data virtuality

Questions & Answers

How data fabric enables a Broad spectrum of modern use cases

The world of data is evolving rapidly with three main drivers:

- Ever increasing complexity of the data landscape
- Importance of agility
- Growing number of more diverse use cases

In most of the cases, data comes from multiple places - two data warehouses, a data lake, an ODS (Operational Data Store) and five APIs. And to integrate all this data efficiently, companies make use of different ways, for example ETL, ELT, CDC or streaming. This is the new "digital" reality. On top of that, other important aspects such as data quality, data governance, metadata management or data lineage need to be considered.

Read in this Q&A Paper written by Noel Yuhanna, VP and Principal Analyst at Forrester how data fabric provides a framework for all this.

Q   What is a Data Fabric?

Data fabric is a modern data management platform that helps accelerate new and emerging business use cases, such as customer 360, customer intelligence, fraud detection, and advanced analytics, by enabling a semantic layer across data silos. It automates the ingestion, curation, transformation, governance, and integration across disparate data in real-time and near real-time. Data fabric includes data management capabilities, such as data catalogs, data preparation, data discovery, data modeling, data security, data pipelines, and data governance, all working together from an end-to-end perspective.

Q   Why is the Data Fabric architecture gaining so much traction?

The key reason for data fabric architecture's strong momentum is that most organizations are struggling to manage the growing silos of data across hybrid and multicloud environments. There is an increased demand for real-time insights and faster time-to-value for new business use cases, resulting in a lack of data consistency across applications and insights. Forrester finds that organizations that invest in data fabric respond more quickly to business needs and competitive threats, grow faster than their competitors, and deliver innovative products and services. Architects usually design the data fabric architecture and are often deployed by a team comprising data engineers and architects and infrastructure, data security, data management, and operational personnel.

Q | What are the key components of this architecture?

Data fabric comprises several data management components that work together to deliver an agile, integrated, secure, and scalable data platform. The six core data fabric layers include:

- **Data management layer.** Data management is a critical end-to-end layer of data fabric that spans the other five data layers. Its components include data security, data governance, metadata, policies, data catalogs, search, data quality, and data lineage. This layer ensures the reliability, security, integrity, integration, and governance of data.

- **Data ingestion and streaming layer.** Data ingestion deals with getting the data sources connected, ingested, streamed, and processed to support the data fabric use cases. It can stream data from devices, sensors, logs, click-streams, databases, applications, and various cloud sources.

- **Data processing and persistence layer.** This layer leverages data lakes, data hubs, data warehouses, and other data components such as data pipelines to process and persist data for use by the fabric. The data fabric can push down code and processing to various data repositories.

- **Data orchestration layer.** Data orchestration is a crucial layer of data fabric that transforms, integrates, and cleans data to support various use cases in real time or near real time.

- **Data discovery layer.** This layer automates the discovery of new internal or external data sources and presents them as a new data asset for consumption by business users. Dynamic discovery includes several components, such as data modeling, data preparation, curation, and virtualization, to deliver a flexible data platform.

- **Data access layer.** This layer provides access to data and insights and includes data caching in-memory to ensure the performance of critical and repeatable insights. You can also embed fabric components in analytical solutions, tools, and dashboards.

Q | Who should consider building a Data Fabric and why?

Data fabric benefits all organizations that have data silo issues. With a data fabric architecture, organizations can support new business use cases with a low- and no-code approach, accelerating time-to-value to achieve competitive advantage. Data fabric provides various technology and business benefits, including:

- **A 360-degree view of customer data.** Data fabric can ingest, transform, and integrate data from multiple data sources to support various customer-facing applications. With a data fabric, you can deliver a more personalized and contextualized 360-degree customer experience by integrating data from multiple sources, including CRM, enterprise resource planning, data warehouses, social media, SaaS (software-as-a-service) applications, clickstreams, and log streams, quickly and with minimal effort.

- **A real-time data platform for applications, insights, and analytics.** Data fabric delivers a real-time data platform that any data consumer can leverage. It allows applications, processes, dashboards, tools, and analytics to securely access trusted business data in real time, regardless of where the data is physically residing. Data fabric enables a real-time common data platform that delivers trusted and consistent data across applications and analytics, creating new business opportunities.

- **Self-service and collaboration among employees, users, and partners.** Traditional databases and big data platforms were mostly used by developers, architects, and data scientists, primarily because of their complexity. Data fabric changes this by delivering self-service and a high degree of automation, enabling even business users and non-technical users to efficiently leverage data to support the new generation of business dashboards, reports, and advanced insights with minimal effort.

Q If you were a data architect and had to realize the Data Fabric architecture, how would you approach this?

Based on our research, we find that data architects are critical in ensuring success in a data fabric initiative. They understand the semantics essential in integrating, transforming, and orchestrating data across data silos to deliver trusted and consistent data for applications and insights. To realize the data fabric architecture, data architects should:

- Build a data fabric team. To succeed with a data fabric strategy, leverage experts from IT and business. Create a team that includes data engineers, architects, developers, data security pros, data stewards, and business analysts.

- Start with a few data sources. Create a strong foundation. Start with a few data sources, adding new sources over time. Leverage data fabric's adaptive intelligence to discover and ingest new sources to accelerate deployment.

- Enforce data security and governance controls from the start. Data governance and security should be a top priority, especially when dealing with sensitive data. Leverage data fabric's centralized access control and governance to regulate data.

- Expect some level of customization and integration efforts. Specific large and complex data fabric deployment will require customization. Look for solutions that offer flexibility to add additional capabilities over time.

- Use caching to support faster processing and real-time access. Consider data caching within the fabric to deliver shorter low-latency response times, especially for repeatable insights, data applications, and data engineering use cases.

Q How the Data Virtuality Platform enables a data fabric architecture?
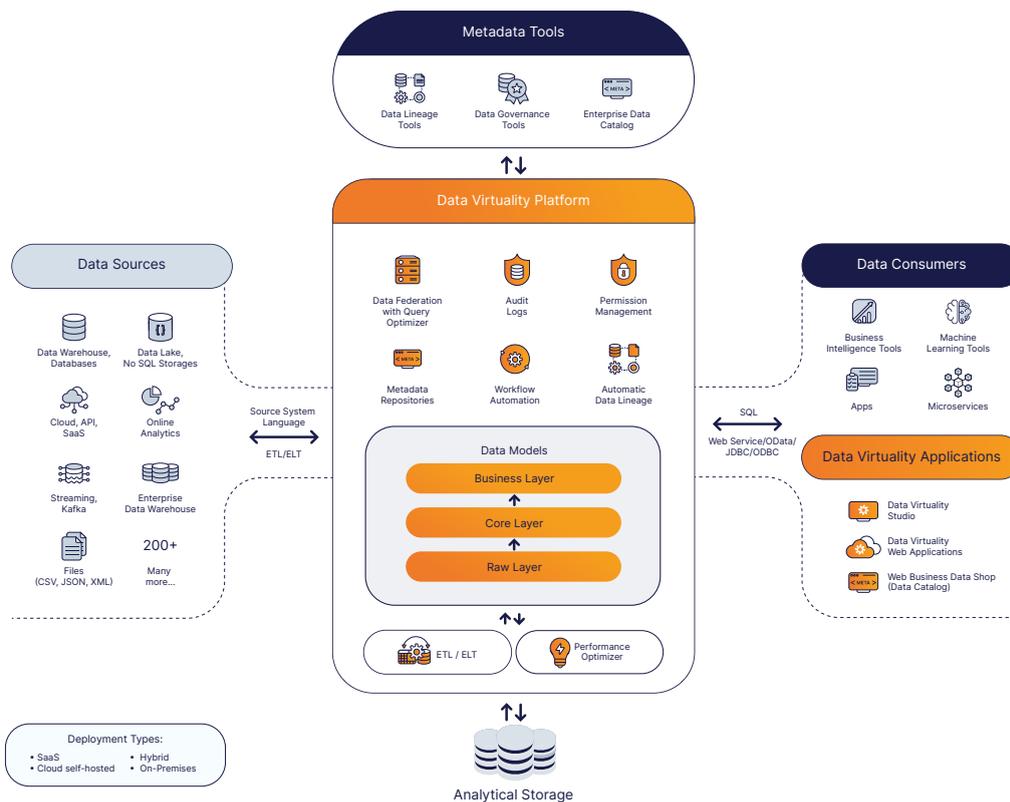
By combining the two technologies, data virtualization and automated ETL, and putting a uniform metadata layer on top of them, the Data Virtuality Platform enables exactly the dynamic and holistic data management architecture needed to transform the data fabric from a concept to reality.

While a data fabric can deliver many use cases, such as real-time analytics, IoT (internet-of-things) analytics, customer intelligence, and fraud detection, two-thirds of all deployments revolve around the customer 360 initiative. Besides, data fabric is also seen accelerating various workloads in healthcare, financial services, intelligence, manufacturing, and retail, offering a next-generation data platform. Data fabric's ability to integrate multiple customer data sources, such as click-streams, log files, and CRM, helps organizations quickly deliver a comprehensive, 360-degree view of the customer. With customer 360, we find organizations leveraging the data fabric to support various up-/cross-sell, targeted marketing, customer intelligence, customer churn analysis, and customer experience initiatives. Personalized customer experiences are becoming more contextual through new data sources such as GPS location, consumer feedback, social data, and interactions with other digital experiences. Data fabric delivers a contextual experience, providing immediate assimilation and processing of customer data. Overall, the use cases are many and will further expand in the coming years.

Noel Yuhanna
VP, Principal Analyst
FORRESTER

Dr. Nick Golovin
Founder and CEO
data virtuality

● About Data Virtuality

● Founded:
2012 by Nick Golovin (PhD) in Leipzig, Germany after 8 years of research

● Offices:
Munich, San Francisco, Leipzig

● Solutions:
Data Virtuality Platform SaaS
Data Virtuality Platform On-Premises
Data Virtuality Pipes Professional
Data Virtuality Pipes

● Acknowledgements:
Honorable Mention in 2022 Gartner Magic Quadrant for Data Integration Tools

● Awards:
Most Innovative Data Management Provider 2022, 2021 and 2019 (A-Team Insights)
2020 and 2019 Deloitte Technology Fast 50