Data Fabric

Idea In Short

Data Fabric is an end-to-end data integration and management solution that helps organizations effectively manage their data. Data Fabric focuses on capabilities that unify diverse and distributed data assets over a unified data management platform architecture. As a result, Data Fabric combines the end-to-end data management processes delivering scale, scope and cost efficiencies.

Data Fabric is a recent innovation in Enterprise data management [EDM] to simplify data access. Historically, Data Access was complicated due to the wide variety of enterprise applications, databases, data models, formats typically found in an enterprise. Data Fabric is an implementation of Data Virtualization principles. Data Fabric helps organizations effectively manage their data through end-to-end data integration and Management. It comprises of an architecture and software system that enables unified data collection, databases and database architectures within an enterprise to provide a unified view of its data assets. Consequently, Data Fabric enables consistent user experience and real-time accessibility. According to Gartner, Data Fabric is:

> a design concept that serves as an integrated layer [fabric] of data and connecting processes. A Data Fabric utilizes continuous analytics over existing, discoverable and inferenced metadata assets to support the design, deployment and utilization of integrated and reusable data across all environments, including hybrid and multi-cloud platforms.

Data Fabric abstracts applications, platforms, and data storage locations. It can be restricted to a single application or extended across the enterprise to collect data across the organization. As a result, Data Fabric removes data access hurdles and enables data sharing within and outside the organization. Usually, Data Fabric comprises of:

- Architecture
- Data Management tools
- Data Integration software, and
- Shared data

What is a Data Fabric?

A Data Fabric is an architecture and set of data services that provide consistent capabilities across the organization. It is a powerful architecture that standardizes data management practices and processes across multiple environments, such as cloud, on premise, and edge devices.

Need for Data Fabric

The volume of data that organizations collect is growing exponentially. These data come in different formats and structures. Managing a variety of data and making them available for use is challenging. Today, most organizations tackle this challenge in silos through point solutions and processes, which result in different ways of managing the data throughout one organization. As a result, making such data available across the organization becomes nearly impossible. As a result, the data remains unused and the organization fails to derive the intended value from data. Lack of comprehensive data access and use results in escalating costs, poor returns on investments in the underlying data infrastructure, reduced data availability to deliver critical business insights, and lower employee productivity. Organizations that aspire to leverage data as a strategic asset and evolve into data-driven entities require a holistic approach of their data assets. However, they should overcome the challenges of time, locations, regulatory requirements, application types, etc. before they can derive value from data. In general, data should be accessible to data consumers who need it, when they need it. Data should not be stashed away in databases or stored piecemeal in a range of locations. Similarly, Data access should not be cumbersome due to stringent firewall rules. Organizations require secure, efficient, unified, and future-proof data environments to thrive in a competitive marketplace. A Data Fabric provides this. Most organizations can efficiently collect data from multiple data sources. However, they face significant hurdles in integrating, processing, curating, transforming, and integrating data with other sources to make data usable. Traditional data integration practices of connecting disparate enterprise systems and databases is no longer meeting new business demands of real-time connectivity, self-service and automation, to name a few use cases. Integrating

the diverse enterprise systems and databases is technically challenging. Furthermore, such integrations are time consuming and expensive, resulting in longer time-to-value and reduced enterprise agility.

Business value drivers

Data Fabric helps break down data silos. By presenting a unified interface, Data Fabric increases employees' productivity by delivering data and information from a single location. Data Fabric affords enhanced visibility into data and insights, while governing data access and control, data protection, regulatory compliance, and security. Furthermore, Data Fabric allows implementing standardized Data Quality Management processes that increases data users' trust in data and helps them make data-driven decisions. Data Fabric integrated an organization's data assets and deliver a holistic view of suppliers, customers, partners, products and financials, which gives the organization a competitive edge.

IT value drivers

The number of applications a typical enterprise uses has significantly increased. Usually, enterprise applications and databases are associated with business functions or user groups. Data Fabric improves integrating disparate applications and databases as well as their portability and availability. As organizations add applications to their inventory, data silos result. Data Fabric helps improve visibility into the full portfolio of enterprise applications and data to improve operational efficiency and empower data consumers. Furthermore, Data Fabric helps overcome data accessibility barriers across locations and regulatory regimes through governed data provisioning and access. In addition, Data Fabric simplifies software development and application deployment by harmonizing data formats and Application Programming Interfaces [APIs]. Data Fabric also helps optimize specific applications compromising data accessibility to other applications or users. It can also help unify data that is already siloed.

How Data Fabric transforms traditional approaches

Today, many organizations use Data Lakes and Data Warehouses to manage their data assets. However, on close scrutiny, these approaches are technology-centric, rather than user-centric. Organizations deploy Data Lakes and Warehouses to collect and store large volumes of data for future use. To a limited extent, these systems deliver the critical insights

that business requires today. Moreover, these systems rely on traditional technologies and data models, so getting a unified view of the data to derive superior business insights is still challenging. Such systems and data processing techniques are not only slow and expensive, but also restrict the insights that users could derive from data, due to limited technical capabilities. With the growing data volumes and time criticality under which which business decision makers in an organization operate, latencies in data access and processing are unacceptable. Data Fabric overcomes these challenges and enables decision-makers with real-time insights by overcoming the limitations that plagued traditional data processing systems.

Implementing Data Fabric

Data Fabric implementation starts with the traditional Online Transaction Processing [OLTP] systems. OLTP systems insert, update or delete detailed information about every transaction in a relational database. Then, the data processing pipelines in the Data Fabric structure, clean and store the data for further consumption. At any point in the Data Fabric, a data user can extract the raw data and manipulate it to derive business insights. As a result, Data Fabric helps organizations leverage their data assets to grow, adapt, and improve. Successful Data Fabric implementation entails:

- **Applications and services**: This is the foundational infrastructure for data acquisition. Organizations should develop applications, services and user interfaces to enable users access and consume data
- **Platforms and integration**: This is the entire data ecosystem for data acquisition, management and storage. Organizations should ensure that no data is lost throughout its journey from source to destination
- **Security**: Data collected from various sources should be managed with proper security policies and comply with the applicable regulatory requirements
- **Lifecycle**: Data should be stored in an accessible and efficient manner. The underlying infrastructure should scale when required. Data should be archived or disposed with the appropriate lifecycle policies when they are no longer required
- **Access**: Data should be accessible from any point in the organization via pre-defined endpoints to allow users derive real-time insights

Benefits of Data Fabric

Data Fabric is ideal for organizations that are geographically diverse, have multiple data sources, and face complex data issues or use cases. Remember, a Data Fabric is not a quick answer to integrate and process your data. For that, you can turn to data virtualization. With continued advancements in hardware capabilities, globalization is expanding into previously unconnected regions. With connectivity speeds rocketing in pace, organizations can be overwhelmed by data from devices and services. While data has been used for quite some time for insights, Data Fabric provides a solution, which encompasses:

- An agile model that allows changes to systems, adapts and adjusts as needed, and works across all operating and storage systems
- Scalable with minimum interference, no investment in massively expensive hardware or highly trained and expensive staff
- Providing maximum integrity and complying to regulations, while maintaining accessibility and real-time flow of information

The massive amounts of data that businesses can access needs to be exploited to derive unique insights. Areas including forecasting, sales and supply chain optimization, marketing, and consumer behavior give the organization a competitive edge and data leadership in its field. Real-time insight derivation can make the organization a cut above the rest. ...

Summary

Data Fabric is not only a combination of traditional and contemporary technologies, but an architectural design concept that transforms human and machine workloads. Data Fabric also optimizes Data Management by automating repetitive tasks, such as data profiling, quality management, data discovery, managing metadata, etc. As a result, organizations can invest in infrastructure solutions that align with their business requirements—without concerns surrounding data service levels, access, and security.