



Informatica®

Faster, Simpler, More Cost-Effective Cloud Data Integration for Analytics

Get more value from your data with ETL, ELT,
and elastic data processing



Contents

Connect Everything, Faster	3	Solve Resource Constraints	8
Cloud Data Management Challenges: Cost, Resource Constraints, and Complexity	4	In Practice: Sunrun Creates a Brighter Future	9
Three Must-Have Capabilities for Cloud Data Integration	5	Simplify Complexity	10
Cost controls: Minimize Cost, Maximize Performance	6	In Practice: Grant Thornton Accelerates Client Service Delivery	12
In Practice: Increasing Student Enrollment Using Analytics	7	Conclusion: Having It All with Informatica	13
		About Informatica®	14

Tip: click to jump straight to any section.



Connect Everything, Faster

Your organization's ability to grow and transform depends on connecting data across your enterprise to generate insights. That imperative puts data integration at the heart of any digital transformation initiative.

And cloud is pivotal to digital transformation. The latest release of the Flexera State of the Cloud Report confirms what many tech leaders are experiencing on the ground: the post-pandemic rate of cloud adoption is accelerating. Specifically, organizations are adopting multiple clouds and integrating them with on-premises systems. According to the report, 92% of enterprises have a multi-cloud¹ strategy and 80% have a hybrid cloud strategy.

Cloud adoption and an increased focus on integration go hand in hand. According to Gartner®, “By 2025, more than 75% of the midsize, large, and global organizations will establish integration strategy empowerment teams to support collaborative integration, up from 40% in 2021.”²

But as data integration grows more important, it has also become a strategic challenge for businesses. A profusion of new technologies, patterns, and use cases can make it difficult to see a clear path forward. Whatever your current scenario, the end goal is the same—gain more value from your data by getting it out of silos and connecting it with other data and with people across your organization who will use it to make better decisions.

In technical terms, this means ingesting and integrating data from various sources and landing it on a cloud data warehouse or lake where it can be used for analytics and data science projects. Your cloud data integration solution should enable you to do this rapidly without requiring cost or performance tradeoffs—and without requiring an army of developers.

Navigating the choices of today's cloud data integration options requires a strategic guide. In this eBook, we will walk you through today's integration challenges and show you how the right approach can help you get started rapidly, gain value from your data faster, and find the most cost-effective method for your needs.

¹ Flexera 2021 State of the Cloud Report

² Gartner, Ensure Your Integration Strategy Supports Modern Integration Trends, Abhishek Singh, Eric Thoo, 23 February 2021
GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.



Cloud Data Management Challenges: Cost, Resource Constraints, and Complexity

Today's cloud data integration landscape can appear crowded and confusing. Navigating it successfully requires a focus on the three challenges that can hold businesses back.

Many organizations that are new to cloud data integration want to get started quickly and are looking for quick wins from new cloud analytics initiatives. Then they realize that they've been approaching integration too narrowly and need more capabilities. They have more data sources and data types to integrate, more complex transformations, multiple clouds involved, and more stakeholders asking for more data.

The problem is, what was doable on a small scale with a few resources becomes exponentially more complex and costly at scale. Upgrades and technology changes—whether changing clouds or processing engines—require heroic efforts. And the opportunity cost becomes steeper.

Other organizations have mature data analytics systems in place. They are modernizing

on-premises data stores and moving to cloud analytics. Companies are expanding their analytics footprint on cloud either by shifting existing projects to cloud or by creating new initiatives on cloud.

These scenarios may seem different on the surface—but on a fundamental level, they reveal three core challenges that define data management today, for both modernization from on-premises systems and for new cloud analytics initiatives.

- **Cost:** Moving large amounts of data in and out of clouds is expensive. The right cloud data integration tools not only make more efficient use of cloud infrastructure for data processing, but they also deliver performance gains—making traditional cost vs. performance tradeoffs a thing of the past.
- **Resource constraints:** Data-led organizations prioritize efficiency and productivity—because the need for integrations is constantly increasing, but resources aren't. Exponential

data growth makes it essential that your strategy includes high-productivity tools that include both easy-to-use, self-service GUIs and intelligent automation.

- **Complexity:** Organizations want to move faster and need simpler tools that won't slow them down. Complexity seeps into every aspect of cloud data management—from complex data landscapes to complicated pricing and hidden cloud fees.

In the past, many data integration tools were limited. And some current solutions are still focused on just a particular use case or solving a single challenge.

But today your cloud data integration solution should be able to do it all—providing a full range of cloud data integration options in a secure, stable, and unified cloud-native solution for multi-cloud and hybrid configurations. In the next section we'll look at how a modern solution can help overcome each of these three challenges.



Three Must-Have Capabilities for Cloud Data Integration

Optimize Costs

The faster you process your data, the fewer compute hours you need. Your cloud data integration solution should be able to accelerate your cloud analytics initiatives while reducing your costs. Auto-tuning and auto-scaling, along with consumption pricing, can help rein in costs. When your workload fluctuates, your data pipeline should have the ability to spin up new instances or take down the unused instances automatically. Auto-scaling helps to build the elasticity in your data pipeline. Auto-tuning, on the other hand, helps to improve the performance by optimizing data processing parameters. Taken together, auto-scaling and auto-tuning give you more granular control over costs.

To optimize costs, you want to be able to send your data processing to the most cost-effective option, whether that's pushing down data processing to the native cloud ecosystem or the cloud data warehouse, processing data

as traditional ETL, or using an elastic Spark processing engine, which drives scalability. Here's a quick rundown of the different capabilities your cloud data integration solution should provide:

- **Extract, transform, load (ETL):** Traditional data processing, where data from any source—such as SaaS applications or on-premises systems—is transformed before ingesting it into a target application such as ERP or a data warehouse for business processes and analytics.
- **Extract, load, transform (ELT):** This approach can take a couple of different forms, including **pushdown optimization**. You may be moving data from a cloud data lake to a cloud data warehouse (public cloud pushdown) or processing data within a cloud data warehouse (cloud data warehouse pushdown). Your cloud data integration solution should use the cloud data warehouse's compute capacity and its native commands for the greatest efficiency.

- **Elastic data processing:** For large data volumes (for instance when ingesting data onto a cloud data lake), **elastic, Spark-based processing** on Kubernetes can process multiple terabytes of data.

For successful cloud analytics, your cloud data integration solution should support all three types of data integration. What's more, you should have the ability to easily choose the type of processing for each workload, with AI-powered intelligence that takes the guesswork out of it.

In addition, your cloud data integration solution should be able to alert you to more cost-effective processing techniques and recommend options. It should enable you to set limits on compute time, automatically adapting limits based on your usage patterns and behavior and optimizing for your location.



Cost controls: Minimize Cost, Maximize Performance

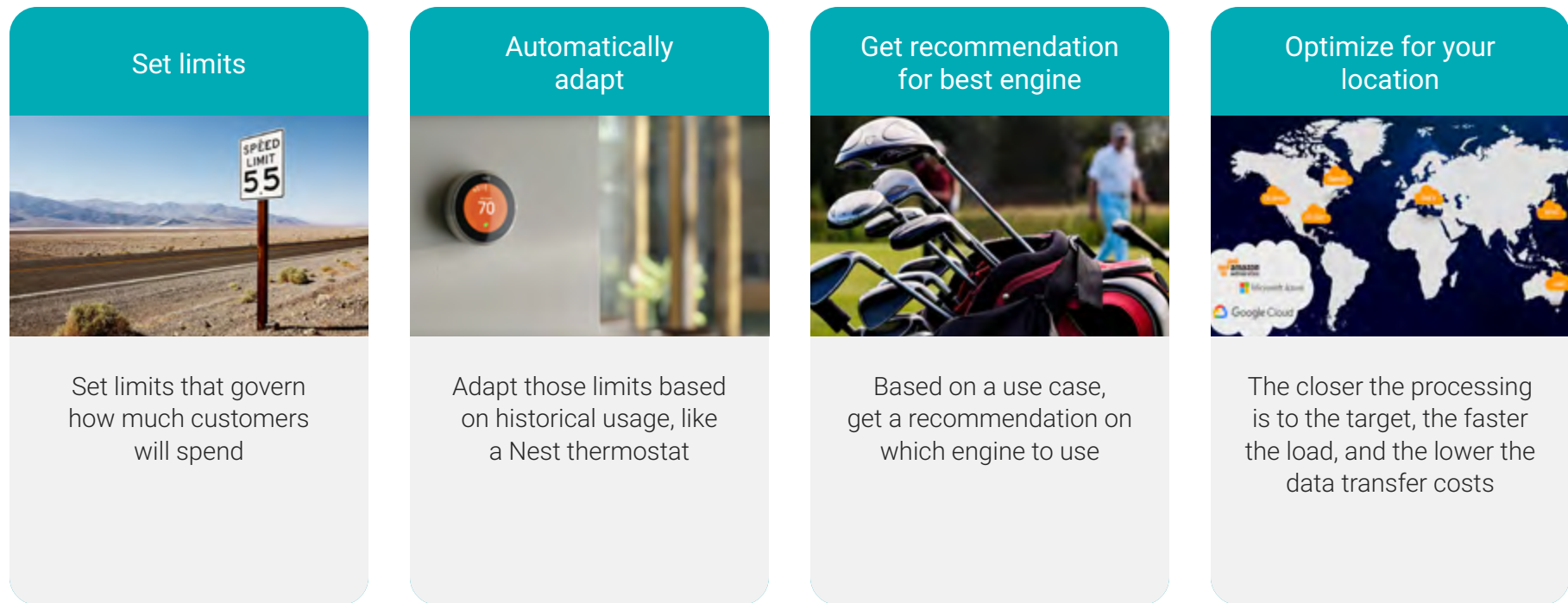


Figure 1. Setting limits and using automated recommendations help rein in costs.

In Practice: Increasing Student Enrollment Using Analytics



The University of New Orleans (UNO) has been ranked third in the nation by The Brookings Institution for providing equal access to higher education for students from all income levels. As an engine of opportunity, the institution wanted to analyze student and employee data from multiple on-premises and cloud sources in a centralized data warehouse. UNO also wanted to improve efficiency and reduce costs by automating manual ETL processes.



The university used cloud-native data integration to extract, transform, and load data from Salesforce, SQL Server, and Workday into a cloud data warehouse. They also used cloud connectors for Snowflake and Workday to connect on-premises and cloud systems without hand coding. UNO can now make data available across the organization for program analytics to help increase student enrollment and retention and has lowered costs by enabling reporting in less time.

“[With cloud-native data integration and data warehousing], we’re democratizing data across the university while improving operational efficiency. We’re making data much more accessible. This enables reporting in less time while helping to improve information security and reduce operating costs.”

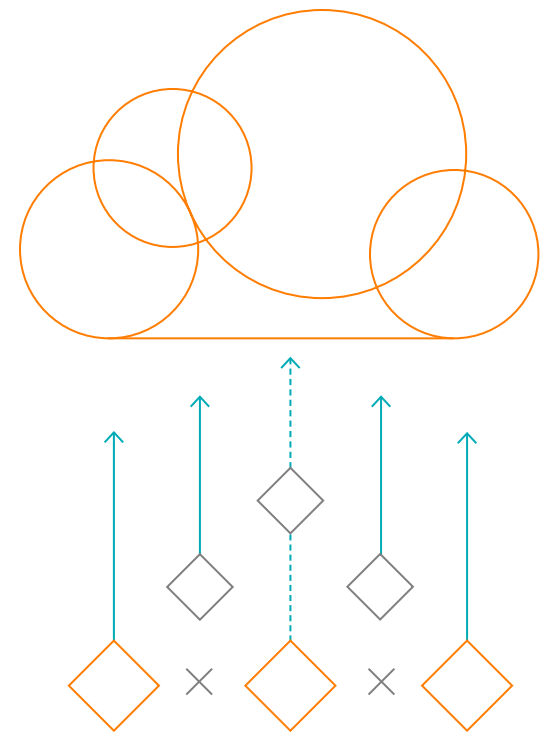
— **Dr. Ray Wang**, CIO,
University of New Orleans

Solve Resource Constraints

According to Daxx, a software development outsourcing company based in Europe, the shortage of developers in the US alone is projected to exceed 1.2 million by 2026.³ Organizations across the globe are similarly struggling to find staff with specialized development skills. The need for hand coding and a lack of automation in many cloud data integration products—combined with a lack of tools for non-technical users—are also holding back many companies.

An AI-powered cloud data integration solution that provides automation and a self-service, graphical user interface (GUI) for simple, role-based tools can help address these gaps. Organizations can reduce development time by up to 80% for everyone from architects to business analysts to IT operations by providing self-service access to no-code/low-code tools with multi-step wizards for data ingestion and integration. A serverless data architecture with no hardware or software to manage can also greatly improve productivity.

Compared to tools that require each job to be built manually, with no reusability, an AI- and ML-powered cloud data integration solution can greatly speed up development, automatically recognizing different file formats and enabling you to deploy many reusable mappings.



³ <https://www.daxx.com/blog/offshore-team/how-hire-offshore-developers-for-startup>

In Practice: Sunrun Creates a Brighter Future

SUNRUN®



With more than 285,000 customers in 22 states, solar company Sunrun is already powering the equivalent of a mid-size city, and it installs a new system approximately every two minutes. The organization wanted to achieve faster, more accurate reporting and visualization of business metrics and analytics. The company also wanted to accelerate cloud migration and data pipeline construction to save time for IT resources and improve refresh cycles for data consumers.

Sunrun migrated from an on-premises infrastructure to a cloud data lake and warehouse and then used cloud-native data integration to ingest data from dozens of sources systems and build hundreds of pipelines into the new cloud infrastructure. The resulting system enables the company to cut reporting and visualization development time from months to weeks, delivering high-quality analytics to business users three times faster. Sunrun also reduced data warehouse design time by 50% and infrastructure building time by 75%.

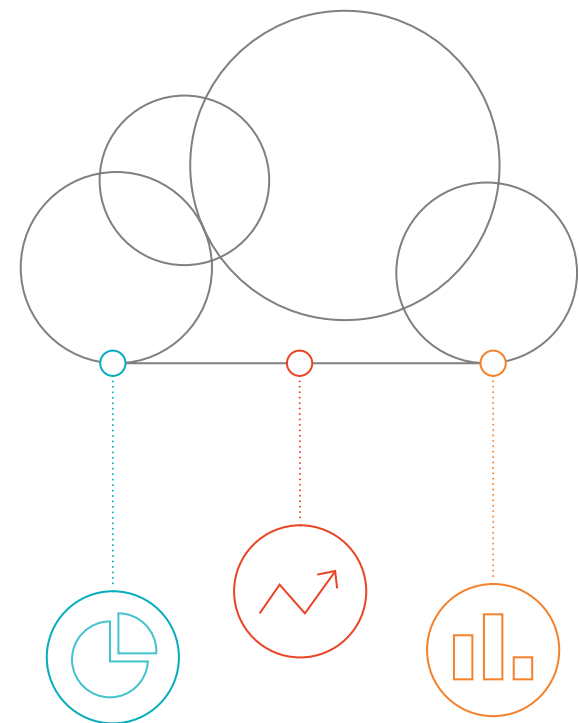
“We needed a robust, cloud-based integration tool to bring data in from our source systems. Refreshing once a day was no longer often enough for some sources; for example, we wanted to load sales and marketing data on an hourly basis.”

— Engineering & Analytics Director,
Sunrun

Simplify Complexity

Disjointed tools for data ingestion, change data capture, data prep, data integration and transformation, and data quality that don't work together present significant hurdles to organizations. In addition, cloud and on-premises deployments often don't interoperate, and a resource-intensive and costly development lifecycle can slow down projects. All these factors increase technical and operational complexities, directly affecting a business's ability to quickly innovate and achieve a competitive edge.

To streamline the process of developing and maintaining data integration pipelines, you need a single, comprehensive cloud data integration solution that provides one set of tools for all your cloud and on-premises data and applications (see Figure 2 on next page). End-to-end cloud-native data management—for everything from data ingestion to data warehouse and lake modernization to application modernization—enables organizations to unify the development and production process for faster, more-effective cloud data integration and a speedier path to cloud analytics.



Simplify Complexity (continued)

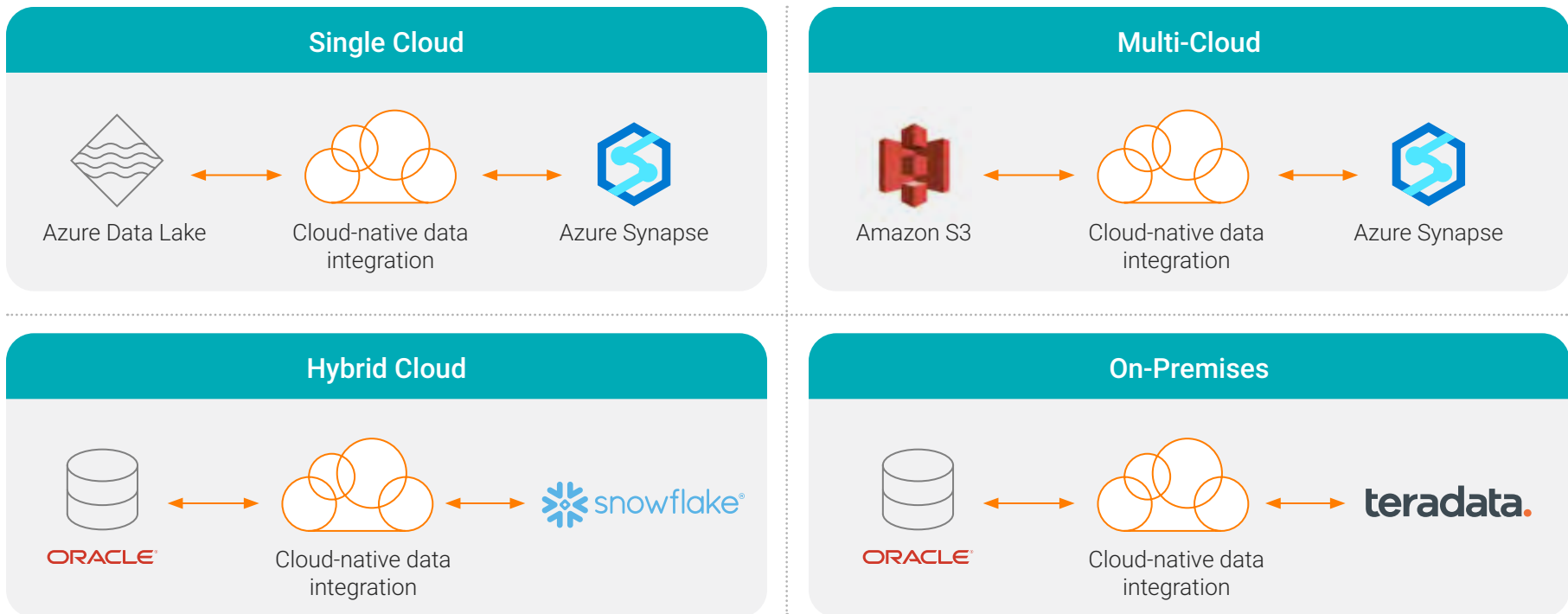


Figure 2. Simplify complexity by using one set of tools for all your cloud and on-premises data and applications.

In Practice: Grant Thornton Accelerates Client Service Delivery



Grant Thornton, one of the world's leading organizations of independent audit, tax, and advisory firms, wanted to transform more customer and financial data, faster. The company also wanted to integrate data from new cloud-based decision-making, financial analytics, and client services applications without spending time developing and maintaining integrations.



To accomplish these goals, Grant Thornton turned to a cloud data integration solution that unifies data from on-premises and cloud applications for analysis. They also used cloud data integration capabilities to deliver accessible, trusted, and secure customer and financial data from most common data sources. As a result, Grant Thornton reduced the data integration development and maintenance lifecycle by 50% and improved business decision-making, financial analytics, and client services with faster access to needed data.

“Our strategy is not just to be cloud ready, but to integrate new cloud solutions. There’s a lot more technology out there that’s playing a part in how our customers do business and therefore how we service them.”

— Senior Director of Enterprise Architecture,
Grant Thornton

Conclusion: Having It All with Informatica

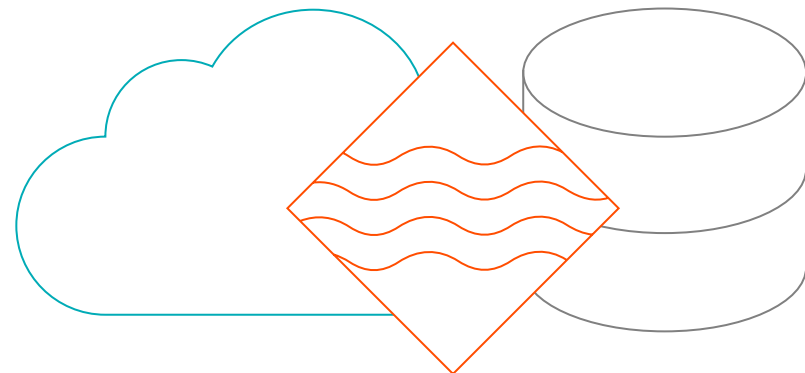
Informatica Cloud Data Integration (CDI) is a comprehensive, cloud-native data integration solution that can ingest, enrich, and transform any data, anywhere, on multiple clouds or on-premises, for pervasive data management and analytics. Built on a modern technology stack, Informatica CDI is designed to simplify data management, democratize data engineering for all users, and support enterprise-level scaling.

It ensures better control of your data integration cost with AI-powered recommendations and consumption-based pricing. CDI offers high-performance ETL, ELT, ingestion, synchronization, and replication for a multi-cloud and serverless world. It covers a diverse set of patterns, use cases, and users, ensuring you have well-architected and seamlessly automated data pipelines that serve as the neurons of your business.

Experience Informatica Cloud Data Integration now with a FREE 30-day trial

See how broad and deep out-of-the-box connectivity, prebuilt advanced transformations, and zero-code orchestrations help you build enterprise integration workloads—fast.

[START TRIAL](#)



About Informatica®

At Informatica, we believe data is the soul of business transformation. That's why we help you transform it from simply binary information to extraordinary innovation with our Informatica Intelligent Data Management Cloud™. Powered by AI, it's the only cloud dedicated to managing data of any type, pattern, complexity, or workload across any location—all on a single platform. Whether you're driving next-gen analytics, delivering perfectly timed customer experiences, or ensuring governance and privacy, you can always know your data is accurate, your insights are actionable, and your possibilities are limitless. Informatica. Cloud first. Data always™.

Worldwide Headquarters

2100 Seaport Blvd, Redwood City, CA 94063, USA

Phone: 650.385.5000

Fax: 650.385.5500

Toll-free in the US: 1.800.653.3871

[informatica.com](https://www.informatica.com)

[linkedin.com/company/informatica](https://www.linkedin.com/company/informatica)

twitter.com/Informatica

[CONTACT US](#)

IN19-1221-4185

© Copyright Informatica LLC 2021. Informatica and the Informatica logo are trademarks or registered trademarks of Informatica LLC in the United States and other countries. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners. The information in this documentation is subject to change without notice and provided "AS IS" without warranty of any kind, express or implied.



Informatica®