

Informatica Cloud Integration Hub

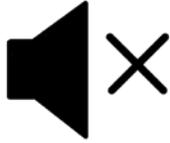
Ragavendra Kinnera

Principal Customer Success Technologist



Informatica™

Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available to view on our **INFASupport YouTube channel** and **Success Portal**. The link will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

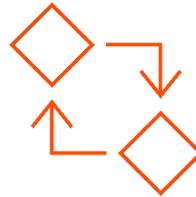
Feature Rich Success Portal



Bootstrap trial and
POC Customers



Enriched Customer
Onboarding
experience



Product Learning
Paths and Weekly
Expert Sessions



Informatica
Concierge with
Chatbot integrations



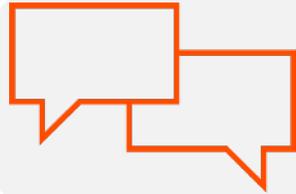
Tailored training and
content
recommendations

More Information



Success Portal

<https://success.informatica.com>



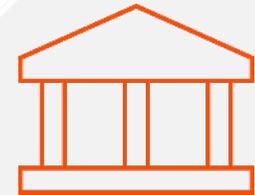
Communities & Support

<https://network.informatica.com>



Documentation

<https://docs.informatica.com>



University

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

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

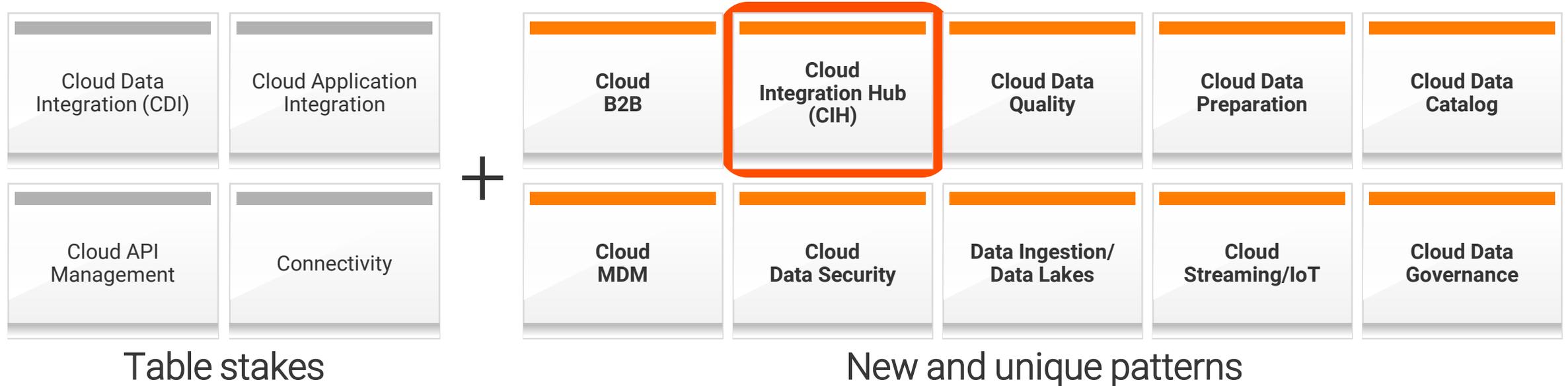
Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Agenda

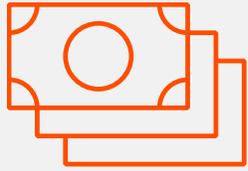
- Introduction to Cloud Integration Hub
- Applications
- Topics
- Publications
- Subscriptions
- Data Integration Tasks
- Rest API's
- Demo
- Q&A

Informatica iPaaS (IICS): Any data, any user, any pattern

Cloud Integration Hub (CIH) is part of IICS



Challenges that companies are facing



Cost per integration going up

- Too many developers hand coding integrations
- Draining resources on point-to-point integration tasks
- Data transfer fees and API call charges going up



Data loss causes business to slow or stop

- Data loss/delay occurs as tightly coupled applications fail and stop communicating
- Downtime getting costlier by the minute
- Network bandwidth is limited

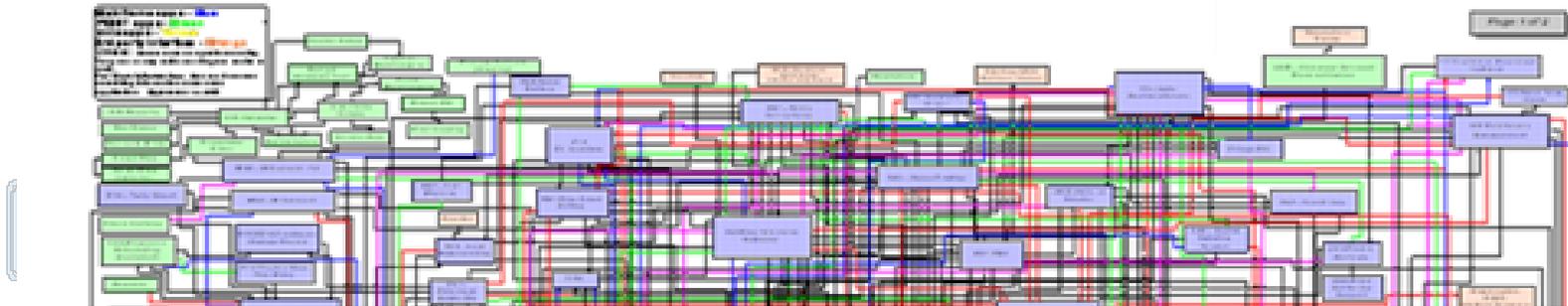


Sub-optimal integration design delaying time to market

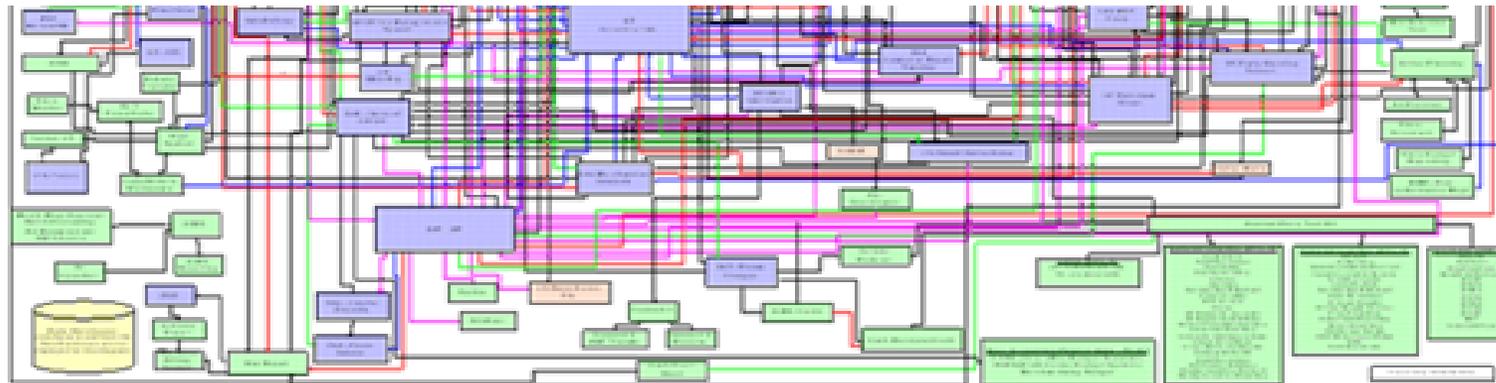
- Business users do not have access to right data
- Not able to bring systems on-board fast enough
- Data stored far from consuming applications

Example 1: Cost per integration – how it's going up

Simple data flow mappings turn into a “hairball” as the integration tasks gets repetitive and complex

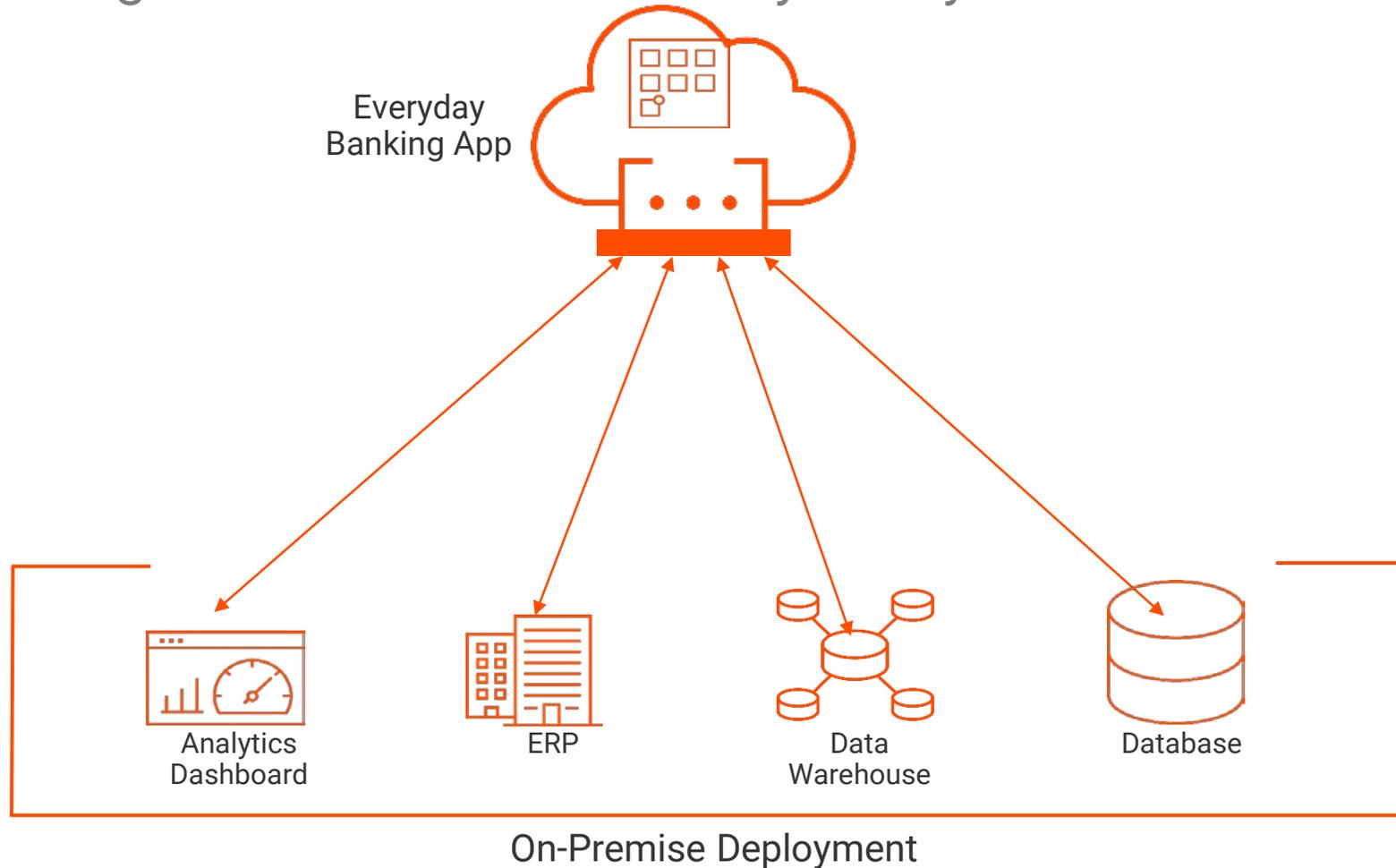


Point to point data integration architecture increases the number of connections and the cost



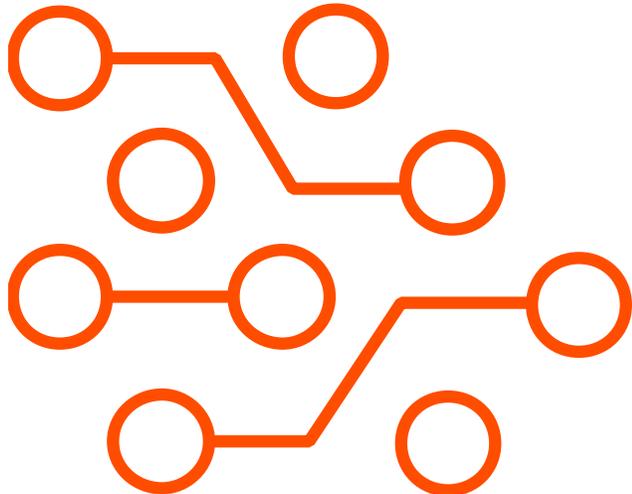
Example 2: Cost is going up due to integration

You pay API charges and transfer fees every time you fetch data from the cloud



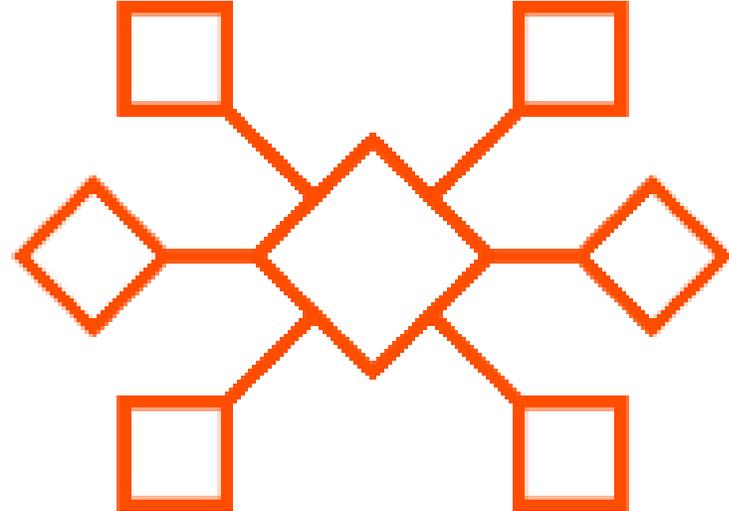
A Hub and a Spoke integration pattern is what's needed

It improves internal efficiency by reducing the number of connections



Point to Point Integration

10 sources and 10 targets mean 100 connections

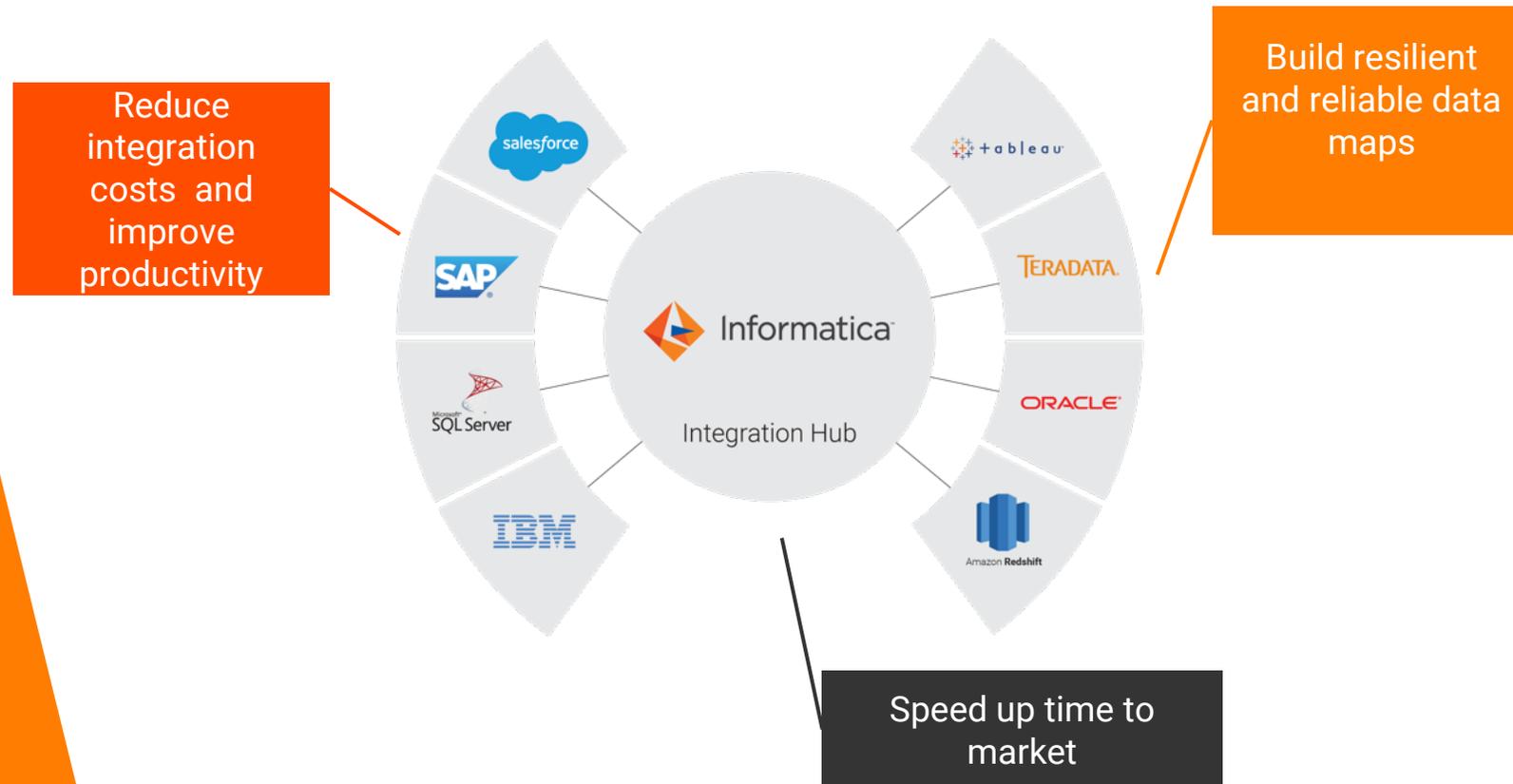


Hub n Spoke Integration

10 sources and 10 targets mean 20 connections

Integration Hub

Accelerate data-driven digital transformation

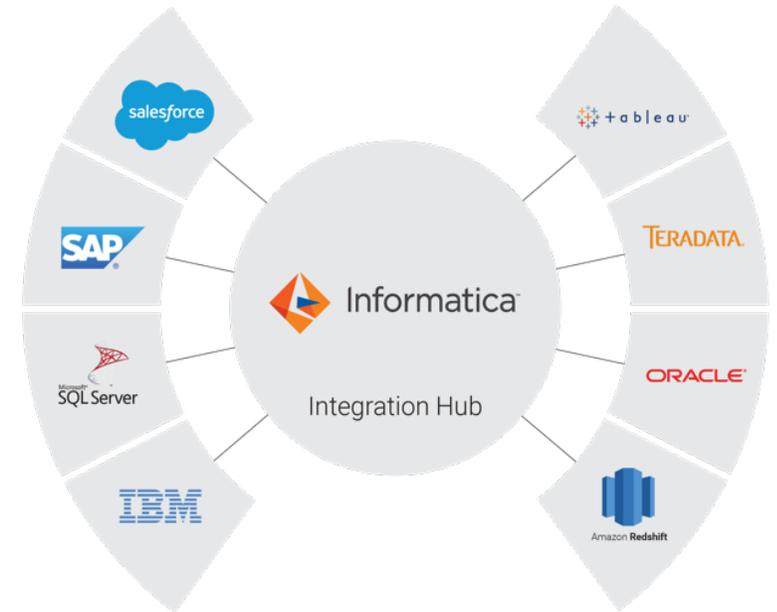


Unify. Govern. Share.

- Single hub to **unify** multiple clouds, big data, streaming sources and any existing systems
- **Self-Service** to increase agility and put data in the hands of the business
- Greater **visibility** and **control** of data
- Reduce the no. of interfaces each application has

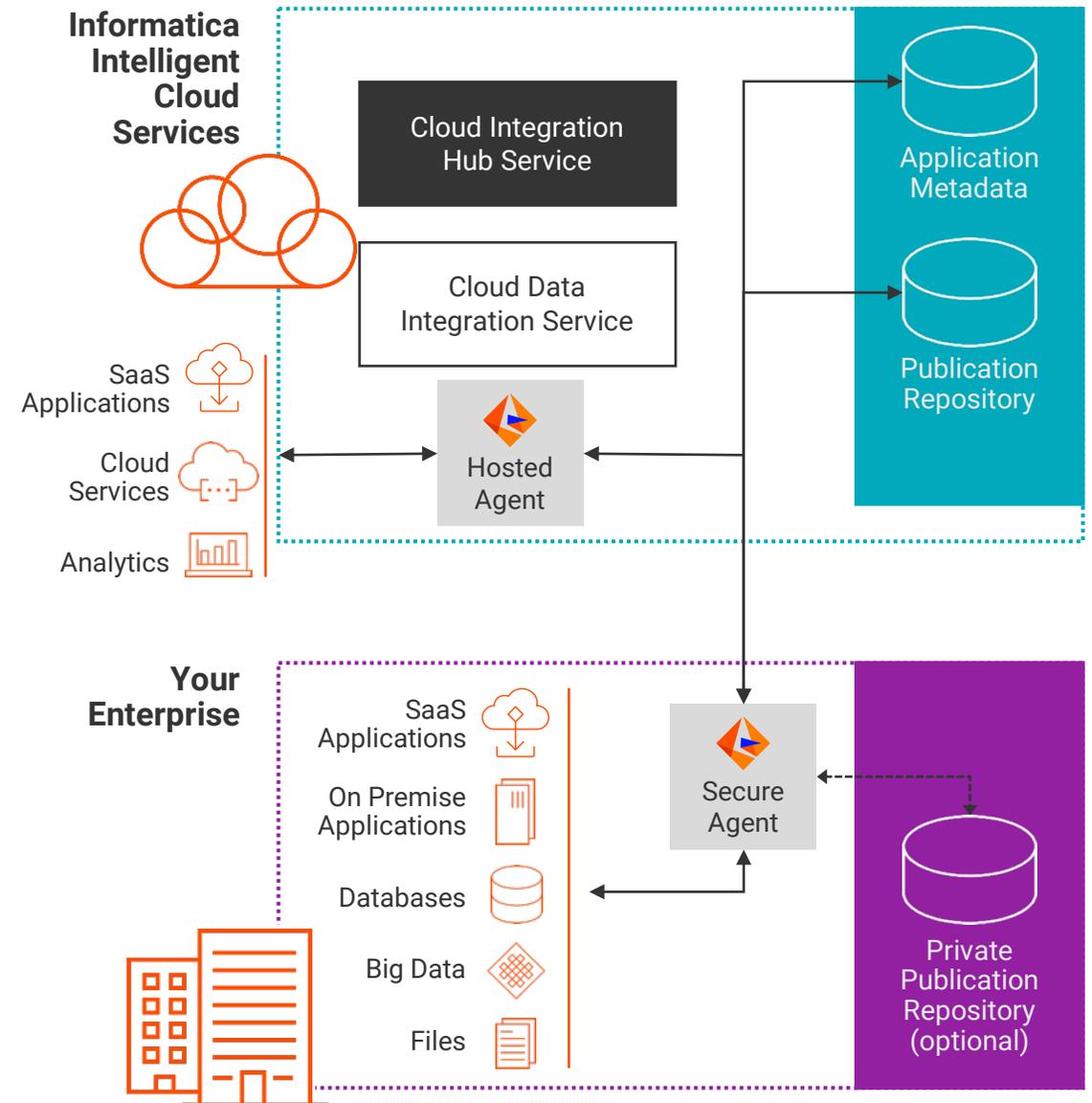
Integration Hub – Features Highlights

- **Decouple** source from target
- **Standardize** data access with canonical Topics
- **Simple** topic creation using metadata files or from a connection
- **Hosted** or **Private** Publication Repository
- Invoke **complex** cloud data integration batch mappings and tasks
- Publish and Subscribe to topics with direct **REST APIs**
- Internal **scheduler** for Pubs/Subs (External scheduler supported)
- View end to end data flow with integration **lineage**
- Track and **monitor** events with granular search
- Analyze existing DI mappings and get Hub pattern recommendations*



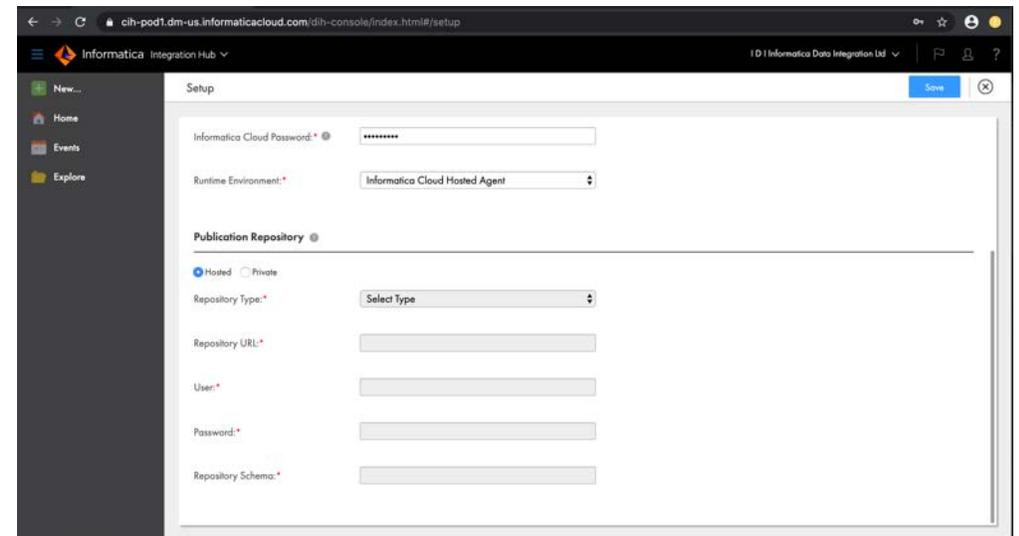
Cloud Integration Hub Architecture

- Hosted on IICS
- Support full Cloud and Hybrid mode
- Hosted or Private Publication Repository
- Secure and Encrypted, SOC-2 Compliant
- Source and Target De-coupling
- Multi-Latency, Batch & APIs
- Connectivity to any service or application



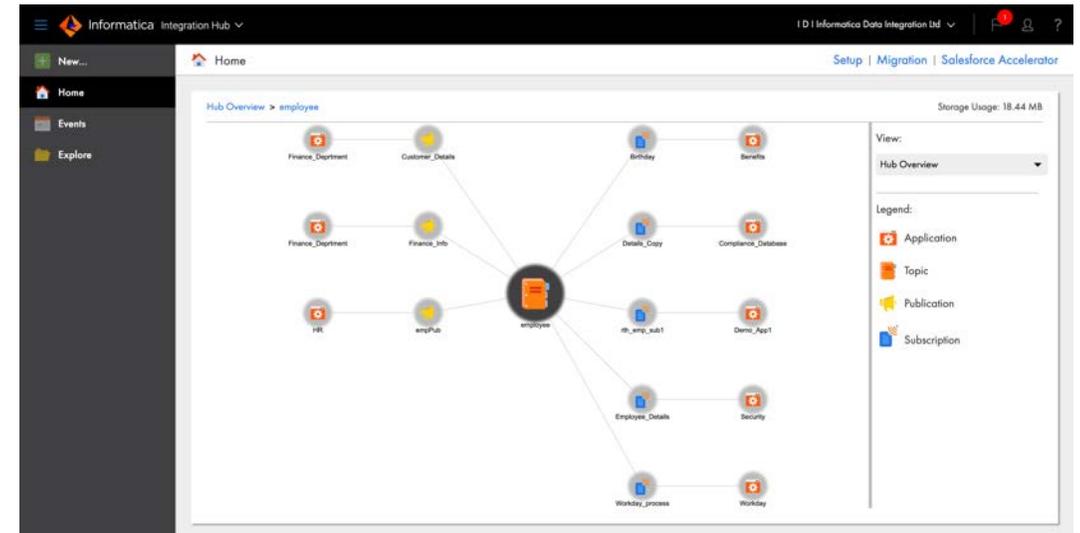
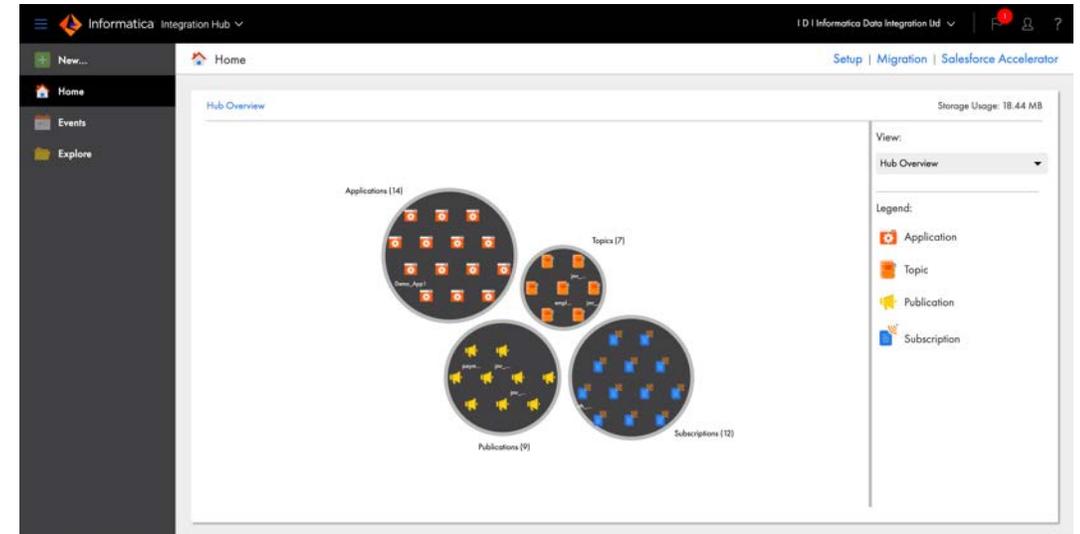
Publication Repository Service (PRS)

- A highly scalable transient data store that hosts the topics in relational structure
- Hosted on the robust Informatica infrastructure
- Can be set to Private repository mode for on-premise, customer managed mode (Oracle, MS-SQL or MySQL)



Cloud Integration Hub – Home View

- Provides an overview of Hub assets and their relationships
- Filters for most used topics, errors etc
- Clickable links for quick access to different asset pages
- Lineage view for end-to-end data flow visibility



Glossary of terms for Cloud Integration Hub

- **Application** – Logical entity that represents an external service that publishes or subscribes data to/from CIH
- **Topic** – Logical object that defines a data subject area or domain
- **Publication** – Provides data to the hub via an Informatica Cloud Integration task or an API
- **Subscription** – Consumes data from the hub via an Informatica Cloud integration task
- **Publication Repository** – The underlying transient data store that hosts the Topics

Applications

- An application is a logical entity that represents any application, DB or other connection that writes to or reads from the hub.
- It is comprised of a name and description and a list of associated publications and subscriptions. It is mainly used to simplify the lineage view and tie in the pub and sub processes with business entities in the real world

The screenshot displays the Informatica Integration Hub interface for configuring an application. The left sidebar shows navigation options: New..., Home, Events, Explore, and Finance_Department (selected). The main content area is titled 'Finance_Department' and includes a 'Save' button. Under 'General Details', the 'Application Name' is set to 'Finance_Department' and the 'Description' field is empty. Below this, the 'Publications (2)' section shows a table with two entries:

Name	Description	Mode	Topic	Last Modified
Customer_Details		Enabled	employee	03/04/2019
Finance_Info		Enabled	employee	24/12/2018

Below the publications table, the 'Subscriptions (0)' section shows a table with no data to display.

Topics

- A logical entity that holds a defined data set for a specific domain.
- The topic defines how data is written to and read from the hub.
- A topic structure is a set of one or more tables that are associated with a certain data domain
- Topic tables can be created manually from a file or from a connection schema
- Topics can be used as a standardization and facilitate consistent data proliferation across the organization
- The retention period defines the number of days that a published data set is persisted in the hub (for consumed and unconsumed data)

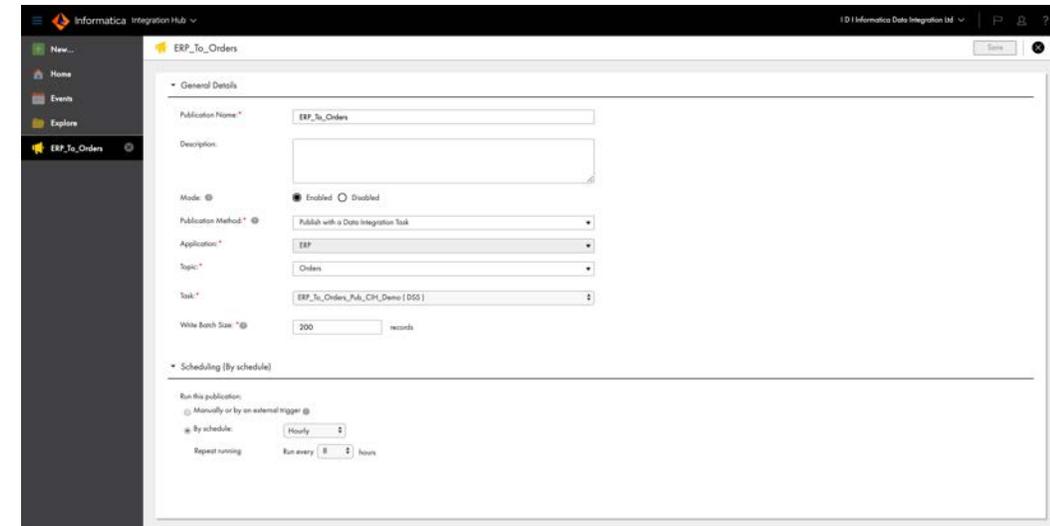
The screenshot displays the Informatica Integration Hub interface for configuring a topic named 'Orders'. The top section shows a Topic Diagram where 'Orders' is the central hub, connected to various source and target systems: ERP, ERP_to_Orders, Sub_Orders_EDW, Enterprise Data Warehouse, Sub_Orders_Finance, Finance, Sub_API_DataMart, and Sales DataMart. Below the diagram, the 'General Details' section includes fields for 'Topic Name' (Orders), 'Description' (Orders), 'Topic Type' (Incremental Load selected), and 'Publication data retention period' (1 days). The 'Topic Structure (2 tables)' section provides a detailed view of the tables within the topic, listing columns, data types, and precision. The bottom section shows 'Publications (1)' and 'Subscriptions (3)', with a table listing their names, descriptions, and status.

Table	Column	Filter Accelerator	Data Type	Precision	Scale	Encrypted
orders	orderNumber	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	orderDate	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	requiredDate	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	shippedDate	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	status	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	comments	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	ORD_customerNumber	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
orders	DH_PUBLICATION_INSTANCE_DATE	<input checked="" type="checkbox"/>	DATE_TIME			<input type="checkbox"/>
orders	DH_PUBLICATION_INSTANCE_ID	<input checked="" type="checkbox"/>	DECIMAL	19	0	<input type="checkbox"/>
order_details	orderNumber	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>
order_details	productCode	<input type="checkbox"/>	STRING	100		<input type="checkbox"/>

Name	Description	Mode	Last Modified	Last Run	Status	Number of consumed rows
Sub_API_DataMart		Enabled	29/09/2019	17/05/2020 11:00 AM	<input checked="" type="checkbox"/>	3
Sub_Orders_EDW		Enabled	10/03/2020	17/05/2020 11:00 AM	<input checked="" type="checkbox"/>	3
Sub_Orders_Finance		Enabled	10/03/2020	17/05/2020 11:00 AM	<input checked="" type="checkbox"/>	3

Publications

- A publication defines a dataflow and schedule for an application to write data into the hub
- A publication can be associated with either an IICS data integration task or an API
- Use API publications for event driven data processes and a DI task for batch-oriented data flow powered by Informatica flexible mapping capabilities.
- Every publication is tied to a topic – which is the target of the publication flow



Subscriptions

- A subscription is very similar to a publication, but the data flow is other way around. It reads data from the hub and writes it to the application.
- Like a publication, you can define a subscription that is based on an API or on a DI task
- If there is no existing DI task, you can generate a simple synchronization task for the subscription from within the subscription form
- Scheduling for DI based subscriptions can be defined for specific time, triggered when the associated publication is done or triggered by an external API
- API based subscriptions also includes an optional notification API

The screenshot shows the 'Sub_Orders_Finance' configuration page in the Informatica Integration Hub. The 'General Details' section includes:

- Subscription Name: Sub_Orders_Finance
- Description: (empty text area)
- Mode: Enabled Disabled
- Consumption Method: Consume data with a Data Integration task
- Unbound Subscription:
- Application: Finance
- Topics: Select a Topic (dropdown) + Add Topic (+)

Topic Name	Description	Type
Orders	Orders	Incremental Load

Below the table, there is a section for 'Task' configuration:

- Task: Orders_to_Finance_Sub_CH_Demo (DSS)
- Read Batch Size: 1000

The 'Scheduling' section is currently empty.

The screenshot shows the 'Sub_API_Datamart' configuration page in the Informatica Integration Hub. The 'General Details' section includes:

- Subscription Name: Sub_API_Datamart
- Description: (empty text area)
- Mode: Enabled Disabled
- Consumption Method: Consume data with an API
- Application: Sales Datamart
- Topics: Select a Topic (dropdown) + Add Topic (+)

Topic Name	Description	Type
Orders	Orders	Incremental Load

Below the table, there is a section for 'Run this subscription' with two options:

- When published data is ready
- Manually or by an external trigger

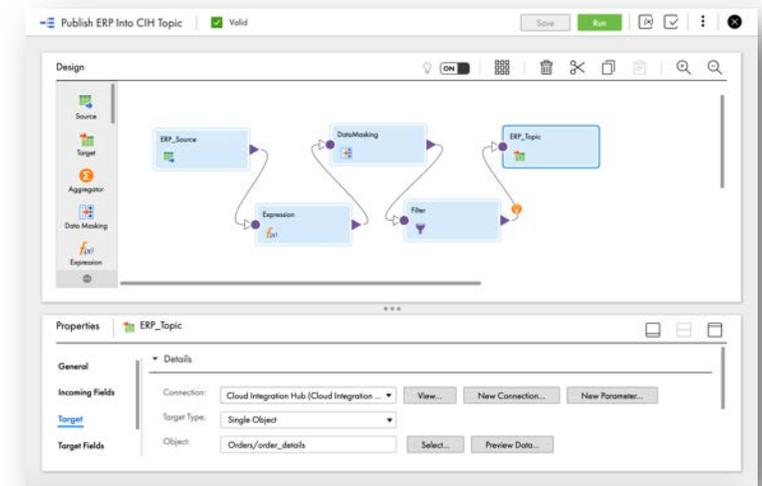
The 'By schedule' option is selected, with a dropdown set to 'Every 5 minutes'.

Additional configuration for API-based subscriptions:

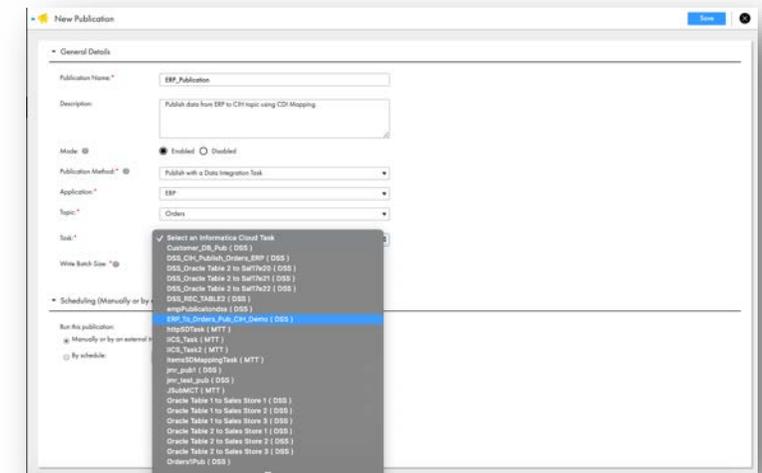
- To run the subscription, use the following URL: https://cjh-pdq1.dms.us.informaticahub.com:443/dh-pub-mpa-service/api/v1/subscriptions/Sub_API_Datamart/data (Copy)
- The Consume Data API returns the Swagger structure for the topics from which the subscription consumes data: https://cjh-pdq1.dms.us.informaticahub.com:443/dh-console/api/v1/subscriptions/Sub_API_Datamart/dmSwagger (Copy)
- Send notification when data is ready to consume:
- Notification URL: <https://hooks.zapier.com/hooks/catch/1584947/132y8nu/>

Integration Hub and Data Integration

- CIH is tightly coupled with CDI and can be perceived as another tier of data integration – one that facilitates the hub & spoke / pub-sub paradigm
- Publications & Subscriptions can invoke Data Integration Mappings and Synchronization Tasks – where the Target or Source of the DI asset is the relevant topic
- Simple DSS subscriptions can be generated from the CIH UI
- CDI assets can be seen in CIH and associated with CIH assets
- The pub/sub parameterization, scheduling, monitoring & reprocessing is managed by the Hub
- The status of the relevant tasks are captured in the Hub event view (including session logs & processing information)



CDI Mapping Designer – CIH Topic is the target



CIH Publication definition using the CDI mapping

Hub Pattern Recommendation Tool

- Analyze existing data integration mappings and patterns
- Get org specific recommendations to move to a hub pattern
- Create all associated assets (DI & CIH) with a single click

The screenshot displays the Informatica Data Integration Integration Hub interface. The top navigation bar includes the Informatica logo, 'Data Integration', and user information. A left sidebar contains navigation options: 'New...', 'Home', 'Explore', 'Bundles', 'Integration Hub', 'My Jobs', and 'My Import/Export Logs'. The main content area is titled 'Integration Hub' and shows 'Analysis Results | Updated on Sep 8, 2019, 11:28 PM Reprocess'. Below this, a summary row displays key metrics: 'DI Tasks: 572', 'Identical Sources: 13', 'Identical Targets: 0', 'Topics: 13', 'Publications: 13', and 'Subscriptions: 572'. A search bar is located below the metrics. The main area displays a grid of hub pattern recommendations, each with a 'Create' button. The first recommendation is 'Topic_REC_TABLE15', which shows a central hub icon connected to three peripheral icons: 'Pub_REC_TABLE15', 'Sub_REC_ORA_14', and 'Sub_REC_ORA_13'. Other recommendations include 'Topic_Account', 'Topic_orderA_txt', and 'Topic_REC_TABLE16', each showing a central hub icon connected to multiple peripheral icons representing various data sources and targets.

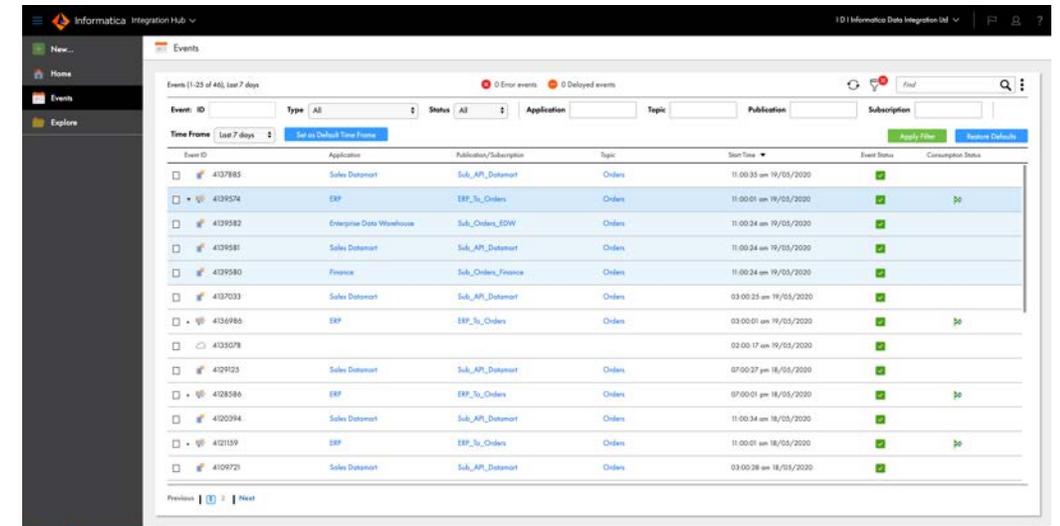
Available for existing Cloud Data Integration customers!

Integration Hub REST API's

- Run Publication Subscription REST API – Starts a publication or a subscription, including disabled publications and subscriptions, and returns the event ID of the publication or the subscription event that Cloud Integration Hub generates
- Publish Data REST API. Publishes data directly to a topic on the Cloud Integration Hub publication repository. Returns the status of a publication process
- Consume Data REST API. Consume data directly from a topic on the Cloud Integration Hub publication repository
- Change Publication Subscription Mode REST API – toggles the mode of a publication or a subscription, that is, enables a disabled publication or subscription and disables an enabled publication or subscription
- Reprocess Event REST API. Reprocesses a publication or subscription event, including events of disabled publications and subscriptions
- Event Status REST API. Returns the status of a publication or subscription event
- Catalog REST API. Extracts data from the Cloud Integration Hub catalog, including topic, publication, and subscription metadata

Events

- The events view shows you a list of all the events, scheduled or manually triggered and their status.
- The status can be completed, errored or delayed. Delayed means that a publication is ready to be consumed but is not consumed yet (didn't trigger yet)
- Additional details on the associated data flow can be accessed – such as timing, number of processed rows, log info etc.
- You can also invoke action through this screen – for example reprocess a failed subscription



Event ID	Application	Publication/Subscription	Topic	Start Time	Event Status	Consumption Status
4137885	Sales Datamart	Sub_AP_Datamart	Orders	11:00:35 am 19/05/2020	Success	
4138574	ERP	ERP_Sales_Orders	Orders	11:00:01 am 19/05/2020	Success	▶▶
4139582	Enterprise Data Warehouse	Sub_Orders_EDW	Orders	11:00:24 am 19/05/2020	Success	
4139581	Sales Datamart	Sub_AP_Datamart	Orders	11:00:24 am 19/05/2020	Success	
4139580	Finance	Sub_Orders_Finance	Orders	11:00:24 am 19/05/2020	Success	
4137033	Sales Datamart	Sub_AP_Datamart	Orders	03:00:25 am 19/05/2020	Success	
4136986	ERP	ERP_Sales_Orders	Orders	03:00:01 am 19/05/2020	Success	▶▶
4133078				03:00:17 am 19/05/2020	Success	
4129725	Sales Datamart	Sub_AP_Datamart	Orders	07:00:27 pm 18/05/2020	Success	
4128586	ERP	ERP_Sales_Orders	Orders	07:00:01 pm 18/05/2020	Success	▶▶
4120394	Sales Datamart	Sub_AP_Datamart	Orders	11:00:34 am 18/05/2020	Success	
4121159	ERP	ERP_Sales_Orders	Orders	11:00:01 am 18/05/2020	Success	▶▶
4120721	Sales Datamart	Sub_AP_Datamart	Orders	03:00:28 am 18/05/2020	Success	

Monitoring Rules

- Allows the user to define rules for alerts
- Create per user, per status
- Send an email to relevant stake holder when the defined rule is engaged
- For example one could define a rule to send an email to a specific department IT lead when a subscription to that department fails for some reason

The top screenshot shows the 'General Details' section of the 'Orders_Publication' rule configuration. The 'Rule Name' is 'Orders_Publication' and the 'Description' is 'Monitor Order Publication Errors'. The 'Mode' is set to 'Enabled'. The 'Content' section contains the following text: 'If the status of Publications: ERP_To_Orders Are in status: Error, Reprocessed, Warning Then: Send email to: lgal@informatica.com, mbolovan@informatica.com'. Below this is the 'Affected Assets' section, which shows a table of publications to be monitored.

Name	Description	Mode	Topic	Last Modified
<input type="checkbox"/> Customer_Details		Enabled	employee	10/24/2019
<input type="checkbox"/> empHub		Enabled	employee	1/15/2020
<input checked="" type="checkbox"/> ERP_To_Orders		Enabled	Orders	5/19/2020
<input type="checkbox"/> Finance_Info		Enabled	employee	6/27/2019
<input type="checkbox"/> prc_pub1		Enabled	prc_topic1	8/26/2019

The bottom screenshot shows the 'Affected Statuses' and 'Actions' sections. In the 'Affected Statuses' section, 'Error', 'Reprocessed', and 'Warning' are selected. In the 'Actions' section, 'Send email notification' is selected, and a table of email recipients is shown:

Username	Email
lgal@informatica.com	lgal@informatica.com
Non-existing User	mbolovan@informatica.com

Performance Tuning Guidelines

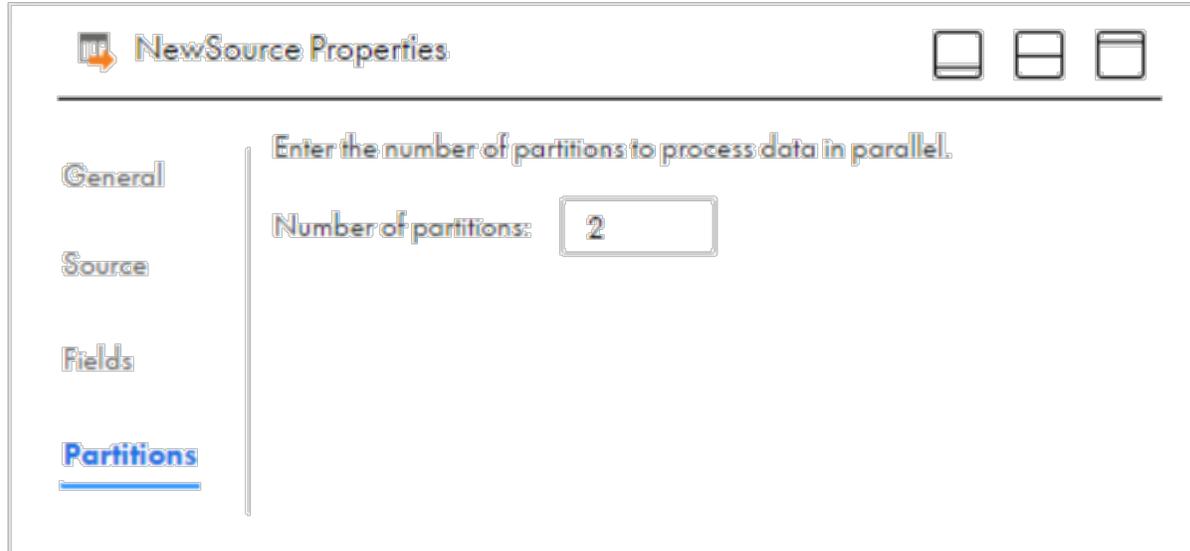
Source Reader Partitioning

<https://network.informatica.com/onlinehelp/IICS/dev/CDI/en/index.htm#page/jj-cloud-transformations/Partitions.html>

Partitioning with a Flat File Source

You have a mapping task that uses a large, 1GB flat file source. You want to specify two partitions in the Source transformation to optimize performance.

On the Partitions tab for the Source transformation, you enter the number of partitions, as shown in the following image:



Write Batch Size

Number of records that the CIH connector writes to the publication repository in a single batch

The screenshot shows the 'New Publication' configuration page in the Informatica Integration Hub. The page is divided into a left sidebar and a main content area. The sidebar contains navigation options: 'New...', 'Home', 'Events', 'Explore', and 'New Publication' (which is currently selected). The main content area is titled 'New Publication' and contains a 'General Details' section. This section includes several fields: 'Publication Name' (text input), 'Description' (text area), 'Location' (dropdown menu with 'Default' selected and a 'Browse' button), 'Mode' (radio buttons for 'Enabled' and 'Disabled', with 'Enabled' selected), 'Publication Method' (dropdown menu with 'Publish with a Data Integration Task' selected), 'Application' (dropdown menu with 'Select an Application' selected), 'Topic' (dropdown menu with 'Select a Topic' selected), 'Task' (dropdown menu with 'Select an Informatica Cloud Task' selected), and 'Write Batch Size' (text input with '200' and 'records' next to it). The 'Write Batch Size' field is circled in black.

Read Batch Size

Number of records that the CIH connector reads from the publication repository in a single batch

The screenshot shows the 'New Subscription' configuration page in Informatica Integration Hub. The left sidebar contains navigation options: 'New...', 'Home', 'Events', 'Explore', and 'New Subscription' (which is active). The main content area is titled 'New Subscription' and includes the following fields and controls:

- Mode:** Radio buttons for 'Enabled' (selected) and 'Disabled'.
- Consumption Method:** A dropdown menu with the selected option 'Consume data with a Data Integration task'.
- Unbound Subscription:** A checkbox that is currently unchecked.
- Application:** A dropdown menu with the placeholder text 'Select an Application'.
- Topics:** A dropdown menu with the placeholder text 'Select a Topic' and an 'Add Topic' button to its right.
- Table:** A table with columns 'Topic Name' and 'Description'. The table is currently empty, displaying 'No data to display'.
- Task:** A dropdown menu with the placeholder text 'Select an Informatica Cloud Task' and a 'Create New Task' button to its right.
- Read Batch Size:** A text input field containing the value '1000'. This field is circled in black in the original image.

Private Publication Repository – JDBC Direct

- Direct JDBC connection from CIH Connector to customer repository DB
- Supported only on Private PRS

The screenshot displays the Informatica Administrator web interface. On the left is a navigation menu with items: Organization, Licenses, SAML Setup, Settings, Users, User Groups, User Roles, Runtime Environments, Connections, Add-On Connectors, Schedules, Add-On Bundles, and Swagger Files. The main content area is titled 'Cloud Integration Hub' and contains the following configuration fields:

- Connection Details**
 - Connection Name: * Cloud Integration Hub
 - Description: (empty text box)
 - Type: * ? Cloud Integration Hub (Informatica LLC)
- Cloud Integration Hub Connection Properties** (?)
 - Runtime Environment: * ? __POD2_DO_NOT_CHANGE_tavr7iicsagnt
 - Cloud Integration Hub Console URL: * https://cih-pod2.dm-us.informaticacloud.com:443
 - Cloud Integration Hub Console Username: * zshemesh@informatica.com
 - Cloud Integration Hub Console Password: * (masked with dots)
 - Publication Repository Service URL: https://cih-pod2.dm-us.informaticacloud.com:443
 - Publication Repository URL: (empty text box)
 - Publication Repository Username: (empty text box)
 - Publication Repository Schema: (empty text box)
 - Publication Repository Password: (empty text box)
 - Do not use intermediate staging for subscription flows:
 - Use JDBC for Private Publication Repository:

Informatica Cloud Integration Hub is designed



- To bring down integration cost
- Increase data and integration reliability
- Improve IT efficiency and speed to market

Right data, in the hands of the right people, right when they need it

Demo



Thank You

