WhereScape® Automation with Streaming

Empower your business with real-time data

WhereScape Automation with Streaming Benefits

- Leverage Internet of Things (IoT), social media and other streaming data sources to identify trends and changing conditions
- Automate data infrastructure development, from data warehouses and vaults to data lakes and marts, to support real-time analytics needs
- Easily ingest and process real-time data streams into an existing data landscape, and enrich with batch-based business data when needed

- Quickly understand new streaming data flows through automated lineage and documentation
- Enable streaming data analysis for data science, anomaly detection, machine learning, operations optimization and other use cases
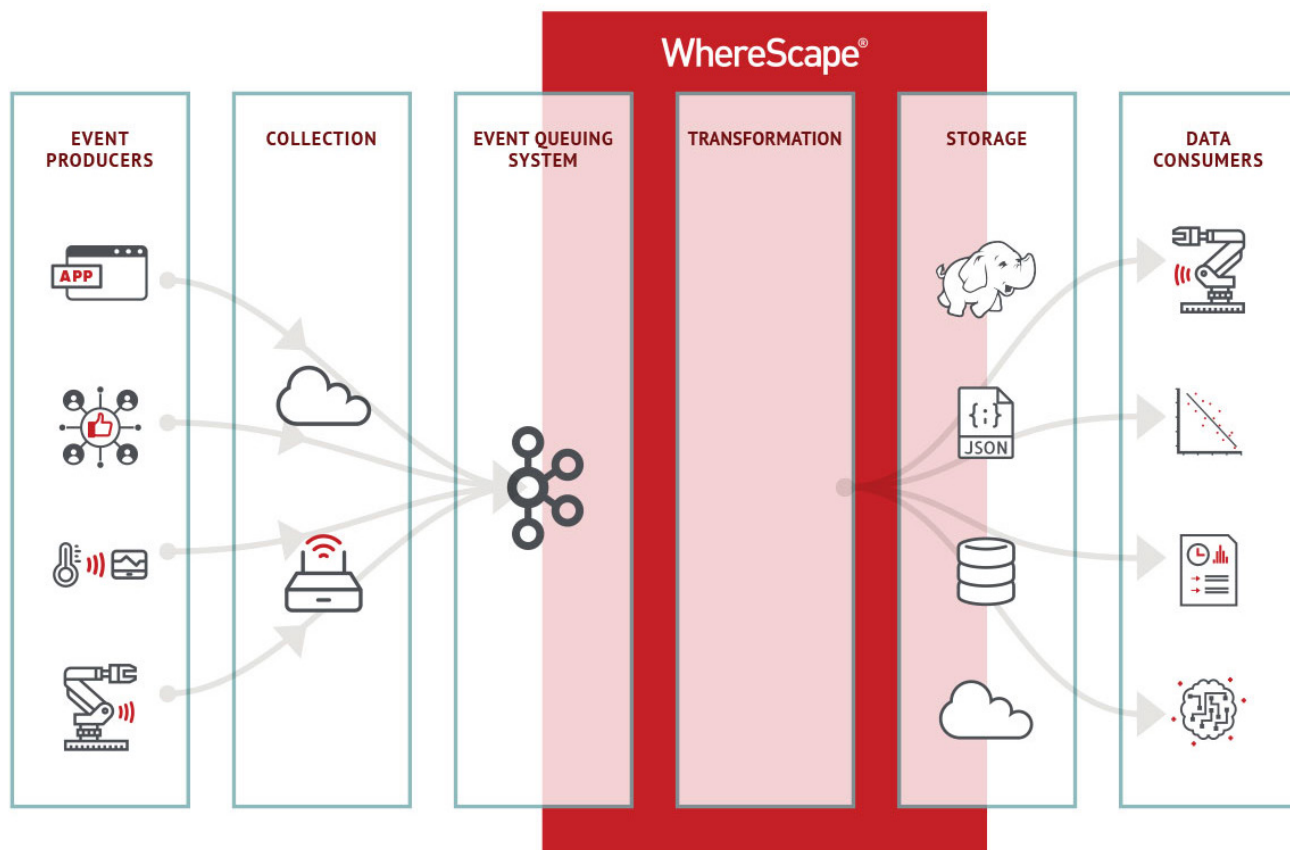- Use real-time data at the speed of your business, leveraging data now or storing it for future use

Analysis of streaming, or real-time, data—sensor data, log files, retail purchases, social media content, or any other continuously generated data—gives companies unprecedented visibility into trends and opportunities as they emerge. The challenge for IT teams is how to incorporate these new data sources and streaming technologies into existing analytic environments to make them quickly and easily accessible to the business.

With WhereScape® automation with Streaming, IT teams can efficiently and cost-effectively leverage real-time streaming data within existing data infrastructure—opening new opportunities for data science, machine learning, anomaly detection and other real-time analytics use cases. Powered by StreamSets' dataflow management technology, WhereScape supports widely used streaming platforms such as Apache Kafka and streaming frameworks such as Apache Spark to minimize the learning curve for your IT teams.

With Streaming, you can easily ingest and process real-time data streams into your company's data infrastructure alongside—or enriched with—traditional, batch-based data sources. Backed by all of the productivity capabilities of WhereScape automation, your IT teams can fast-track data delivery, broaden organizational analytics and adopt emerging data trends and technologies at the speed of the business.

Streaming is an optional add-on available for use with WhereScape® RED and WhereScape® Data Vault Express™ to fast-track the development of data infrastructure that leverages real-time data.

Streaming IoT Integration: Sensor Information Collection

Streaming Features

Ingests real-time data flows
Ingest and automatically assign metadata to real-time Kafka data streams.

Discovers and documents sources
Easily understand available data, format, and structure to speed up data infrastructure integration and minimize project risk.

Enriches incoming streaming data
Deliver enriched data for complete real-time analysis, anomaly detection or long-term persistence.

Processes data streams
Process and transform incoming data streams before persisting for downstream analysis.

Persists to leading data platforms
Persist real-time data into multiple data platforms in parallel.

Creates and maintains documentation
Use automatically created documentation and lineage to easily understand streaming data flows.

Populates Data Vaults
With WhereScape Data Vault Express, use data vault-specific processing, such as hash key generation, in-flight before persisting to a data vault.

Generates Apache Spark™ code
Automatically generate native, optimized Apache Spark™ code and support for Spark stream-based processing.

About WhereScape
WhereScape helps IT organizations of all sizes leverage automation to design, develop, deploy, and operate data infrastructure faster. More than 700 customers worldwide rely on WhereScape automation to eliminate hand-coding and other repetitive, time-intensive aspects of data infrastructure projects to deliver data warehouses, vaults, lakes and marts in days or weeks rather than in months or years. WhereScape has global operations in the USA, UK, Singapore and New Zealand.

WhereScape.com