Information Management software

IBM InfoSphere Streams: Advanced analytics for a smarter planet

Highlights

- *Perform complex analytics on data in-motion*

- *Handles multiple structured and unstructured data types*

- *Handles massive data volumes*

Data volumes are expected to double every two years over the next decade. The global economic slowdown is resulting in organizations seeking to become more nimble with their operations and more innovative with their decisions. In the face of exploding data volumes and shrinking batch time windows, these organizations are struggling to make 'truly' real time decisions and beat the competition. Existing tools and technologies that aid decision making by the Line of Business first require data to be recorded on a storage device and run queries after the fact to detect actionable insights. Savvy businesses are fast realizing that the time lost in this process leads to missed opportunity that might be the difference between success and failure. InfoSphere Streams addresses this gap effectively by providing a futuristic technology that can detect insights within data streams still in motion.

IBM InfoSphere Streams enables continuous and extremely fast analysis of massive volumes of data-in motion to help gain insights and drastically improve decision making. It provides a state of the art infrastructure to support highly complex heterogeneous data analysis with exceptional performance and efficiently interoperates with existing application infrastructures. In addition, it can scale from a single server to hundreds of compute nodes.

InfoSphere Streams radically extends the state of the art in information processing by simultaneously addressing several technical challenges, including:

- Sub millisecond response time to events and changing requirements
- Continuous analysis of data at rates significantly higher than existing systems
- Simple development of streaming applications with the ability to seamlessly extend existing applications with new analysis types

InfoSphere Streams has demonstrated initial successes with a number of commercial and scientific applications across a spectrum of industries. In Financial Services, an InfoSphere Streams based prototype is being used to analyze and correlate over 5 million market messages per second to execute algorithmic option trades with an average latency of 150 microseconds. In Radio Astronomy, InfoSphere Streams is being used to correlate information from multiple radio telescopes to deliver enhanced imagery. In Healthcare, InfoSphere Streams is being used to correlate information from multiple sensors in a Neonatal Intensive Care Unit with a goal to detect medical conditions up to 24 hours earlier than experienced ICU nurses. In law enforcement and security, InfoSphere Streams is being used to correlate information from multi-modal surveillance systems and deliver real-time intelligence to security personnel. In the Water Management arena, applications based on InfoSphere Streams are being developed to interpret signals from acoustic sensors to better understand and manage natural ecosystems. Other emerging applications of InfoSphere Streams are in areas such as wildfire monitoring and control, intrusion detection, fraud prevention, call detail record extraction / transformation / loading, and manufacturing test analysis.

InfoSphere Streams delivers a wide array of analytic operators, an advanced set of development tools and an autonomic execution environment that could scale to 1000s of compute nodes. Applications are built using a simple declarative stream processing language, and the run-time environment manages execution across a cluster of computers. It can handle on-the-fly processing including aggregations and correlations over time-based windows. Analytic operators supplied with the system can perform simple arithmetic functions like weighted averages and also text analytics. Customers can extend the supplied analytic operators with user defined operators, such as custom trading algorithms.

InfoSphere Streams can leverage strong synergies with existing IT infrastructure in organizations. For example streaming data such as stock market quotes can be enriched with an existing high speed in memory database such as solidDB. By adding information to stock market ticks such as industry or interested traders, more complex analysis and customized alerting can be achieved. As Streams works to analyze information in real time, it can determine which data is most important to store in an existing Data Warehouse for subsequent analysis.

InfoSphere Streams includes:

- Streams Studio: An Integrated Development Environment (IDE) based on Eclipse 3.4 which may be used by developers within the organization and deployed on desktop PCs
- Streams runtime: May be deployed on a single server or a cluster of servers
- Tools and adapters: Includes StreamSight, ODBC drivers, high speed solidDB drivers, WebSphere Front Office, file I/O, and TCP/IP communication adapters

Runtime Operating Environment:

- Operating systems supported: Red Hat Enterprise Linux (RHEL) Version 5.2 64 bit or RHEL Version 4.4 32 bit
- Hardware supported: Intel/AMD x86 architecture (32 or 64 bit)

Development Operating Environment:

- Operating systems supported: Red Hat Enterprise Linux (RHEL) Version 5.2 64 bit or RHEL Version 4.4 32 bit
- Hardware supported: Intel/AMD x86 architecture (32 or 64 bit)

For more information

To learn more about IBM InfoSphere Streams, please contact your IBM marketing representative or IBM Business Partner, or visit **ibm.com/software/data/infosphere/streams**