

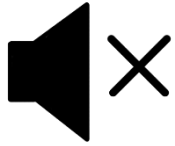
11 Nov, 2025

Unlock SAP Data: A Flexible, Low-Code Approach for Real-Time BI and AI

- Bharath Raghavendran, Principal Product Manager, R&D
- Faisal Ishaq, Senior Product Manager, R&D

Where data & AI come to 

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 on our [Success Portal](#) - where you can download the **slide deck** for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

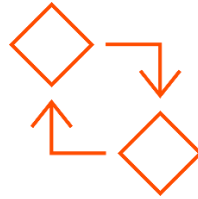
Feature Rich Success Portal



Bootstrap trial and
POC Customers



Enriched Customer
Onboarding
experience



Product Learning
Paths and Weekly
Expert Sessions

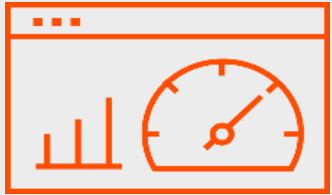


Informatica
Concierge



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

1 CDIR Overview

2 SAP Replication offerings

3 Fall Release highlights

3 Serverless Offering

5 Demo

6 Q&A

CDIR Overview

Cloud Data Ingestion and Replication

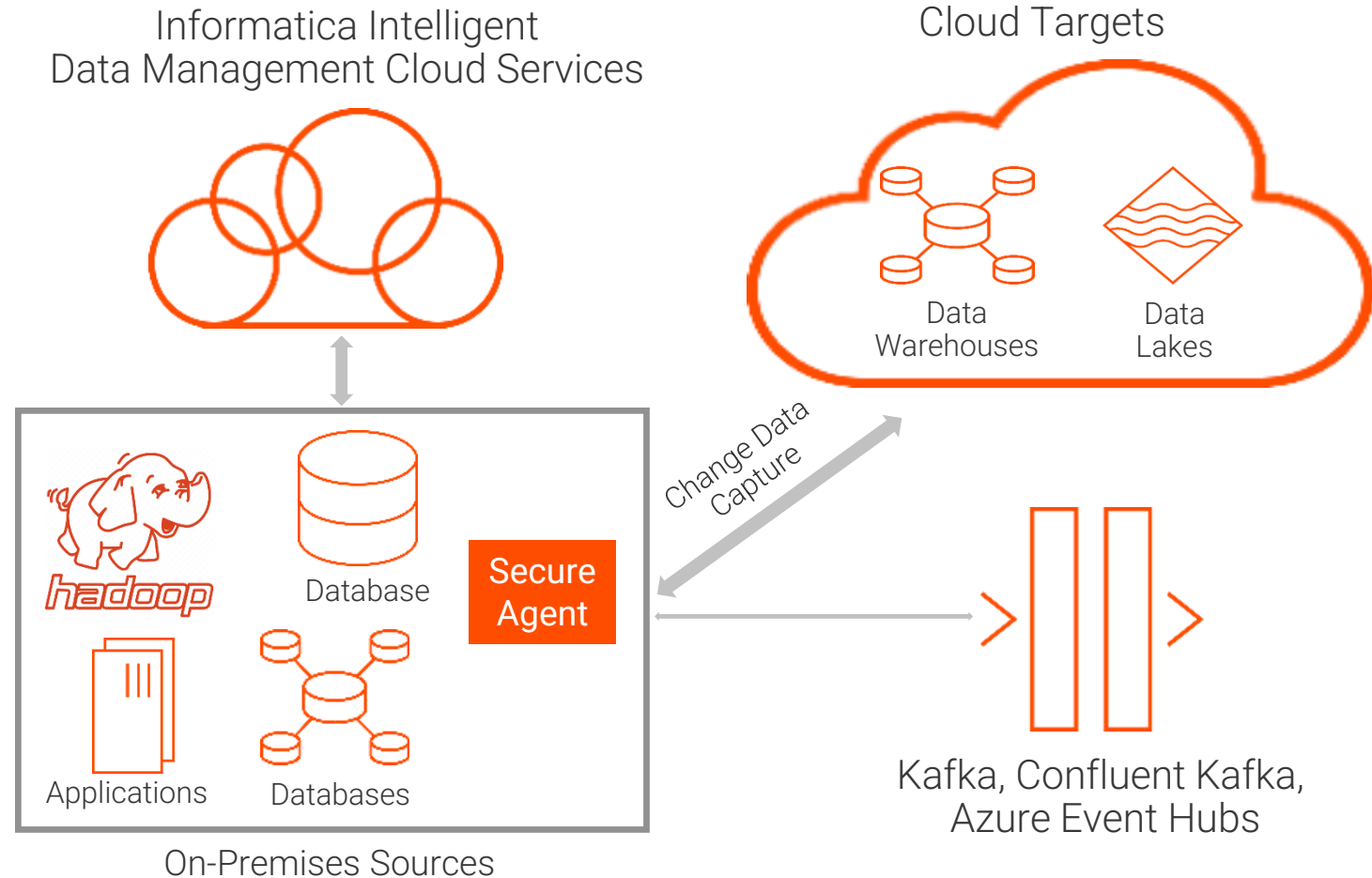
Apps, files, streaming, DBs and CDC

Provides wizard approach to quickly ingest data into the cloud (or to wherever the Secure Agent can reach)

Maintain synchronization between applications and sources as they continue to change

Real-time monitoring of ingestion jobs with lifecycle management and alerting in case of issues

Orchestrate data ingestion in **hybrid/cloud** as **managed** and **secure** service



Market Drivers for SAP Replication

Unprecedented growth in cloud is pushing the replication of operational data from SAP to cloud data warehouses and data lakes

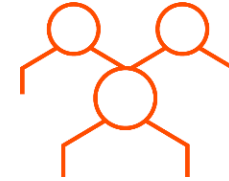


World's largest provider of **ERP** platform



SAP BW was used as analytics platform

Move to cloud



Data engineering **democratization**



Cloud used as consumption layer for analytics

Initial use case

Current state

Demo Overview

Use Case: Breaking the Data Barrier - MegaMart's Dual Pipeline Edge

MegaMart, a global retail giant, needs to integrate high-volume transactional data from SAP ERP with real-time customer interaction data from Oracle systems to ensure flawless Black Friday execution.

Business challenge

- Need to process and analyze massive volumes of orders and customer interactions
- Lack of unified view between operational data (SAP ERP) and customer engagement data (Oracle systems)

What do they want to do?

- Implement dual data pipelines combining SAP ERP transactional data and Oracle CDC streams into a centralized data warehouse
- Create automated monitoring and alerting system for supply chain optimization and dynamic pricing adjustments

SAP Replication Offerings



1

Application Layer

In this approach we can go to the application layer and get the data ingested from SAP source to any of the supported cloud target.

2

Database Layer

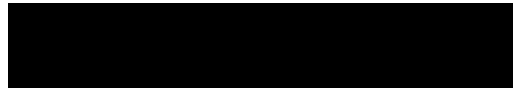
We can also replicate using the underlying database.

- Log-based replication
- Trigger-based replication (SAP HANA DB)

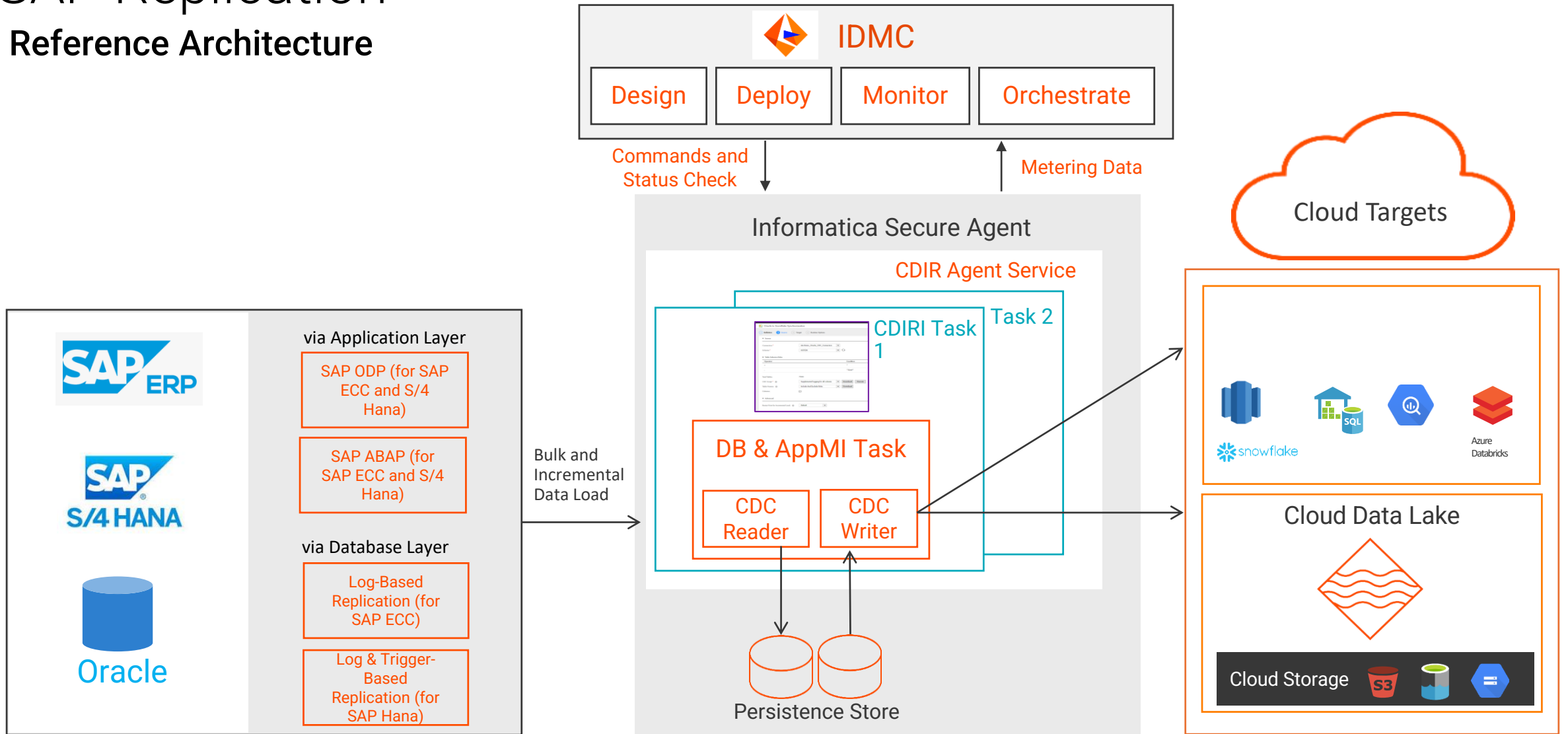
3

Hybrid (App + DB)

Use application layer for doing the bulk load and then use the DB layer log-based capture for CDC.



SAP Replication Reference Architecture



Data Replication using Application layer

SAP ODP Layer

- Uses Operational Data Provisioning (ODP) to read SAP objects data via their data sources
- Interacts with SAP application layer
- Leverages ODP “delta” capability to extract CDC information; probes every **CDC Interval** configured by the user
- CDIR supports bulk, incremental and combined load using this approach

SAP ABAP

- Uses ABAP for extracting data from SAP sources
- Interacts with SAP application layer
- CDIR supports bulk load using this approach

SAP OData V2

- Uses ODATA standards to read data from ODP data sources
- CDIR supports bulk, load using this approach

Data Replication using Database layer

SAP ECC

- Using Oracle/SQL Server/DB2 source and Database Mass Ingestion task
- Connects to underlying database for bulk load and its **log records** to get CDC data

SAP S/4 Hana [SAP HANA]

- Using HANA Mass Ingestion connector and Database Mass Ingestion service
- Connects to underlying Hana database to get initial data and its **log records** to get initial and CDC data
- Supports both HANA **triggers and log-based approach** to capture CDC data

Data Replication using Hybrid Approach

SAP ECC [Oracle]

- Using SAP Mass Ingestion connector and Application Mass Ingestion service
- Uses ABAP based connector for initial load and log-based CDC for incremental data

SAP S/4 Hana [SAP HANA]

- Using SAP HANA source and Database Mass Ingestion task
- Connects to underlying Oracle database and its **log records** to get initial and CDC data

Benefits

- Best of breed approach from both Application and Database layer
- Performance improvement over ODP for fetching incremental data
- Increased coverage of SAP ECC objects with the support of pool and cluster tables

Demo

Fall Release feature highlights

CDC Data staging



Key Highlights

- **Read log data once** and use it for multiple CDC tasks
- **Optimized** read from source with intermediate staging of required log/CDC-table data
- Define a **CDC data staging group** and associate it with multiple tasks
- Provide staging location on cloud data lakes
- Observability : Track the staging job as well as the replication jobs
- Supports migrating existing jobs to use CDC Data Staging



Benefits

- **Efficiency** – Access and read log data once and use it for multiple tasks/targets
 - Reduce the overhead on source log
- **High Performance Ingestion** – optimized read performance
 - CDC data is ready to be used by multiple tasks

New Staging Group



Overview

Custom Properties

Group Name: *

Staging Location Connection: *

Runtime Environment: *

Enable Alternate Connection for Reading Logs:

Alternate Connection for Reading Logs:

CDC Method:

Include LOBs:

Row Flush Threshold:

Flush Interval: M: S:

Save

Cancel

Row-level filtering



Key Highlights

- Filter data to be replicated according to values in specific columns.
- Use simple and Advanced filtering capabilities to define the filter conditions
- Ability to change filter conditions for a task
- Only relevant rows gets replicated to target

Benefits



- Security and regulations – prevent replication of sensitive and unauthorized data by filtering it out according to different filtering rules
- Time savings and improved productivity with replicating only the needed data
- Cost-effective – No need to replicate and store data that is not needed on target

Tables	Transformation	Filters
<input type="checkbox"/> > 0Audit0_SRC_ALLCHAR2_8K_Tet		
<input checked="" type="checkbox"/> > APPML15_ORA_EmailTemplate_Marketo		
<input type="checkbox"/> > APPML15_ORA_Folder_Marketo		
<input type="checkbox"/> > APPML15_ORA_Folder_Marketo_LOG2		
<input type="checkbox"/> > APPML15_ORA_LandingPageTemplate_...		

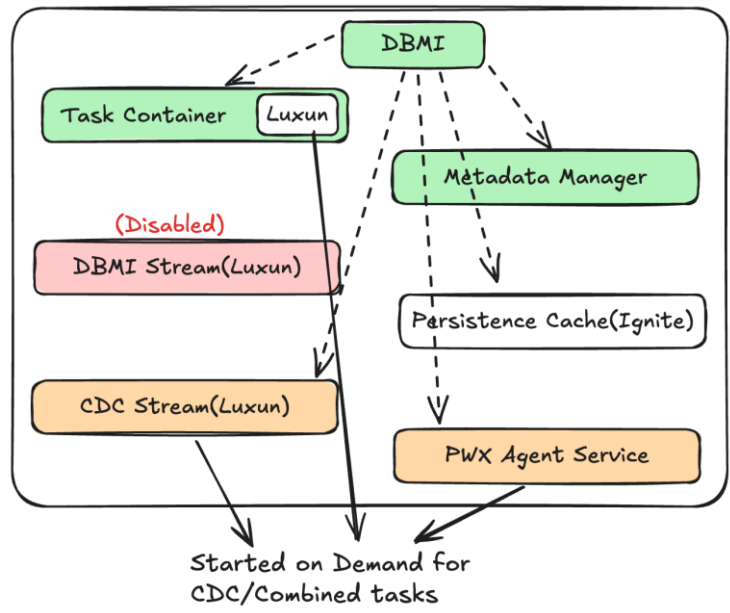
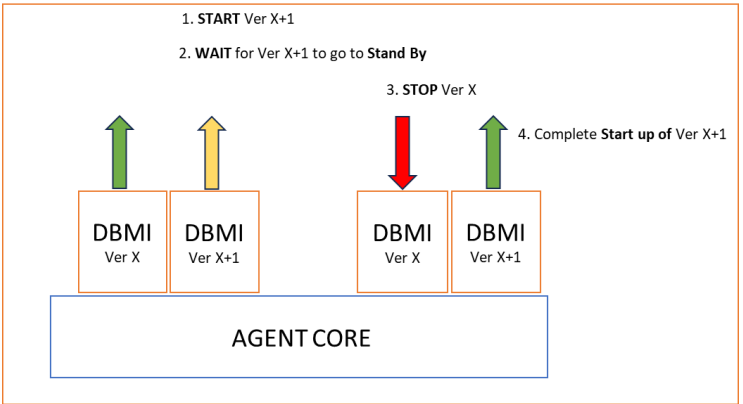
Column Name	Operator	Value
id (NUMBER)	>	0

Agent-ZDT



Key Highlights

1. Employs a **blue-green methodology of agent upgrade** where a newer version of the DBMI agent gets downloaded, runs essential services and as it reaches standby, old agent is shut down and jobs migrated to the new agent, resulting in **near-zero downtime**
2. Implements a **light-weight version** of the DBMI agent ensuring only key services are started during start-up
3. **Completely removes any dependencies on local cache persistence**, migrating existing caches to the cloud ensuring multiple agent versions could be started in parallel
4. Refactors code to **prevent multiple instances of services simultaneously accessing singleton agent storage**, allowing multiple concurrent agent versions to be available
5. All DBMI Agent logs such as Metadata Manager logs, Task Container logs, CDC stream logs etc **include their respective version information** to allow easy identification.

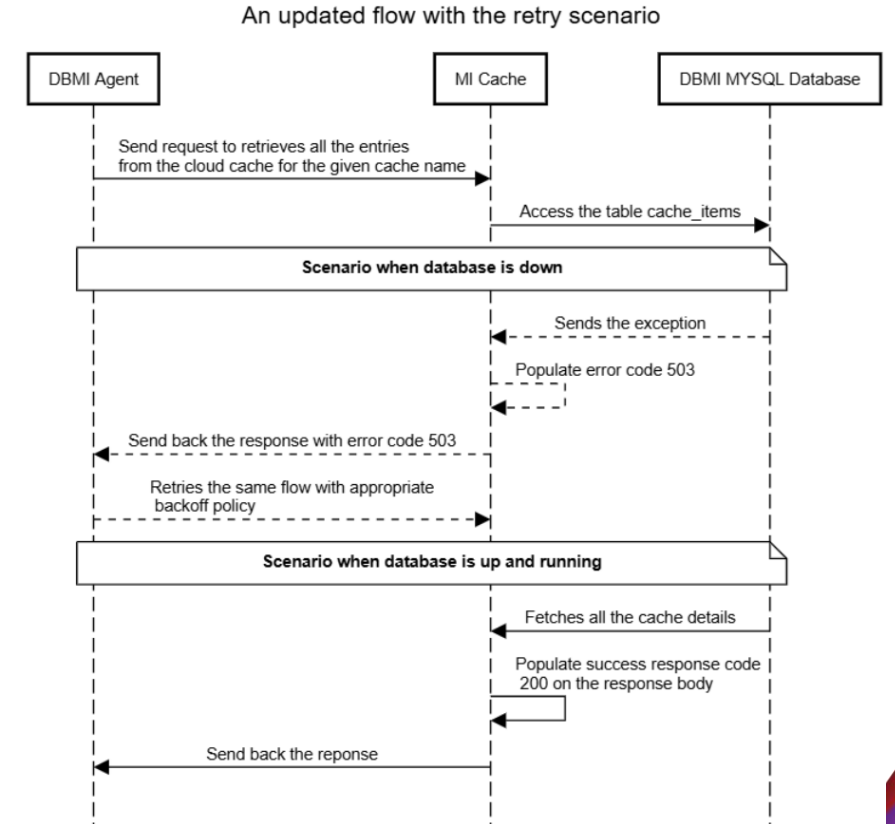


Database-ZDT



Key Highlights

1. Achieves near ZDT during database upgrades by implementing **retries across communication interfaces** between DBMI Agent, DBMI cloud service and other SaaS services
2. Implements **scenario-specific retries** across inter-service communication interfaces to achieve service resiliency and minimize disruption
 1. Retries implemented only for transient error codes
 2. Retries avoided in cases where the client request could have timed-out
3. Rationalizes retries implementation through **exponential back-offs** and **maximum retry attempts** to reduce impact on downstream services and prevent infinite retry loops
4. Extends the **retry mechanism to scheduled jobs** to ensure seamless deployment through database upgrades
5. Implements changes to **write metering information to MySQL** database to benefit from platform's ZDT implementation on MySQL



Custom Properties in UI

Key Highlights

- Exposes **key custom properties** in the UI for easy user configuration
- Provides improved **in-line help** for clear understanding of property meanings and implications.
- These custom properties are **thoroughly documented**, providing users with detailed definitions and usage guidelines.
- Implements **framework-level out-of-the-box defaults** for key custom properties, which have been observed to work best. Users can optionally override defaults
- Custom property definitions are also available for **natural language query via Doc LLM** in DI copilot
- Implements **section specific custom properties** available across source, target and runtime sections to reduce incorrect placements

The screenshot displays the Informatica Data Integration (DI) user interface. The top navigation bar shows 'atica Data Integration' and 'DBMLCDCGRPQA'. The main content area is divided into two sections. The upper section, titled 'orc_sflake_ini_01', shows a progress bar with steps: Destination (checked), Source (checked), Task Details (active, showing 'Target Details'), Transform (checked), and Let's Go! Below this, there are sections for 'Table Renaming Rules', 'Data Type Rules', and 'Custom Properties'. The 'Custom Properties' section is expanded, showing a table with columns for 'Property Name' and 'Value'. A dropdown menu is open under 'Property Name', listing options: 'Writer Distributor Count' (with a help icon), 'Writer Helper Thread Count' (with a help icon and a tooltip: 'The number of distributor threads that can write data in parallel to the target, when multiple distributors are used. System name is writerDistributorsCount'), 'Writer Unload Multiple Distributors' (with a help icon), and 'Custom' (with a help icon). A yellow gift box icon is visible to the right of the dropdown. The lower section, titled 'Replication_Task_20250902110808', shows a progress bar with steps: Destination (active), Source, Task Details, Transform, and Let's Go!. Below this, the 'Where is your data going?' section displays a grid of destination connection options, including PostgreSQL, Oracle Database Ingestion, Amazon S3 v2, Snowflake Data Cloud, and Databricks. A 'Learn More' sidebar is visible on the right, providing instructions on how to configure the destination connection.

Task Pipeline Insights

Key Highlights

- Capture **source, target, network** and **Secure Agent system** and **pipeline metrics**
- **Easy to read Insights** based on performance analysis jobs execution
- Ingest only a **small fraction of rows** to collect insights

Benefits

- **Simplify and expedite** identifying performance bottlenecks for better and faster performance improvements
- **Consolidated Insights** across multiple task processes
- **Quickly test iterative changes** by unloading a small subset of data
- Drive **best possible performance** on a given environment resources



Property Name	Value
enableTaskPipelineInsights(Custom)	true

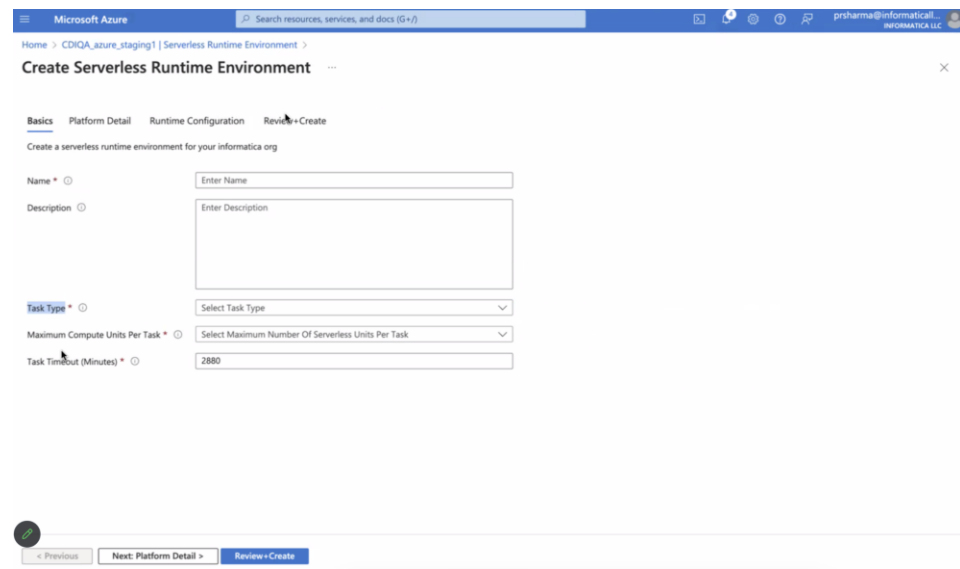
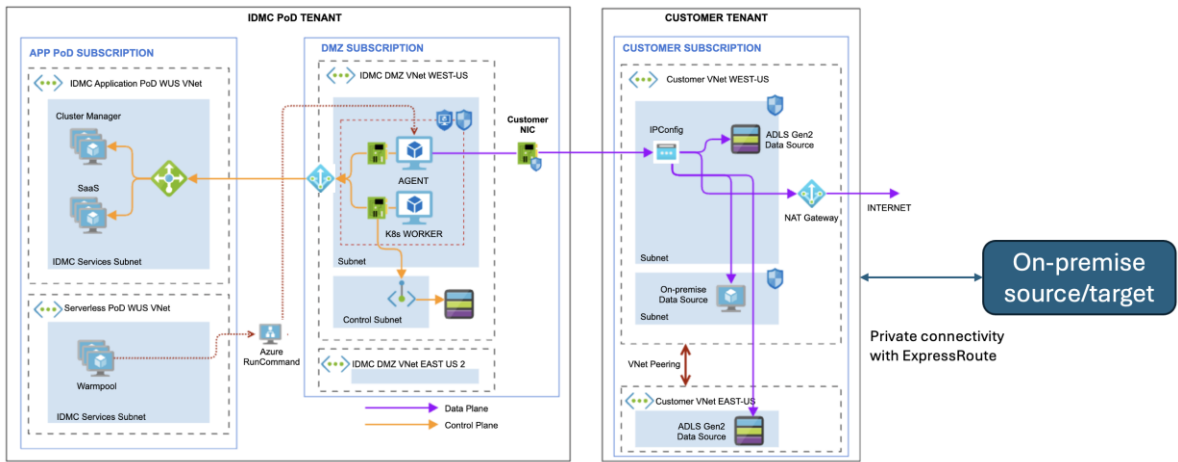
```
[[===== Task Pipeline Insights =====]].
[[***** System Insights *****]].
[[----- System Insights -----]].
[[OS_MEMORY_STATS: Total Physical Memory : [63881 MB], Free Physical Memory : [41976 MB]].
[[OS_PROCESSORS_STATS: Available Processors for this JVM : [32], CPU Load on this OS : [871.4%]].
[[JVM_HEAP_INFO_PRE_TASK: JVM Max Heap Size : [16384 MB], Current Heap Size : [3440 MB], Current Free Heap Size : [2665 MB]].
[[JVM_HEAP_INFO_POST_TASK: JVM Max Heap Size : [16384 MB], Current Heap Size : [12776 MB], Current Free Heap Size : [8496 MB]].
[[***** Source Insights *****]].
[[----- Source Insights -----]].
[[SOURCE_DATABASE_LATENCY: <192391> nanos].
[[SOURCE_NETWORK_LATENCY: <159664> nanos].
[[UNLOAD_EVENT_BATCH_SIZE: <7813> events].
[[JDBC_FETCH_SIZE: <181>].
[[PARTITION-1: RECORDS_READ: 1000000, QUERY_EXECUTION_TIME: <46.0> millis, FETCH_RATE: <111111.11> events/sec, GET_OBJECT_TIME: <1.2> minutes,
SOURCE_RECORDS_PROCESSING_TIME: <2.0> minutes].
[[PARTITIONS_TOTAL_PROCESSING_TIME: <2.0> minutes].
[[SOURCE_TOTAL_PROCESSING_TIME: <2.0> minutes].
[[***** Target Insights *****]].
[[----- Target Insights -----]].
[[TARGET_DATABASE_LATENCY: <152.8> millis].
[[TARGET_NETWORK_LATENCY: <9.1> millis].
[[OBJECT_DISTRIBUTOR-1: RECORDS_WRITTEN: 250016, FILE_CREATION_TIME: <5.1> seconds, BYTES_WRITTEN: <2998918481> bytes, AVERAGE_ROW_SIZE: <11994> bytes,
FLUSH_PREPARATION_TIME: <9.0> minutes, FILE_UPLOAD_TIME: <18.4> seconds, CYCLE_DURATION: <9.3> minutes].
[[OBJECT_DISTRIBUTOR-2: RECORDS_WRITTEN: 249952, FILE_CREATION_TIME: <4.6> seconds, BYTES_WRITTEN: <2998150802> bytes, AVERAGE_ROW_SIZE: <11994> bytes,
FLUSH_PREPARATION_TIME: <8.9> minutes, FILE_UPLOAD_TIME: <18.3> seconds, CYCLE_DURATION: <9.3> minutes].
[[OBJECT_DISTRIBUTOR-3: RECORDS_WRITTEN: 250016, FILE_CREATION_TIME: <4.7> seconds, BYTES_WRITTEN: <2998903549> bytes, AVERAGE_ROW_SIZE: <11994> bytes,
FLUSH_PREPARATION_TIME: <9.3> minutes, FILE_UPLOAD_TIME: <18.1> seconds, CYCLE_DURATION: <9.6> minutes].
[[OBJECT_DISTRIBUTOR-4: RECORDS_WRITTEN: 250016, FILE_CREATION_TIME: <4.8> seconds, BYTES_WRITTEN: <2998916064> bytes, AVERAGE_ROW_SIZE: <11994> bytes,
FLUSH_PREPARATION_TIME: <9.0> minutes, FILE_UPLOAD_TIME: <18.6> seconds, CYCLE_DURATION: <9.3> minutes].
[[COPY_STATEMENT_EXECUTION_TIME: <7.7> minutes].
[[TARGET_TOTAL_PROCESSING_TIME: <17.3> minutes].
[[***** Pipeline Insights *****]].
[[----- Pipeline Insights -----]].
[[METADATA_FETCH_TIME: <1.1> seconds].
[[DBMI_TASK_INITIALIZATION_TIME: <12.8> seconds].
[[JDBC_READER_EVENT_PROCESSING_RATE: <8384.90> events/sec].
[[UNLOAD_HELPER_EVENT_PROCESSING_RATE: <1802.58> events/sec].
[[WRITER_SCHEMA_CACHE_HELPER_EVENT_PROCESSING_RATE: <1851.87> events/sec].
[[WRITER_PAYLOAD_HELPER_EVENT_PROCESSING_RATE: <1848.15> events/sec].
[[OBJECT_DISTRIBUTOR_EVENT_PROCESSING_RATE: <1828.57> events/sec].
[[UNLOAD_DML_EVENT_BUILD_TIME: <36.2> minutes for <1000000> raw events].
[[DML_TO_TARGET_PAYLOAD_CONVERSION_TIME: <5.5> minutes for <1000000> DML events].
[[RING_BUFFER_CONTENTION_DETECTED: false].
[[DBMI_TASK_TOTAL_TIME: <18.0> minutes].
[[*****]]].
```

Serverless Offerings

Serverless Runtime

Key Highlights

- Creates an **elastically scalable runtime** for processing massive workloads concurrently
- **Highly secure interaction** between SRTE configured in INFA's tenant and the data resources in the customer tenant using subnet delegation, which creates a secondary NIC for INFA DMZ within the customer's subnet
- Supports connection to **on-premise sources/targets** via setting up private connectivity between on-premise resource and customer's VNET
- Uses resource groups to **logically isolate** Azure VMs, application and network security groups for a customer within INFA's tenant
- Supports **configurable runtime environment properties** to govern the CPU, RAM requests for individual task tables
- Supported for **DBMI and APPMI**



Post-replication transformation

Key Highlights

- Seamlessly trigger **Taskflow** for post replication **transformations** from DBMI or App MI task
- Support for **complex transformations** post full load or incremental load
- Efficient **event-based architecture** with support for only-once processing

Benefits

- End-to-end EL&T (ingest and transform) flow - **Efficiently detect updates and apply transformations**
- Simplification – apply **complex post-replication transformations** only on updated records
- **Near real-time processing** by automatically triggering taskflow for incremental data



Path	Condition	Value
Path 1	Contains	PRODUCT_REVIEWS
Path 2	Otherwise	



Demo





Q&A





Thank You

