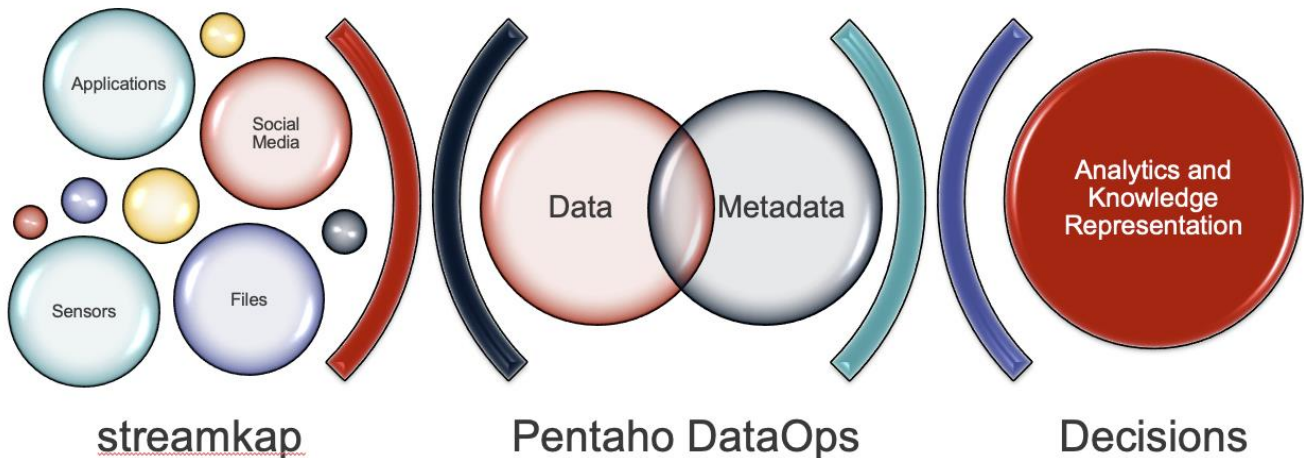




Streaming into the Future: Harnessing the Power of Real-time Data

Streaming Data: Pentaho DataOps Platform & Streamkap

Streaming data is crucial for the Federal Government because it enables real-time high-velocity low-latency applications monitoring, data analysis, and decision-making across various domains, including national security, public safety, public health, disaster management, and resource allocation. Integrating streaming data with existing systems and batch data is often challenging and the data can be generated in various formats, such as text, images, audio, and sensor data in different time frequencies.

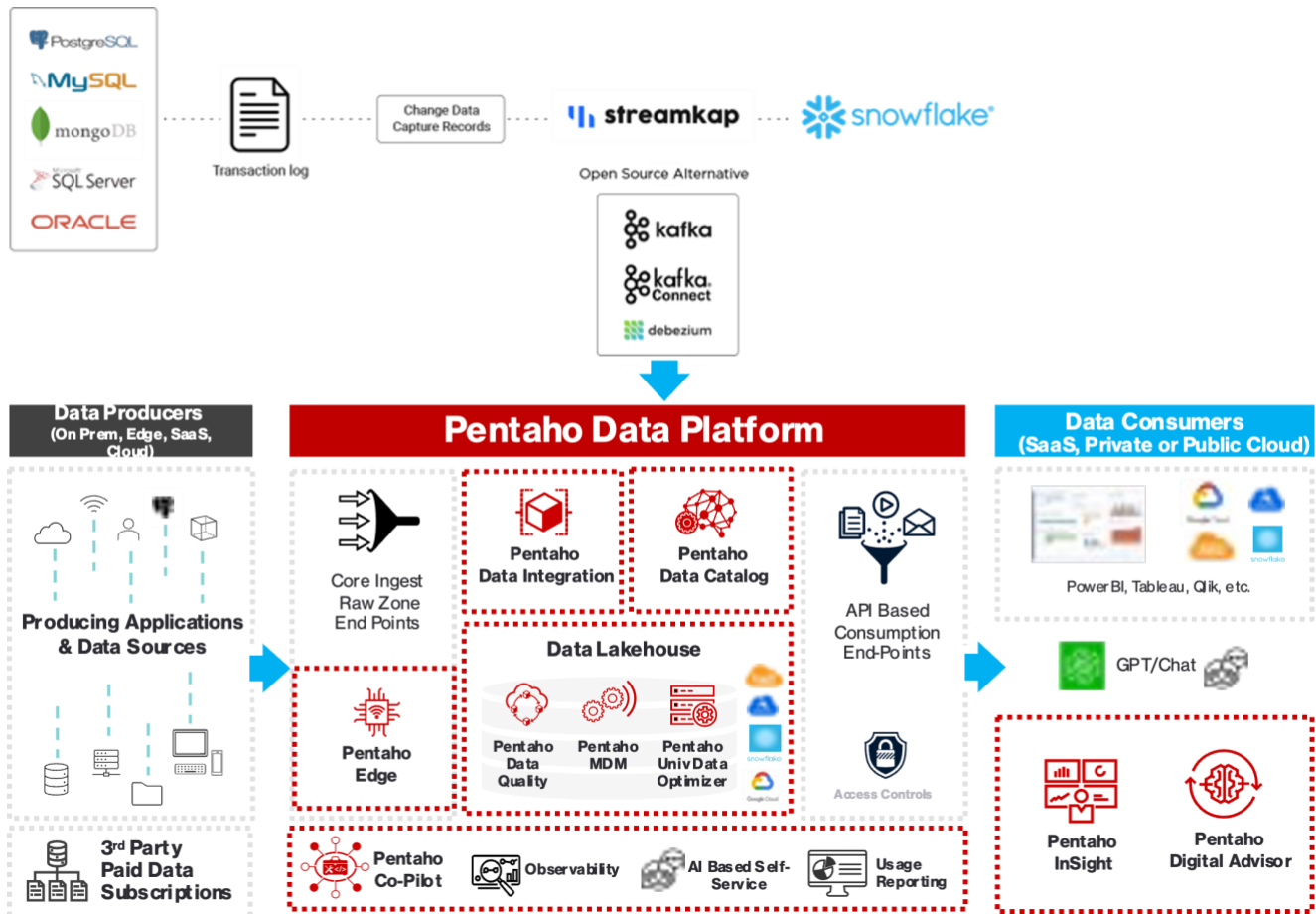


The fundamental tenets of streaming data that are different from typical data operations and require special handling via solutions such as Pentaho DataOps Platform and Streamkap are –

- **Real-Time** or near-real-time, enabling change data capture irrespective of the temporal frequency
- **High Velocity** where data flows continuously and at a high speed
- **Low Latency** with minimal tolerance for delay in data processing and analysis
- **Infinite (Big) Data** necessitates scalable and efficient storage and processing solutions
- **Event-driven** with data generated in response to specific triggers, different from batch processing
- **Unbounded Data Sets** without a defined endpoint
- **Immutability** for temporal consistency and ease of auditability
- **Disparate data formats**, including text, images, audio, sensor data, and more

Streamkap utilises the database source transaction log file using change data capture methods for data replication to read the data in real-time without any load being put upon the database. The data is then streamed into the data warehouses or data lakes using real-time inserts with an option to transform structured and unstructured data and extract metadata using Pentaho DataOps Platform for richer evolved analytics.

Streaming Data: Pentaho DataOps Platform & Streamkap



Both public sector and industry are capturing, storing, and analyzing more data than ever before. This presents non-trivial challenges to overcome, as the data growth rates are exploding, and majority of the new data is complex and unstructured. A globally connected and information-driven decision-dominant warfighting environment or a citizen service platform needs this strong data-enriched foundation to remain functional, even in a disconnected, denied, intermittent and/or limited (DDIL) environment. It is critical to overcome these constraints as a key design component of the next generation systems for decision dominance irrespective of the task being tactical, operational, and/or strategic.

The open architecture foundation of this combined solution enables ease of integration with the existing data infrastructure irrespective of modality (structured tabular data, character separated values (csv), text, audio, video, image, log extracts, social media, RSS Feeds, emails, application APIs, and more), language, media format and geography.

This architecture allows data scientists, analysts, and users to reduce the time it takes to process captured data by accelerating metadata discovery and categorizing data. It permits data curators and stewards to manage and control access to sensitive data. It additionally enables IT administrators to eliminate unintended redundant data and configure data tiering (including dynamic and virtualized).

About Hitachi Vantara Federal

Hitachi Vantara Federal is the most reliable data solutions provider with the most complete set of hybrid cloud and data fabric solutions in the Federal IT market. We're a collaborative, full-service company empowering data-driven insight with a deep bench of integrated partners advancing Federal customer missions regardless of their data maturity levels.

We're more than just a leader in digital infrastructure. We're constantly evolving our data management, analytics, and Internet of Things (IoT) solutions to service Federal leaders' mission needs today, and in the future. All of our Federal customers get the platinum treatment as a standard offering. From civilian agencies to the Department of Defense (DoD) and Intelligence Community (IC), we're known for our agility, dependability, and excellence.

Balancing a century of tradition with constant innovation, we stand for agility, dependability, and excellence—guaranteed. We have a TS facility clearance and we are familiar and adhere to the NIST AI Risk Management Framework, DoD Ethical Principles for Artificial Intelligence and the fast-emerging Responsible and Trustworthy AI Guidelines from various Federal agencies as part of our solution implementations. Our analytics products are part of the DoD CDAO Tradewinds Marketplace as well as part of DIA Needipedia (ID# = 06APR2023_01MAR02023_HITACHI-VANTARA_6-1-3_ENG_SUM).

Get in Touch

WEB: hitachivantarafederal.com/about/contact-us/

EMAIL: info@hitachivantarafederal.com

PHONE: 1-703-787-2900

About Streamkap

The Streamkap team has 20+ years of experience building real-time and batch data replication solutions. We built tools for the modern data stack and saw how it easy it is now to build batch data pipelines. We weren't satisfied with the complexity of streaming data systems and the months of complex DevOps and engineering required for production grade streaming pipelines and knew we could bring a simple serverless experience to building data pipelines for streaming.

Get in Touch

WEB: <https://streamkap.com/>

CONTACT: <https://streamkap.com/contact-us>