



# Delivering Powerful Analytics Using MapR Distribution and High-speed Ethernet

## EXECUTIVE SUMMARY

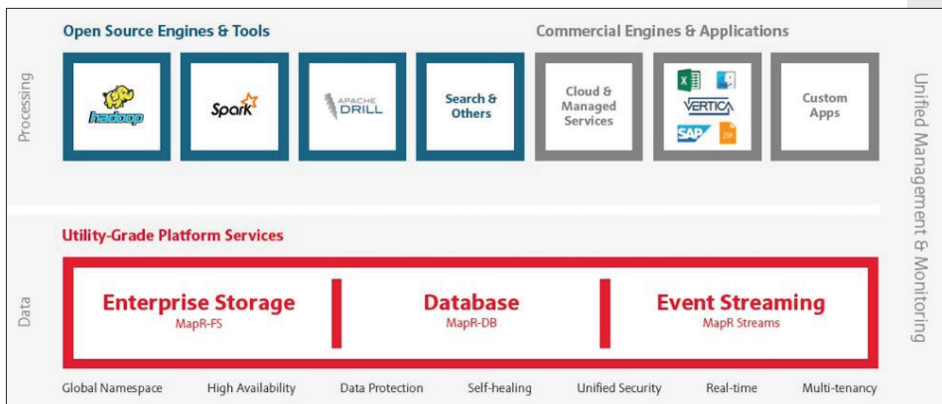
The rise of mobile technology and the Internet of Things (IoT) has resulted in sensors in cars, industrial equipment, and medical devices to name a few. Do to this, an abundance of data is being created form areas that have never produced digital data. This data is increasingly unstructured in nature and is typically stored in Hadoop and NoSQL systems. New actionable insights can be gleaned from these to fuel new business growth and the spoils of this new found data can spur a competitive advantages for companies that can successful analyze this data. However, for big data to deliver on the promise of its vast potential, technology must be in place to enable organizations to not just capture and store data but also quickly gain new insights that could be used in many ways to improve business. This joint solution between MapR and Mellanox discusses the data analyze advantages of pairing a high-speed Ethernet network with the MapR Converged Data Platform to solve problems involving massive amounts of data and computation.

## KEY BUSINESS BENEFITS

- Conduct analysis and formulate new hypotheses quicker than before
- Continuous operation with Zero-Touch deployment
- Accelerate Hadoop deployment with enterprise grade reliability from MapR and Mellanox
- Lowest TCO by maximizing system resources and supporting multiple workloads with the most economical choice for building a Hadoop cluster

“The MapR Converged Data Platform integrates the power of Hadoop, Spark, NoSQL, and streaming for developing and running innovative data applications. This partnership between MapR and Mellanox helps customers get even more output and value from their big data deployment.”

- Steve Wooledge VP of Product Marketing  
MapR Technologies



## THE MapR CONVERGED DATA PLATFORM

As a technology leader in Hadoop, MapR Technologies provides a converged platform that includes a complete distribution of Apache Hadoop. MapR continues to be the fastest, and leads in features and capabilities among major Hadoop distributions. The MapR Converged Data Platform integrates Hadoop and Spark with global event streaming, real-time database capabilities, and enterprise storage for developing and running analytics applications. The MapR Platform Services layer, which includes the MapR file system and NoSQL database (MapR- FS and MapR-DB, respectively), provides several enterprise grade features including high availability, data security, and easy integration with NFS. The platform supports a wide range of applications that process structured and unstructured data stored in files as well as NoSQL database. This single data platform for all big data workloads provides the best ROI for hardware use.

## HIGHLY SCALABLE PLATFORM OPTIMIZED FOR PERFORMANCE

The high-performance data platform from MapR enables business analysts to formulate new business values quicker than before using different big data applications. The workloads could be SQL-based (using Hive, Impala, Spark SQL, etc.), the newer class of NoSQL-based (using MapR-DB), or the more advanced applications using machine learning, graph, and streaming analytics. Due to the advantages of MapR-FS and the increasing usage of flash storage in production grade Hadoop environments, network bandwidth often becomes the bottleneck during data ingestion and analytics. Investing in faster servers and flash storage for MapR does not make sense if performance is restricted by the network. Mellanox provides the most efficient end-to-end Ethernet network tailored for big data applications at 10/25/40/50 and 100G Ethernet. Mellanox Ethernet switches feature consistently low latency and can support a variety of non-blocking, lossless fabric designs. Furthermore, big data applications utilizing TCP or UDP over IP transport can achieve the highest throughput and application density using hardware-based stateless offloads and flow steering engines in Mellanox ConnectX® Ethernet adapters. These advanced stateless offloads reduce CPU overhead in IP packet processing, allowing completion of heavier analytic workloads in less time in the big data cluster.

## EASE OF END-TO-END MANAGEMENT

In large scale-out software ecosystems such as Hadoop and NoSQL, delivering high availability and low downtime becomes a crucial challenge for administrators when deploying a wider ecosystem of big data software such as Spark, Drill, and Sqoop, in addition to MapReduce applications.

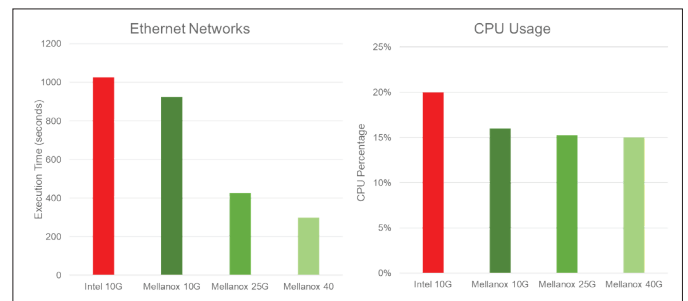


350 Oakmead Parkway, Suite 100  
 Sunnyvale, CA 94085  
 Tel: 408-970-3400 • Fax: 408-970-3403  
[www.mellanox.com](http://www.mellanox.com)

MapR adds to the ease of management with capabilities not found in other Hadoop distributions that simplify business continuity, tiered storage, security, and provisioning. With MapR Control System, the browser-based administration tool, administrators can manage and monitor the MapR deployment to ensure a reliable, high performance system. Mellanox Onyx™ provides the most advanced network orchestration, automation and monitoring platform for IT administrators to get 360 view of the entire network. Furthermore, with the One-Click feature in Onyx, data center administrators can easily configure their large scale-out network in simple templates thereby allowing them to automate replication of network state whenever new data nodes are added in the cluster.

## CONCLUSION

Although Apache Hadoop offers a powerful tool for analyzing large and diverse data sets, deploying a successful, reliable and high-performance infrastructure can be a daunting challenge. The MapR distribution for Hadoop provides critical technology advances to make Hadoop implementation easy, dependable, and fast for production deployments. By taking advantages of Mellanox’s advanced offloads, IT organizations no longer have to choose between performance and reliability. With these integrated capabilities, businesses can achieve the competitive advantages of big data analytics faster with less risk and with the confidence of an enterprise grade ecosystem.



### About Mellanox

Mellanox Technologies is a leading supplier of end-to-end InfiniBand and Ethernet interconnect solutions and services for servers and storage. Mellanox interconnect solutions increase data center efficiency by providing the highest throughput and lowest latency, delivering data faster to applications and unlocking system performance capability. Mellanox offers a choice of fast interconnect products: adapters, switches, software, cables and silicon that accelerate application runtime and maximize business results for a wide range of markets including high-performance computing, enterprise data centers, Web 2.0, cloud, storage and financial services.

To find out more, visit our website: [www.mellanox.com](http://www.mellanox.com)