Den store digitale teknikkdagen
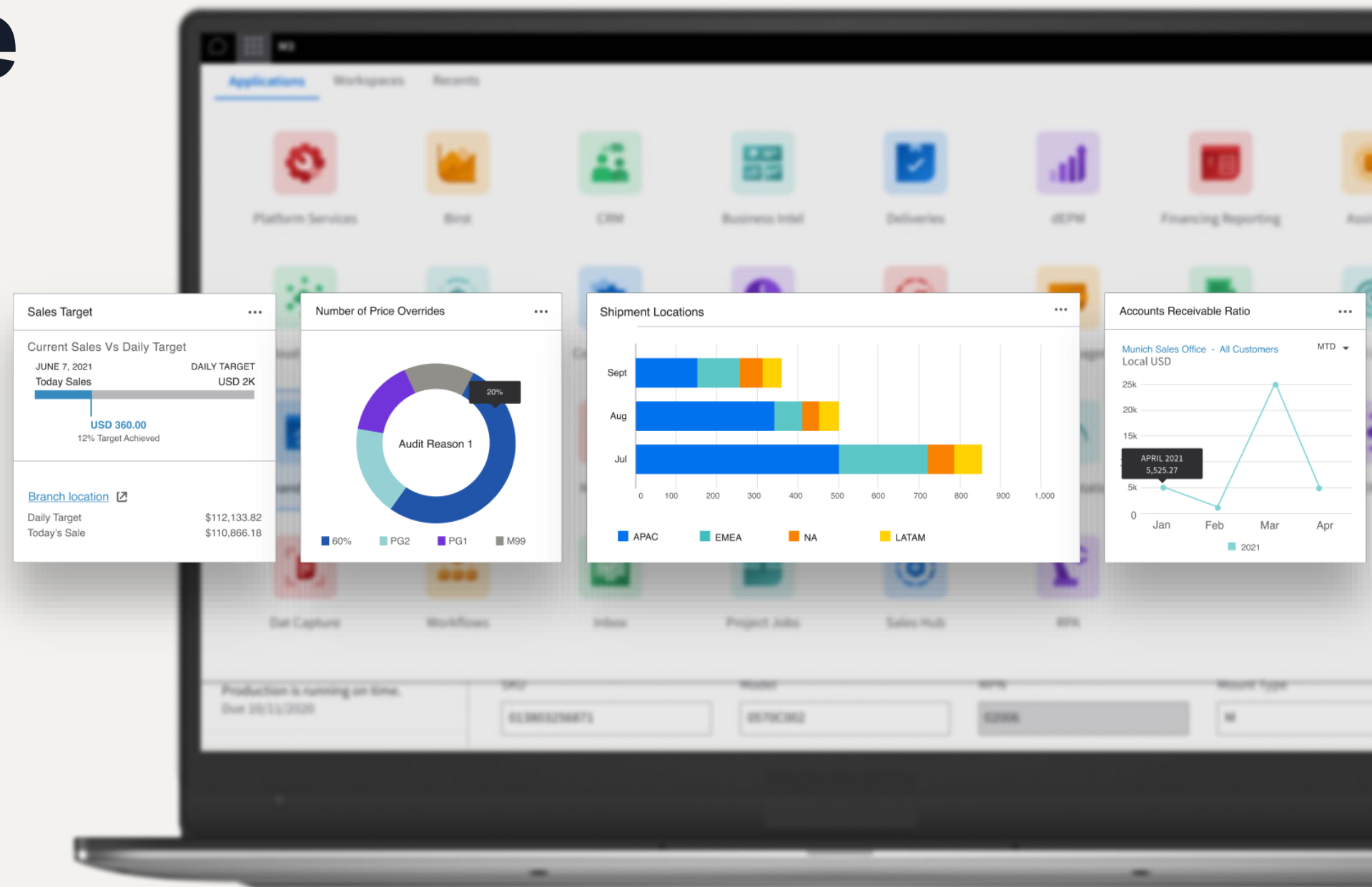
Muligheter med Streaming og M3CE med Data Fabric

Per Melander
Principal Architect, Infor Development

April 2023
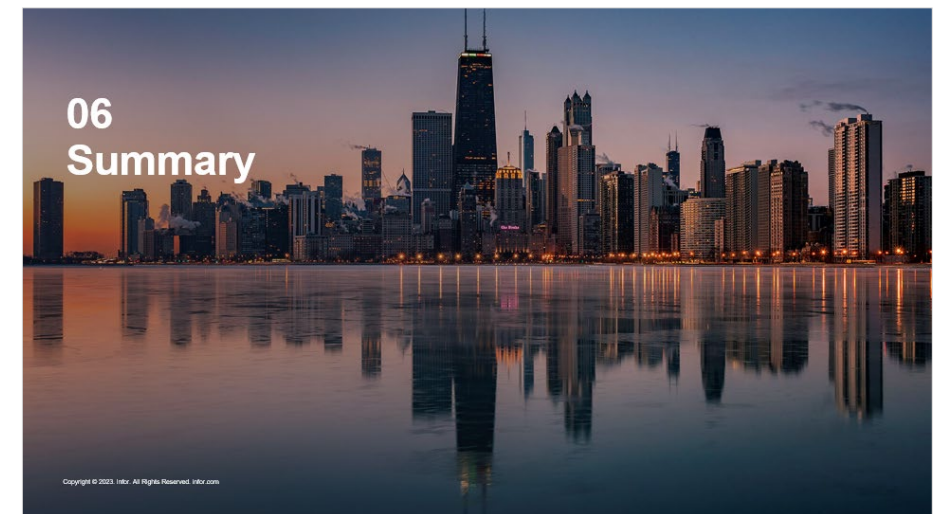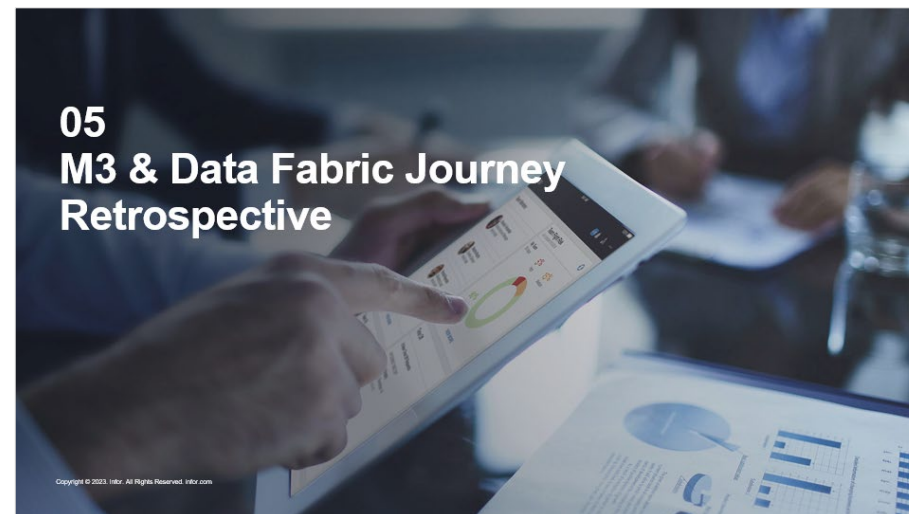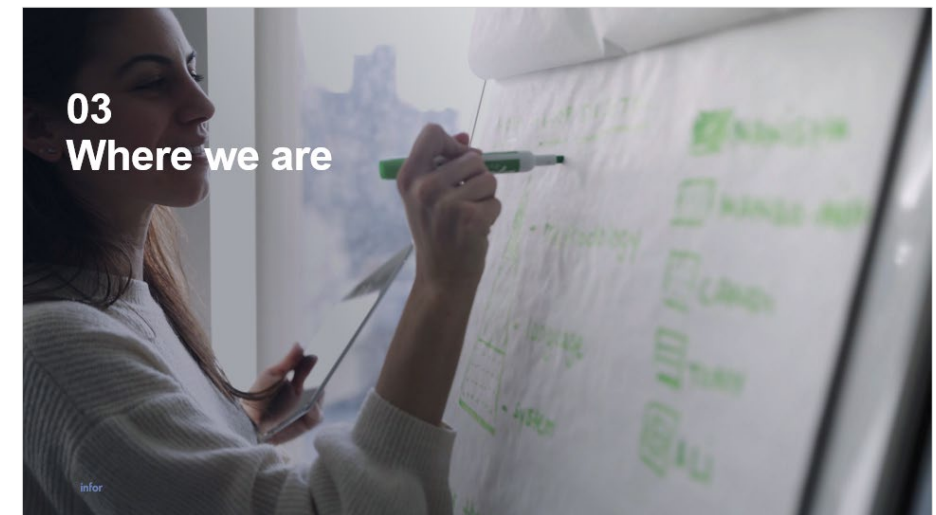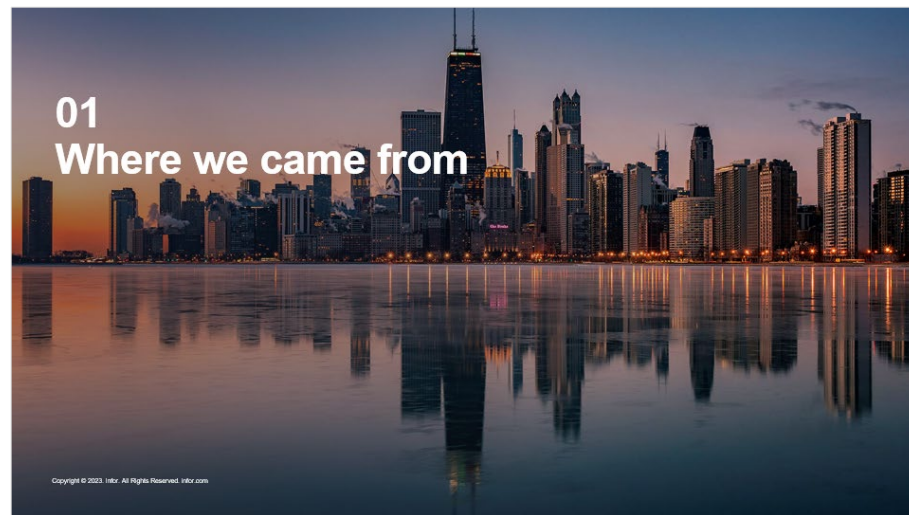
Infor

ERP Simplified:
Smart. Preconfigured. Modern.

Disclaimer

This presentation reflects the direction Infor may take with regard to the products or services described herein, all of which is subject to change without notice. This presentation is not a commitment to you in any way and you should not rely on any content herein in making any decision.

Infor is not committing to develop or deliver any specified enhancement, upgrade, product, service or functionality, even if such is described herein. Many factors can affect Infor's product development plans and the nature, content and timing of future product releases, all of which remain in the sole discretion of Infor. This presentation, in whole or in part, may not be incorporated into any agreement. Infor expressly disclaims any liability with respect to this presentation.

infor

Agenda

01
Where we came from
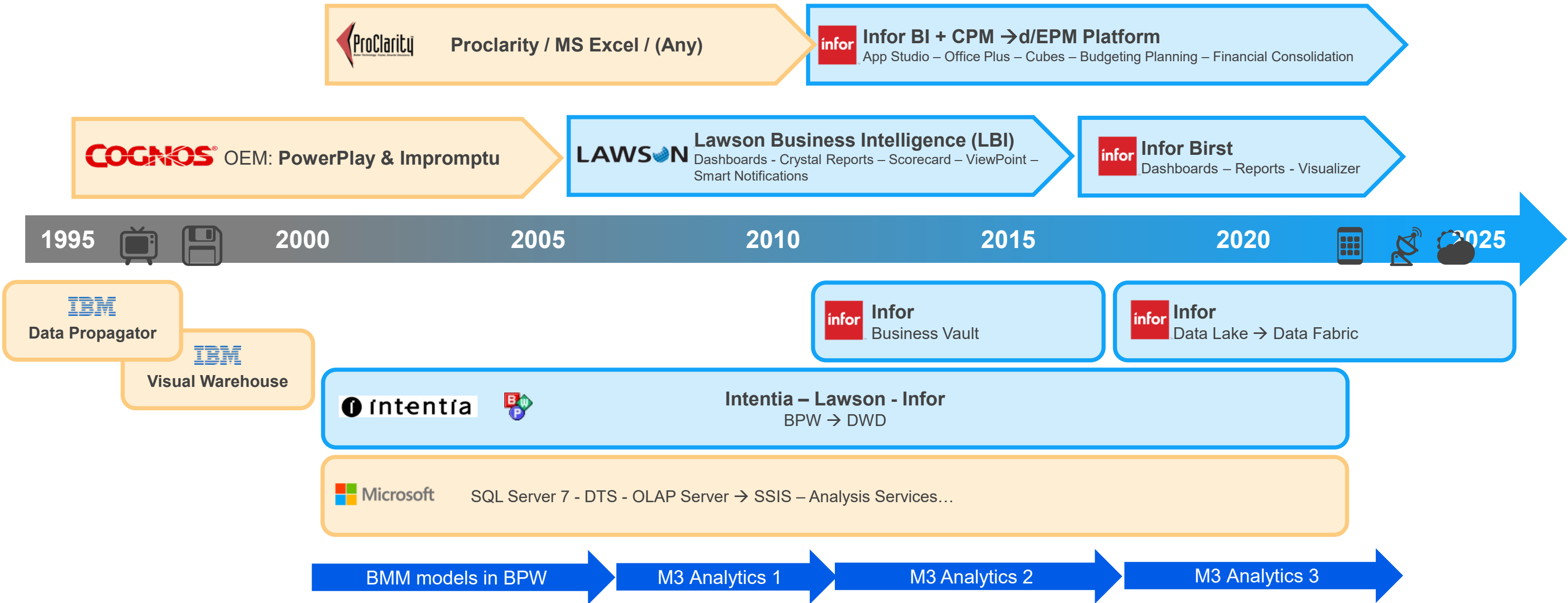
02
The last 5 years

03
Where we are

04
Future Direction

05
M3 & Data Fabric Journey Retrospective

06
Summary

01
Where we came from
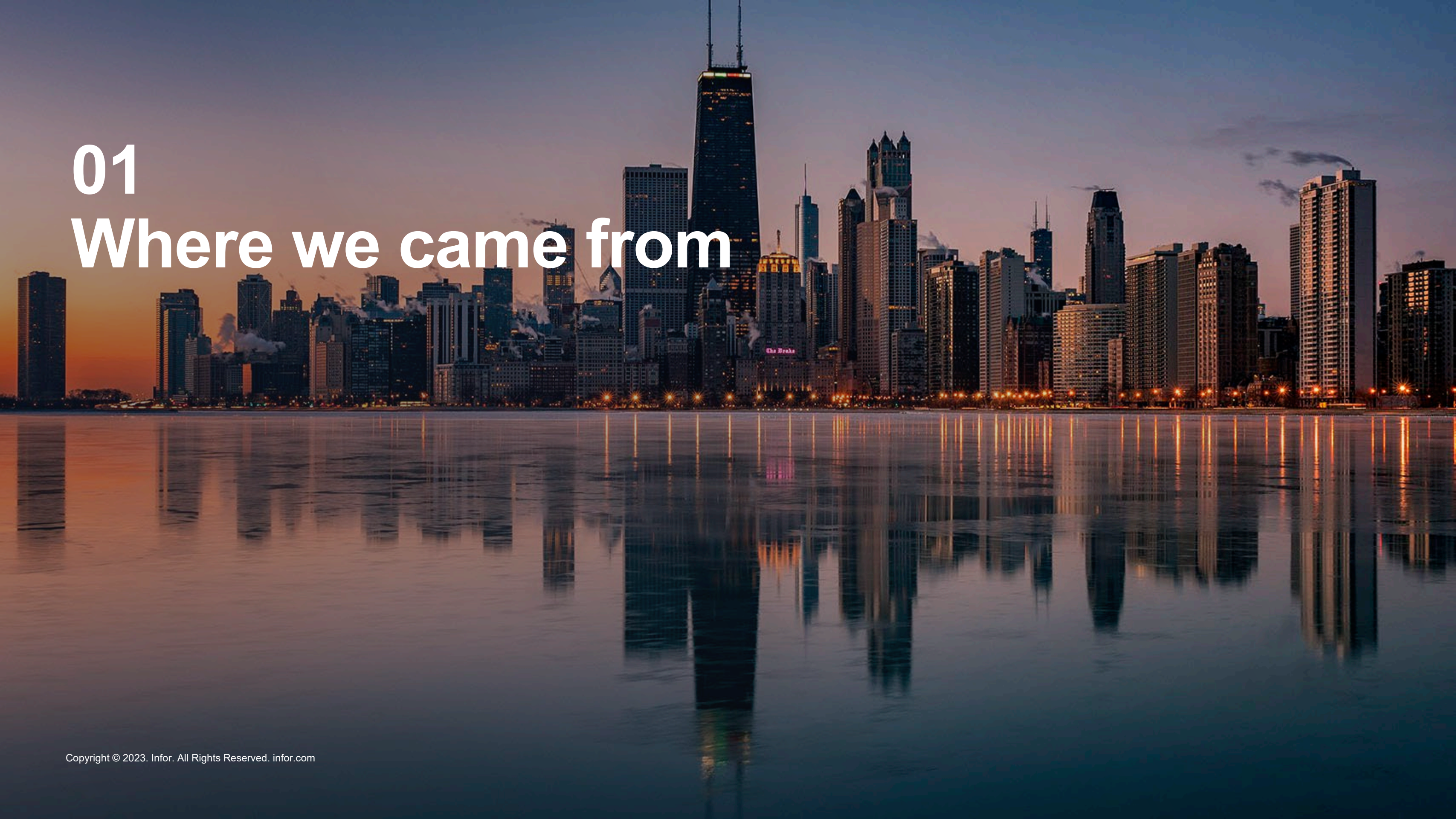
Analytics – BI and Data Warehousing platforms M3 Evolution

Proclarity / MS Excel / (Any)

Infor BI + CPM →d/EPM Platform
App Studio – Office Plus – Cubes – Budgeting Planning – Financial Consolidation

COGNOS OEM: **PowerPlay & Impromptu**

LAWSON **Lawson Business Intelligence (LBI)**
Dashboards - Crystal Reports – Scorecard – ViewPoint – Smart Notifications

Infor Birst
Dashboards – Reports - Visualizer

| 1995 | 2000 | 2005 | 2010 | 2015 | 2020 | 2025 |

IBM
Data Propagator

IBM
Visual Warehouse

Infor
Business Vault

Infor
Data Lake → Data Fabric

intentia **Intentia – Lawson - Infor**
BPW → DWD

Microsoft SQL Server 7 - DTS - OLAP Server → SSIS – Analysis Services…

BMM models in BPW ➤ M3 Analytics 1 ➤ M3 Analytics 2 ➤ M3 Analytics 3

02
The last 5 years

How does it work?

Infor
CloudSuite
Applications

Push

DATALAKE

MITMAS
FGLEDG
CMNDIV
OOHEAD
OOLINE
CSYTAB
OCUSMA
CFACIL

Sync.SecurityUserMaster

Sync.SourceSystemGLMovement

Access

M3 Analytics

d/EPM

Coleman AI

GRC

Data Warehouse
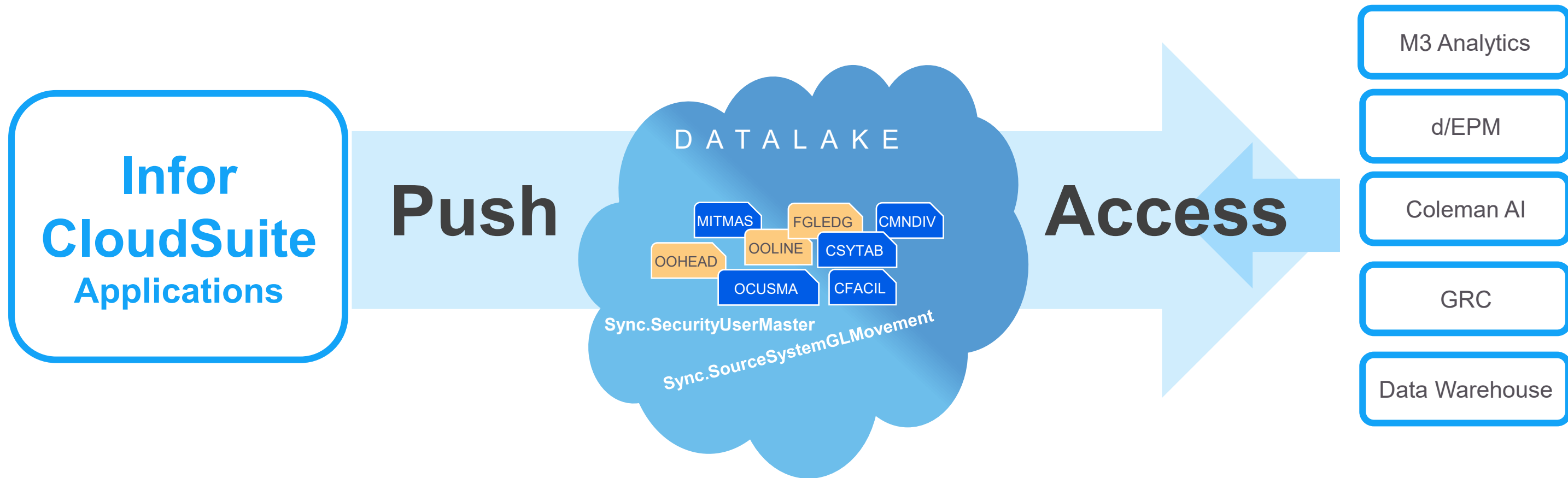
M3 and Data Lake (current implementation 2017-2022)

M3 Analytics

Infor M3 Apps and **M3 Technology**

Infor OS Data Fabric & ION

Infor Birst Platform

M3 Database

Event Hub

Data Lake Publisher

Publishing events in near real time

ION

Data Lake

Data Lake Flow

JDBC Driver

Incremental Extraction

M3 Analytics powered by Infor Birst

End User Semantic Layer

Extensions

Dimensions

Facts

Stage

Processing data
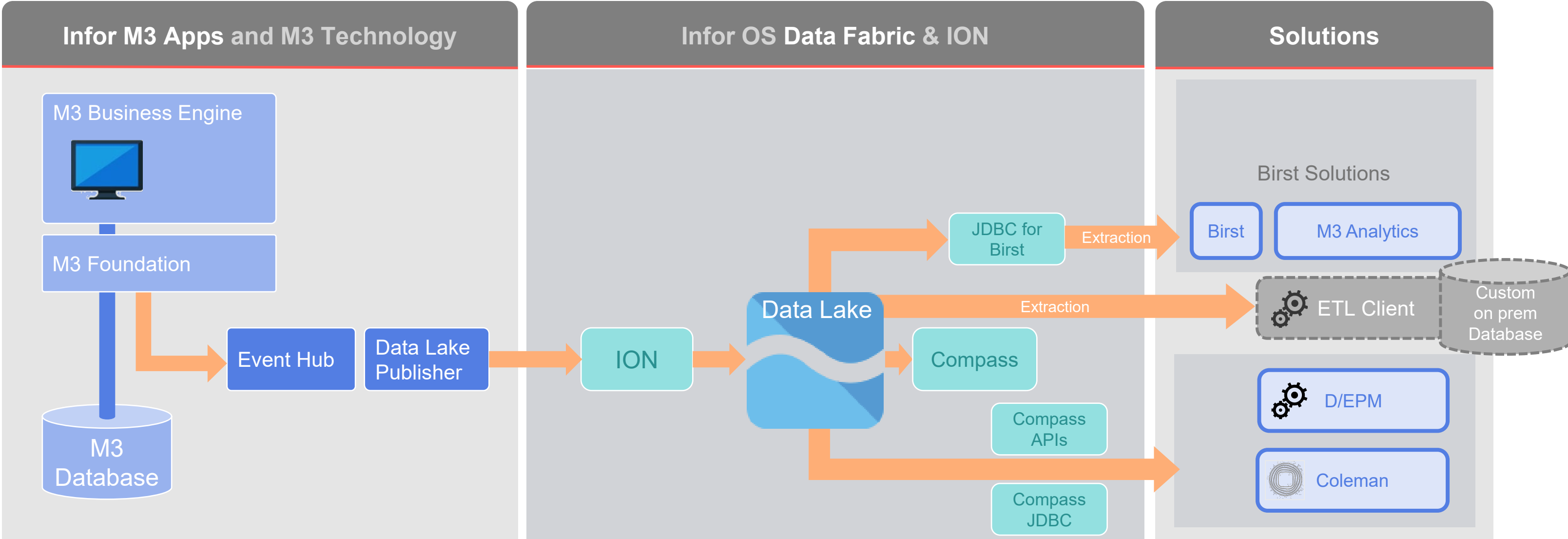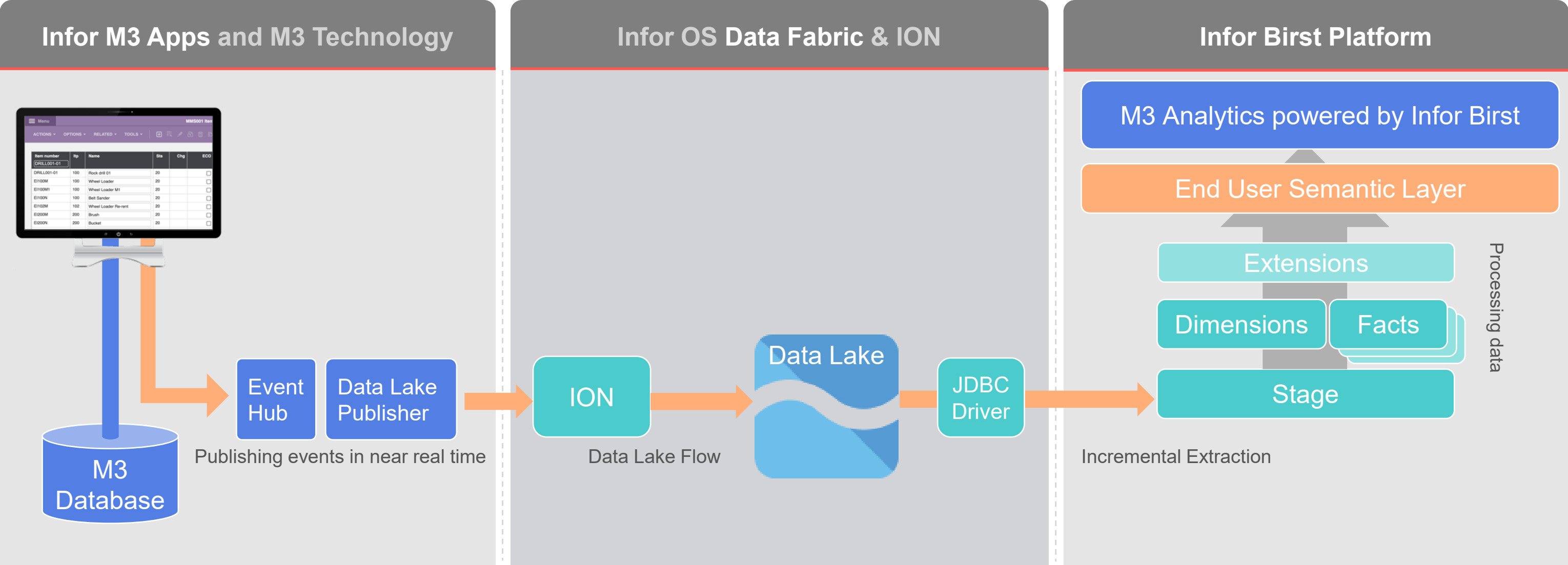
M3 Solutions using Infor Data Lake

## M3 Analytics using the Infor Birst platform

- M3 is feeding Infor Data Lake in near real time – initial load option also available.

- M3 Analytics was the first solution to integrate to Infor Data Lake via Birst.

- First early adopters enabled in August 2019 – General Available in March 2020

M3 CE integration with d/EPM CE

- This two-way integration is enabling Infor d/EPM to send an approved budget back to M3 for the control of actual costs and revenues. M3 data is published to Infor Data Lake and Infor d/EPM is loading into staging- and integration tables to feed multidimensional structures.

- This new M3 CE integration was enables in June 2020

- Financial Consolidation integration was enabled in September 2021

M3 CE integration with Coleman AI Platform

- Delivered example: Sales Price Recommender

  – *How much discount should I give to this customer for this product if I want to win the order?*

- Training material on Campus

- Delivered via KB Article: Data Lake Queries, Quests, Widget and M3 IC App instructions

- Available as content in June 2020
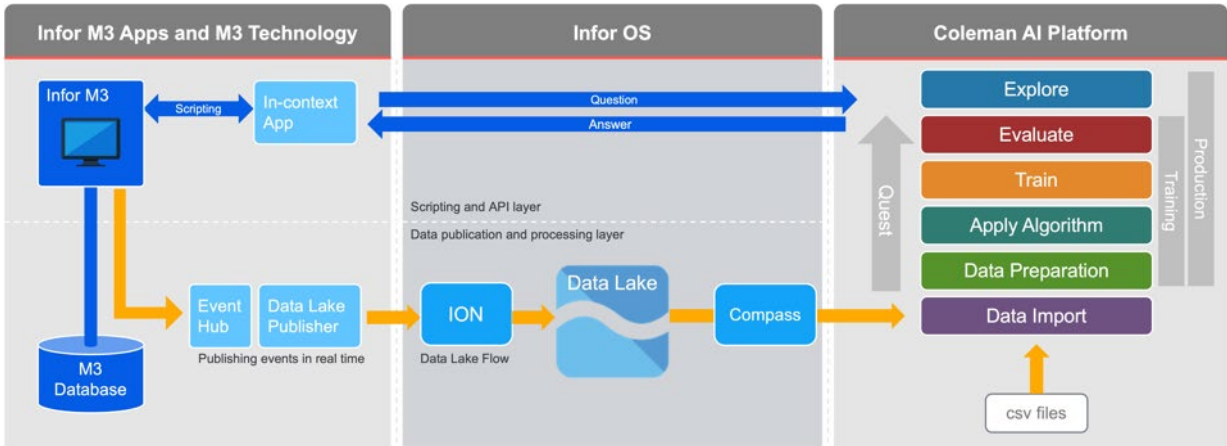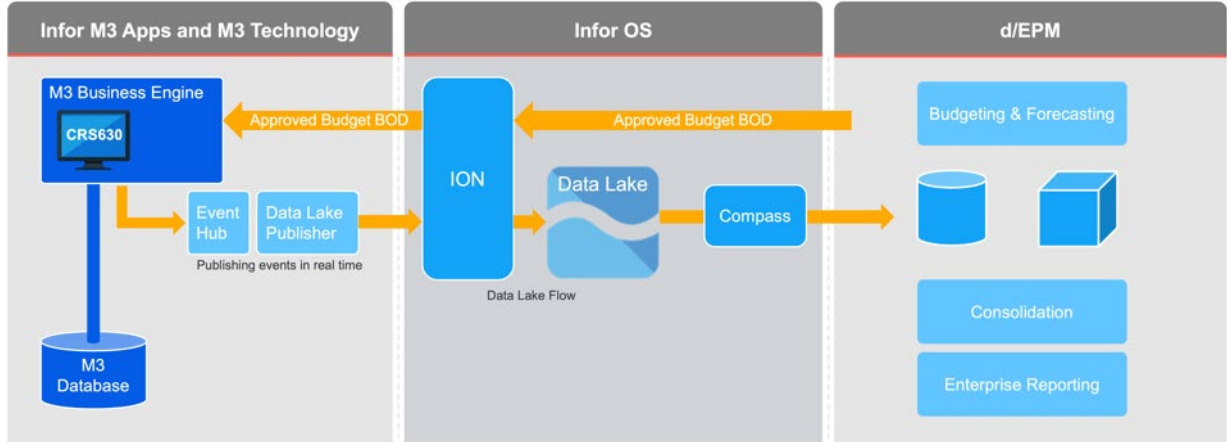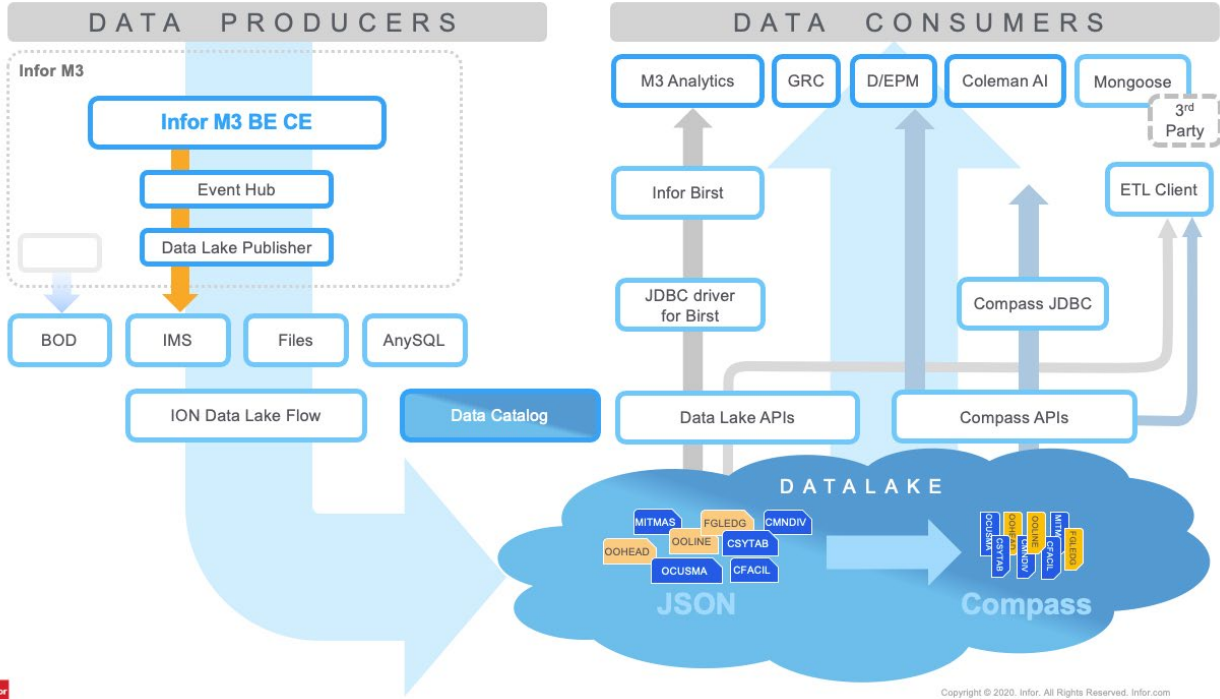
Additional Use cases on current architecture

- XtendM3 Dynamic tables publishing to Data Fabric

- M3DMP publishing data to Data Fabric using batch ingestion

- M3 Archiving to Data Lake

- Field Audit Trail publishing to Data Lake

- PLM publishing to Data Lake

- CLM publishing to Data Lake as source for Analytics.

Master KB Article

Main resource for Knowledge Transfer

- How to Setup and Configure M3 with Data Fabric
- Administration
- Data Validation & Troubleshooting
- How access and make use of Data Lake

Master KB Article for DL Integration - > **2186646**

M3 Integration with Infor Data
Master KB Article

This article provides links to all KB articles describing configuration, guides and other useful information ..

In order to receive a notification on updates for this KB articles, please click and follow this Sign-Up link, then click 'Sign-Up'

Visit https://docs.infor.com for the following documents

- Infor M3 CE Core Administration Guide
- Infor M3 Cloud Configuration Guide
- Infor ION Development Guide - Cloud Edition
- Infor Data Fabric User Guide
- Infor ION Technology Connectors Administration Guide - Cloud Edition

Setup & Configuration	Source	Updated	Link
Configure M3 with Infor Data Lake (using ION)	M3 Dev	2020-07-13	2106513
Configure M3 Streaming to Data Fabric	M3 Dev	2023-04-05	2294200

Administration	Source	Updated	Link
Initial Load or Partial load of Infor Data Lake	M3 Dev	2021-11-04	2188356
Refresh invalid M3 data in Infor Data Lake	M3 Dev	2021-08-27	2210110
Align Data Lake with M3 Database changes	M3 Dev	2021-02-05	2180699
Manual updates of M3 metadata in Infor Data Catalog	M3 Dev	2021-06-23	2203008

Validation & Troubleshooting	Source	Updated	Link
Resolve M3 duplicate data in Infor Data Lake	M3 Dev	2021-03-09	2186562
Synchronizing Compass data with Data Lake data	Infor OS	2020-12-28	2149400

Using Infor Data Lake	Source	Updated	Link
Extract data from Infor Data Lake	M3 Dev	2021-01-01	2180691
Data Lake aspects of updates in M3 table OPRICL	M3 Dev	2020-12-17	2172397
Access M3 data in Infor Data Lake	M3 Dev	2020-03-12	2106860

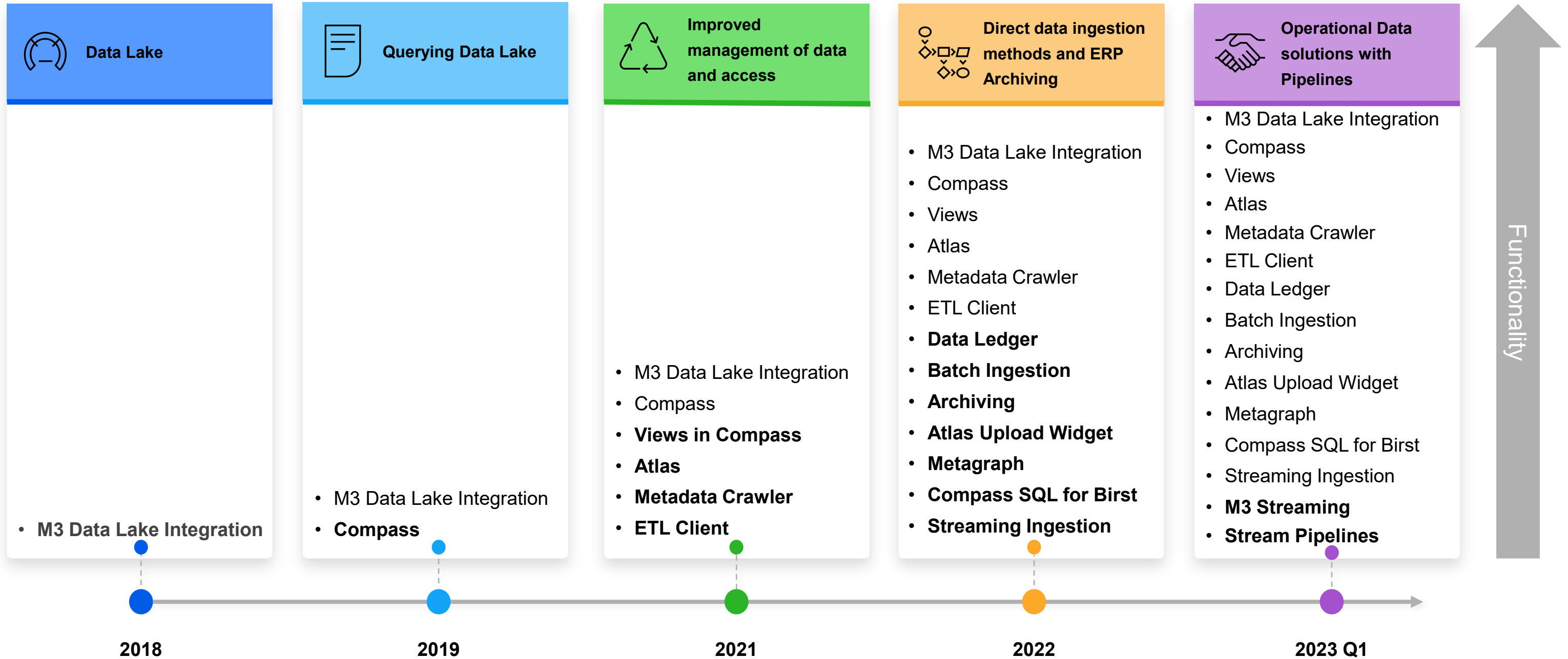| Related		Source	Link
M3 Analytics Master KB Article		M3 Analytics	2022298
Infor M3 CE CloudSuites - Release Information Overview		M3	1956330
M3 CE Integration and Infor OS content - KB Articles		M3 Integration	2017962

Archived	Source	Updated	Link
M3 Date field formatting in DL Compass 2020-07-09	M3 Dev	2020-07-09	2142796
Revert changes in M3 Date formats in Data Lake Compass after M3 April 2020 release	M3 Dev	2020-05-26	2133791

03
Where we are

infor

Evolution of Data Fabric with M3

Data Lake	**Querying Data Lake**	**Improved management of data and access**	**Direct data ingestion methods and ERP Archiving**	**Operational Data solutions with Pipelines**
			• M3 Data Lake Integration	• M3 Data Lake Integration
			• Compass	• Compass
			• Views	• Views
			• Atlas	• Atlas
		• M3 Data Lake Integration	• Metadata Crawler	• Metadata Crawler
		• Compass	• ETL Client	• ETL Client
		• **Views in Compass**	• **Data Ledger**	• Data Ledger
		• **Atlas**	• **Batch Ingestion**	• Batch Ingestion
		• **Metadata Crawler**	• **Archiving**	• Archiving
		• **ETL Client**	• **Atlas Upload Widget**	• Atlas Upload Widget
	• M3 Data Lake Integration		• **Metagraph**	• Metagraph
• **M3 Data Lake Integration**	• **Compass**		• **Compass SQL for Birst**	• Compass SQL for Birst
			• **Streaming Ingestion**	• Streaming Ingestion
				• **M3 Streaming**
				• **Stream Pipelines**

2018   **2019**   **2021**   **2022**   **2023 Q1**

Functionality

CloudSuite Data Platform

M3 & Data Fabric

Delivering Big Data into the enterprise

M3 and Data Fabric come together in a cohesive solution architecture that captures the entirety of your data's life cycle from cradle to grave - and all of it accessible through Compass SQL, an ANSI SQL interface that exposures intelligent queries to satisfy analytics, operational, data science, and exploratory use cases.

Predicated on a streaming architecture, real-time data processing is layered into Data Fabric to help unlock immediate insights and workload processing in the Cloud.
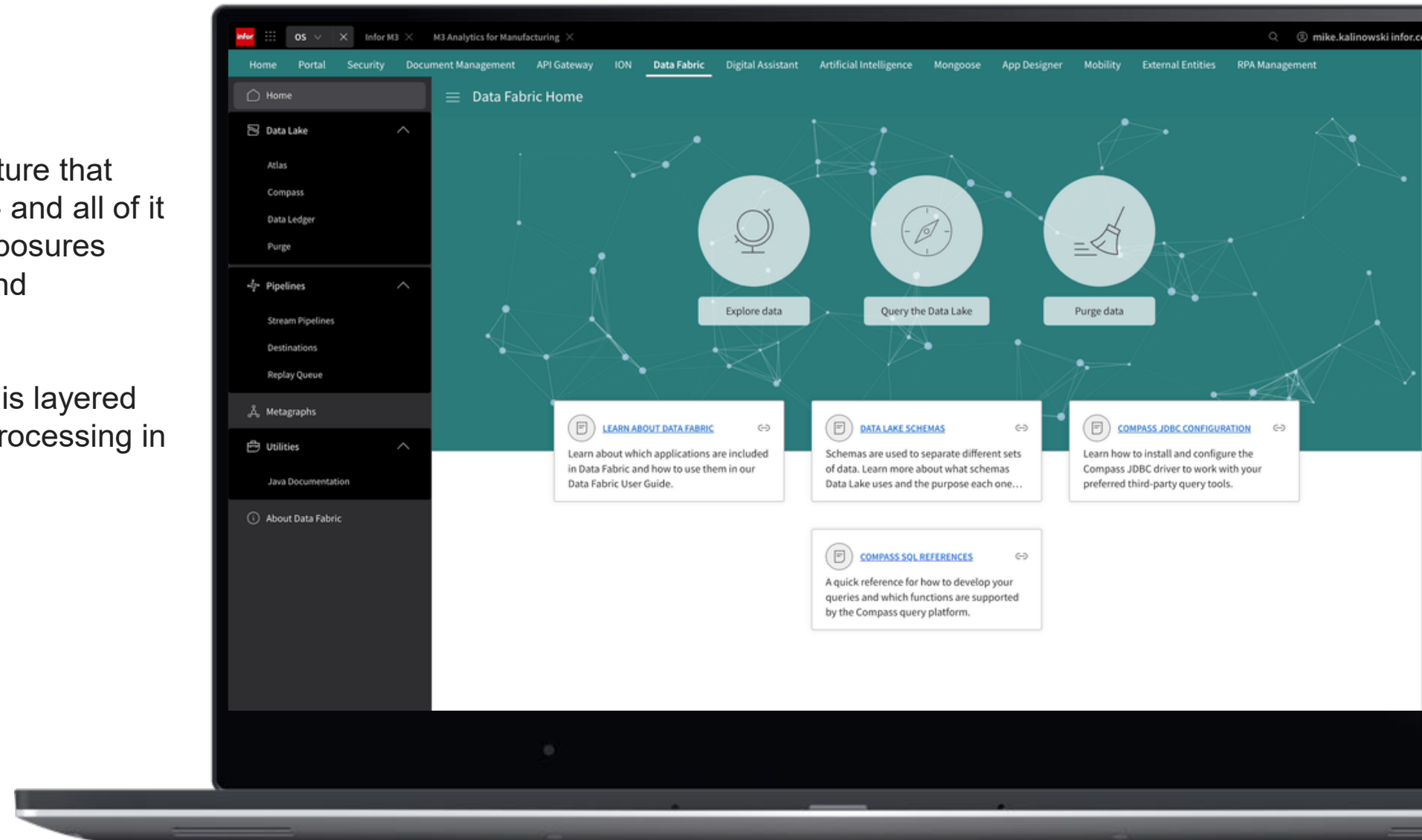
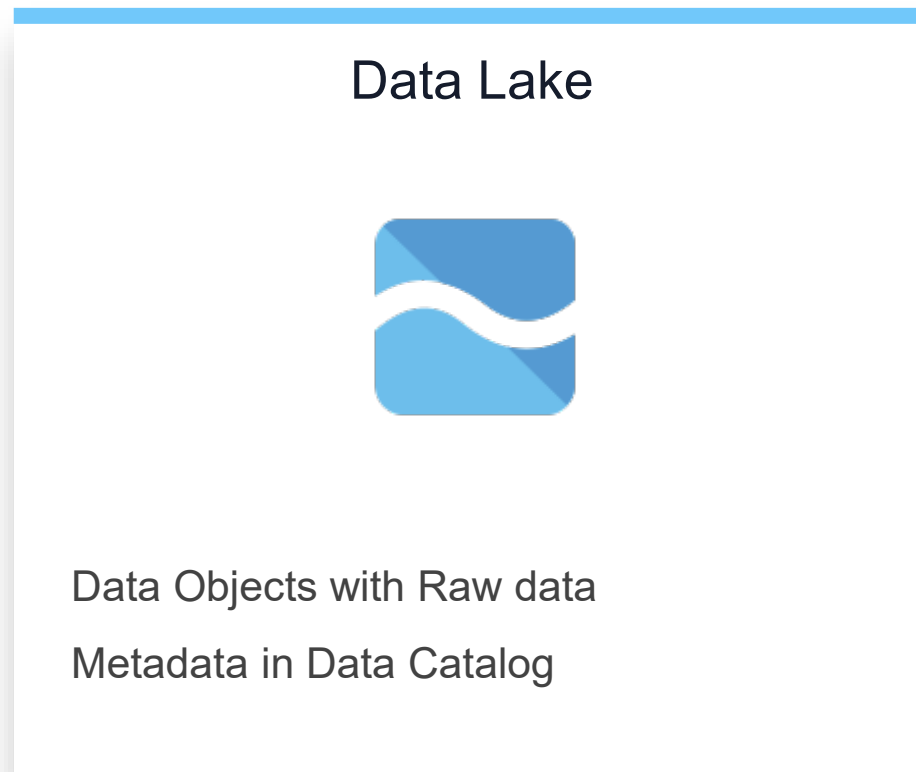# CloudSuite Data Platform

M3 & Data Fabric

M3 will ultimately build an open content framework for efficient discovery and utilization in Lakehouse, Infor's Data warehouse-as-a-service solution, to support both operational data use cases as well as rich analytics solutions in an open and flexible platform providing predefined content building blocks for Analytics, EPM, Coleman AI and 3rd party BI platforms.
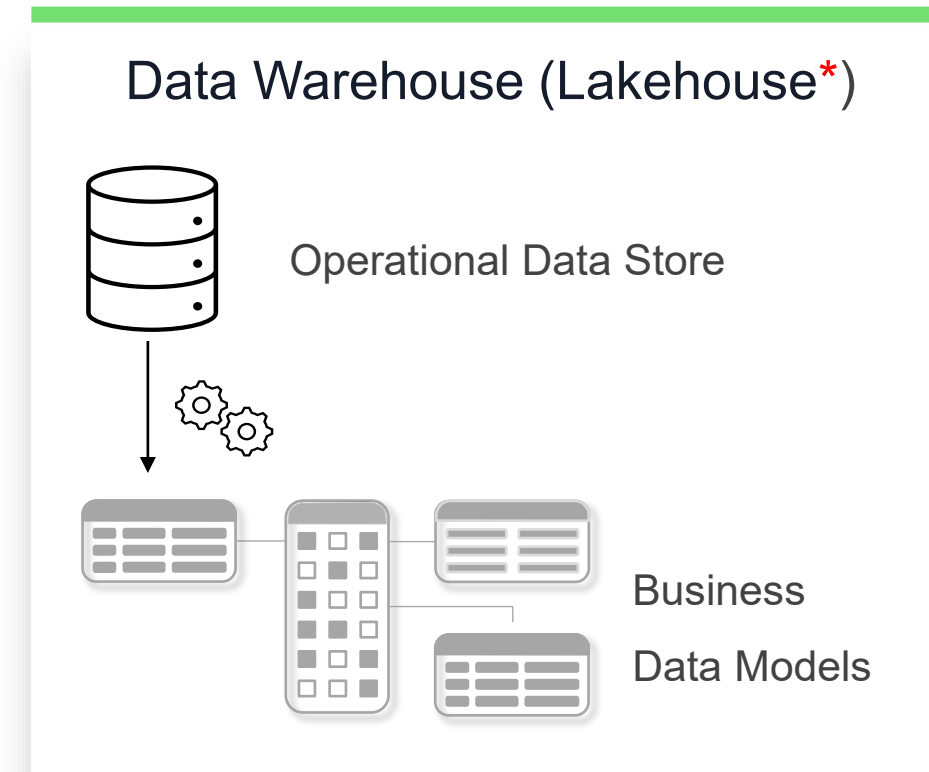
Key features

- Simple point & click data publishing and data integration management
- Compass SQL – ANSI SQL query interfaces (RESTful APIs & JDBC drivers)
- Consolidated metadata management in Data Catalog
- Query across time to explore how data has evolved with data versioning
- Real-time and batch-based ingestion platform
- Stream Pipelines for real-time delivery and operational needs
- Seamless data-driven integration ecosystem (Birst, Coleman AI/ML, EPM, Mongoose & more)
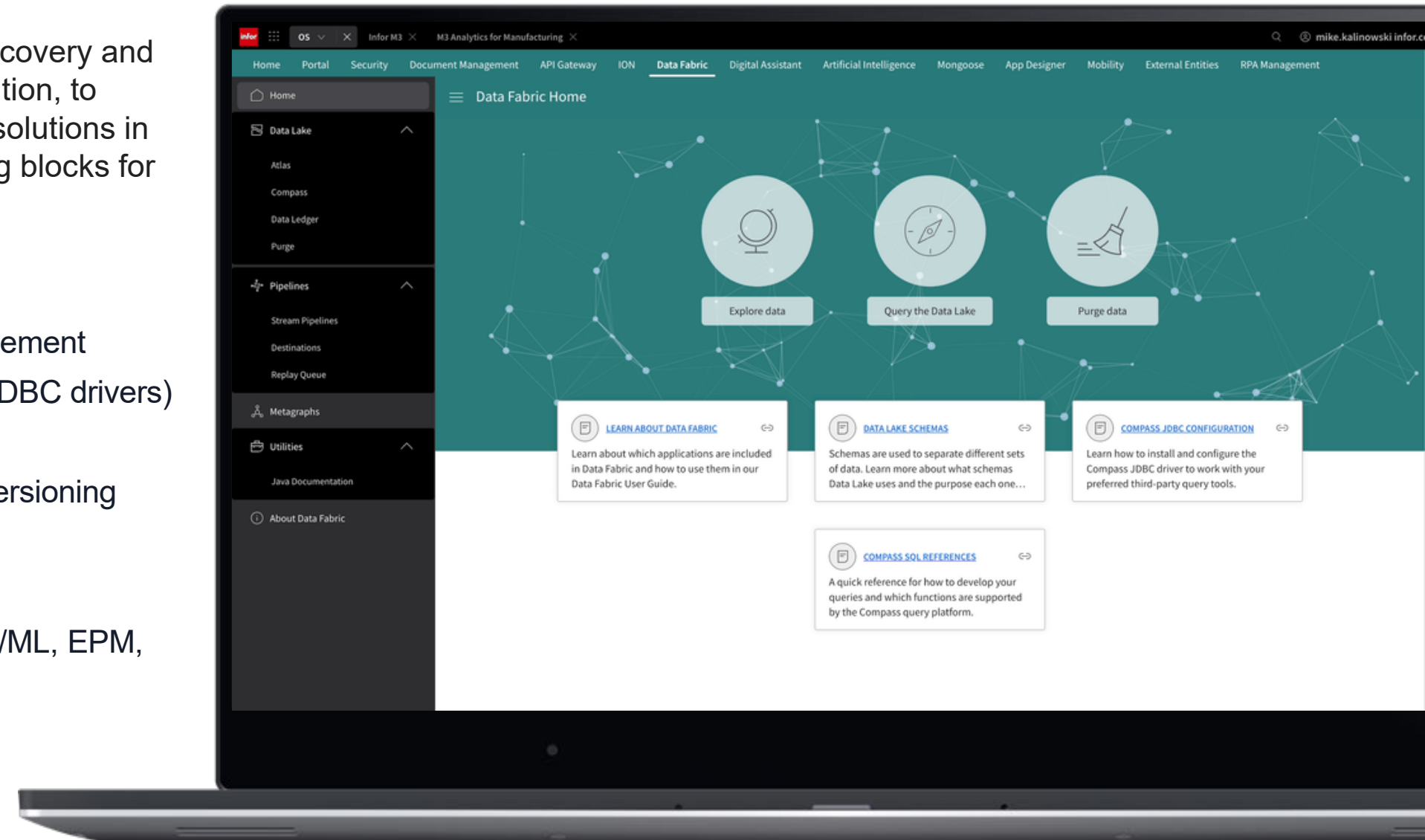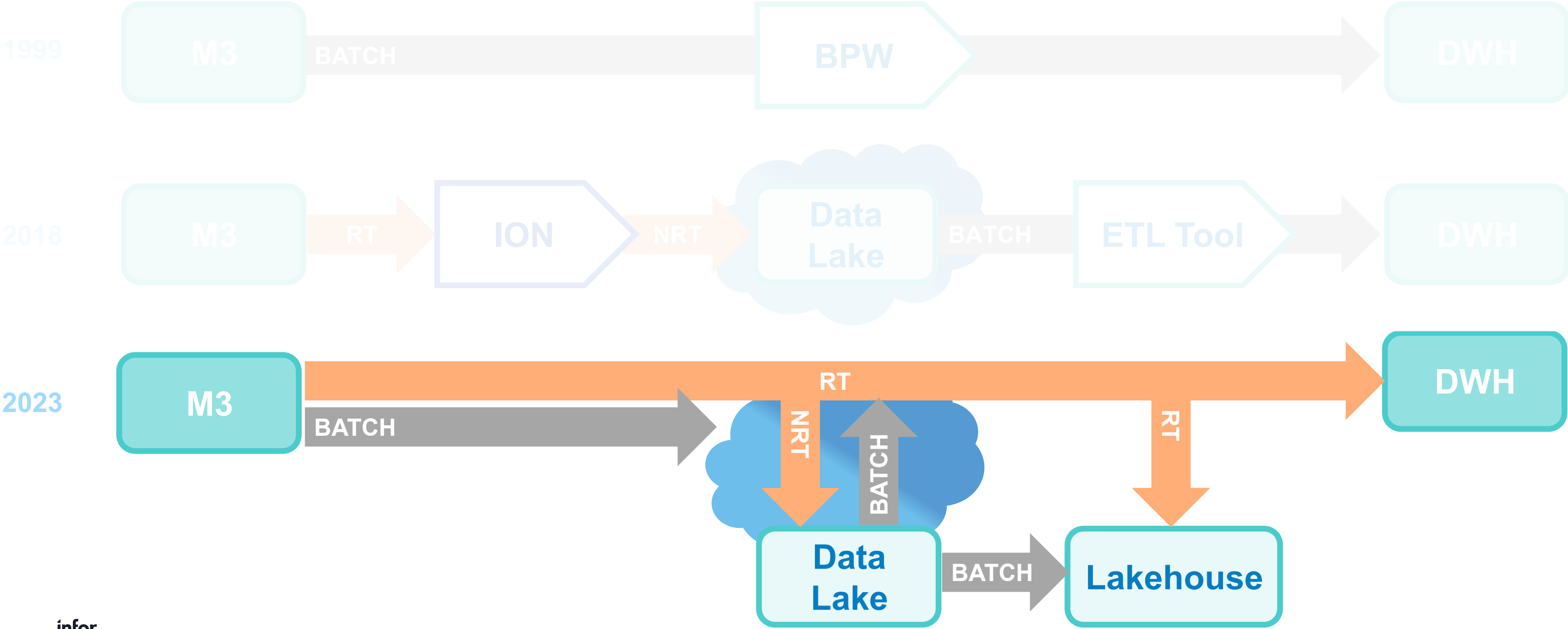
Data Lake versus Data Warehouse

Data Lake

Data Objects with Raw data

Metadata in Data Catalog

Versus

Data Warehouse (Lakehouse*)

Operational Data Store

Business

Data Models

Use Cases and Considerations

- Lightly governed, raw content
- Schema on-read – *data transformation on-demand & as-needed*
- Data Scientists - Data & content developers - Business analysts - Ad-hoc discovery
- Unstructured, semi-structured, & structured datasets
- Machine Learning - Analytics – Governance

- Fully governed, curated data models
- Schema on-write – *conforming data to well-defined target structures*
- Business intelligence & operational reporting consumers
- Structured, relational systems
- Business Intelligence - Operational Reporting

04
Future Direction

Architecture Development

1999 — M3 — BATCH — BPW — DWH

2018 — M3 — RT — ION — NRT — Data Lake — BATCH — ETL Tool — DWH

2023 — M3 — BATCH / RT — Data Lake — NRT / BATCH — BATCH — Lakehouse — RT — DWH

infor

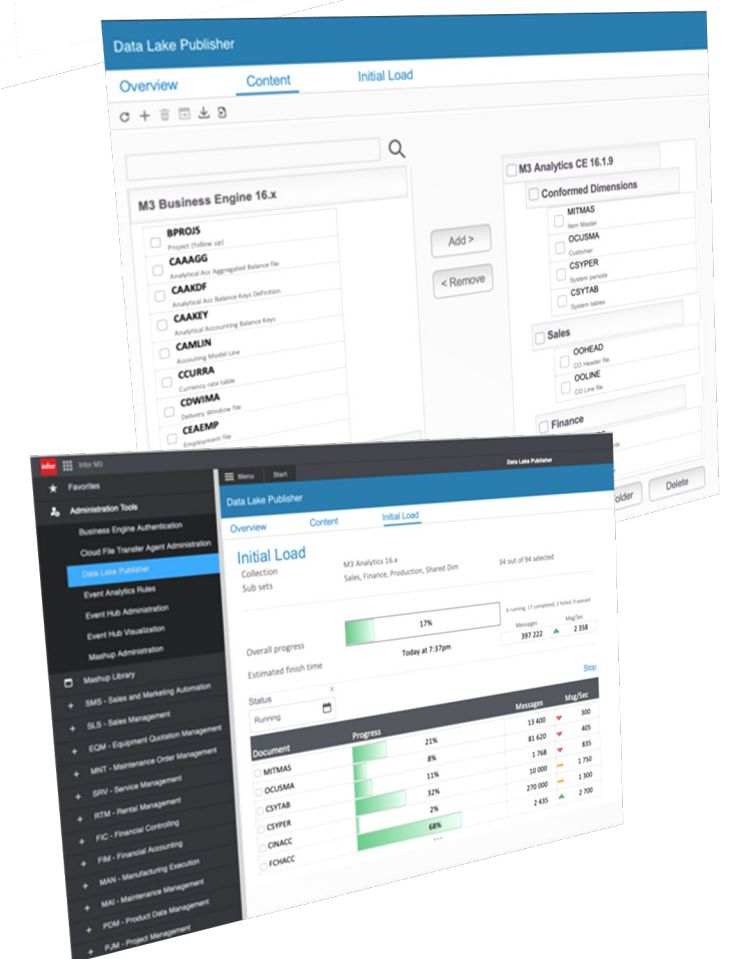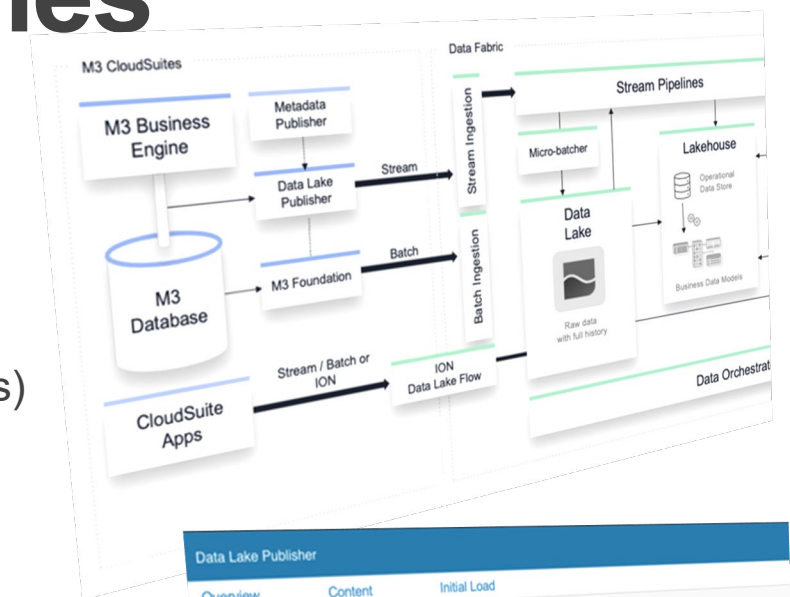# M3 integration to Data Fabric – Roadmap themes

Improve Operability

- Investigate different options for avoiding M3 BE to send all table updates to Data Fabric – including M3 BE redesigned structures
- Define alternative architectures for feeding data from M3 BE to Data Fabric - (initial load, streaming, frequently updated tables)
- Simplify the Data Flow architecture
- Enable real-time data in the database for applications and operational reporting use cases
- Build up a central data warehouse platform by including non-M3 BE data exchange from and to Lakehouse as a central hub
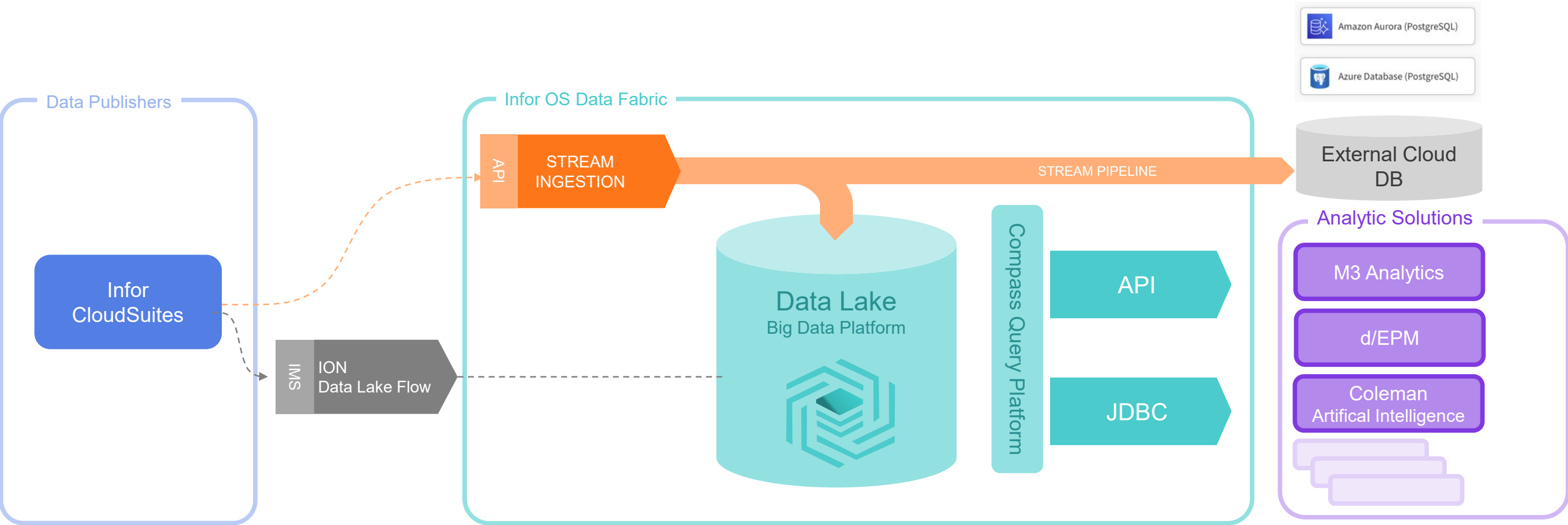
Improve Usability by enhancing content management

- Build content grouping definition in M3 DLP to allow loading subject area data into Data Fabric and simplify content updates
- Keep track of Infor-defined solutions required table data and support table grouping to enhance initial/re-load

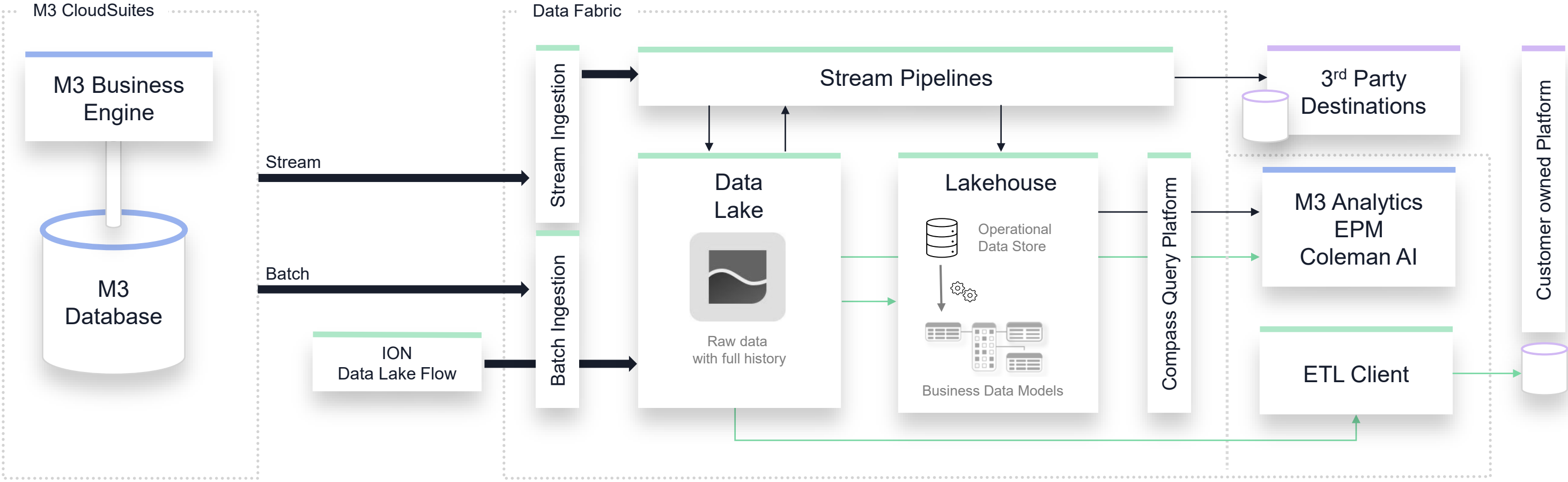## Improve Visibility and Control of the overall solution

- Build a user interface for initial load in M3 Data Lake Publisher with real-time monitoring capabilities to give control to customers
- Implement a heartbeat function to make it possible to immediately alert when infrastructure issues happen
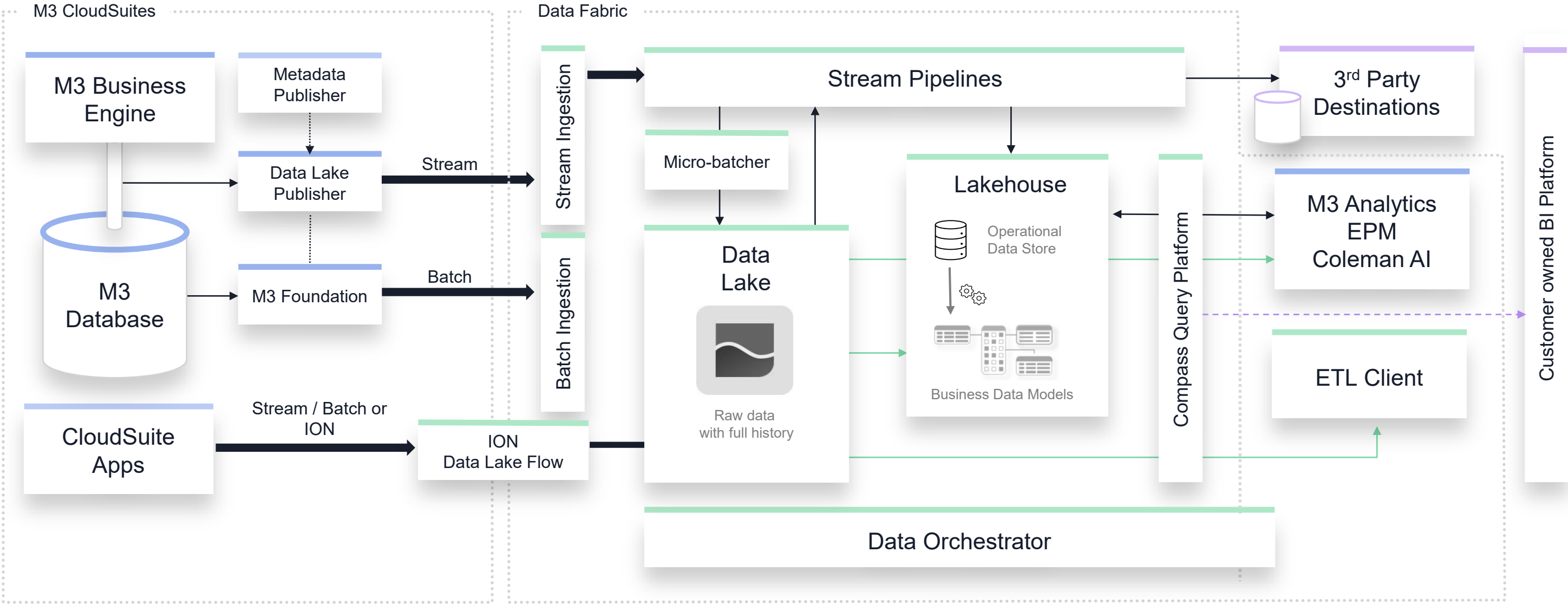
M3 and Data Fabric Overview

Data Publishers

Infor OS Data Fabric

Amazon Aurora (PostgreSQL)

Azure Database (PostgreSQL)

External Cloud DB

Analytic Solutions

Infor CloudSuites

API — STREAM INGESTION

STREAM PIPELINE

IMS — ION Data Lake Flow

Data Lake
Big Data Platform

Compass Query Platform

API

JDBC

M3 Analytics

d/EPM

Coleman
Artifical Intelligence

M3 and Data Fabric 2023

M3 CloudSuites

M3 Business Engine

M3 Database

Stream

Batch

ION Data Lake Flow

Data Fabric

Stream Ingestion

Batch Ingestion

Stream Pipelines

Data Lake

Raw data with full history

Lakehouse

Operational Data Store

Business Data Models

Compass Query Platform

3rd Party Destinations

M3 Analytics EPM Coleman AI

ETL Client

Customer owned Platform

M3 and Data Fabric 2023

M3 CloudSuites

Data Fabric

- M3 Business Engine
- Metadata Publisher
- Data Lake Publisher
- M3 Database
- M3 Foundation
- CloudSuite Apps

Stream — Stream Ingestion
Batch — Batch Ingestion
Stream / Batch or ION — ION Data Lake Flow

- Stream Pipelines
- Micro-batcher
- Data Lake — Raw data with full history
- Lakehouse
  - Operational Data Store
  - Business Data Models
- Compass Query Platform
- 3rd Party Destinations
- M3 Analytics EPM Coleman AI
- ETL Client
- Customer owned BI Platform
- Data Orchestrator

…och nu blir det Produktdemonstration

05
M3 & Data Fabric Journey
Retrospective

M3 and Data Lake

Access M3 data with full record lifecycle

Data Lake is the file storage component of Data Fabric, which stores all data sent from M3 in a similar format as the M3 Business Engine database.

Based on what you select to publish, changes to the M3 database are sent to Data Lake which gives you a full history of changes in data over time.

The variation aspect of all data makes it possible to build scientific and analytical solutions on top of the data lake that can analyze data over time and processes.

Data Lake makes it possible to query a representation of the M3 database and track changes on record level over time

Main features are:

- M3 Data Lake Publisher lets you select which table to publish to Data Lake
- XtendM3 Dynamic Tables can also be published in the same way
- Data Lake storage has a very low cost compared to database storage
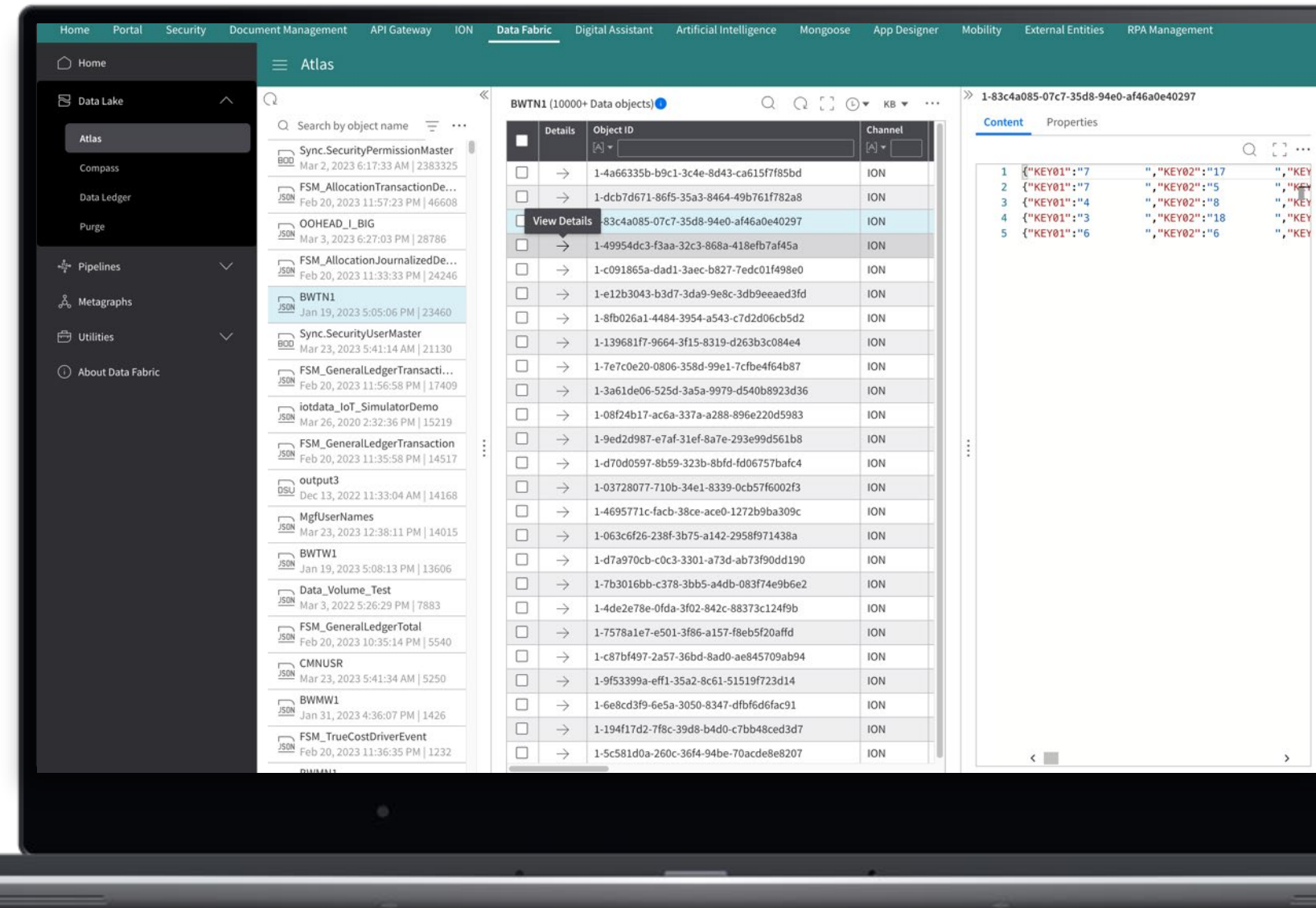
Compass

**Access Data Lake using SQL with variation logic**

Compass on top of Data Lake gives you full control over data using SQL like it was your transactional M3 database.

Data Lake is not and has not have the performance of a database but Compass caches data in a structure to support high-performance queries on the files stored in Data Lake.

With Compass you don't have to use asynchronous APIs on Data Lake to access data

Main features are:

- Data access using familiar ANSI-SQL standard
- JDBC and API access to Data Lake using Compass
- Data Management and administration in easy-to-use experience in Data Fabric application
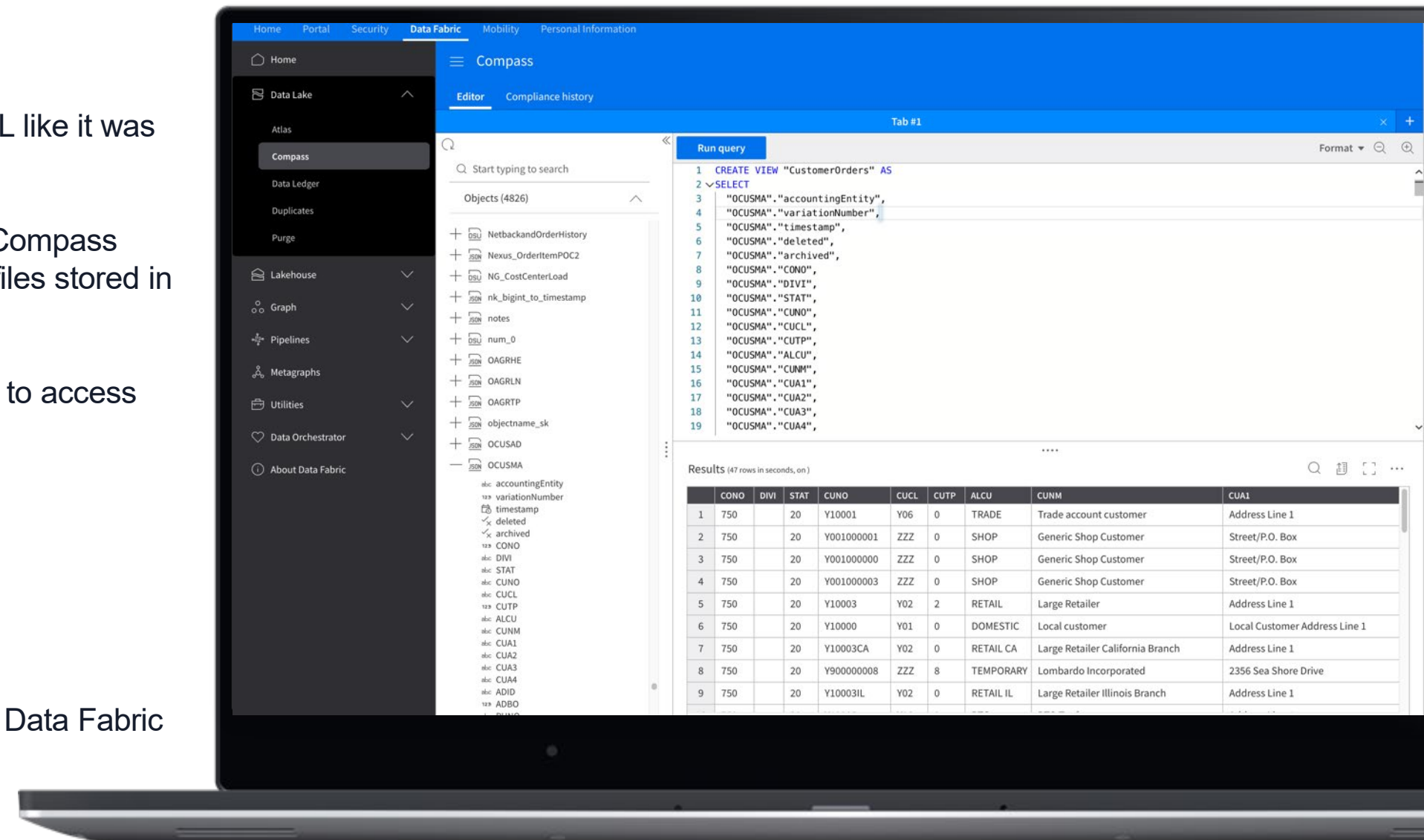
M3 and Data Ledger

Data Ledger monitors the M3 data publishing and storage in Data Lake

To reconcile Data Lake with M3, Data Ledger is monitoring what M3 sends to Data Lake and what is being stored.

This is the first step in monitoring whether data is published and available in Data Lake in an asynchronous architecture as with ION. A second tab dashboard shows streamed data into Data Fabric stored in Data Lake.

Main features are:

- M3 sends periodically information about table row count in M3 Business Engine Database as of a timestamp
- Data Ledger looks into Data Lake for row count as of the same timestamp
- Statistics presented in Data Ledger with red or green cards indicating successful or deferred or failed data delivery
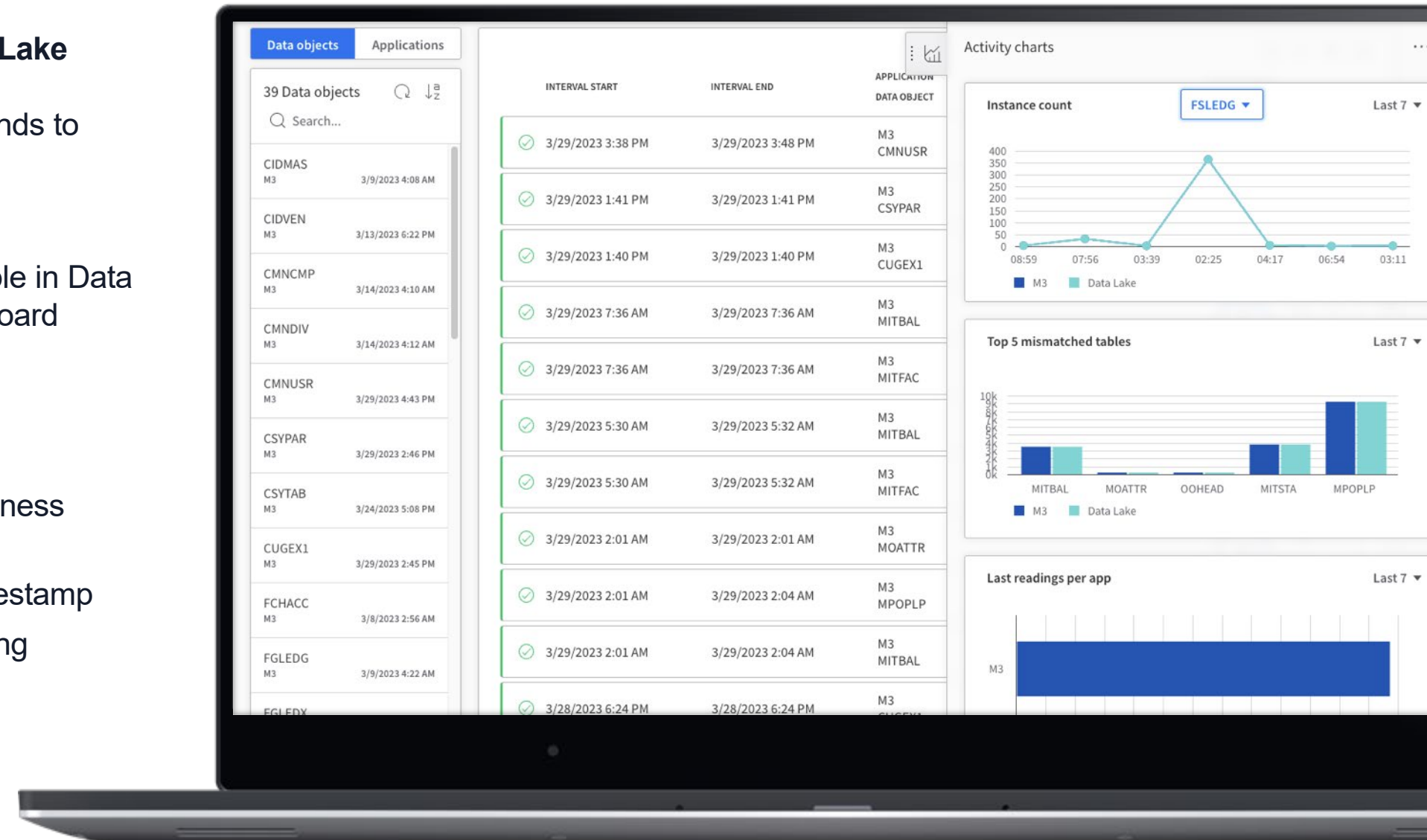
Metagraph

Accelerating the content development curve.

Metagraphs alleviate headaches and reduces guesswork with Infor-provisioned and user-created Metagraphs that provide domain-specific canonical models representing functional application modules, screens, and reports.

Graph-led representation of your metadata footprint is easily converted and published as a Data Lake View to accelerate developer time in producing meaningful content.

Main features are:

- Better understand your data at the metadata level with graph-oriented designs describing relationships
- Drag & drop modeling allows users to compose and extend Metagraphs as domains evolve
- Operationalize your Metagraph with the Publishing Wizard and publish Views straight to Data Lake within 3 steps
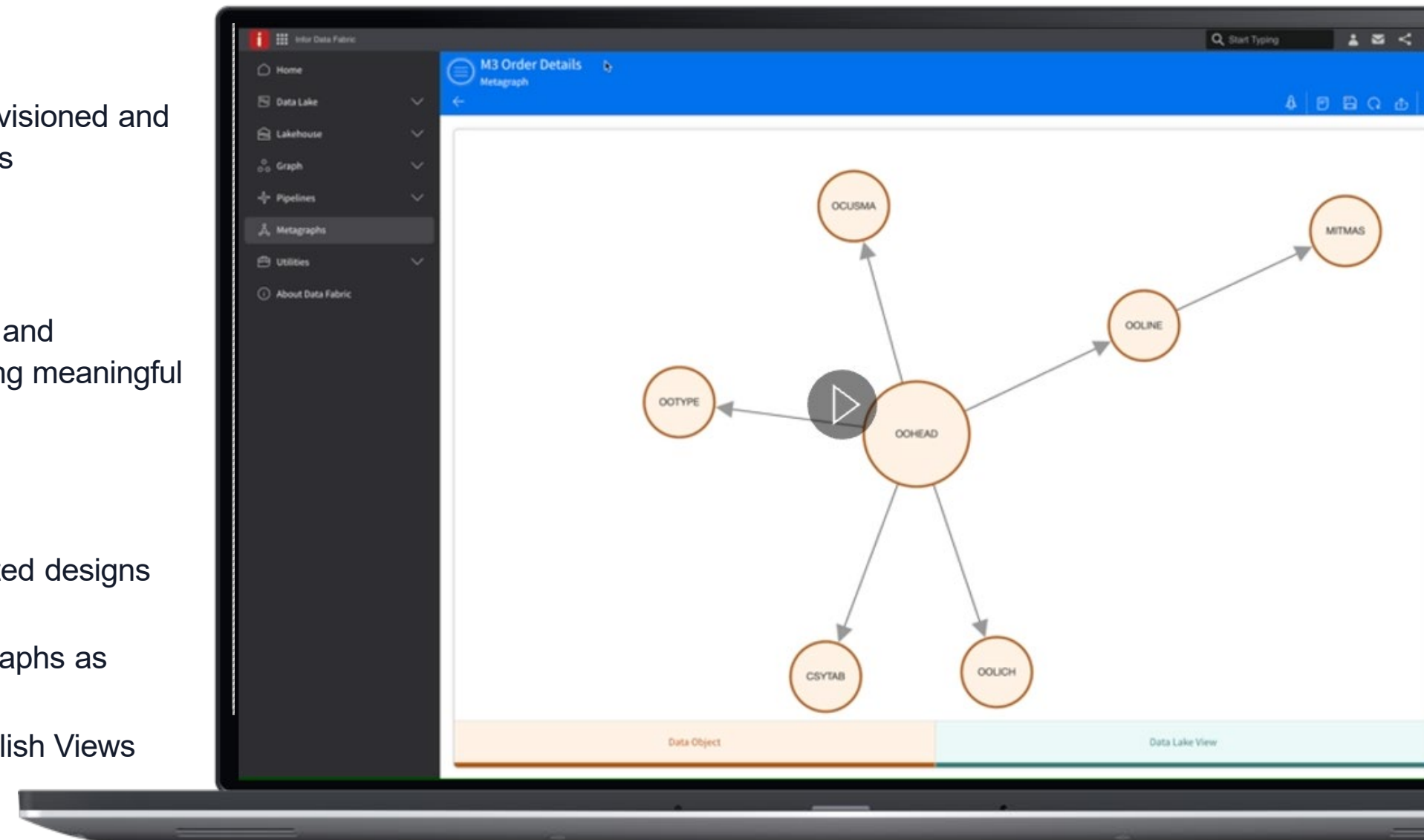
Atlas Upload Widget

Quick & easy method for users to upload data in Data Lake using an Infor OS widget interface.

Transfer local files to Data Lake without setting up data flows to extract from a source or having to integrate with the Data Fabric Ingestion APIs.

- Available in Ming.le Homepages and Infor Portal (V2)
- Empowers non-technical personas to contribute to Data Lake
- Deploy data from unintegrated data sources
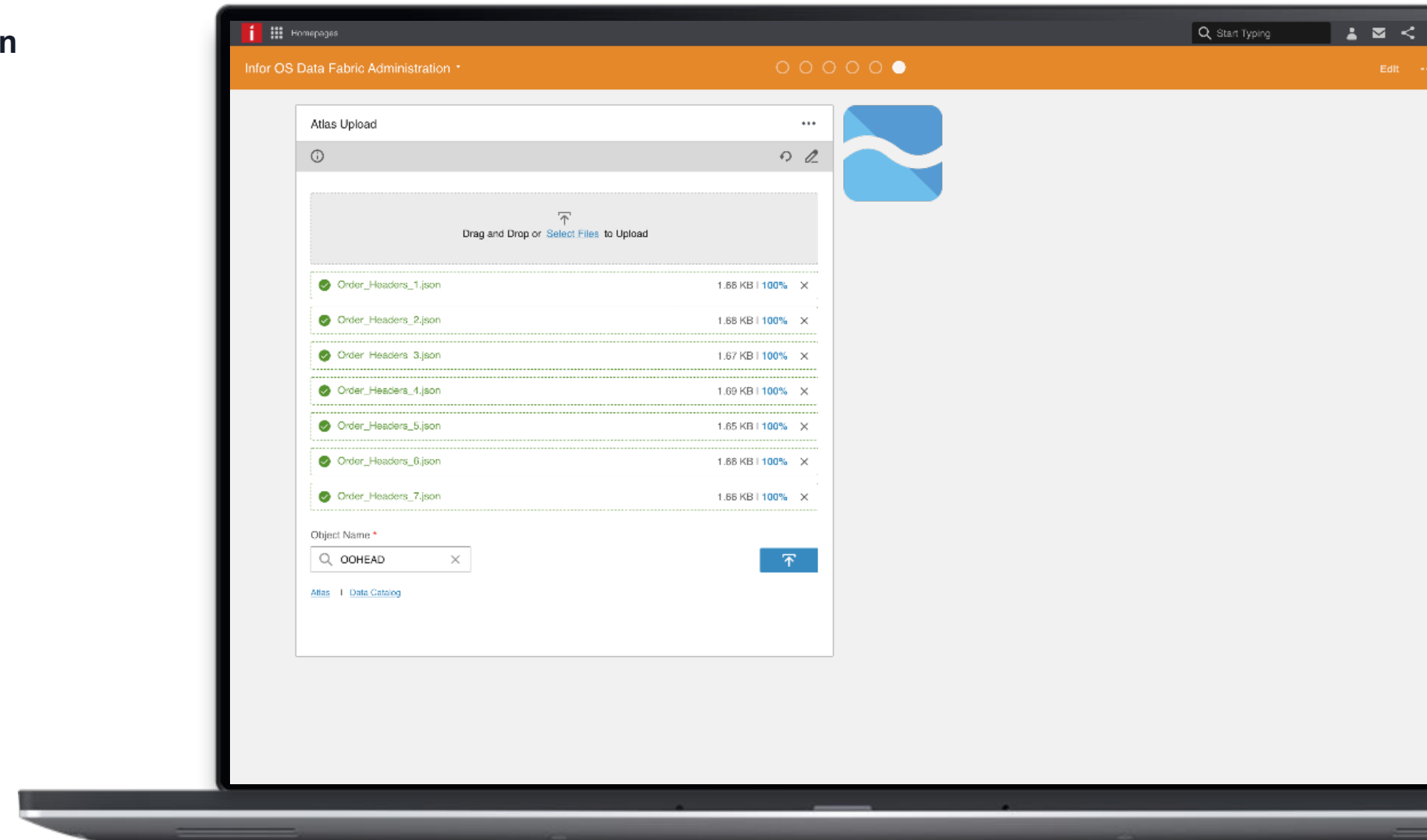
M3 Archiving to Data Lake

Offload and save space in your transactional database by moving archived data to Data Lake

Archiving to Data Lake is a functionality in M3 Business Engine that takes archived data into Data Lake.

This optional process of archiving is reducing the cost for M3 database storage by moving data into much lower-cost storage.

You can query the archived data in Data Lake by using Infor-specific query hints and present the data in query tools or ETL the data to an on-prem database.

Archived data has a specific flag set that protects it from being purged. Every record in the archived data also has a record-level archive flag set.

Main features are:

- M3 archive functionality works the same
- An additional (optional) step transfers archived data from the M3 archived schema into Data Lake
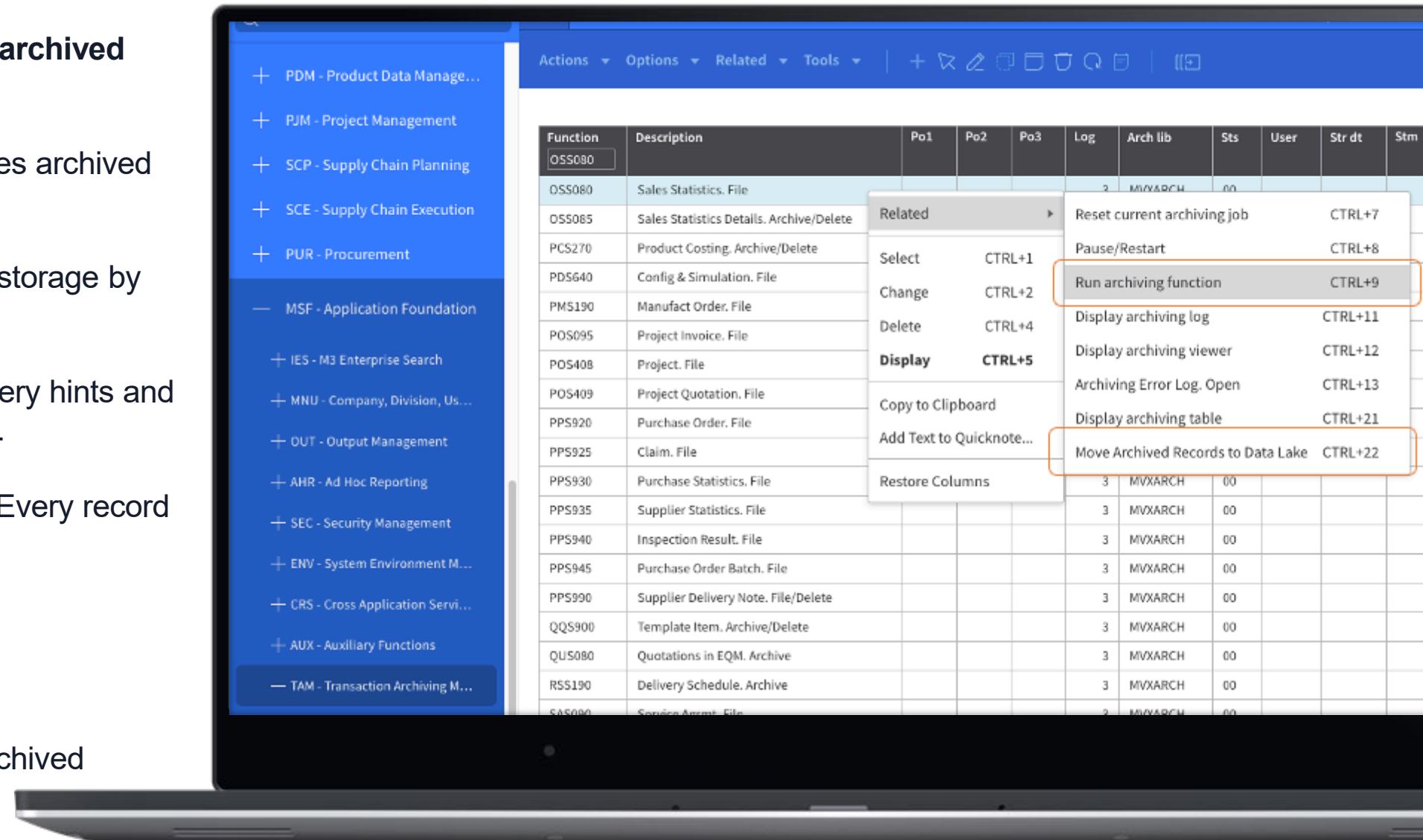
M3 Streaming to Data Fabric

Real-time data streaming to Data Fabric to support operational data scenarios

M3 streaming makes it possible to use Stream Pipelines into a cloud destination database for real-time data consumption. This brings data faster to downstream applications and solutions than classic architecture using ION Data Flows.

This makes massive data streams go direct to Data Fabric and let ION focus on application integrations.

Data Lake Publisher and The EventHub platform sends data to Data Fabric without micro-batching it in advance.

Main features are:

- Easy configuration in Data Lake Publisher
- No need to configure ION after adding new tables in Data Lake Publisher
- Reduced risk for duplicates with direct ingestion
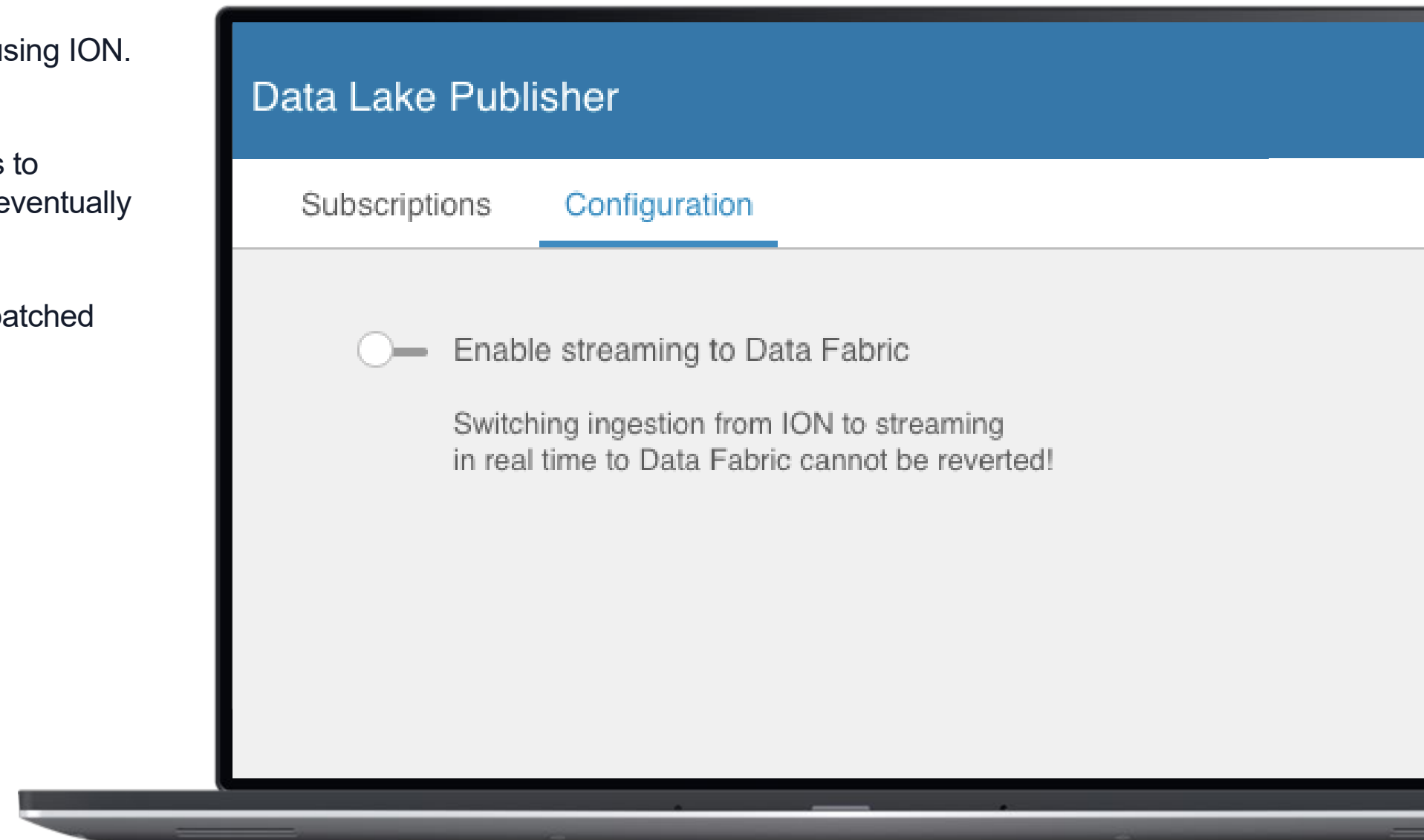
M3 Streaming Toggle

Streaming is a new way to publish data to Data Fabric directly without using ION. When switching to Streaming, you cannot go back to ION!

M3 Streaming is intended to be used with Data Fabric Stream Pipelines to support operational data solutions with real-time requirements. Data is eventually micro-batched by Data Fabric into Data Lake after 10 minutes or 5MB.

Without Stream Pipelines, data will be streamed and eventually micro-batched the same way.

+ Support real-time solutions with Stream Pipelines
+ Data Lake files get an optimized structure

- 10 min latency in Data Lake

Data Lake Publisher

Subscriptions   Configuration

⊙— Enable streaming to Data Fabric

Switching ingestion from ION to streaming
in real time to Data Fabric cannot be reverted!

M3 using Stream Pipelines

Stream Pipelines will take your streamed data from M3 and feed your cloud database.

When you have operational use cases and want real-time data, the Data Fabric Stream Pipelines can be defined to tap into the Stream Ingestion feature and pipe the data in real-time to the destination.

All data hasn't necessarily real-time requirements and can be accessed in Data Lake using Compass, but for the same tables in M3 there could be a need for that.

Stream Pipelines bring M3 streamed data immediately into a destination but can also replay existing Data Lake data through an initial load by the pipeline.

Main features are:

- First enabled destinations are PostgreSQL on AWS and Azure, managed by you as a customer
- Batch ingested data will also flow through the pipeline
- Additional destination types will be available soon

infor
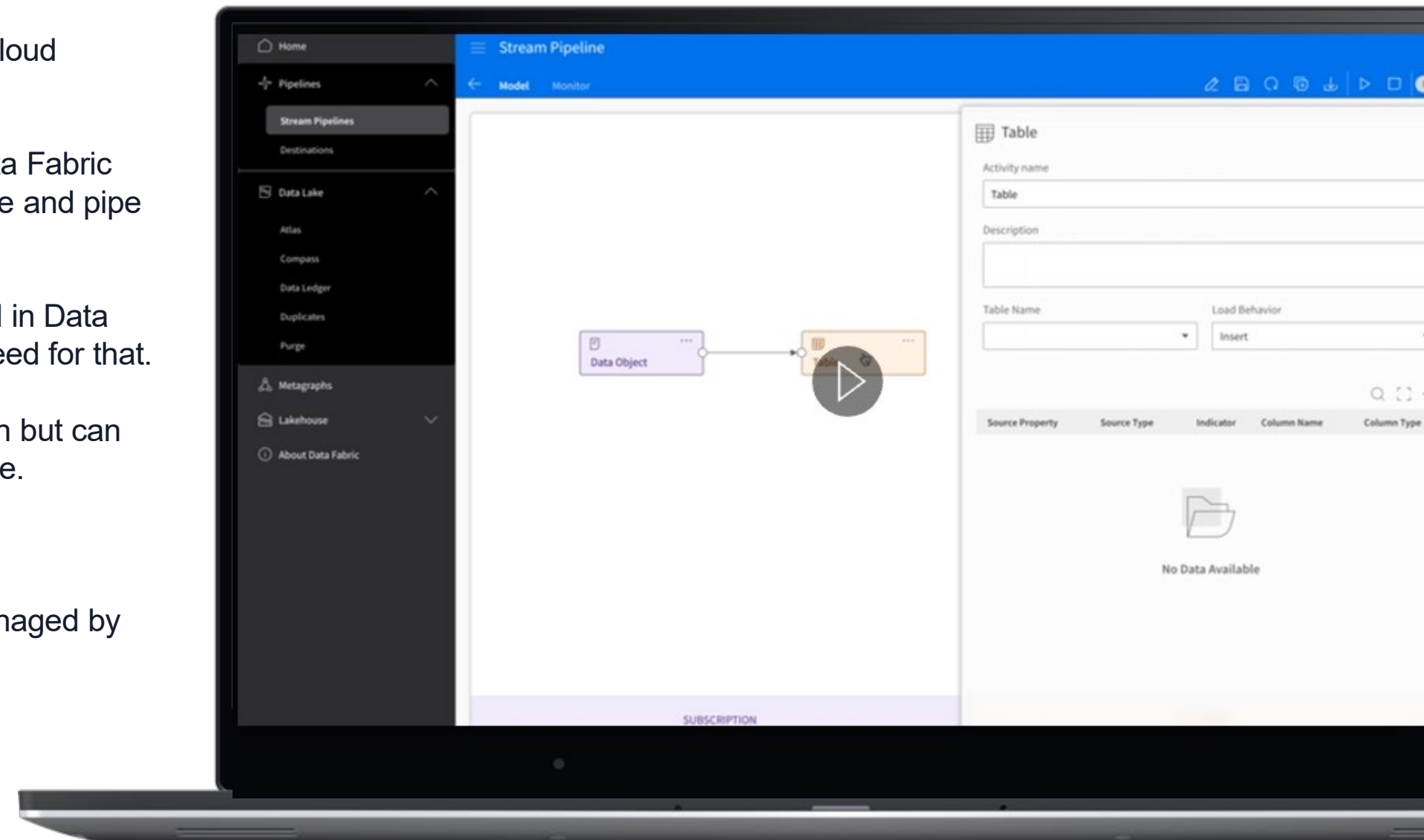
ETL-Client

**Extract and transform data from Data Lake and load it into your on-prem relational database**

The ETL-Client supports incremental extraction of data ingested to Data Lake and loads it into your database.

This enables a database representation of M3 that can let your existing BI platforms and applications access the data in a similar way as M3 on-premises.

You install and manage the ETL-Client on your own premises environment and incrementally extract data from your Infor CloudSuite Data Lake.

Main features are:

- Scheduled periodic incremental data extraction
- Functionality to transform data before loading into the database.
- Supports several database platforms
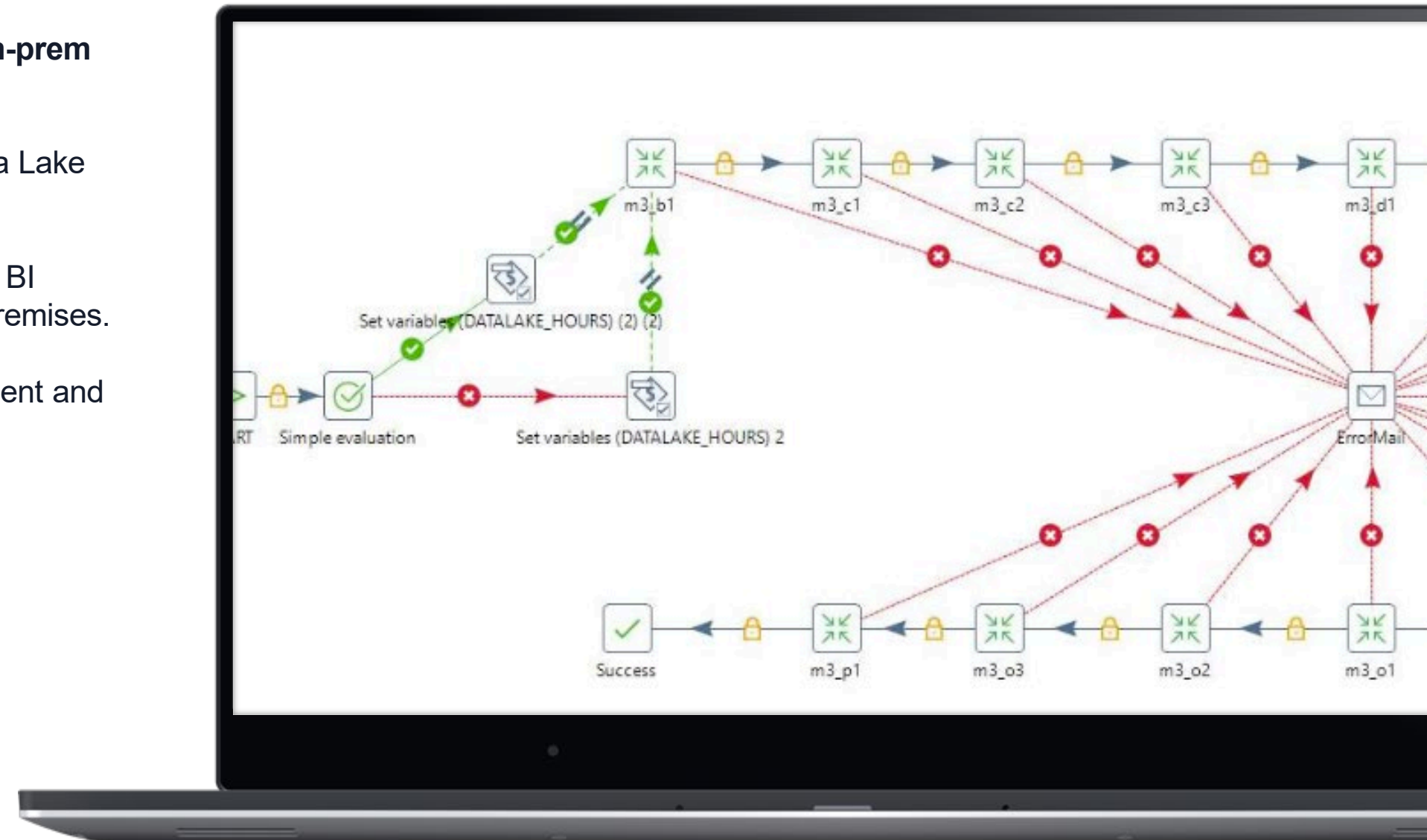
Pipelines 1ˢᵗ release features

Pipeline designs and operation

Build & Operate

Data pipelines are created and deployed to runtime with few easy steps using a visual modeling interface

Live monitoring

Stream Pipeline's live monitoring dashboards enable you to gain immediate insight in processing, delivery and exception capturing activities

Data Sources

Live data

Subscribe to Data Objects to process live data from streaming sources, as well data coming from batch publishers

Initial load

Feed Stream Pipelines with historical data that is already stored in Data Lake

Data Delivery

Destinations

Define, manage and use delivery locations in pipelines. Supported technology: PostgreSQL

Insert and Upsert

Choose either to store every historical version of a record or have exactly once consistency with only the latest version as in the source system

Replay Queue

Failed deliveries are caught in the Replay Queue so you can identify and resolve the issue before re-streaming the records back into the pipeline

infor

Infor Data Fabric – Stream Pipelines

Destinations

Pipelines enables fast data delivery to various technologies, relational databases, analytics warehouses, streaming platforms and storage locations.

The Destinations component is used for defining and managing the connection to these locations where Stream Pipelines can offload data in real-time processing.

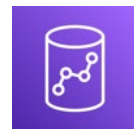|			
Amazon Aurora PostgreSQL	Azure Database for PostgreSQL	Data Fabric Lakehouse (2H 2023)	Snowflake (2H 2023)
(supported in first release)			
SQL Server (2H 2023)	Amazon Redshift	Streaming platforms	Storage

Destinations:
PostgreSQL requirements

In terms of setting up, hosting and connecting your PostgreSQL data warehouse to Stream Pipelines, you have the following options:

- In your cloud vendor, create the database with a public endpoint that is accessible on the Internet

- Configure the database to have password authentication

- Configure your firewall to allow access from Infor IP addresses

- Create a database user for Stream Pipelines with access grants to read, write and create tables

- Create a Pipeline Destination for an Amazon Aurora PostgreSQL or Azure Databases PostgreSQL data warehouse

06
Summary

M3 and Data Fabric - Summary

M3 has chosen to collaborate closely with Infor OS Data Fabric

- M3 started early to adopt **Data Lake** back in 2018 as the data source for M3 Analytics, using Birst and JDBC against Data Lake. M3 sent data in micro-batches to an ION Data Lake Flow for storage in Data Lake.

- **Compass** on top of Data Lake made customers start querying Data Lake with standard SQL and integrated their downstream applications to work similarly to how SQL Queries were used in the on-prem platform.

- When business-critical applications were using Data Lake data, it was necessary to reconcile data in Data Lake, and **Data Ledger** was developed by the Data Fabric team.

- To make Data Fabric support direct ingestion without ION, **batch** ingestion, and **stream ingestion** methods were built. M3 is the first Infor solution to use streaming and several M3 solutions are using the batch ingestion method.

- Several customers had on-prem SQL Servers for their BI platforms and the **ETL Tool** was added to Data Fabric to support incremental ETL from Data Lake.

- To support operational data use cases with real-time requirements, **Data Fabric Stream Pipelines** feature was built and M3 can now stream data in real-time to an AWS Aurora PostgreSQL or Azure PostgreSQL database.

- **Lakehouse** is the next area for M3 to adopt and build solutions for Infor's Data Warehouse as a Service for open **Decision Support** solutions and **operational** real-time data access solutions.
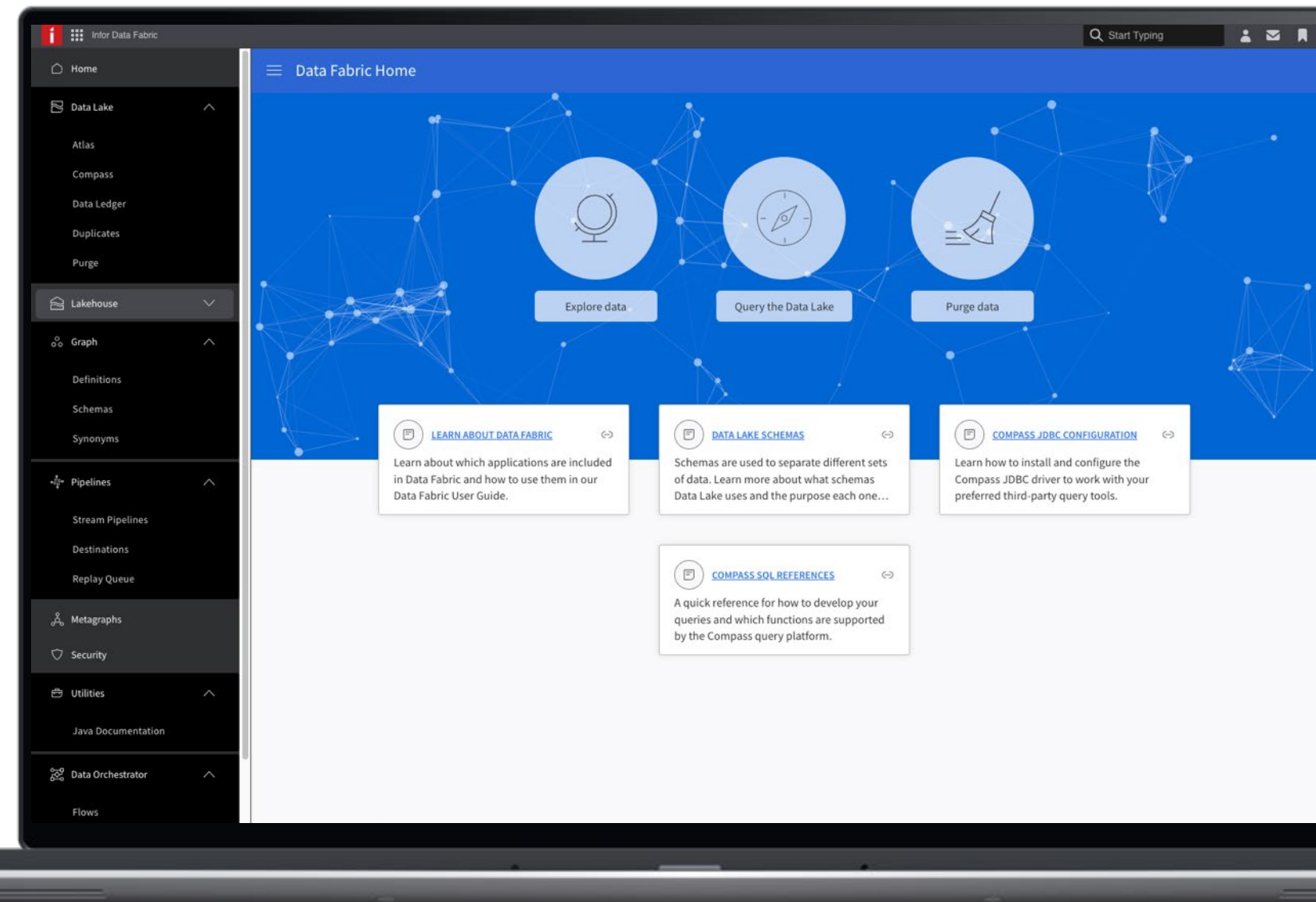
Infor

Thank you

Infor is a global leader in business cloud
software specialized by industry.

infor.com

infor | Smart. Preconfigured. Modern.