# Informatica Cloud Integration Hub
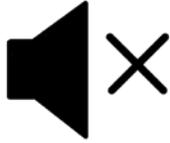
Ragavendra Kinnera

*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.
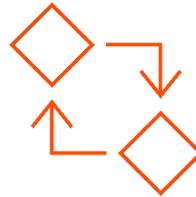
Informatica

# Feature Rich Success Portal

Bootstrap trial and POC Customers

Enriched Customer Onboarding experience

Product Learning Paths and Weekly Expert Sessions

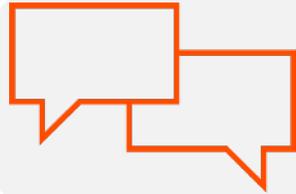Informatica Concierge with Chatbot integrations

Tailored training and content recommendations

Informatica®

# More Information
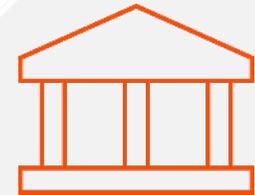
## Success Portal

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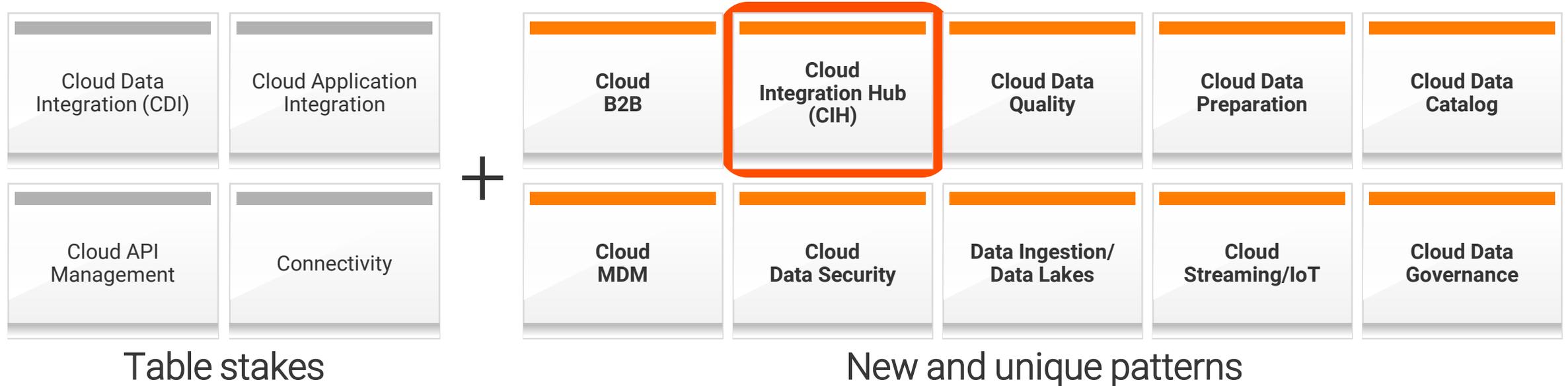
Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Informatica™

# Agenda

- Introduction to Cloud Integration Hub

- Applications

- Topics

- Publications

- Subscriptions

- Data Integration Tasks

- Rest API's

- Demo

- Q&A

Informatica

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

## Cloud Integration Hub (CIH) is part of IICS

| Cloud Data Integration (CDI) | Cloud Application Integration |
|---|---|
| Cloud API Management | Connectivity |

**+**

| Cloud B2B | Cloud Integration Hub (CIH) | Cloud Data Quality | Cloud Data Preparation | Cloud Data Catalog |
|---|---|---|---|---|
| Cloud MDM | Cloud Data Security | Data Ingestion/ Data Lakes | Cloud Streaming/IoT | Cloud Data Governance |

Table stakes

New and unique patterns

**Informatica™**

# Challenges that companies are facing
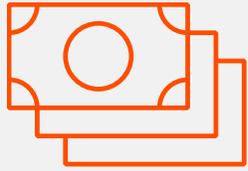
## Cost per integration going up

- 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

**Informatica**®

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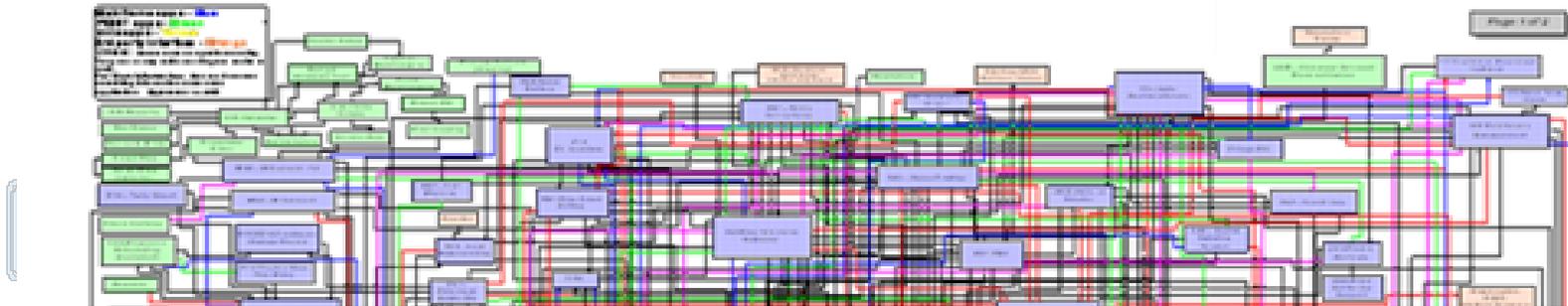
Point to point data integration architecture increases the number of connections and the cost

© Informatica. Proprietary and Confidential.

Informatica
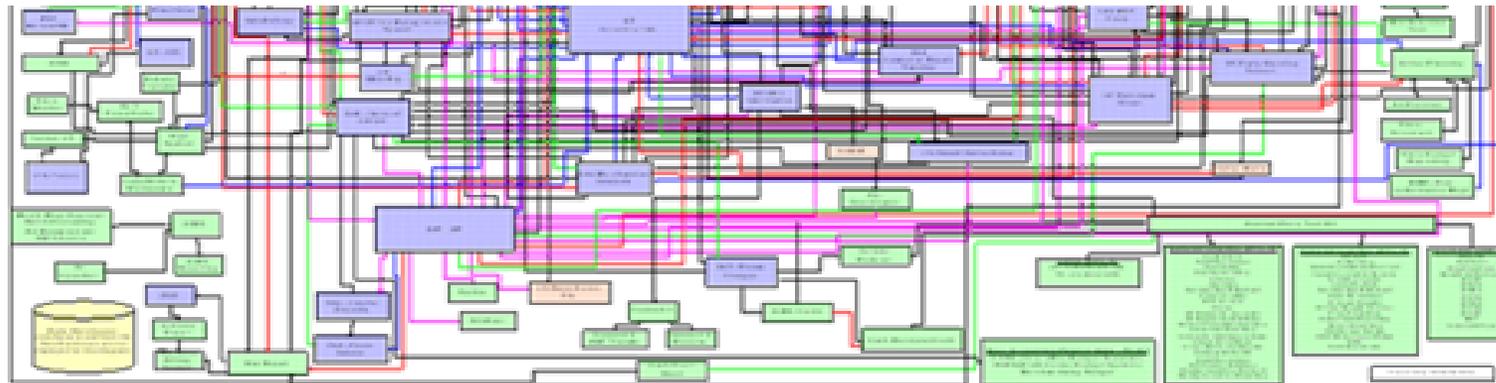
# Example 2: Cost is going up due to integration

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

Everyday
Banking App

Analytics
Dashboard

ERP

Data
Warehouse

Database

On-Premise Deployment

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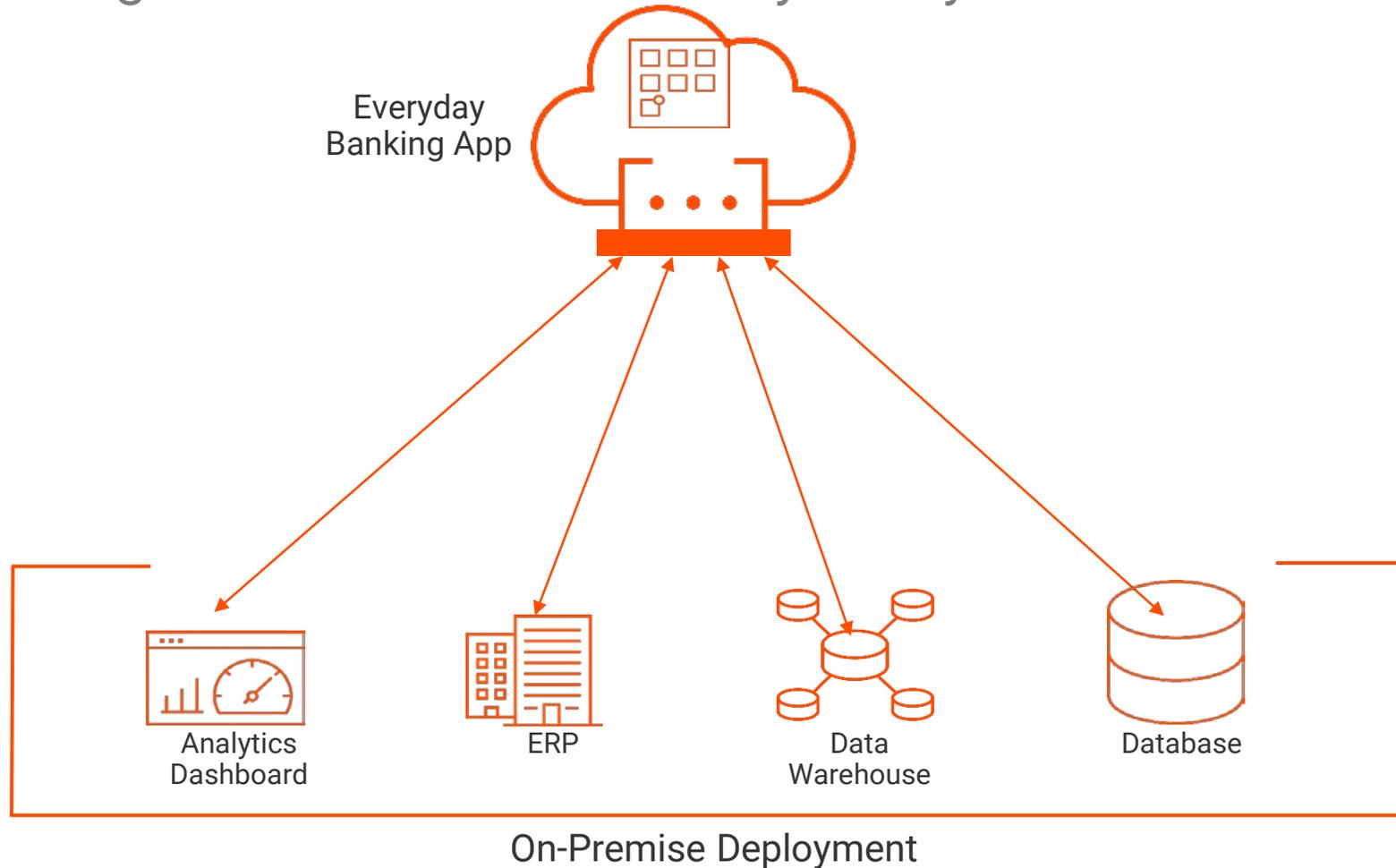
# A Hub and a Spoke integration pattern is what's needed

It improves internal efficiency by reducing the number of connections

Point to Point Integration
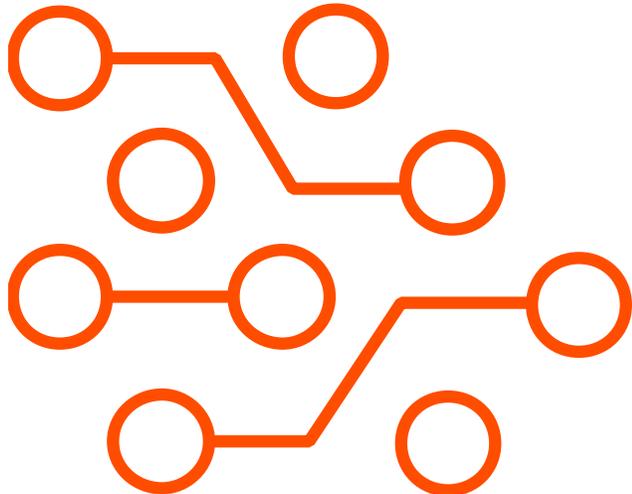
Hub n Spoke Integration

10 sources and 10 targets mean 100 connections

10 sources and 10 targets mean 20 connections

Informatica

# Integration Hub

## Accelerate data-driven digital transformation

Reduce integration costs and improve productivity

Build resilient and reliable data maps

Speed up time to market
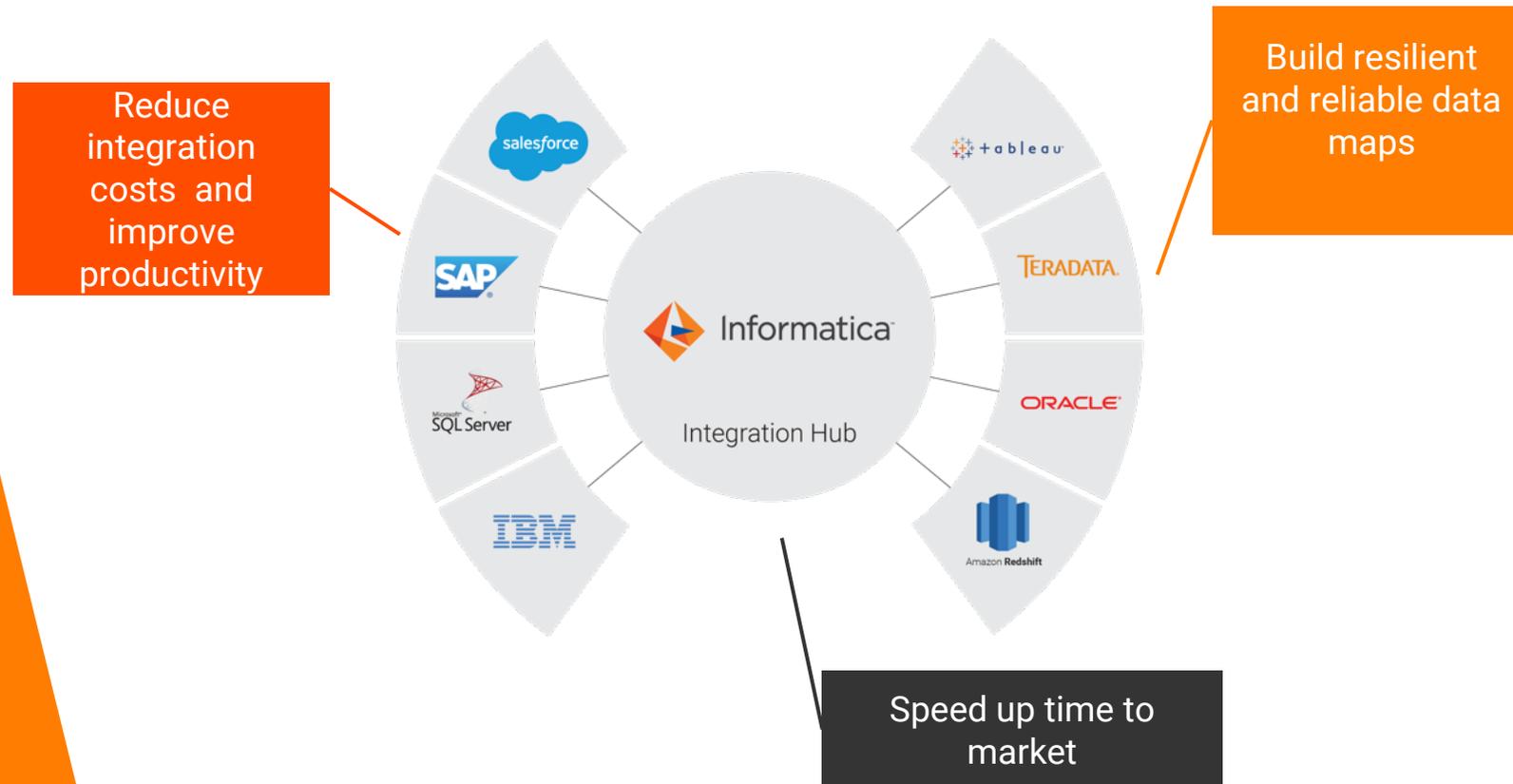
### Unify. Govern. Share.

- 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 API**s

- 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*
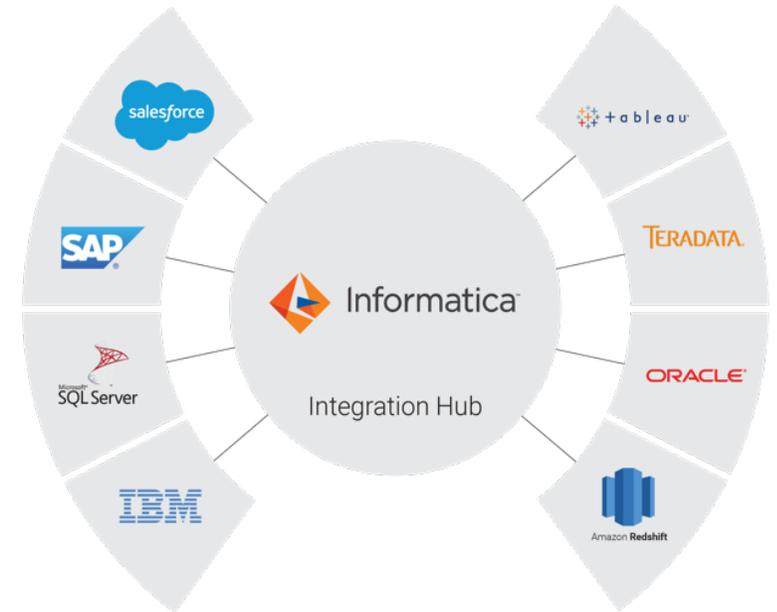
# Cloud Integration Hub Architecture

- 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
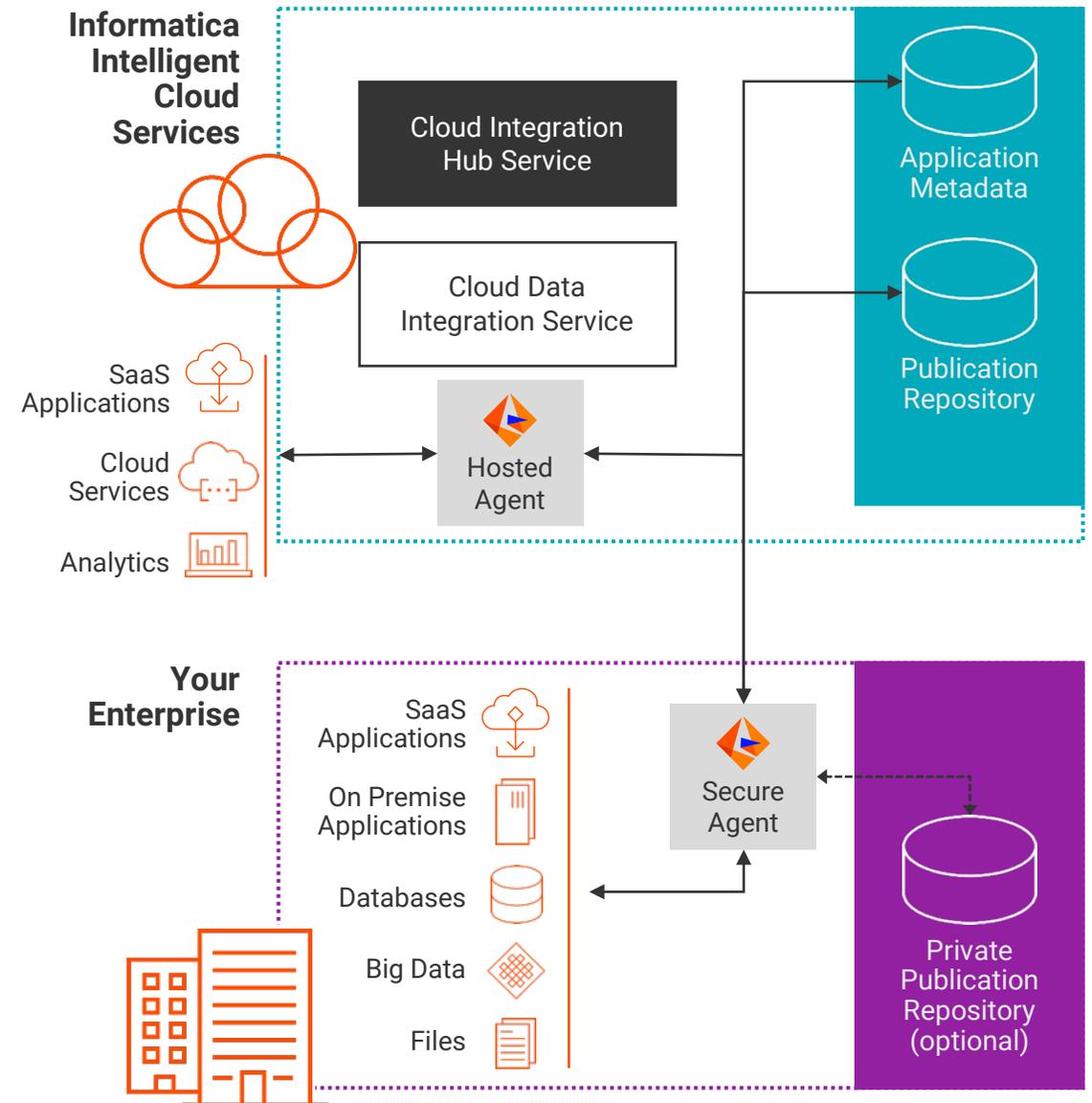
# Publication Repository Service (PRS)

- 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)
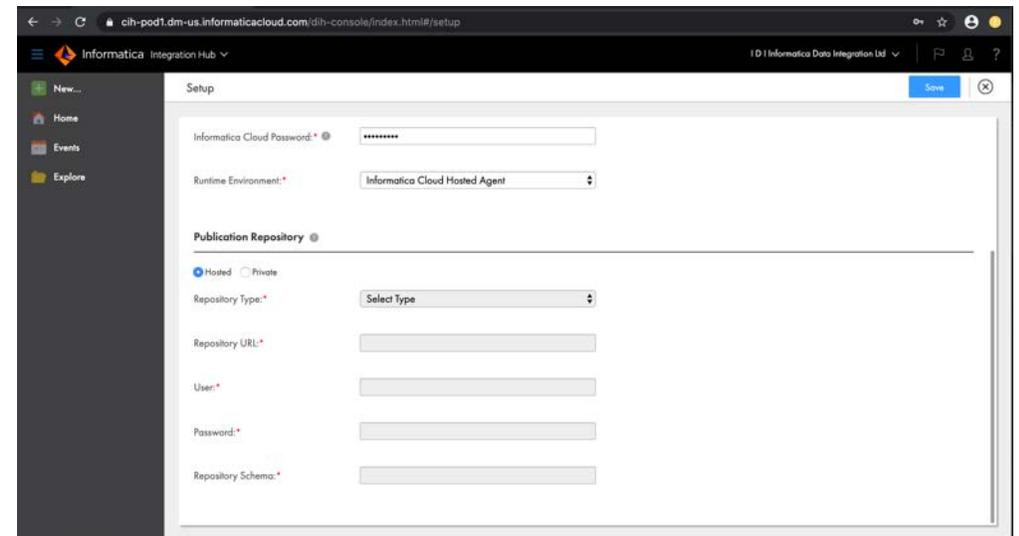
# Cloud Integration Hub – Home View

- Provides an overview of Hub assets and their relationships

- Filters for most used topics, errors etc

- Clickable links for quick access to different assest pages

- Lineage view for end-to-end data flow visibility
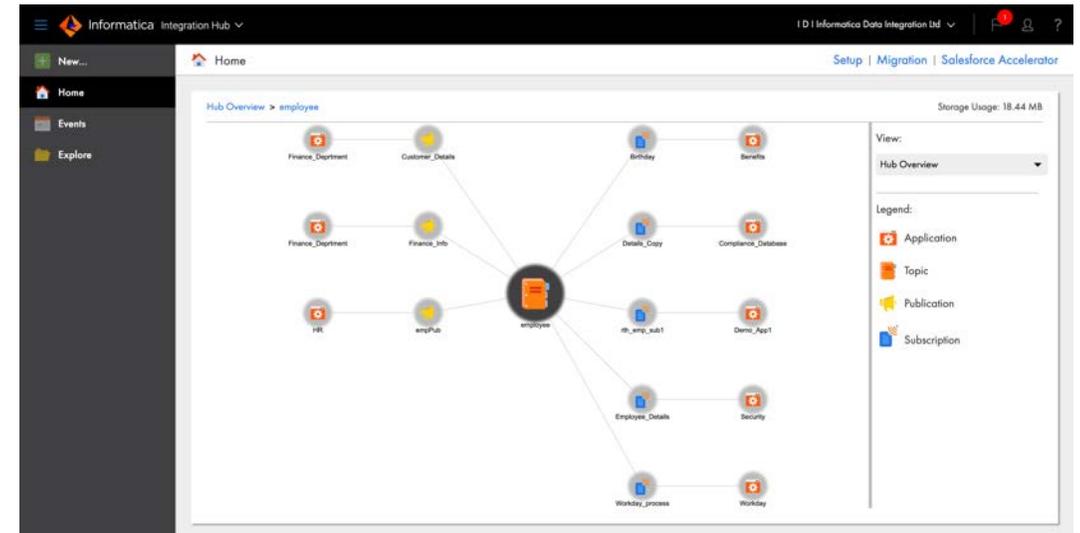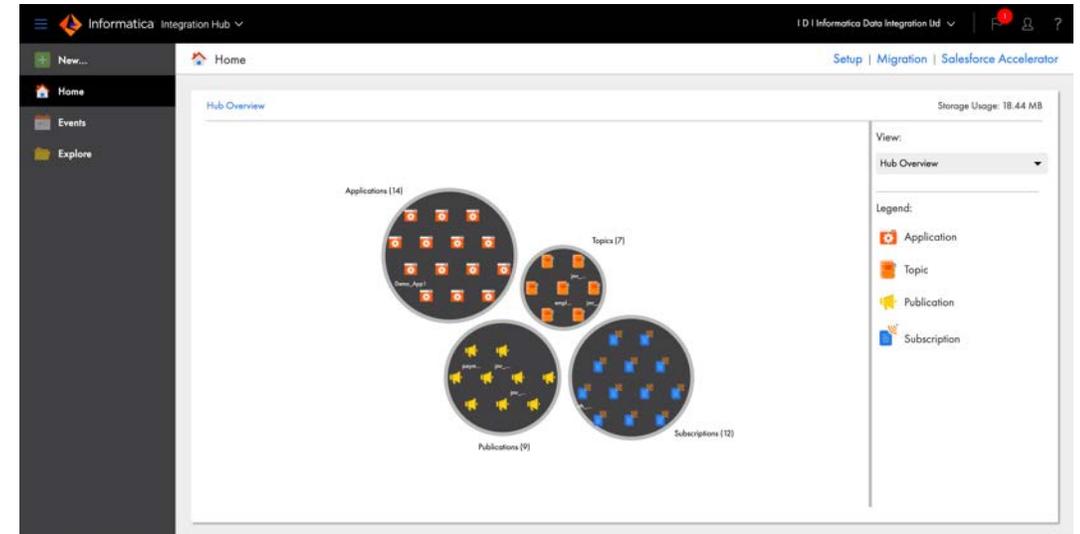
# Glossary of terms for Cloud Integration Hub

- **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

# 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)

# 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
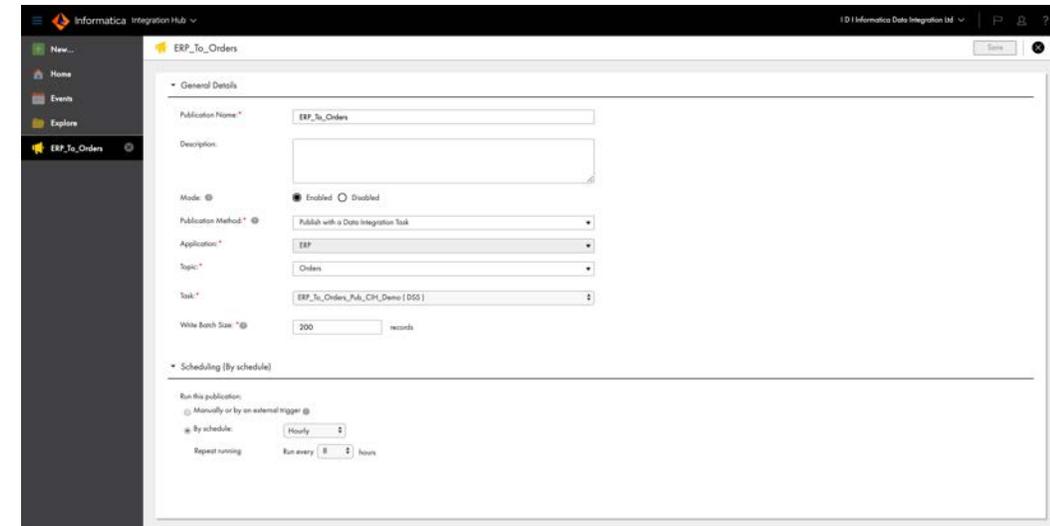
# Subscriptions

- 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

# 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)
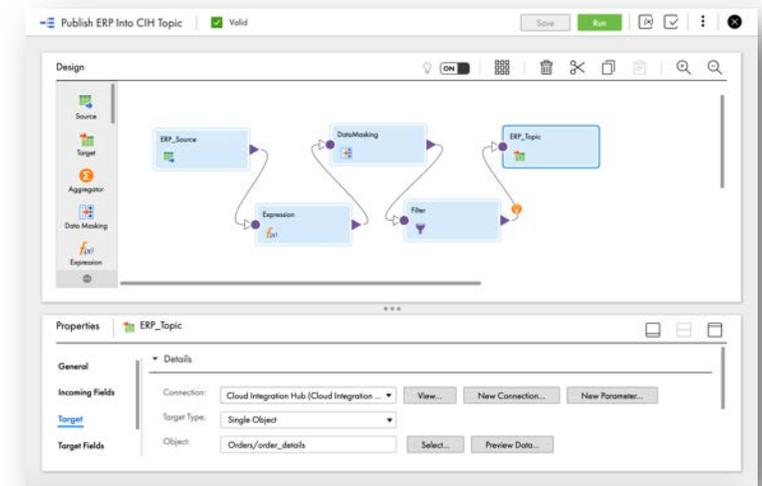


CDI Mapping Designer – CIH Topic is the target
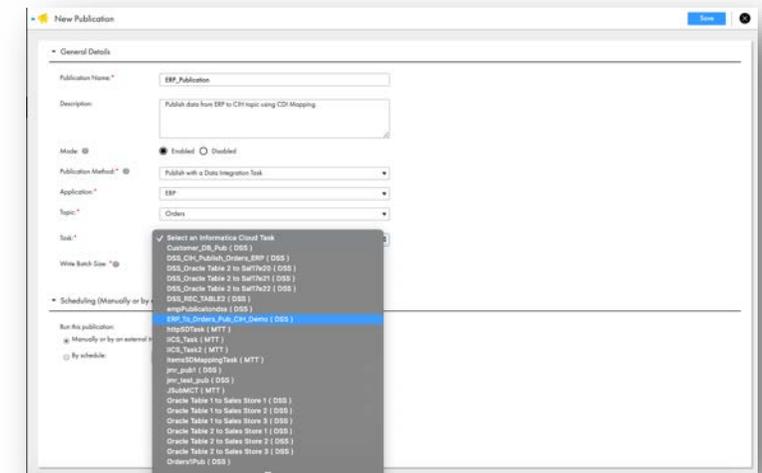


CIH Publication defintion using the CDI mapping

# 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



**Available for existing Cloud Data Integration customers!**

    *Tech Preview – Enabled on demand*

**Informatica**

# 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

Informatica

# 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

# 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

# Performance Tuning Guidelines

Informatica™

# Source Reader Partitioning

**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:

NewSource Properties

General

Source

Fields

**Partitions**

Enter the number of partitions to process data in parallel.

Number of partitions: 2
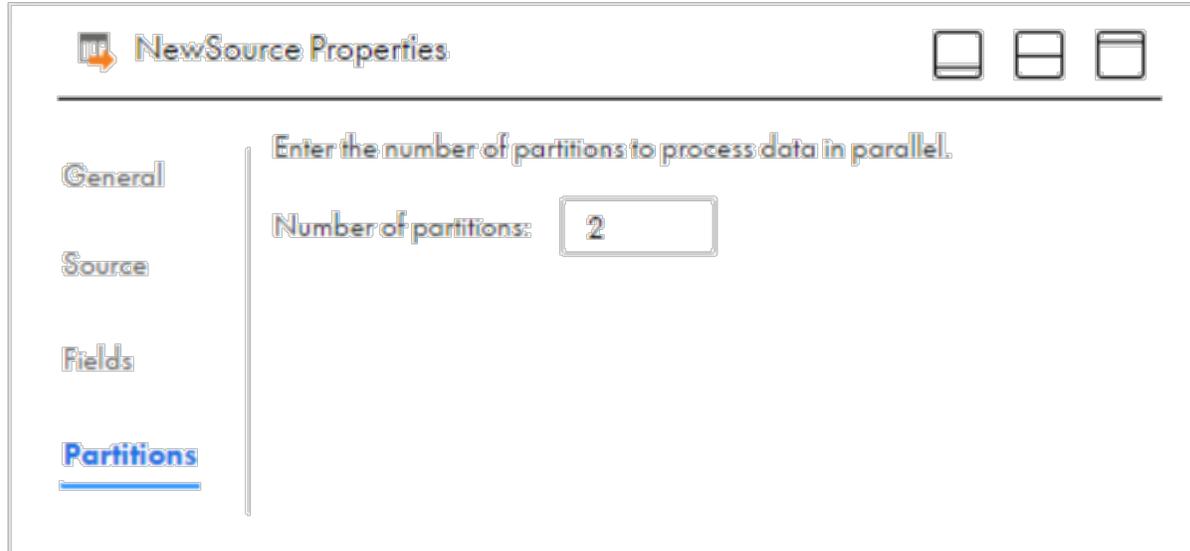
Informatica

# Write Batch Size

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

© Informatica. Proprietary and Confidential.

# Read Batch Size

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

© Informatica. Proprietary and Confidential.

# Private Publication Repository – JDBC Direct

- Direct JDBC connection from CIH Connector to customer repository DB

- Supported only on Private PRS

# 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

Informatica®

Demo

Informatica

Thank You