Intelligent Data Streaming

**Benefits**

- Enable real-time operational intelligence with big data streaming analytics
- Reduce time-to-value with increased productivity and rapid deployment on-premises or in the Cloud
- Deliver information at any latency with one flexible platform
- Simplify configuration, deployment, administration, and monitoring of real-time streaming
- Minimize risks associated with complex and evolving open source technologies

**Deliver real-time information at the speed of business.**

Businesses today have an unprecedented opportunity to gain insight from a steady stream of real-time data—for example, transactions from databases, clickstreams from web servers, application and infrastructure log data, geo-location data, and data coming from sensors or agents placed on the almost endless variety of devices and machines making up the Internet of Things.

This continuous flow of messages and events can increase the effectiveness, agility, and responsiveness of decision making and operational intelligence. However, as data flows in at high rates, it accumulates quickly into large volumes. Organizations can derive maximum value from data only if they can gather and analyze it immediately and at an ever-increasing scale.

**Modern scalable architecture for streaming analytics**

Informatica Intelligent Streaming allows organizations to prepare and process streams of data and uncover insights while acting in time to suit business needs. It can scale out horizontally and vertically to handle petabytes of data while honoring business service level agreements (SLAs).

Intelligent Streaming provides pre-built high-performance connectors such as Kafka, HDFS, Amazon Kinesis, NoSQL databases, and enterprise messaging systems and data transformations to enable a code-free method of defining your data integration logic. Productivity and ease-of-maintenance is dramatically improved by the automatic generation of whole classes of data flows at runtime based on design patterns.

**Multi-latency data flows built to last as technologies evolve**

Informatica Intelligent Streaming builds on the best of open source technologies in an easy-to-use enterprise-grade offering. It primarily uses Spark Streaming, one of today’s more vibrant open source technologies, under the covers for stream processing and supports other open source projects such as Kafka and Hadoop. As new technologies inevitably evolve, Informatica Intelligent Streaming adapts, using the same data flows, so you don’t have to rebuild them. And data flows can be scheduled to run at any latency (real time or batch) based on the resources available and business SLAs.
Secure streaming data management on a foundation of best-in-class open source technologies

Informatica Intelligent Streaming includes an extensive library of prebuilt transforms running natively on Spark Streaming to process all types of data at scale. In addition to running on Spark Streaming, it uses:

- Secured Kafka (with Kerberos) as the data transport across mappings and data replay for recoverability
- HDFS as a highly available persistence store for recoverability data
- In-memory capabilities for speedy lookup to avoid continuous database lookups.

Consistent approach handles all latencies

Informatica Intelligent Streaming anticipates the rapidly evolving technology landscape by providing a single, consistent data-processing approach for all latencies. Developers design data streams once and deploy them only once. Existing data pipelines are easier to maintain and face less risk, as Spark Streaming evolves or if a new stream-processing engine is adopted. As a result, data streams and new innovations are implemented faster with less impact and risk to production systems.
Simple, centralized configuration, administration, and monitoring
Informatica Intelligent Streaming is built on the Informatica Intelligent Data Platform, leveraging a unified set of tools and services to help you effectively administer, monitor, and manage your deployment. The administrator tool lets you easily manage and monitor your system, users, and deployed mappings.

Cloud-ready streaming
More and more data is moving to the cloud as a cost-effective, scalable, and agile platform to store, process, and manage real-time streaming data. Informatica Intelligent Streaming fully supports Amazon Kinesis Streams as a source, Amazon Kinesis Firehose as a target, and Amazon EMR in streaming mode, making it easy to collect, deliver, and process large amounts of real-time data efficiently.

Process streams of virtually all types of data
In the world of fast data there are many different data formats and types produced by machines and IoT devices. Informatica Intelligent Streaming processes all types of data including complex hierarchical data objects in a variety of formats (e.g., JSON, XML, Avro, CSV) and types (e.g., Array, Struct, Record and Maps, Nested HTYPE).

Cloud-ready streaming
More and more data is moving to the cloud as a cost-effective, scalable, and agile platform to store, process, and manage real-time streaming data. Informatica Intelligent Streaming fully supports Amazon Kinesis Streams as a source, Amazon Kinesis Firehose as a target, and Amazon EMR in streaming mode, making it easy to collect, deliver, and process large amounts of real-time data efficiently.

High availability, scalability and architectural flexibility
Informatica Intelligent Streaming supports high availability, automated failover configuration on commodity hardware (with no need for a shared file system), and guaranteed delivery of data. This is required for uninterrupted processing of streaming data to ensure data is never lost and SLAs are met. Increasing horizontal and vertical scalability is as easy as deploying more Spark nodes. The flexible architecture supports changing business requirements with sources and targets connected in any pattern.

The Informatica Intelligent Streaming visual development environment provides upto five times the productivity of hand coding.
Key Benefits

Increase the ROI of your big data initiatives
Enable real-time operational intelligence with a single streaming analytics solution that can capture, transport, refine, enrich, process, and distribute streaming data in real time. Combine real-time data from sensors, devices, and machine logs with other enterprise data such as transactions, customer, product, and reference data to discover and respond to actionable insights at the speed of business.

Future-proof your investment with one versatile platform
Optimize your stream and batch processing based on available system resources and business SLAs. Data processing can range from sub-second stream processing on Spark Streaming to batch processing on Hadoop without having to redesign or rebuild data pipelines. You can build data pipelines once and run them at any latency to keep costs down by optimizing data processing across both existing data platforms and evolving technologies.

Reduce time-to-value with rapid deployment
Time-to-value measures how quickly you can progress from design, build, and test to deploy and maintain. Informatica Intelligent Streaming increases development productivity up to five times over hand coding. Using a visual development environment and prebuilt dynamic templates, developers can build data streams without specialized knowledge of Spark Streaming concepts and languages and rapidly deploy data streams into production with simple configuration parameters. This level of abstraction between the visual development environment and the underlying processing engine enables data streams to be deployed anywhere whether on-premises or in the cloud.

Minimize risks associated with complex and evolving open source technologies
Informatica Intelligent Streaming minimizes risks associated with rapidly evolving technologies such as Spark and Spark Streaming. IT can make one investment that continues to work with the changing technology landscape, providing a single, consistent data processing approach for all types of data at all latencies. Data pipelines are easier to maintain as emerging technologies continue to evolve and change which means your development is future proofed to quickly adopt the latest innovations in real-time streaming.