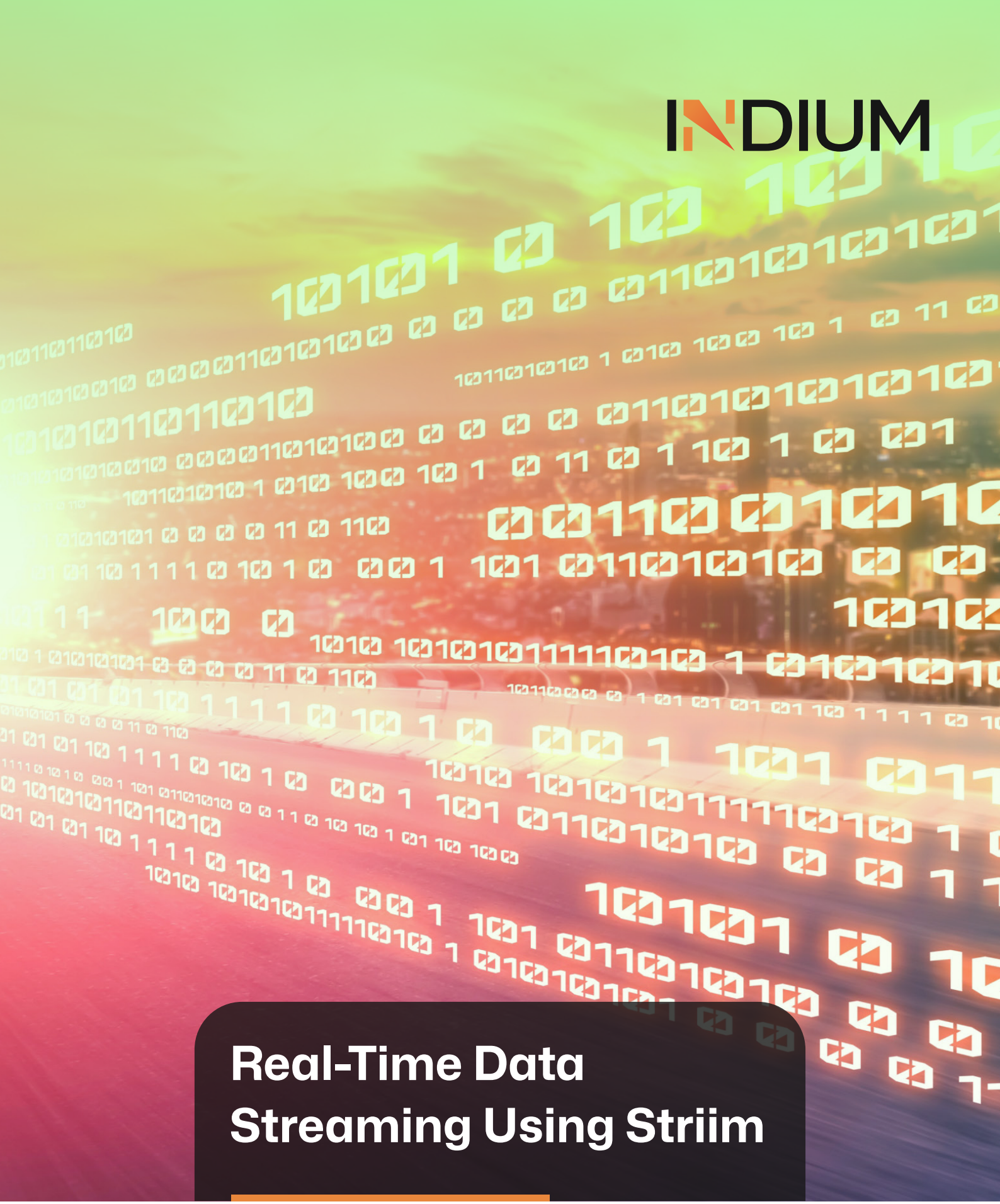


INDIUM

Real-Time Data Streaming Using Striim

A Whitepaper

Real-time data streaming empowers businesses to make an insightful decision as the need arises. Striim provides the platform to enable real-time data integration across platforms and real-time querying, dashboards as well as visualization.

Data Streaming Enabled by Digital Transformation

The adoption of digital transformation technologies such as cloud, Artificial Intelligence (AI), big data analysis and the Internet of Things (IoT) not only improve the e°ciency of business processes but also provide data that can give insights into user behaviour and preferences.

They also enable access to data in real-time in addition to historical data, thereby keeping businesses on the cutting edge, responding to customers in a timely manner. Accessing real-time data can help improve decision making by enabling accurate forecasting.

A report from Marketsandmarkets estimates the streaming analytics market size to grow from USD 10.3 billion in 2019 to USD 35.5 billion by 2024 at a Compound Annual Growth Rate (CAGR) of 28.2% in this period. Streaming analytics leverages real-time data for:

- Analytical and transactional systems
- Extract data from a variety of sources including log files, databases, sensors and messaging systems to provide a comprehensive view of the business
- Enable timely insight for better operational decision making

Need for Streaming Data

Microservice is an architectural style which structures an application as a group of services. The services are loosely coupled and hence it's easy to maintain and they can be deployed independently. Microservices are widely accepted new age architecture used by large enterprises for its agility.

A modern, streaming data architecture such as Striim's provides the foundation to easily integrate all structured, semi-structured and unstructured data, on-premises or in the cloud, in near real-time. Data across Cloud, Big Data and IoT devices are integrated into a next-generation infrastructure instead of getting confined to a single topology. Pre-built integration and wizards-based development accelerate the building of streaming data pipelines on Striim.

Why Striim

Some of the advantages of using Striim include

- Ease of replicating data
- Real-time streaming analytics and data transformations
- Easy to implement and user friendly
- Web/Email Alerts in case of any change in the trends
- Secured data migrations between source and target (many types of encryption available)
- Ease in data recovery in case of failures

Agent-based approach for highly secured databases.

Striim, an end-to-end, enterprise-grade platform o , er s non-intrusive, real-time change data capture capabilities with in-flight data processing and data visualization moves to transforms data with sub-second latency. Using an SQL-like language, it allows for the adoption of future-proof, smart data architecture for accelerated innovation.

Real-time data can be extracted from a wide variety of sources, including:

- Databases
- Log files
- IoT devices
- Message queues

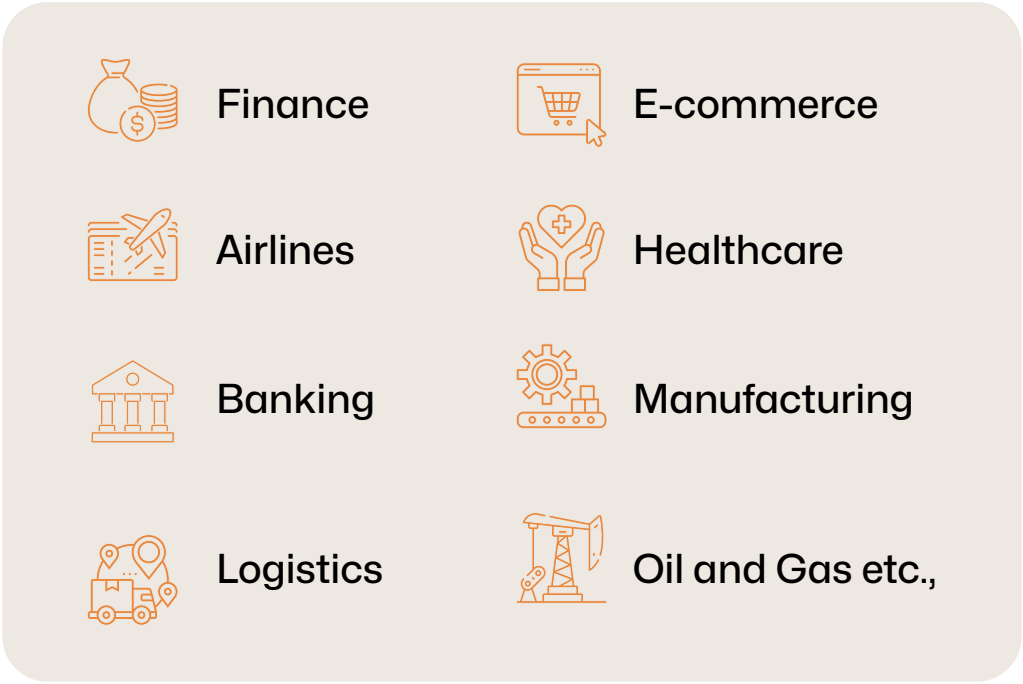
This is possible for di , erent data types such as:

- JSON
- XML
- Delimited
- Binary
- Free text
- Change records

Benefits

Real-time integration and streaming of data enable reporting and analysis in real-time as well. Striim not only makes it easy to move data from on-prem to the cloud but also makes it possible to create real-time dashboards and enable visualization.

Striim based real-time data integration has wide application across industries such as:



Data transformation and data enrichment can be done within the data pipeline and extraction happens through transaction logs. This way, contact with the source database can be avoided and read loads on the source database can be reduced. One of the greatest advantages is that Striim needs no component on source/target servers to be installed.

Requirements

Striim can be installed on any VM / physical instance and the process is very straight-forward. It needs a minimum of 4 core and 16 GB RAM. It can operate on major OS like Linux, Windows & Mac, and requires the JDBC drivers of the respective database provider. Meta Data Repository can be in-built Derby, Oracle or PostgreSQL..

Challenges that Striim Addresses

Often, businesses face the challenge of data migration between heterogeneous databases and across the platform. Striim enables this through non-intrusive change data capture (CDC) from one database to another, No-SQL, file systems, on-prem to clouds and vice-versa. It has an in-built architecture to handle data recovery in case of failures. It also enables data streaming from unconventional sources such as weblog, social media, sensors, etc.

In the case of transactional databases, Striim filters, transforms, aggregates enriches and analyzes the data-in-motion by running queries continuously and delivering it to virtually any target with sub-second latency. It can do batch-processing if required.

Striim Best Practices

Data Movement requires a reliable and scalable model and the data often requires to be transfoemed and enriched through stream processing to add value. This requires constant decision-making regarding matters of architecture and technology at every step during the run time and not just during designing. Some of the best practice for real-time data movement include:

Taking a Streaming-First approach using technology such as change data capture (CDC) and file tailing where the initial collection of all daa is continous and in real-time since you cannot ensure up to date data with batch or microbatch data collection.

Process data for analytics by removing superfluous fields and redundant data, summarization and change detection. Data enrichment by denormalizing IDs in database records or correlating data from multiple sources to create structure in a destination can add value. Performing these tasks in memory at the time of ingestion rather than after it has been stored can reduce storage costs in the target and maximize analytics.

Minimal disk I/O by moving data at scale and with low latency will prevent slowing down the architecture. Use of intermediate topics on a persistent messaging system such as Kafka can achieve this.

Multiple benefits of real-time streaming data include optimizing data flows and minimizing resource usage by processing the same data in different ways and delivered simultaneously and in real-time to different endpoints such as Kafka, Hadoop on-premises and cloud storage.

Avoid custom coding for building data pipelines and streaming data as that will need teams of developers, reduce flexibility and make maintenance a nightmare. The Striim platform lets Business Analysts and Data Scientists work directly with the data using SQL-based queries. This speeds up development and automatically makes it scalable and reliable.

Continuous real-time processing is made possible by incorporating intelligence, looking for anomalies in data formats, volumes, or seasonal characteristics to support reliable mission-critical data flows.

In a Striim platform, these best practice are in-built. To further enhance its efficiency, hybrid cloiud, real-time application and fast data use cases ensure real-time data movement and stream processing.

Indium Approach

Indium is an authorized Striim partner with more than a decade of experience in Big Data and Analytics. It has experience and expertise in the following:

- End to End implementation and training the client
- Setting up Striim node /cluster with HA / Multi-node architecture
- Customer Support for queries or use cases (POCs)
- Professional service for app development and maintenance

Some of the use cases for Indium's real-time streaming on Striim include:

- On-Prem Oracle to Postgres Cloud (For a Columbian Bank)
- Oracle/MSSQL to KAFKA (Health Insurance)
- Real-Time Event generation and Sending Alerts (Airlines USA)
- Oracle / MSSQL server to Azure SQLDB and ADLS gent (For a leading bank in India)
- Alibaba cloud to GCP (PostgreSQL) (For a leading eCommerce company)
- CDC from MSSQL in AWS-RDS to GCP Big Query along with the audit records

Indium measures the success of the projects by ensuring there is no data loss in migrations, there is high throughput in data migration and end to end support throughout the period. Of course, the final measure is customer satisfaction.

If you would like real-time data streaming using Striim for your organization, contact us **www.indium.tech I ©Indium**

INDIUM

About Indium

Indium is an AI-driven digital engineering company that helps enterprises build, scale, and innovate with cutting-edge technology. We specialize in custom solutions, ensuring every engagement is tailored to business needs with a relentless customer-first approach. Our expertise spans Generative AI, Product Engineering, Intelligent Automation, Data & AI, Quality Engineering, and Gaming, delivering high-impact solutions that drive real business impact.

With 5,000+ associates globally, we partner with Fortune 500, Global 2000, and leading technology firms across Financial Services, Healthcare, Manufacturing, Retail, and Technology–driving impact in North America, India, the UK, Singapore, Australia, and Japan to keep businesses ahead in an AI-first world.

USA

Cupertino | Princeton
Toll-free: +1-888-207-5969

INDIA

Chennai | Bengaluru | Mumbai
Hyderabad | Pune

Toll-free: 1800-123-1191

UK

London
Ph: +44 1420 300014

SINGAPORE

Singapore
Ph: +65 6812 7888

www.indium.tech

For Sales Inquiries
sales@indium.tech

For General Inquiries
info@indium.tech