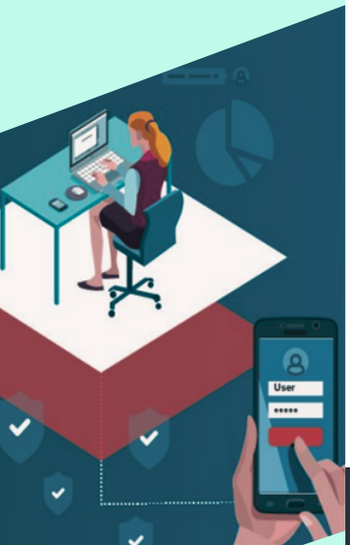
DATA MESH : TRANSACTIONAL VS ANALYTICAL WORKLOAD

Michel Thai
Data Architect Lead

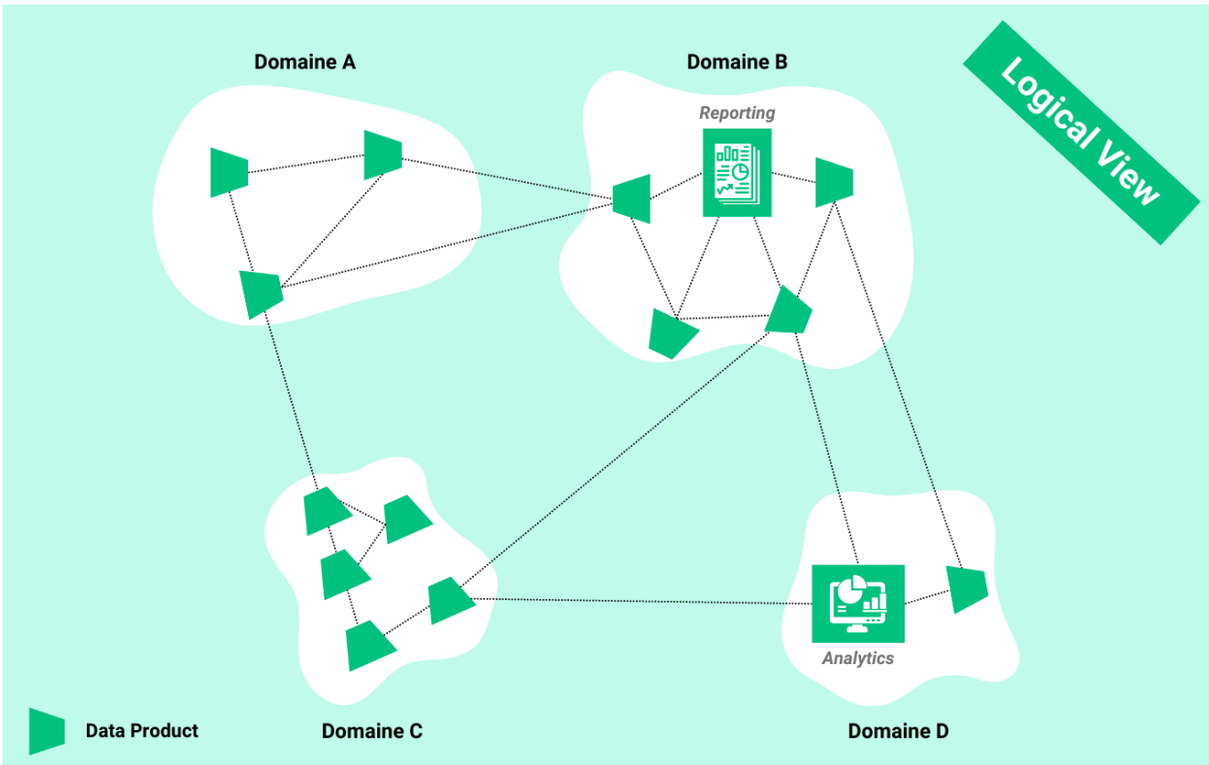
“By understanding these fundamental differences, organizations can design data solutions tailored to their specific needs, thereby maximizing the value of their data within a Data Mesh.”



In the past few decades, the way we handle data has changed a lot, moving from a centralized model to a distributed and scalable approach. One of the new ideas that came out of this is called Data Mesh. This new implementation pattern requires managing both transactional and analytical workloads. This article explores the differences between these two types of workloads within a distributed and scalable approach to data management.

1 Understanding the Data Mesh

Data Mesh is a new way of organizing data that different teams within an organization have their own data to control, making it easier to manage and use effectively. It's an approach to managing decentralized and distributed data, organized by business domain, where data is treated as products used by all consumers within the organization.



With a Data Mesh architecture, this new paradigm requires managing data in motion, ensuring real-time processing of streaming data, while also effectively handling data at rest in various storage systems, allowing for comprehensive data analysis and insights.
In simple terms, Data Mesh combines existing paradigms such as Domain-driven Design, Data Marts, Microservices, Event Streaming, and Data Virtualization.

2 Transactional Workloads

Transactional workloads focus on real-time data management. They are primarily associated with CRUD operations of data. Common examples of transactional workloads include financial transactions, real-time data updates, Online Transaction Processing (OLTP) database operations, and more.

Key characteristics of transactional workloads within a Data Mesh include:

1. **High Concurrency: Transactional workloads often generate a large number of concurrent queries, requiring efficient lock management and rapid response to requests.**
2. **Low Latency: Responses need to be nearly instantaneous as users expect real-time results.**
3. **Frequent Data Changes: Data is constantly updated, added, or deleted, requiring careful data consistency management.**
4. **Emphasis on Data Integrity: Transactional workloads prioritize data integrity and transaction management.**
5. **Data in Motion: Data is processed and correlated continuously while new events are fed into the platform. Business logic and queries execute in real-time.**

3 Analytical Workloads

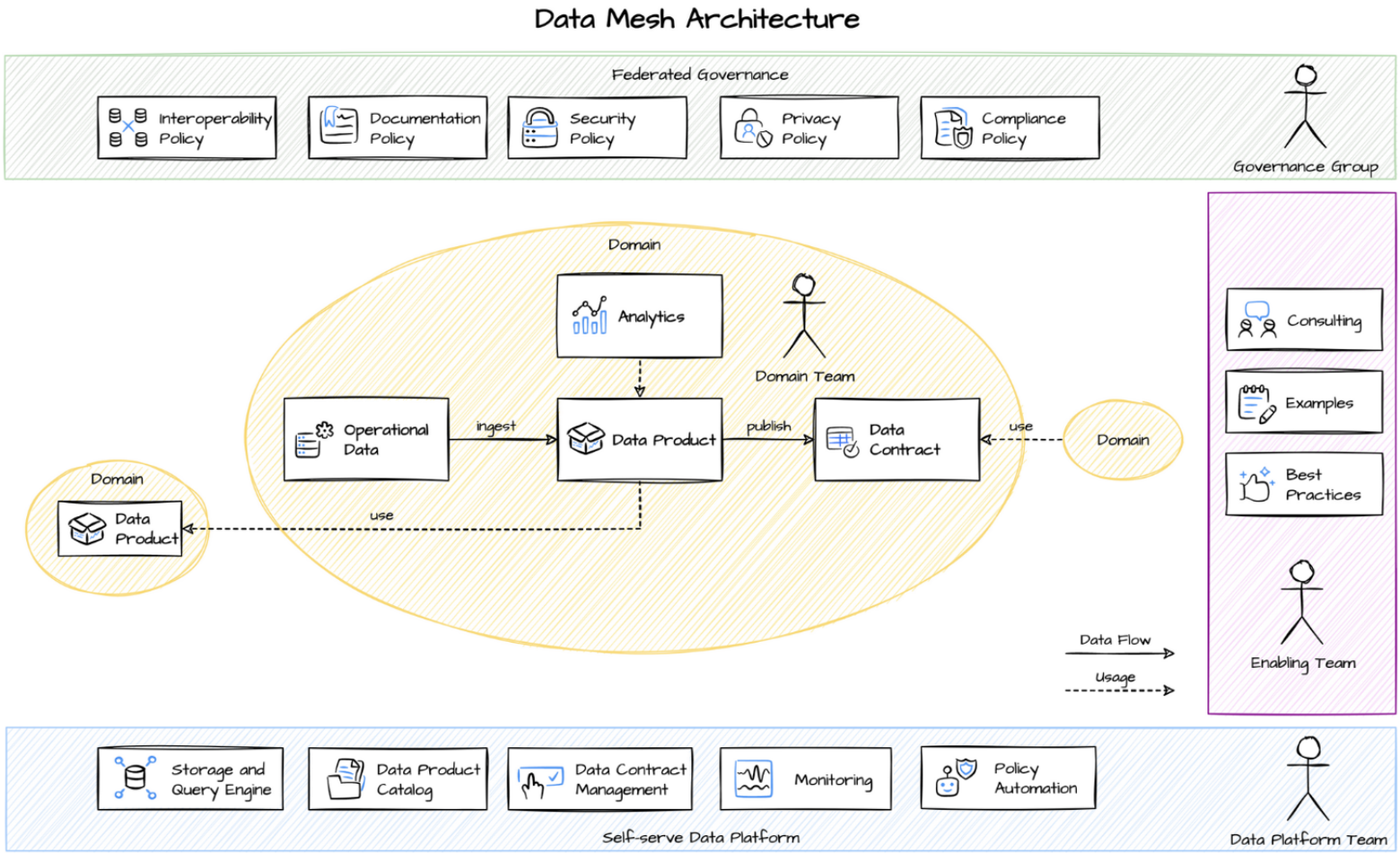
Analytical workloads, on the other hand, focus on large-scale data analysis. They typically involve complex querying, data aggregation, report generation, and data exploration. Common examples include data warehouses, predictive analytics, and big data analysis.

Key characteristics of analytical workloads within a Data Mesh include:

1. **Processing Massive Data Volumes: Analytical workloads require the capacity to handle large amounts of data, sometimes in the petabytes.**
2. **Query Complexity: Analytical queries can be complex and involve operations such as joins, filtering, aggregation, and more.**
3. **Tolerant Latency: Unlike transactional workloads, analytical workloads can tolerate higher latency as they do not require instant responses.**
4. **Ad-hoc queries: The primary goal of analytical workloads is to discover insights from data, which may involve intensive calculations, many of which will only be run once.**
5. **Data at Rest: Data is ingested and stored in a storage system (database, data warehouse, data lake). Business logic and queries execute against the storage.**

4 Conclusion

As part of the Data Mesh architecture, it is important to understand the workloads of each domain. Indeed, transactional domains are focused on real-time transaction processing operations, while analytical domains are used for processing data operations in delayed time or for large-scale data analysis.



SOURCES *datamesh-architecture.com*

Fastcube is a data consultancy and expertise firm, encompassing data integration and transformation, through to analysis and decision-making. We are committed to turning data into a tangible driver of value creation and enhancing our clients' business strategy.

To learn more, please visit: fastcube.fr | Email: contact@fastcube.fr