the CSV elements for the Events by Current Status report:

Element	Description
EVENT_STATUS_NAME	Name of the event status.
EVENT_STATUS_ID	Identifier for the event status in Data Integration Hub.
PUB_COUNT	Number of publications events that Data Integration Hub generated for the event status during the selected time frame. If the selected event type is Subscriptions, the value of PUB_COUNT is zero.
SUB_COUNT	Number of subscription events that Data Integration Hub generated for the event status during the selected time frame. If the selected event type is Publications, the value of SUB_COUNT is zero.
TOTAL_COUNT	Number of events that Data Integration Hub generated for the event status during the selected time frame, for both publications and subscriptions.
Other file elements are for internal use and can be ignored.	

The following table describes the CSV elements for the Event Distribution and Error Event Distribution reports:

Element	Description
ROW_NR	Numeric identifier for the time interval during which Data Integration Hub processed the events.
DAY/TIMESLICE	Start date and time of the time intervals during which Data Integration Hub processed events. The DAY column appears for reports that you view in a day scale. The TIMESLICE column appears for reports that you view in an hour scale.
TOTAL_COUNT	Total number of events that Data Integration Hub processed during each time interval.

The following table describes the CSV elements for the Event Average Processing Time report:

Element	Description
ROW_NR	Numeric identifier for the time interval during which Data Integration Hub processed the events.
DAY/TIMESLICE	Start date and time of the time intervals during which Data Integration Hub processed events. The DAY column appears for reports that you view in a day scale. The TIMESLICE column appears for reports that you view in an hour scale.
TOTAL_COUNT	Total number of events that Data Integration Hub processed during the each time interval.
TOTAL_TIME	Overall processing time for all of the events that Data Integration Hub processed during each time interval.

Managing the Dashboard

You use the Dashboard in the Operation Console to view and analyze events that Data Integration Hub generates during a defined time frame. You personalize your view of the Dashboard and add or remove panels and tabs as needed. If you require more screen space to view the Dashboard, you can collapse the Navigator.

You can only view the Dashboard if the Data Integration Hub administrator installed and configured it and if you have dashboard privileges.

1. In the Navigator, select **Dashboard**.

The Data Integration Hub administrator can set the Dashboard to open by default when you log in to the Operation Console.

2. To add a tab to the Dashboard, perform the following steps:

- a. Click the Add Tab icon, choose the panels that you want to add in the **Add Panels** window, and then click **Done**.
- b. Click the Settings icon in the tab, choose **Change Layout**, choose the number of panel columns in the **Tab Layout** dialog box, and then click **Done**.
- c. Click the Settings icon in the tab, choose **Rename Tab**, and then enter a name for the tab.

3. To further personalize the appearance of the Dashboard, use one or more of the following options:

- To add panels to a tab, click the Settings icon in the tab, choose the panels that you want to add in the **Add Panels** window, and then click **Done**. Each panel contains one report.

Note: You can add only one instance of each panel to a single tab. All instances of a report display identical data. For example, if you change the local report filter on one tab, all instances of the same report in all tabs filter the data accordingly.

- To change the layout of a tab, click the Settings icon in the tab, choose **Change Layout**, choose the number of panel columns in the **Tab Layout** dialog box, and then click **Done**.
- To rename a tab, click the Settings icon in the tab, choose **Rename Tab**, and then enter the new name.
- To remove a tab, click the Settings icon in the tab and then choose **Remove Tab**.
- To rename a panel in the tab, click the Settings icon in the panel, choose **Rename**, and then enter the new name. The panel retains its original name in the Dashboard catalog.
- To remove a panel from the tab, click the Settings icon in the panel and then choose **Remove**. The panel remains available in the Dashboard catalog.
- To restore the default Dashboard and remove any additional panels or tabs that you added, click **Restore Default Dashboard**.

4. In the **Filter** pane, define the filters that you want to use and click **Apply Filters**.

Note: The filters apply to all reports except for the reports that show unresolved events: the unresolved errors view in the Errors by Application report and the Events by Current Status report.

5. To change the time frame filter in the reports that show unresolved events, click **Customize...** below the **Filter** pane and define the time frame that you want to view.
6. Navigate to the panel that you want to view and set the display preferences as needed.
7. To export data from a specific panel to a CSV file, click **Export to CSV** in the specific panel and save the file.

CHAPTER 15

Glossary

aggregated subscription

A subscription that consumes multiple data sets from the same topic with a single batch workflow. An aggregated subscription can use an automatic mapping or a custom mapping to process data. When you use an automatic mapping, the subscription sorts the data according to the publication date and time of the publication instances.

application

An entity that represents a system in your organization that needs to share data with other systems. An application can be a publisher and a subscriber. An application can publish multiple data sets.

automatic Data Integration Hub mapping

A mapping that Data Integration Hub automatically generates to process a data set with a basic transformation logic that maintains the structure of the data. When you create a publication or a subscription with an automatic Data Integration Hub mapping, Data Integration Hub also creates the PowerCenter sources, targets, metadata folders, and connections to process the data set.

batch workflow

A workflow that runs once and stops after completion. The workflow reads from a file, a database, or another source and writes to a target. Use a batch workflow to process Data Integration Hub publications and subscriptions with a PowerCenter workflow or a Data Engineering Integration mapping.

big data

A set of data that is so large and complex that it cannot be processed through standard database management tools.

catalog

A list of all available topics in Data Integration Hub. You can subscribe to any topic in the catalog. You can also drill down to view the topic structure, and to view and edit the associated publications and subscriptions.

child event

An event within the hierarchy of another event that acts as a parent event. The child event is a subsidiary of the parent event. A child event in Data Integration Hub represents a subscription.

compound subscription

A custom subscription, either a custom batch subscription or a custom cloud subscription, that consumes data sets from multiple topics. All topics to which the subscription subscribes must be of the same topic type. For example, all the topics are managed on a relational database publication repository.

custom Data Integration Hub mapping

A mapping that processes a data set. The mapping includes the data sources and targets, metadata folders, and connections to process the data.

A custom mapping uses a PowerCenter workflow, a Data Engineering Integration mapping, or an Informatica Cloud task to process data. Workflows and tasks can perform complex transformations on the data.

A custom mapping that uses a PowerCenter workflow or a Data Engineering Integration mapping can include parameters.

Data Integration Hub repository

A relational database table set that contains the metadata required to process publications and subscriptions in Data Integration Hub. It also contains the events that Data Integration Hub generates while it processes publications and subscriptions.

Data Integration Hub server

A service that manages publication and subscription processing in Data Integration Hub. The Data Integration Hub server triggers batch workflows and sends and receives notifications from PowerCenter.

Data Integration task

A Data Integration task is a process that you configure to analyze, extract, transform, and load data. In Data Integration Hub, a Data Integration task is a task that reads from a file, a database, or another source and writes to a target. Use Data Integration tasks to process Data Integration Hub publications and subscriptions with Informatica Intelligent Cloud Services.

When you use a Data Integration task to process publications, you use the Data Integration Hub cloud connector as the publication target. When you use a Data Integration task to process subscriptions, you use the Data Integration Hub cloud connector as the subscription source.

document store

File directory where Data Integration Hub stores files that are published with a publication that is configured with the **When the file is ready to be published** scheduling option. When the application publishes from a file source with a publication with a **When the file is ready to be published** schedule, Data Integration Hub copies the file from the source directory to the document store to process the file contents. The document store directory must be accessible to the Data Integration Hub server, the Apache Tomcat server, and the PowerCenter Integration Service with the same file path.

event

An occurrence of a publication or subscription at each stage of processing. The Data Integration Hub server generates the event and updates the event status while it processes the publication or subscription.

Operation Console

Web interface to manage applications and topics, run and monitor publications and subscriptions, and administer user access in Data Integration Hub. Use a web browser to access the Operation Console.

parent event

An event at the top level of a hierarchy of events. A parent event in Data Integration Hub represents a publication. The parent event shows the status of both the parent event and of the child events that represent the subscribers to this publication.

publication

An entity that defines data flow from a data source to the Data Integration Hub publication repository and the data publishing schedule. The publication publishes the data to a topic that defines the structure of the data in the publication repository. When a publication runs, Data Integration Hub extracts the data set from the application, processes the data, and writes the data to the publication repository. You can then create one or more subscriptions to process and write the published data set to target applications.

publication repository

A relational database table set that contains published data sets that subscribers can consume. After all subscribers consume the data and the retention period expires, Data Integration Hub deletes the data set from the publication repository.

real-time workflow

A PowerCenter workflow that is scheduled to run continuously and does not need to be started by Data Integration Hub. The workflow reads data from real-time sources and writes the data to Data Integration Hub. For example, data that is published through web-service providers or Java Message Service (JMS) queues. Use a real-time workflow to read data from real-time sources and write the data continuously to Data Integration Hub. When publishing through a real-time workflow, you group the published data into single publications at predefined time intervals.

subscription

An entity that defines the type, format, and schedule of data flow from the Data Integration Hub publication repository to a data target. When a subscription runs, Data Integration Hub extracts the data set from the publication repository, processes the data, and writes the data to the target application. You can subscribe to one or more topics. Each topic to which you subscribe can contain data from multiple publishers.

topic

An entity that represents a data domain that applications publish and consume through Data Integration Hub. A topic defines the data structure and additional data definitions such as the data retention period. Multiple applications can publish to the same topic. Multiple applications can consume data from the same topic.

unbound subscription

A subscription that is not restricted to specific publication instances. It subscribes to all the data that a publication publishes and consumes the data based on the subscription filter, regardless of when or in what batch the data was published.

INDEX

A

- add table from file
 - wizard [61](#), [62](#)
- Add Table from File wizard
 - definition [61](#)
 - structure [62](#)
- add topic table
 - create [63](#)
 - from file [61](#)
 - from PowerCenter workflow [62](#)
- add topic tables
 - from database [60](#)
- application
 - definition [26](#)
 - managing [29](#)
 - permissions [21](#)
 - properties [26](#)
- architecture
 - components [17](#)
- automatic publication
 - data-driven [72](#)
 - flat file source [70](#)
 - pass-through file source [71](#)
 - relational source [69](#)
- automatic subscription
 - data-driven [125](#)

B

- batch workflow
 - publication process [75](#)
 - subscription process [128](#)
- big data
 - description [18](#)

C

- catalog
 - subscribe to topic [25](#)
- cloud
 - custom publication [73](#)
 - custom subscription [126](#)
- cloud task
 - publication process [76](#)
 - subscription process [128](#)
- create topic table
 - properties [63](#)
- creating
 - topic [37](#), [39](#)
- custom publication
 - cloud [73](#)
 - multi-latency [73](#)

- custom subscription
 - cloud [126](#)
 - workflow [125](#)

D

- Dashboard
 - CSV file structure [197](#)
 - Error Event Distribution report [195](#)
 - Errors by Application report [191](#)
 - Event Average Processing Time report [196](#)
 - Event Distribution report [194](#)
 - Events by Current Status report [193](#)
 - reports [191](#)
- Dashboard and reports
 - managing [199](#)
 - overview [190](#)
- data-driven
 - publication process [76](#)
 - subscription process [129](#)
- database topic table
 - properties [60](#)
- deleting
 - topic [37](#)
- delivery
 - subscription [132](#)

E

- editing
 - topic [37](#)
- event
 - monitoring [182](#)
 - tracking [182](#)
- event monitoring
 - overview [177](#)
- events
 - actions [178](#)
 - advanced search properties [180](#)
 - basic search properties [179](#)
 - managing [178](#)
 - overview [177](#)
 - publications and subscriptions [181](#)
 - searching for [179](#), [180](#)
 - types and statuses [181](#)
- events page
 - how to access [178](#)
 - how to search [179](#), [180](#)

F

- field mapping
 - join table [115](#), [116](#)

- field mapping (*continued*)
 - map join tables [115](#)
 - map source fields [116](#)
 - map source tables [115](#)
 - map target fields [170](#)
 - map target tables [170](#)
 - source field [116](#)
 - source table [115](#), [116](#)
 - target field [171](#)
 - target table [170](#)
- file repository
 - data retention [65](#)
- flat file source
 - configure [111](#)

G

- glossary
 - of terms [200](#)

H

- hadoop repository
 - data retention [66](#)
- Hub Overview
 - diagram [19](#)
- Hub Overview diagram
 - description [19](#)

M

- managing
 - topics [37](#)
- map join table
 - properties [115](#), [116](#)
- map source field
 - properties [116](#)
- map source table
 - properties [115](#), [116](#)
- map target field
 - properties [171](#)
- map target table
 - properties [170](#)
- modular publication
 - cloud source [74](#)
- modular subscription
 - cloud target [127](#)
- monitor
 - events [182](#)
- monitoring
 - overview [177](#)
 - rules [182](#)
- monitoring rule
 - creating [183](#)
 - deleting [183](#)
 - disabling [183](#)
 - editing [183](#)
 - enabling [183](#)
 - properties [185–189](#)
- monitoring rule properties
 - general [185](#)
 - select actions [188](#)
 - select entity [186](#)
 - select status [187](#)
 - summary [189](#)

- monitoring rules
 - managing [183](#)
- monitors page
 - how to access [183](#)

O

- operation console
 - description [19](#)
- overview
 - description [14](#)

P

- permissions
 - applications [21](#)
- post-processing
 - subscription [129](#)
- pre-processing
 - publication [76](#)
- privileges
 - topics;applications;publications;subscriptions [21](#)
- publication
 - creating [79](#)
 - definition [21](#)
 - deleting [79](#)
 - disabling [79](#)
 - editing [79](#)
 - enabling [79](#)
 - field mapping [115](#), [116](#)
 - filter [77](#)
 - filter conditions [77](#)
 - flat file source [111](#)
 - overview [68](#)
 - pre-processing [76](#)
 - process [74](#)
 - properties [107–110](#), [113](#), [114](#), [117–120](#)
 - running [79](#)
 - schedule [78](#)
 - source [77](#)
 - type [68](#)
- publication filter
 - advanced expression [118](#)
 - basic expression [117](#)
- publication process
 - batch workflow [75](#)
 - cloud task [76](#)
 - data-driven [76](#)
 - real-time [75](#)
- publication processing
 - Informatica Cloud mapping [109](#)
 - workflow [108](#)
- publication properties
 - data-driven publication [120](#)
 - general [107](#)
 - join [114](#)
 - processing [108](#), [109](#)
 - target [114](#)
- Publication properties
 - field mapping [114](#)
 - filter [117](#)
 - schedule [118](#)
 - source [109](#)
 - summary [119](#)
- Publication Source
 - definition for flat file [110](#)

Publication Source (*continued*)
 definition for pass-through file [113](#)
 definition for relational database [109](#)
publications
 event types and statuses [181](#)

R

real-time workflow
 publication process [75](#)
relational database repository
 data retention [65](#)
rule
 monitoring [182](#)

S

schedule
 publication [78](#)
 subscription [131](#)
source
 flat file [77](#), [130](#)
 pass-through file [77](#), [130](#)
 publication [77](#)
 relational [77](#), [130](#)
subscription
 creating [132](#)
 custom [125](#), [126](#)
 definition [21](#)
 deleting [132](#)
 delivery [132](#)
 disabling [132](#)
 editing [132](#)
 enabling [132](#)
 field mapping [170](#)
 filter [131](#)
 filter conditions [131](#)
 overview [121](#)
 post-processing [129](#)
 process [127](#)
 properties [162](#), [164–169](#), [172](#), [173](#), [175](#)
 running [132](#)
 schedule [131](#)
 target [130](#)
 type [122](#)
 unbound [129](#)
subscription filter
 advanced expression [172](#)
 basic expression [172](#)
subscription mapping
 automatic [164](#)
 custom [165](#)
 Informatica Cloud mapping [164](#)
 Informatica Intelligent Cloud Service task [165](#)
 workflow [165](#)
subscription process
 batch workflow [128](#)
 cloud task [128](#)
 data-driven [129](#)
subscription properties
 data-driven subscription [175](#)
 general [162](#)
 join [166](#)
 processing [164](#), [165](#)
 source [166](#)

Subscription properties
 field mapping [169](#)
 filter [172](#)
 schedule [173](#)
 summary [175](#)
 target [166](#)
Subscription Target
 definition for flat file [167](#)
 definition for pass-through file [168](#)
 definition for relational database [166](#)
subscriptions
 event types and statuses [181](#)

T

table mapping
 flat file [116](#)
 relational database [115](#)
target
 subscription [130](#)
topic
 create [37](#)
 creating [39](#)
 definition [30](#)
 delete [37](#)
 edit [37](#)
 field names [36](#)
 managing [37](#)
 overview [20](#)
 properties [57](#), [58](#), [60](#), [65](#), [66](#)
 publication repository [30](#)
 relational database [39](#)
 structure [32](#), [34](#), [41–44](#), [48](#), [50](#), [51](#), [60–63](#)
 table names [36](#)
 table relations [37](#)
 tables [34](#), [60–63](#)
topic properties
 data retention [65](#), [66](#)
 general [58](#)
 permissions [66](#)
 structure [60](#)
 summary [66](#)
 table relations [65](#)
topic structure
 create table [44](#), [51](#)
 from database [41](#), [48](#)
 from file [42](#), [48](#)
 from PowerCenter workflow [43](#), [50](#)
topics
 catalog [23](#)
 details [23](#)
 subscribing from catalog [25](#)
track
 events [182](#)

U

unbound
 subscription [129](#)
user roles
 operator [22](#)

W

workflow

custom subscription [125](#)

workflow topic table

properties [63](#)