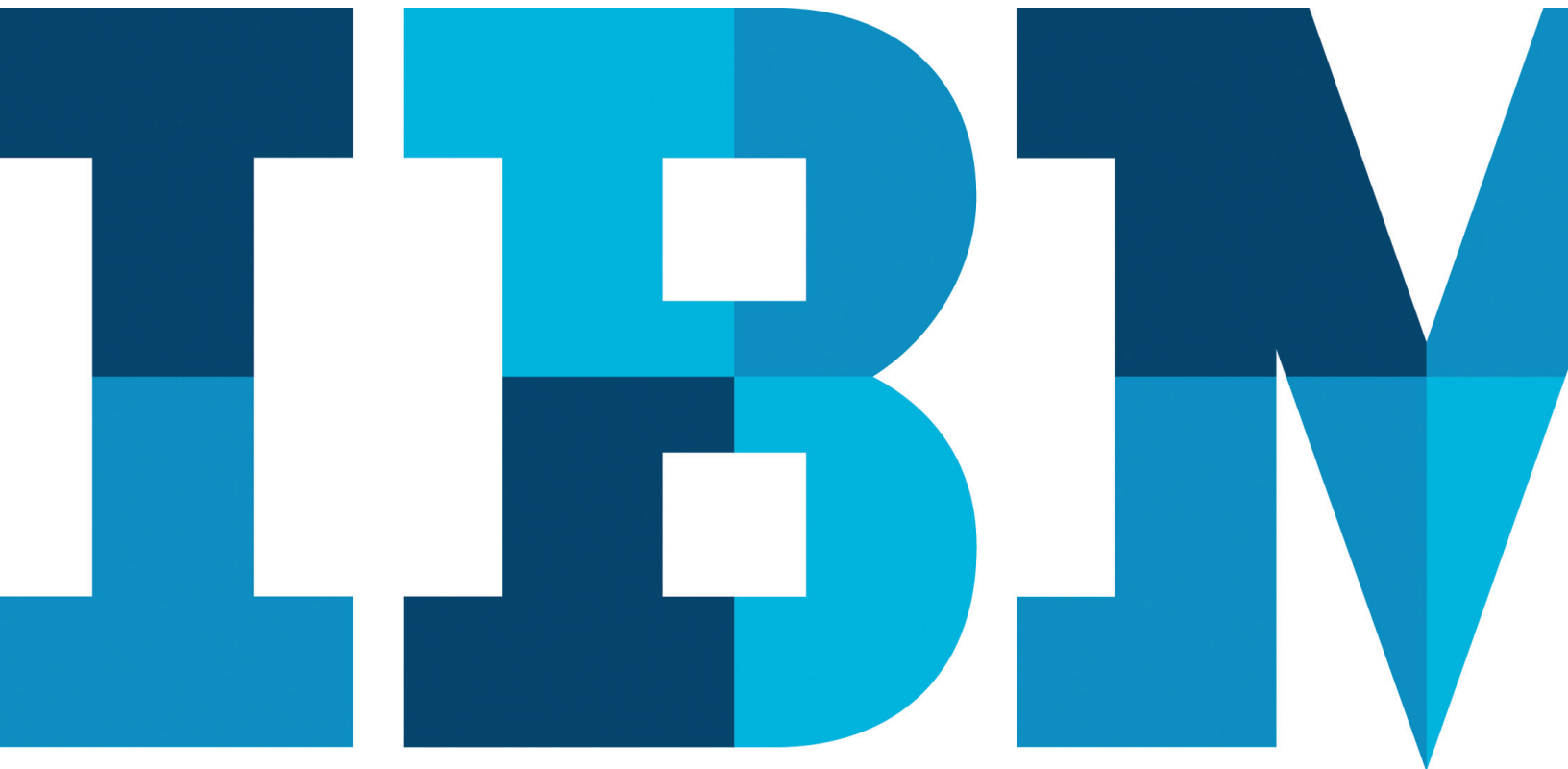


IBM InfoSphere BigInsights for Hadoop Quick Start Edition

Hadoop for the enterprise: Quick, easy, free



For many organizations wanting to take the plunge into big data analysis, Apache Hadoop is a natural starting point. This open source software project enables distributed processing of large data sets across clusters of commodity servers. Designed to scale and able to run on commodity hardware, the resiliency of these Hadoop clusters comes from the ability to detect and handle failures at the application layer.

It sounds straightforward, so why is Hadoop such a big deal when it comes to big data? Hadoop changes the economics and dynamics of large-scale computing by enabling a solution that is:

- **Scalable:** Add new nodes as needed without changing data formats, how data is loaded, how jobs are written or the applications on top.
- **Cost effective:** Hadoop brings massively parallel computing to commodity servers. The result is a significant decrease in the cost per terabyte of storage, which in turn makes it affordable to model all your data.
- **Flexible:** Hadoop is schema-less, and can absorb any type of data, structured or not, from a number of sources. Data from multiple sources can be joined and aggregated in arbitrary ways, enabling deeper analyses than any one system can provide by itself.
- **Fault tolerant:** When you lose a node, the system redirects work to another location of the data and continues processing without missing a beat.

Considering industry estimates that 80 percent of the world's data is unstructured, organizations that use Hadoop are able to tap into data that many others simply have to discard. This ability to analyze new data types can lead to more informed, faster decision making.

If Hadoop was easy, everyone would be using it

Hadoop clearly provides an advantage to organizations that want to access all of the data available to them. So why aren't more organizations using it?

Hadoop comes with its own set of challenges. It typically requires Hadoop experts to be able to work with it and develop big data applications—and many organizations are reluctant to invest in the skills required for fear of turnover.

Enterprise-level businesses need an enterprise-ready solution packaged to fit seamlessly into their existing IT architecture. IBM® InfoSphere® BigInsights™ for Hadoop is that solution.

InfoSphere BigInsights for Hadoop: Bringing Hadoop to the enterprise

IBM recognizes the value in open source Hadoop, but also knows that by itself, open source is incomplete. IBM has built InfoSphere BigInsights on top of the open source components of Hadoop, adding value-added features like built-in analytics, visualization and security to take the complexity out of Hadoop.

Calling all experimenters

In 2013, IBM introduced the latest addition to the InfoSphere BigInsights for Hadoop family with the Quick Start Edition, a *free, downloadable, non-production version* that lets you check out enterprise-grade big data features from IBM.

Regardless of your experience level, you can use the Quick Start Edition to begin experimenting with Hadoop today. You get access to hands-on learning with a set of tutorials designed to guide you through the Hadoop experience. Plus, there is *no data capacity or time limit*, so you can work with large data sets and explore different use cases at your convenience.

InfoSphere BigInsights for Hadoop Quick Start Edition does not come with a support option. To explore your support options, visit the InfoSphere BigInsights for Hadoop Enterprise Edition page at ibm.com/software/products/us/en/infobiginteedit

InfoSphere BigInsights for Hadoop Quick Start Edition features

- **Text analytics:** Sophisticated text analytics unique to InfoSphere BigInsights that include a vast library of extractors that draw actionable insights from large amounts of native textual data
- **IBM BigSheets:** Web-based analysis and visualization tool with a familiar, spreadsheet-like interface that allows analysis of sizeable volumes of data and helps design and manage long-running data collection jobs
- **IBM Big SQL:** New, native SQL query engine that enables SQL access to data stored in InfoSphere BigInsights, leveraging MapReduce for complex data sets and direct access for smaller queries
- **Workload optimization:** Adaptive MapReduce adjusts to user needs and system workloads automatically to improve performance and simplify job tuning, while a workload scheduler provides optimization and control of job scheduling based on user-selected metrics
- **Development tools:** Familiar, Eclipse-based development environment for building and deploying analytic applications, along with a set of developer tools, extractors and editors for fast adoption and reduced coding and debugging
- **Management capabilities:** Auditing helps tighten security and access control, while monitoring offers the ability to control all applications from a centralized dashboard

InfoSphere BigInsights for Hadoop Enterprise Edition-only features

While the Quick Start Edition provides the majority of robust enterprise-level features that InfoSphere BigInsights for Hadoop is known for, some of them come only with the Enterprise Edition. These include:

- Production support
- IBM General Parallel File System (IBM GPFS™) File Placement Optimizer (FPO)
- Machine Data Analytics Accelerator
- Social Data Analytics Accelerator
- Limited-use licenses for IBM Cognos®, InfoSphere Data Explorer and InfoSphere Streams

Expanding the Quick Start family: InfoSphere Streams Quick Start Edition

InfoSphere BigInsights for Hadoop is not the only Quick Start Edition available; there's one for InfoSphere Streams as well.

InfoSphere Streams Quick Start Edition is a free, downloadable version of the real-time analytic processing software. InfoSphere Streams is an advanced computing platform that can quickly ingest, analyze and correlate information as it arrives from thousands of real-time sources. Because it can handle high throughput rates, InfoSphere Streams can analyze millions of events per second, enabling sub-millisecond response times and instant decision making. This capability lets you analyze massive data volumes quickly, often in real time, and turn data into actionable insight.

InfoSphere Streams comes bundled with InfoSphere BigInsights for Hadoop Enterprise Edition; however, it doesn't come with the InfoSphere BigInsights for Hadoop Quick Start Edition. You can download either Quick Start Edition to work with data at rest and data in motion. Then, when you're ready for production, each product—InfoSphere Streams and InfoSphere BigInsights for Hadoop—comes bundled with a limited-use license of the other.

Download Quick Start today

InfoSphere BigInsights for Hadoop Quick Start Edition is appropriate for all levels of Hadoop expertise, but those with little to no Hadoop experience are especially encouraged to download Quick Start. You will find the education and guidance you need to begin developing big data applications in no time.

There are two download options for the Quick Start Edition: native software installation and VMware image. If you choose the native software option, you will download Quick Start directly onto your machine. With the VMware image, you will be presented with a single node or cluster option. The single node VMware image is for smaller data sets, while the cluster version allows you to add multiple nodes for larger volumes of data. Regardless of which option you choose, the features and capabilities are the same.

If you've been looking to experiment with Hadoop, or don't have any Hadoop projects planned right now but would like to learn, give the InfoSphere BigInsights for Hadoop Quick Start Edition a try. You'll be amazed at how easy it is to get started.

For more information

Download your free copy of InfoSphere BigInsights for Hadoop Quick Start Edition today at ibm.com/infosphere/quickstart

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



© Copyright IBM Corporation 2013

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
October 2013

IBM, the IBM logo, ibm.com, BigInsights, Cognos, GPFS, and InfoSphere 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

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.



Please Recycle
