



Highlights

- Accelerates time-to-value with scalable, factory-integrated solutions
 - Performs low-latency analysis on large volumes of stored or streaming data
 - Analyzes data in its native format using IBM InfoSphere software
 - Uses deep architectural services for rapid deployment and tuning
-

IBM System x reference architecture solutions for big data

Easy-to-implement hardware, software and services for analyzing data at rest and data in motion

Big data is here—in a big way. Mind-boggling volumes of data are created every second of every day. What's more, the number and types of sources generating data is growing exponentially from both people and electronic devices. More data should lead to better decisions. But because data growth is so fast, many enterprises are analyzing a smaller percentage of data than before. So what is the answer? You don't have the time, resources or capital to experiment with what works in big data analytics and what doesn't. The challenge you face is to manage the volume, variety and velocity of data, apply analytics for better insight and use this insight for making better business decisions faster than the competition.

Reference architecture solutions for big data analytics from IBM are easy to order, easy to implement solutions. They include IBM hardware, software and services along with a standardized blueprint created from IBM enterprise experience and expertise. You can quickly deploy these cost-effective remedies to analyze data at rest, which is stored data such as databases, or data in motion, like streaming data, or both. The solutions are built around powerful, affordable, scalable IBM® System x® servers and IBM System Networking solutions combined with IBM InfoSphere® BigInsights™ and InfoSphere Streams software, so you can deploy proven solutions quickly.



The benefits of deploying a reference architecture solution from IBM

Big data projects are not all alike. What works for one company, may not work for another. Instead of re-inventing the wheel to apply business analytics to big data, it makes business sense to take advantage of what has already been proven. This is especially true when considering a reference architecture developed by IBM, a company with experience based on thousands of IT infrastructure engagements in wide-ranging environments for companies of all sizes in virtually every industry.

Reference architectures for big data analytics provide technical blueprints that include a well-defined scope, a complete listing of requirements, and architectural decisions proven in the field including IBM Intelligent Cluster™ that helps make complex solutions simple. A reference architecture provides the value of synergy amongst each of the solution building blocks, with the flexibility needed to meet your requirements. Take a look at some of the advantages of using a reference architecture for big data from IBM:

Ease of integration: Deployment using a reference architecture helps ensure the new solution works with what you already have in place and are likely to add later, such as data warehouse, stream compute engines, internal and external storage devices and more.

Flexibility and simplicity: A reference architecture from IBM helps strike a cost-effective balance between requirements of the solution and time-to-value, offering flexibility for your enterprise to adapt as your big data requirements evolve.

Faster deployment: A reference architecture specifying IBM pre-integrated components can also include Intelligent Cluster, an integrated, factory-built and factory-tested cluster

solution that you can roll into your data center and put to work right away. Add to this an optimized software stack that relies on open source code, and you can get the total solution up and running faster, with fewer headaches and much less trial and error.

Quicker access: Every business is looking for fast time-to-value. There is no letup in data volumes, economic pressures or competitive demands. The accelerators and pre-built analytics tools for developers and business users incorporated in the reference architecture mean solutions are working harder sooner.

Simplify complex solutions with IBM Intelligent Cluster

Intelligent Cluster takes advantage of decades of IBM experience to reduce the complexity of deployment with pre-integrated, delivered and fully supported solutions that match best-in-industry components with optimized solution design. With Intelligent Cluster, you can focus your efforts on maximizing business value, instead of consuming valuable resources to design, optimize, install and support the infrastructure required to meet business demands.

Intelligent Cluster solutions are built, tested, delivered and installed by IBM, and supported as a single solution instead of being treated as hundreds of individual components. IBM provides single point of contact, solution-level support that includes both IBM and third-party components to deliver maximum system availability throughout the life of the system, so you can spend less time maintaining systems and more time delivering results.

Reference architecture for big data at rest from IBM

InfoSphere BigInsights takes open source Apache Hadoop and adds the enterprise-class functionality and integration necessary to help meet your critical business requirements. It provides a thoroughly tested and integrated solution that combines the benefits of leading-edge technologies with mature, enterprise-ready features. Organizations can run large-scale, distributed analytics jobs on clusters of cost-effective server hardware. This infrastructure leverages Hadoop's MapReduce framework to tackle very large data sets by breaking up the data across many nodes and coordinating data processing across a massively parallel environment.

While the power of Hadoop is easy to understand, it can be difficult to implement due to the wide variety of workload characteristics that must be identified before the IT infrastructure can be designed—variables such as workload type, size of MapReduce worker and task, compression, scaling requirements and many more. Relying on an IBM reference architecture can help speed time-to-value with faster deployment and less guesswork. In addition, IBM Systems Lab Services can help tune your architecture to accomplish your goals quicker.

System x M4 servers, like the x3630 M4 and x3550 M4, help enhance performance and reduce power consumption. The x3630 M4 offers a cost-effective option to achieve higher performance for data intensive workloads in a two-socket, 2U design while the x3550 M4 is a dense, affordable and easy-to-use rack solution that delivers a pay-as-you-grow design to help lower costs and manage risks.

System x servers are intelligent, industry-leading enterprise x86 systems designed to reduce costs and complexity.

System x solutions are the result of a design that builds proven IBM innovation around industry standard components. Solutions range from feature-rich, scalable enterprise x86 servers to rack optimized servers designed for business productivity to entry level tower servers. They have been created to achieve three primary objectives:

- **Maximize performance:** System x servers are designed to extract the most compute power possible from x86 CPUs and deliver it to your workloads.
- **Minimize cost:** System x helps lower acquisition costs of enterprise servers, while preserving the ability to expand to much greater capabilities if and when you need to.
- **Simplify deployment:** With single points of management and remote administration, these systems are built for simpler qualification, faster deployment and easier administration.

Reference architecture for big data in motion from IBM

InfoSphere Streams provides a state-of-the-art computing platform to help companies turn burgeoning data volumes into actionable information and business insights. As a critical element of the IBM platform for big data, InfoSphere Streams delivers a highly scalable, agile software infrastructure that enables businesses to perform in-motion analytics on a wide variety of relational and non-relational data types at unprecedented volumes and speeds, from thousands of real-time sources. With InfoSphere Streams, your enterprise can capture and act on key business data just in time, all the time.

InfoSphere Streams provides a development platform and runtime environment that can ingest, filter, analyze and correlate massive volumes of continuous data streams. These data streams could include text, spreadsheets, images, video and audio recordings, email, instant messaging, financial transactions, customer records and many others. InfoSphere Streams fuses these heterogeneous data types onto a powerful computing platform that enables complex data analysis with exceptional performance and impressive response times.

Stream processing requires resilient, memory-rich servers to ensure continuous delivery of insight. The core reference architecture is built with the x3550 M4 and x3690 X5. The x3630 M4 offers a cost-effective option to achieve higher performance for data-intensive workloads in a two-socket, 2U design, while the x3690 X5 is a two-processor server that features the latest enterprise IBM X-Architecture® technology (eX5) for extreme flexibility and enterprise-class availability and reliability at a lower total cost of ownership.

Designing for stream workloads requires an understanding of core design characteristics such as ingest streaming rates, data packet sizes, output rates required, latency and QoS requirements. These have an impact on architectural design considerations such as node design, processor speed, network bandwidth and even server redundancy for mission critical applications. The reference architecture approach addresses these requirements to ensure faster, lower risk implementations.

As mentioned above, System x intelligent, industry-leading, enterprise x86 systems are designed to reduce costs and simplify your IT infrastructure. Managing energy in the data center is a

growing concern due to the increasing numbers of servers, the incremental heat they generate and ultimately, the rising cost of energy. System x servers not only help increase performance per watt, but also help you budget, plan and control power usage.

Get the help you need with analytics from IBM Systems Lab Services

IBM Systems Lab Services and Training helps carve a path to analytics on IBM solutions. Each big data analytics engagement uniquely focuses on implementing a solution that will help clients understand their data and identify opportunities where analytics solutions will make the biggest impact to their business.

Here are just a few of the available services that can help make small work of implementing your big data solutions:

- **Big data analytics assessment:** Helps with understanding the data clients have today, data they could be collecting or need to collect to identify big data opportunities.
- **Big data application proof of concepts (POCs):** This can offer an overview of the BigInsights, Hadoop or Streams application development process, along with assistance with application architecture and development.
- **Training:** Helps the client organization understand the product and solution implemented to get the biggest usability of the solution.

Why IBM?

The IBM commitment to understanding, exploring and refining analytics also includes the investment of more than USD14 billion in analytics acquisitions of such market leaders as IBM SPSS®, IBM Cognos®, IBM Unica®, IBM Coremetrics®, IBM Platform Computing™ and many others. Consider the IBM investment in analytic intellectual property within Netezza® or the IBM Smart Analytics System. The fact that IBM has a business analytics and optimization (BAO) division represents the kind of long-term capabilities IBM is committed to delivering for analytics in its big data platform. In May, 2011 IBM announced a USD100 million investment in massive scale analytics focused primarily on refining Hadoop. It is also likely that no other vendor can talk about and deliver analytics for both big data at rest and big data in motion together.

IBM can make this scale of commitment in large part because of a century-old track record of successful innovation. IBM has the single largest commercial research organization in the world, with more than 200 mathematicians at work developing breakthrough algorithms. Today, IBM has more than 9,000 experienced strategy, analytics and technology experts and consultants. IBM provides thought leadership and

practical insights from the IBM Institute for Business Value and offers proven solutions and use cases across almost every industry and functional area.

For more information

To learn more about IBM System x reference architecture solutions for big data, please contact your IBM representative or IBM Business Partner, or visit the following website:

ibm.com/systems/x/bigdata

Additionally, IBM Global Financing can help you acquire the IT solutions that your business needs in the most cost-effective and strategic way possible. We'll partner with credit-qualified clients to customize an IT financing solution to suit your business goals, enable effective cash management, and improve your total cost of ownership. IBM Global Financing is your smartest choice to fund critical IT investments and propel your business forward. For more information, visit: ibm.com/financing



© Copyright IBM Corporation 2012

IBM Corporation
IBM Systems and Technology Group
Route 100
Somers, NY 10589

Produced in the United States of America
September 2012

IBM, the IBM logo, ibm.com, System x, InfoSphere, BigInsights, Intelligent Cluster, SPSS, Cognos, Unica, Coremetrics, Platform Computing, and X-Architecture 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

Netezza is a registered trademark of IBM International Group B.V., an IBM Company.

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