



---

## Highlights

- Continuously analyze data in motion across real-time sources to deliver actionable insight
  - Immediately understand context in the business moment to dynamically adjust decisions and change course of action
  - Make real-time predictions and discoveries as data arrives to enhance and improve analytic models and cognitive systems
  - Deploy best-in-class real-time analytics to satisfy industry-specific requirements
  - Help increase IT productivity with highly visual, drag-and-drop development tools that support faster time to deployment and simple production management
  - Detect and respond to critical events 10 times faster compared to traditional techniques<sup>1</sup>
  - Choose cloud or on-premises deployment to manage costs and further speed deployment
- 

# IBM InfoSphere Streams

*Store less, analyze more and speed actions  
with real-time analytics in context*

Every day, consumers and businesses generate data at a speed that would have seemed unprecedented just a few years ago. For example, the ICU at Emory University Hospital generates 100,000 real-time data points per second, running up to 50 algorithms on each patient in real time.<sup>2</sup> Telecommunications providers are analyzing 7 billion call data records (CDRs) per day.<sup>3</sup> And social media contributes to massive data growth. Every second, on average, around 6,000 tweets are sent on Twitter, which corresponds to over 350,000 tweets per minute, 500 million tweets per day and around 200 billion tweets per year.<sup>4</sup>

According to an IBM Institute for Business Value Report,<sup>5</sup> several significant shifts occurred in the realm of big data—shifts business executives cannot afford to ignore. The most significant change: velocity is now the value driver for big data. Organizations can significantly impact business performance and market position through speed-driven data and analytic practices. The components of the most differentiating organizations are those capable of creating an agile and flexible infrastructure for analyzing fast data.

Organizations need to spot risks and opportunities in high-velocity data—opportunities that often can be detected and acted on only at a moment's notice. Flows of data from real-time sources such as market data, the Internet of Things, mobile devices, sensors, clickstreams and even transactions remain largely un-navigated. It's time to unlock this data to optimize decision making.



## Experience the power of now

IBM® InfoSphere® Streams enables this new generation of fast data analytics. InfoSphere Streams can continuously integrate and analyze data in motion to deliver real-time analytics and enhanced, more accurate analytic and cognitive systems. Organizations now have the ability to dynamically act and adjust in the business moment.

InfoSphere Streams is a context-aware stream computing platform with privacy built in. It consists of a development environment, a high-speed runtime architecture and a set of real-time analytic toolkits—including text, entity analytics, machine learning and more—that enable organizations to detect insights in high-velocity data (see Figure 1). Real-time analytics put data in context, delivering actionable insight when and where it is needed.



Figure 1. InfoSphere Streams real-time analytical toolkits.

## Comprehensive tools for an easy-to-use development environment

With InfoSphere Streams, developers can easily build real-time analytic applications. InfoSphere Streams Studio, an Eclipse-based integrated development environment (IDE), supports rapid development with editors, wizards, application flow graphs and runtime monitoring to simplify the process of building and managing InfoSphere Streams applications. These tools are designed to help organizations quickly and easily extend application capabilities with sophisticated, real-time data analytics:

- **Team development** is offered through tools such as IBM Rational Team Concert™ and Subversion.
- **The drag-and-drop graphical editor** allows users to build applications while automatically synchronizing graphical and Streams Processing Language (SPL) source code views. This round-tripping process lets developers create an application with a text or graphical editor, and update it later using the other editor.
- **An instance graph**, available in both InfoSphere Streams Studio and the management console, provides visual monitoring of application health and metrics and lets users quickly identify issues using customizable views.

---

### Join the InfoSphere developer community

InfoSphere has a thriving community of developers and architects. Join now at <https://developer.ibm.com/streamsdev>

---

Developers write InfoSphere Streams applications in the Streams Processing Language. A Streams application defines the data processing result that you want. The *how* portion of that execution is determined by the Streams runtime environment (which includes the scheduler), the Streams Processing Language Compiler, and to a lesser extent, a number or property and resource files. This enables highly optimized data processing and fast results.

## Scale-out architecture

InfoSphere Streams helps organizations extend their current IT investments without a massive infrastructure overhaul. It can scale from a single node to a virtually unlimited number of nodes to process high volumes of data—from terabytes to zettabytes. InfoSphere Streams provides a clustered runtime environment that can handle up to millions of events per second with microsecond latency.<sup>6</sup> Actionable results can be achieved with near-zero latency.

The advanced compiler fuses portions of the application for increased performance. The fused components can then be spread across one or more nodes. The compiler also supports multiple high-speed transports, including Ethernet and InfiniBand. Existing applications can be dynamically extended with new applications to read the same streams of data, further leveraging current investments.

A web-based management console makes it easy to configure and manage the running time and applications, including automatically placing features and deploying application components. Users can monitor applications and their individual elements for status and performance metrics to help ensure the organization attains its service-level agreements.

InfoSphere Streams includes several features that enhance high availability and redundancy. Administrators can add or remove processor nodes to or from the cluster dynamically. This enables administrators to perform necessary maintenance without downtime, helping to improve overall availability of the environment. Also, if the management nodes that control the system fail, the application nodes continue to run the InfoSphere Streams applications.

Programmatic manipulation of streaming data makes it easy to create highly available applications by using redundant application components. Dynamic allocation of host pools, and more location and isolation constraints of nodes in the runtime cluster, can be used to isolate redundant application components to separate nodes during the runtime.

## Sophisticated analytics with toolkits and accelerators

InfoSphere Streams includes integrated analytic toolkits that facilitate development of solutions for particular industries, functions or use cases. For example:

- **A financial services organization** can use InfoSphere Streams to combine incoming stock market quotes with industry data or information held in an internal database. Adding context to stock market ticks helps the organization achieve more complex levels of analysis. In addition, InfoSphere Streams can determine, in near-real time, which data should be stored in a database for future analysis, thereby helping to reduce storage and administration costs.
- **An insurance agency** can monitor unique sensor data such as temperature, wind and wave height, and scale up sophisticated analytics resources as needed—for example, during hurricane season. With up-to-the-minute information on storm severity and location, the agency can proactively issue alerts to customers with suggestions for protecting their families and properties. It can also quickly send response teams to the most affected areas. More people can be kept safe while costs stay under control.
- **Local governments and public safety organizations** can scale up resources to increase surveillance monitoring and create social media watch lists during large public concerts, protests or community events. The near-real-time processing can help alert teams to potential problems, medical incidents or even find criminals hiding in the chatter—enabling faster reaction times and better-informed first responders.

### Best-practice-based application accelerators

InfoSphere Streams also includes an application accelerator designed exclusively for the telecommunications industry. Complete with sample code, the IBM Accelerator for Telecommunications Event Data Analytics leverages best practices to help speed up implementation and turn streaming data into business value. It provides business logic, data processing and UI/visualization capabilities, analytic operators or sample applications that can be used as design templates for new applications, helping improve time to value.

The IBM Accelerator for Telecommunications Event Data Analytics includes mediation and analytics for telecommunication company call and event-detail records, which can boost revenue assurance and fraud detection. This enables telecommunications providers to strengthen billing insights based on services, vendors and business lines. It also supports efforts to differentiate services, strengthen customer loyalty, reduce churn, offer targeted services, personalize billing and improve the customer experience. Developers can also write their own accelerators for specific use case requirements.

---

### The open source IBMStreams project

Recently, IBM created an open source project for some InfoSphere Streams analytic toolkits to help speed the development of applications and harness the energies of the development community. In future releases, IBM expects to incorporate new functions from the open source project into the source product. IBM believes that a mix of open source and closed source development is the best way to encourage adoption in the marketplace, as seen by the success of open source offerings such as Apache Web Server and Eclipse where IBM was an early participant. To learn more about the IBMStreams open source project, visit <https://github.com/IBMStreams>

---

### Built-in context, performance and privacy

InfoSphere Streams includes real-time analytics out of the box, so organizations can:

- **Put data in context with real-time analytics:** InfoSphere Streams delivers context in real time, at the point of data ingestion, to make sense of big data.
- **Support privacy and security:** Built-in privacy means companies can analyze and share data without compromising identities or exposing sensitive information about people, places and things.
- **Speed analytic performance:** Organizations can capture the time value of data with in-memory stream computing. InfoSphere Streams outperforms Apache Storm by 2.6 to 12.3 times in terms of throughput while simultaneously consuming 5.5 to 14.2 times less CPU time.<sup>7</sup> Furthermore, the throughput and CPU time gaps widen as data volume, degree of parallelism, and/or number of processing nodes grows.

### Flexible deployment options

With cloud-based, real-time analytic offerings, InfoSphere Streams allows organizations to tap into fast data without the burden of managing all infrastructure operations in house. It enables organizations to rapidly develop real-time analytic applications in the cloud and respond quickly to changing business environments by analyzing larger volumes of data more cost-effectively than traditional approaches. InfoSphere Streams is offered as a service on IBM Cloud marketplace. InfoSphere Streams on Cloud offers:

- **Real-time analytics** when and where you need it
- **Faster and simplified** systems management for IT users
- **Bring your own license** (monthly or perpetual)
- **Fully managed** hardware, with InfoSphere Streams preinstalled

### Enterprise integration: Extending the power of InfoSphere Streams

Out-of-the-box integration with other products in the IBM big data platform, together with built-in XML and JavaScript Object Notation (JSON) support, gives InfoSphere Streams access to a wide variety of data sources and systems for handling enterprise applications:

- **IBM InfoSphere BigInsights™** lets you store streaming data in an enterprise-class Apache Hadoop environment for additional analysis or historic retention.
- **Support for XML** allows developers to fuse a broader range of traditional and untraditional data.
- **Database and warehouse support** lets users fuse and analyze streaming and unstructured data together with structured sources, including IBM DB2®, IBM Informix®, Microsoft SQL Server, Oracle Database, HP Vertica, IBM PureData® Systems, Teradata and more.

- **Message queues with IBM WebSphere® MQ, IBM MessageSight and IBM Integration Bus** provide a flexible system for efficient transport of messages and data.
- **IBM SPSS®** provides a state-of-the-art environment for understanding data and producing predictive models. The IBM SPSS Analytics Toolkit for InfoSphere Streams lets you integrate the predictive models designed and trained in IBM SPSS Modeler with your IBM InfoSphere Streams applications.
- **IBM Decision Management** is the IBM business rules engine. Using InfoSphere Streams, organizations can extend, identify, create and codify business rules with real-time insights.

---

### Real-time analytic processing at your fingertips: Free to download, quick to start

To help clients experience real-time analytic processing of data in motion, the IBM Quick Start program gives users fast access to market-leading big data technology. Available through a free download, the enterprise-grade big data platform allows users to experiment with stream computing in their own unique, nonproduction environment. There is no data limit or time limit. Learn more on the InfoSphere Streams Quick Start Edition website:

[ibm.com/software/data/infosphere/streams/quick-start](http://ibm.com/software/data/infosphere/streams/quick-start)

---

## InfoSphere Streams V4.0: Features and enhancements

The latest version of InfoSphere Streams includes a wide range of features and functionality enhancements to increase business agility and consumability, empower IT developers and administrators, and support application resiliency.

### Business agility

Microsoft Excel is one of the most widely used business intelligence (BI) tools. Until the release of InfoSphere Streams V4.0, Microsoft Excel could analyze only data at rest. InfoSphere Streams makes streaming data available directly in Excel by allowing you to drag and drop information onto a worksheet (see Figure 2). The result is rapid prototyping of real-time applications, which can increase agility and time to market.

For example, a business analyst can quickly search streaming data for key performance indicators, such as number of web page visits. The analyst can also detect patterns for a market-basket analysis or spot suspicious activity, such as rapid credit purchases on the same account in different time zones.

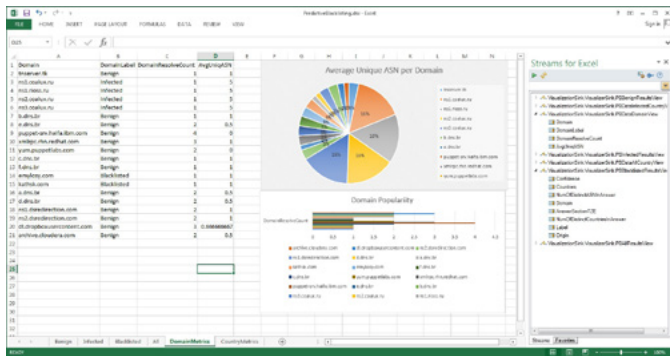


Figure 2. Streaming data visually presented in Excel.

### Application resiliency

With the emergence of new engagement models such as cloud, mobile and social applications, organizations are bringing streaming data sources into mission-critical, front-end, revenue-generating applications, including customer care, security and risk management, and operations analysis. Resiliency of streaming applications is now a priority, yet the complexity of data is increasing. To meet service levels, deliver better customer care and ensure optimal action, organizations need to guarantee all streaming data is processed in a timely and accurate manner. InfoSphere Streams V4.0 is enhanced with simple operators and annotations to help prevent data loss.

### Simplicity for IT administrators and developers

As streaming applications play an increasingly important role in mission-critical solutions, the need for simplicity has grown as well. InfoSphere Streams V4.0 empowers IT users of all types and skill levels to gain deeper insights into operations and application performance. In today's engaged world, a few seconds could mean the difference between success and failure. To address these needs, a new administration console, a Java Management Extensions (JMX) management and monitoring application programming interface (API), simplified security and adoption of Apache Zookeeper are now available in InfoSphere Streams V4.0.



## For more information

To learn more about InfoSphere Streams and how it can help your business gain a competitive advantage, visit:

[ibm.com/software/data/infosphere/streams](http://ibm.com/software/data/infosphere/streams)

Additionally, IBM Global Financing can help you acquire the software capabilities that your business needs in the most cost-effective and strategic way possible. We'll partner with credit-qualified clients to customize a financing solution to suit your business and development goals, enable effective cash management, and improve your total cost of ownership. Fund your critical IT investment and propel your business forward with IBM Global Financing. For more information, visit: [ibm.com/financing](http://ibm.com/financing)

© Copyright IBM Corporation 2015

Software Group  
Route 100  
Somers, NY 10589

Produced in the United States of America  
February 2015

IBM, the IBM logo, [ibm.com](http://ibm.com), BigInsights, DB2, Informix, InfoSphere, PureData, Rational Team Concert, SPSS, and WebSphere are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at [ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml)

Microsoft, Excel, and SQL Server are trademarks of Microsoft Corporation in the United States, other countries, or both.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.

Actual available storage capacity may be reported for both uncompressed and compressed data and will vary and may be less than stated.



Please Recycle

---

<sup>1</sup> "Of Streams and Storms—A Direct Comparison of IBM InfoSphere Streams and Apache Storm in a Real World Use Case." [https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=sw-infomgt&S\\_PKG=ov24826&S\\_CMP=is\\_bdwp36](https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=sw-infomgt&S_PKG=ov24826&S_CMP=is_bdwp36)

<sup>2</sup> "Emory University Hospital uses IBM streaming analytics to gain life saving insights on patients." <http://www.ibmbigdatahub.com/video/emory-university-hospital-uses-ibm-streaming-analytics-gain-life-saving-insights-patients>

<sup>3</sup> InfoSphere Streams and Telecommunications. [ibm.com/software/data/infosphere/stream-computing/smarter-telco.html](http://ibm.com/software/data/infosphere/stream-computing/smarter-telco.html)

<sup>4</sup> Internet Live Stats: Twitter Usage Statistics. [www.internetlivestats.com/twitter-statistics](http://www.internetlivestats.com/twitter-statistics)

<sup>5</sup> "Analytics: The speed advantage." IBM Institute for Business Value Report. October 2014. <http://www-935.ibm.com/services/us/gbs/thoughtleadership/2014analytics>

<sup>6</sup> "Of Streams and Storms—A Direct Comparison of IBM InfoSphere Streams and Apache Storm in a Real World Use Case." [https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=sw-infomgt&S\\_PKG=ov24826&S\\_CMP=is\\_bdwp36](https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=sw-infomgt&S_PKG=ov24826&S_CMP=is_bdwp36)

<sup>7</sup> "Of Streams and Storms—A Direct Comparison of IBM InfoSphere Streams and Apache Storm in a Real World Use Case." [https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=sw-infomgt&S\\_PKG=ov24826&S\\_CMP=is\\_bdwp36](https://www14.software.ibm.com/webapp/iwm/web/signup.do?source=sw-infomgt&S_PKG=ov24826&S_CMP=is_bdwp36)