

MAY 2023

Are Organizations Solving Their Data Needs the Right Way?

CTI Data & Confluent: A Data Streaming Framework for the Modern Data Warehouse

Stephen Catanzano, Senior Analyst

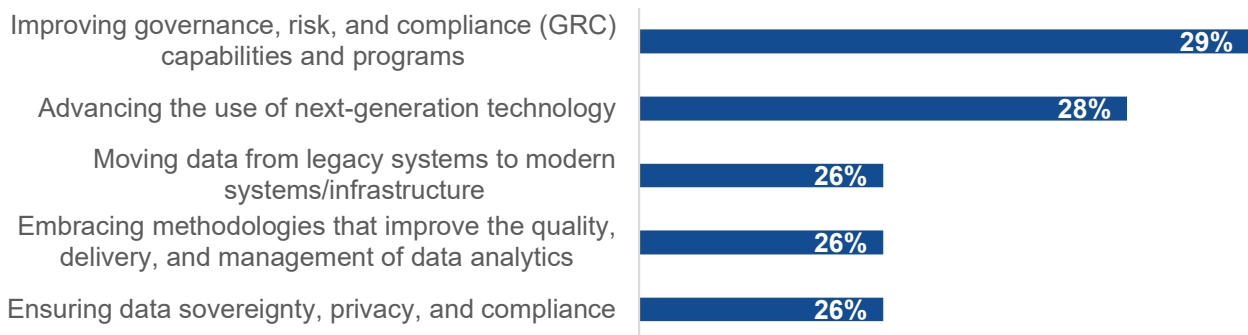
Abstract: Creeping data chaos is accelerating and undermining digital transformation programs as organizations deal with hundreds, if not thousands, of batch and real-time data pipelines across their internal and external processes, but it doesn't need to be this challenging. By implementing analytics, and a data streaming framework, organizations can turn data warehouses into analytics powerhouses. A modern data warehouse improves performance, scalability, data integration, data quality, security, and analytics while reducing costs. CTI Data (CTI) with Confluent provides expertise in implementing frameworks and services to enable organizations across industries to organize data and accelerate time to value.

Improving Data Management and Analytics is a Top Priority

As shown in Figure 1, TechTarget's Enterprise Strategy Group asked research participants about the priorities that matter most to their organization in support of their top data initiatives. 29% of respondents reported that improving governance, risk, and compliance capabilities and programs was one of their top priorities. 28% indicated that they prioritize using next-generation technologies, and 26% look to embrace methodologies that improve the quality, delivery, and management of data analytics.¹ All strongly support improving the use of data across an organization, including analytics, real-time data management, and warehouse modernization.

Figure 1. Top 5 Priorities Related to Data Initiatives

What priorities matter most to your organization in support of new and/or ongoing data initiatives? (Percent of respondents, N=501, three responses accepted)



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

¹ Source: Enterprise Strategy Group Complete Survey Results, [2023 Technology Spending Intentions Survey](#), November 2022.

This Enterprise Strategy Group Showcase was commissioned by Confluent and is distributed under license from TechTarget, Inc.

Data Chaos Is Real

Many organizations are accumulating and managing a growing number of pipelines (aka Kafka "topics"), which rapidly become unsustainable, exploding costs without a proper framework. For organizations just starting their digital transformation efforts and implementing data streaming, now is the time to get control of it.

Data chaos is manifested in challenges to managing data pipelines, including identifying which pipelines are doing what, which ones are redundant, which are going nowhere and are no longer relevant, how they are interdependent, what would break if the data was changed, which pipelines need to be changed to address a workflow change, and whether they are being managed in a compliant and secure way.

Many organizations question how they got to this state of chaos. In the past, data was much narrower in scope, and it was there to underpin a transaction. Now it reflects what is happening in the environment, massively amplifying its breadth and complexities. Meanwhile, system architectures have evolved to be ever more fragmented (loosely coupled), speeding development at the cost of system-wide data standards that were part and parcel of the old monolithic architectures. This kind of data chaos is not a new phenomenon, but what is new is the scale of its disruptive impact. Organizations need to modernize their data warehouses and the data platform to get back in control.

This data chaos leads quickly to higher costs, inefficiencies, potential compliance violations, security challenges, and an inability to maximize the value of data, which is the costliest in terms of value. Organizations that modernize and manage data right create competitive advantages, can innovate faster, and react to unforeseen changes that affect the business.

Digital Transformation Cannot Succeed in a World of Data Chaos

Historically the scope of the business was based on the transaction and how to get things done faster. But now the scope has changed and is focused on the value chain and how to get things done smarter. A big part of "smarter" is turning data warehouses into analytics powerhouses that can deal with batch and event-driven, real-time analytics to revolutionize how an organization utilizes its data.

Organizations should be empowered with in-line analytics focused on enabling people and systems to do more faster. Two key framework parameters are to distinguish between workflow time and real-time.

- **Workflow time:** Analytics used to drive people's productivity. Examples include:
 - **Operations leader:** The system predicts which patients have future needs and flags them for follow-up.
 - **Client representative:** The system predicts a client's long-term value and sets up CRM reminders.
 - **Visiting nurse:** The system identifies patients' in-home needs and plans out weekly schedules.
 - **Senior manager:** The system predicts likely scenarios and lets managers run what-if adjustments.
 - **Executive:** The system correlates and projects leading indicators that impact lookahead performance.
 - **Analyst:** Users can find all customers who meet a criterion and use the system to predict an action.
 - **Scientist:** Can create an algorithm that assesses a system to optimize preventive maintenance
- **Real-time:** Analytics used to create immediate system responses and actions. Examples include:
 - **Customer interaction optimization:** Response analytics drives "next-best action."
 - **Fraud detection:** Transaction analytics kicks out risky/illicit requests before they are processed.
 - **Logistics automation:** Device telemetry analytics continuously optimizes place-pick-pack.
 - **Threat prevention:** Log analytics shuts off risky activity just in time, before attacks occur.

- **Resource allocation:** Performance analytics dictates the on-the-spot reallocation of resources.

The examples above can be applied to any organization going through a digital transformation to become a data-driven organization. There are hundreds of other examples, and implementing the right ecosystem of technology for an organization's unique needs is a crucial part of the equation.

Modernizing a Data Warehouse with CTI with Confluent

Confluent is widely considered an industry leader in real-time data streaming and management and is delivered as a fully managed SaaS solution. CTI is a Confluent partner with a long history of transforming organizations to maximize the value of data. Together CTI and Confluent can modernize organizations to address the critically important need to turn data in-motion into insights that power organizations today and for the future.

Data streaming with Confluent has become a one-size-fits-all solution. In the past, systems were built around three main integration types: application, data, and message:

- Application integration is one system passing data to another system in context/state (event/API-based).
- Data integration is the movement of data between systems with no state dependency (batch/queue-based).
- Message integration is the immediate delivery of data (messages) to any number of registered "listeners" (bus/broker-based).

The Confluent data streaming platform can handle all these needs in one solution. When streaming, there is no need to use extract, transform, and load tools; API gateways; or message buses.

Some of the key benefits of Confluent include:

- Common integration tooling, infrastructure, and skills.
- Ability to evolve from one pattern to another, e.g., data integration to the data bus.
- Single metadata repository, which reduces data chaos.

Applying this to a modern data warehouse can transform a system designed mainly for batch data processing into a real-time unified platform to integrate, store, and analyze data from various sources.

Some of the benefits achieved from modernizing the data warehouse include:

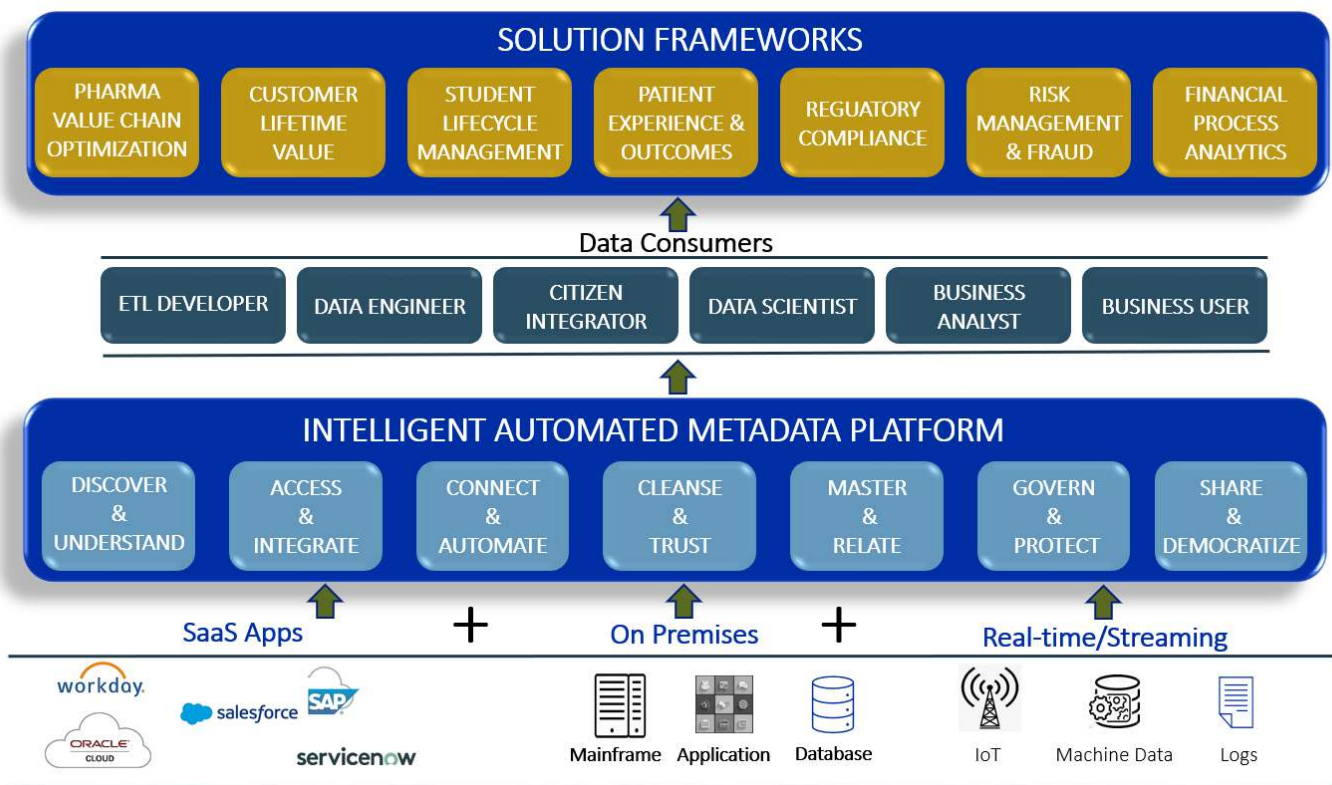
- **Improved data quality:** By integrating data from various sources in real time, organizations can ensure that the data they use for decision-making is up to date and accurate.
- **Faster insights:** Real-time data management allows organizations to analyze data in real time and make decisions quickly without waiting for slower batch data processing.
- **Scalability:** Modern data warehouses are designed to handle large quantities of data and can scale up or down depending on the organization's needs.
- **Cost-effectiveness:** With the use of cloud-based data warehousing solutions, organizations can reduce their infrastructure costs and pay only for the resources they use.
- **Data governance:** Modern data warehouses provide data governance capabilities, allowing organizations to manage data quality, access, and security.

Overall, a modern data warehouse is a powerful tool for organizations looking to manage real-time data and gain insights that can drive business decisions.

CTI: Accelerating Business Value with Data and Analytics

CTI is a Confluent partner, and a modern consulting firm focused on strategy, technology, and digital transformation while leveraging data's hidden value. As shown in Figure 2, CTI's enterprise data integration solutions connect and manage all of an organization's data, no matter where it lives. CTI connects virtually any data source with any environment, in the cloud or on-premises, with a business-driven, integrative approach to optimizing existing investments and realizing the benefits of healthy IT transformation. With a strategy-led approach, CTI builds and manages innovative solutions that align IT teams with business initiatives, including what to buy, build, manage, and operate, and secure mission-critical enterprise environments.

Figure 2. CTI Services Areas



Source: CTI Data

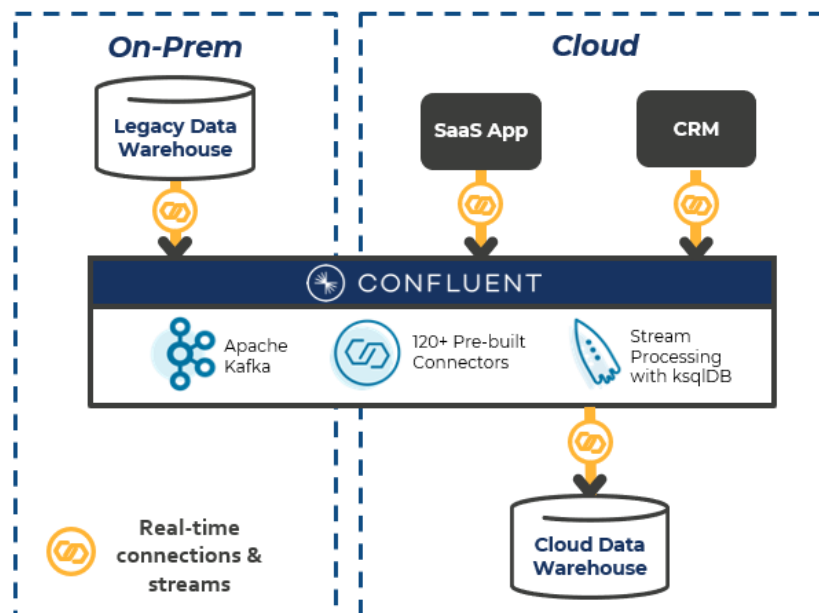
CTI helps organizations build modern data platforms with a multi-disciplinary, holistic team approach that includes the following:

- **Data Fabrics:** Add new data types, products, and apps to the mesh without any impact on established data services. CTI enables organizations to access data from media, the cloud, the web, databases, real-time sources such as IoT devices, and static files generated from applications.
- **Data Pipelines:** Deliver actionable data to consumers in real time using a repeatable process capable of handling batch or streaming jobs and that is compatible with the cloud or big data platforms. Ingest data from any source, organize, and share it at scale.
- **Automation:** Automate data flow and move big data across various sources and destinations. Enable the ability to scale and respond quickly to business demands and opportunities without using valuable engineering time on manual processes.

Confluent's Data In-Motion Solution

Confluent was founded by the creators of Apache Kafka, a widely adopted open source technology for data streaming, and has since become a leading provider of commercial Kafka solutions. Confluent manages real-time data in-motion from any legacy or modern data system into cloud data warehouses to empower analytics, business intelligence, and development teams to use the freshest data for real-time, data-driven decisions and actions. Data can be organized and accessible so analytics systems and data-dependent teams can innovate on the most up-to-date information. An outline of the workflow for modernizing a data platform is seen in Figure 3.

Figure 3. Confluent Real-time Data Warehouse Overview



Source: Confluent

Confluent is a fully managed SaaS solution of Kafka, developed by the creators of Kafka, and continues to be the largest contributor to Kafka open-source. As a SaaS service, Confluent offers scale and flexibility, automated updates and maintenance, rapid deployment, and cost-effective a pay as you grow model.

Confluent's data in-motion solution has three key components:

- **Connect:** Cluster Linking and 120+ pre-built connectors are available to stream data from hybrid and multi-cloud sources to cloud data warehouses like Snowflake, Redshift, Synapse, or BigQuery. Workloads are migrated from on-premises systems like Teradata or Cloudera while maintaining consistent, real-time data for analysis.
- **Process:** Real-time views of business metrics can be built at scale. Native stream processing is used to continuously join, enrich, and aggregate numerous data streams in flight while optimizing them before they move to cloud data warehouses.
- **Govern:** Governance can be applied across streaming data to meet internal and external compliance while ensuring data quality for a cloud data warehouse.

A few key benefits of Confluent's solution include the following capabilities:

- **Power analytics with high-quality data streams:** Maintain high-fidelity data that continuously flows, evolves, and is processed with real-time data streaming and in-flight processing.

- **Boost development productivity with self-service data products:** Deliver ready-to-use, trustworthy data as a product to teams by shaping data into multiple contexts on the fly and enabling self-service search and discoverability.
- **Reduce total cost of ownership of hybrid and multi-cloud data pipelines:** Reduce the complexity of data integration and the risk of downtime or data loss with a fully managed cloud-native service and an incremental, cost-effective cloud migration strategy.

Conclusion

Organizations are dealing with a massive growth in data pipelines and data management challenges, which results in many causes of data chaos due to the new methodologies in data management, an increase in data sources, and more data users demanding faster and better data for decision-making.

CTI understands these challenges and works with organizations to implement solutions designed to break down the chaos and streamline data management across an organization. One core solution is Confluent, by the creators of Kafka, which offers a fully managed Kafka-based SaaS solution with advanced tools and services to manage data in motion at scale.

Enterprise Strategy Group highly recommends that organizations looking to modernize their data warehouse or data platform consider CTI with Confluent to maximize the value of data, accelerate deployment, and be ready to address today's challenges with partners for the future. Contact CTI to understand our framework for addressing data chaos in the modern streaming data warehouse landscape.

©TechTarget, Inc. or its subsidiaries. All rights reserved. TechTarget, and the TechTarget logo, are trademarks or registered trademarks of TechTarget, Inc. and are registered in jurisdictions worldwide. Other product and service names and logos, including for BrightTALK, Xtelligent, and the Enterprise Strategy Group might be trademarks of TechTarget or its subsidiaries. All other trademarks, logos and brand names are the property of their respective owners.

Information contained in this publication has been obtained by sources TechTarget considers to be reliable but is not warranted by TechTarget. This publication may contain opinions of TechTarget, which are subject to change. This publication may include forecasts, projections, and other predictive statements that represent TechTarget's assumptions and expectations in light of currently available information. These forecasts are based on industry trends and involve variables and uncertainties. Consequently, TechTarget makes no warranty as to the accuracy of specific forecasts, projections or predictive statements contained herein.

Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of TechTarget, is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact Client Relations at cr@esg-global.com.

About Enterprise Strategy Group

TechTarget's Enterprise Strategy Group provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

 contact@esg-global.com

 www.esg-global.com