

# Lenovo Big Data Validated Design

for IBM BigInsights SQL Analytics

Lenovo™

**The Industry's First Extreme Performance Big Data SQL Analytics Offering**

## The Big Data Challenge with Apache Spark SQL

Big data is more than a challenge. It is an opportunity to find new insights in data to make your business more agile and to answer questions that were previously beyond reach. To open the door to a world of possibilities IBM employs the latest big data technologies with the faster in-memory processing capability unleashed with Apache Spark.

This Lenovo Big Data Validated Design for IBM BigInsights SQL Analytics is certified by IBM and provides a thoroughly tested and integrated solution which combines the benefits of leading-edge technologies with mature, enterprise-ready features. This solution can be leveraged to tackle very large data sets by breaking up the data into “chunks” and coordinating the processing of the data across a massively paralleled environment.

This solution is designed and developed to deliver extreme SQL analytics performance on IBM BigInsights. This system architecture provides the flexibility and options to select from three different configurations of processor, memory and storage based on customer performance needs.

## A SQL Analytics Solution that Scales up to 100TB and Beyond

SQL Analytics on Big Data clusters is now an essential workload for enterprise data centers. Consequently, Spark SQL is gaining adoption both for its features and its performance. Spark SQL enables organizations to realize the benefits of real-time analytics and gain faster insights in today's competitive business environment. And with this solution, it provides significant performance and cost advantages for analyzing massive datasets with its in-memory processing engine and support for running SQL analytics. By deploying Apache Spark SQL or IBM Big SQL on a distributed Lenovo server cluster using Intel NVMe drives and a high speed data network the following key capabilities of this solution can be realized:

- Obtain big data nodes that are pre-configured, validated and delivered for scalable SQL analytics
- Deploy a high performance solution for Apache Spark SQL or IBM Big SQL analytics
- An energy efficient optimized cluster delivering extreme performance in a small data center footprint
- Achieve faster time to analytics and leverage scalability from 1 to 100 TB data sets and beyond

### Highlights

- Unleash extreme Spark SQL cluster performance for faster analytics with sub-millisecond I/O latency
- Support SQL analytics with in-memory processing for with Apache Spark SQL 2.1 or IBM Big SQL
- Optimize SQL analytics with a validated and tested solution architecture with a reliable infrastructure stack

## The Lenovo System x3650 M5 server and x3550 M5 server for a high capacity, high performance and cost effective Big Data solution



Lenovo System X3650 M5 Server and  
Lenovo System X3550 M5 Server

## Support for Spark SQL or IBM Big SQL

Both Apache Spark SQL and IBM Big SQL analytics are fully integrated components of the Lenovo Big Data Validated Design for IBM BigInsights SQL Analytics. Either can be deployed on this platform. Take advantage of Apache Spark's distributed in-memory storage for high performance processing across a variety of use cases including batch processing, real-time streaming for advanced modeling and analytics.

## Configuration Details

This Lenovo big data solution employs the Lenovo System x3650 M5 and System x3550 M5 servers for Data and Management nodes, respectively. For Data nodes, three configurations of the System x3650 M5 server are employed to yield three different nodal types. These nodal types support three different SQL analytics performance levels. By having three options to choose from allows businesses to select the configuration most appropriate for their performance requirements. Once the performance level is determined, add the required number of nodes in this scale out architecture to build the required cluster size. This solution also employs various networking options up to 100Gbps and beyond. Data networking elements provide the low latency and scalability required to support 100TB clusters and beyond for high performance SQL analytics.

## Why Lenovo

Lenovo is a leading provider of x86 servers for the data center. Featuring rack, tower, blade, dense and converged systems, the Lenovo server portfolio provides excellent performance, reliability and security. Lenovo also offers a full range of networking, storage, software, solutions, and comprehensive services supporting business needs throughout the IT lifecycle. With options for planning, deployment, and support, Lenovo offers expertise and services needed to deliver better service-level agreements and generate greater end-user satisfaction.

## For More Information

To learn more about the Lenovo Big Data Validated Design for IBM BigInsights SQL Analytics, contact your Lenovo Business Partner or visit:

[www.lenovo.com/systems/solutions](http://www.lenovo.com/systems/solutions)



© 2017 Lenovo. All rights reserved.

**Availability:** Offers, prices, specifications and availability may change without notice. Lenovo is not responsible for photographic or typographical errors. **Warranty:** For a copy of applicable warranties, write to: Lenovo Warranty Information, 1009 Think Place, Morrisville, NC, 27560, Lenovo makes no representation or warranty regarding third party products or services. **Trademarks:** Lenovo, the Lenovo logo, System x, ThinkServer are trademarks or registered trademarks of Lenovo. Microsoft and Windows are registered trademarks of Microsoft Corporation. Intel, the Intel logo, Xeon and Xeon Inside are registered trademarks of Intel Corporation in the U.S. and other countries. Other company, product, and service names may be trademarks or service marks of others.

CRN: BGDIB01XX72

06/2017