Data and Analytics Essentials: Data Fabric

Published 13 July 2021 - ID G00745844 - 5 min read

Ehtisham Zaidi

Initiatives: Data Management Solutions

Data fabric is an emerging data management design that enables augmented data integration and sharing across heterogeneous data sources. This presentation will assist data and analytics leaders in raising awareness among stakeholders and in progressively building a comprehensive data fabric architecture.

The Challenge

Most organizations struggle with a comprehensive data management design that allows them to optimally connect/collect, integrate, and deliver distributed data from heterogeneous data sources and applications for their data and analytics (D&A) use-case demands. The number of data and application silos have surged in the last decade, while the number of skilled personnel in data teams has either stayed constant or has even dropped. The result: The time deficit between when the request for integrated data was raised to when this request is fulfilled is now at an all-time high. The growing popularity of cloud for data management has only increased the complexity of creating and maintaining a consistent data management architecture that caters to all existing and upcoming requests for data integration.

The Promise of Data Fabrics

Data fabrics have emerged as an increasingly popular design choice to simplify an organization's data integration infrastructure and create a scalable architecture that reduces the technical debt seen in most D&A teams due to the rising integration challenge. The intent is to not only address cost or create elegant design, but to gradually introduce data utilization, context analysis and alignment.
A data fabric is an emerging data management design concept for attaining flexible, reusable and augmented data integration pipelines and services in support of various operational and analytics use cases delivered across multiple deployment and orchestration platforms. Data fabrics support a combination of different data integration styles, and utilize active metadata, knowledge graphs, semantics and machine learning to augment data integration design and delivery.

The promise of the design concept is that once implemented, data fabrics would significantly eliminate manual data integration tasks and augment (and in some cases completely automate) data integration design and delivery. However, it is important to note that data fabrics are still an emergent design concept (see Hype Cycle for Data Management, 2020). No single vendor currently delivers all the mature components (in an integrated manner) that are needed to stitch together the data fabric (which means that data fabrics cannot be yet purchased form a single vendor). Data fabric also consists of a mix of mature and less mature technology components. Organizations will therefore have to carefully mix and match composable technology components as their use cases evolve, and would need to know where and how to begin with their data fabric design planning and implementation.

The Challenges Organizations Face When Planning For and Implementing Data Fabrics

Beyond just the technological components of the data fabric, a significant number of Gartner client inquiries have revealed that data and analytics leaders looking to implement data fabrics often struggle and stumble at the very early stage of their journey. Their challenges are as follows:

- D&A leaders struggle to define the data fabrics in a way that is well-understood by supporting and participating business teams and key stakeholders. They often find it difficult to connect the promise of a data fabric with defined business outcomes, and as a result, they don’t get stakeholder support and buy-in.

- The next biggest challenge that they face is that they don’t understand how to start with creating a data fabric and the main steps that they need to follow to create a comprehensive data fabric. D&A leaders tell us that they need support in creating the right teams with the necessary roles and skills to enable the data fabric.
And finally, D&A leaders realize that the data fabric is a technology-enabled design, and that they would need the right mix of tools/platforms to assist them. For this, they need to understand which categories of tools and vendors to investigate and adopt.

How This Presentation Helps

This presentation is a direct response to the above stated challenges. This presentation is carefully divided into three logical segments to address the main challenges around data fabrics:

1. The first segment goes into detail about the what and the why. It defines the data fabric in a way that business teams and key stakeholders can understand its need, its benefits and why it is essential for them to adopt the data fabric design. This section lists out the main reasons that make the data fabric a must-have design pattern for any organization struggling with data silos (that need to be integrated). It also explains how data fabric benefits organizations that need automation in their data management tasks to support overburdened data integration and data engineering teams.

2. The second segment of the presentation addresses the common question, “How do I stitch together the data fabric, and what are the components needed to do so?” This part of the presentation helps D&A leaders in their strategic roadmap assessment to understand the key composable components and technology parts needed to deliver their data fabric. It also touches on the discipline (in this case, DataOps) and the team structures and skills needed to make the data fabric more usable. This section next guides D&A leaders through three paths to production for their SLA requirements. These paths to production help them get started (and get some important wins under their belt to keep the business interested) for common use cases. They can then progress to more advanced paths that support use cases that need knowledge graphs for highly interconnected datasets. Finally, the data fabric can move into the final frontier, which is the automation path that supports automation of various tedious, error-prone and repetitive data management tasks by activating metadata and analyzing it.
3. The final part of this presentation will allow data and analytics leaders to navigate the complex vendor landscape to select mature technology components from established vendor markets. This will help them to understand which types of technology choices are needed and how they can be assembled. Finally, the presentation provides some cautions (based on the mistakes that Gartner has noticed early stage adopters make) and a representative list of vendors that can be investigated for implementing data fabrics.

These presentation slides will be most useful to data and analytics leaders who are working to educate their stakeholders. Gartner clients working to get buy-in for their business case to modernize the data management infrastructure can use this simple presentation format to help executives understand the role and importance of each of these concepts.

Please review the Notes Section for each slide, as it explains and builds on the ideas represented in each slide in detail.

These presentation slides will be most useful to data and analytics leaders who are working to educate their stakeholders. Gartner clients working to get buy-in for their business case to modernize the data management infrastructure can use this simple presentation format to help executives understand the role and importance of each of these concepts.

These presentation slides include content delivered at Gartner’s Data and Analytics Summit virtual event held in May 2021.

**Recommended by the Authors**

What Is Data Fabric Design?

Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration

Demystifying the Data Fabric

Emerging Technologies: Data Fabric Is the Future of Data Management

Top Trends in Data and Analytics for 2021: Data Fabric Is the Foundation