A SUCCESSFUL JOURNEY TO AI, ANALYTICS, HYBRID CLOUDS, AND CONTAINERIZATION DEPENDS ON THE RIGHT ENABLER: A NEW DATA PLATFORM

EXECUTIVE SUMMARY

Enterprises are aggressively pursuing strategies to harness the power of data, analytics, and AI. Simultaneously, there is a major shift in infrastructure strategies with hybrid clouds and Kubernetes containerization being adopted at unprecedented rates. This journey is a multi-step process with many different crucial decisions along the way. MapR makes it easy with its founding vision: to create a new data platform that ingests, stores, and manages data on a vast scale, while keeping it readily available to new computational techniques and tools and enabling the use of Kubernetes and hybrid clouds at production scale and reliability.

THE CASE FOR A NEW DATA PLATFORM

Data is becoming a company’s most important asset. Forward-leaning companies understand this new data economy: when they leverage data more effectively than their competitor, they win market share. Last generation and even newer point product technologies can’t solve the new demands. Furthermore, data is being democratized. The data scientist and developer are now more efficient when they can pick and choose their own tools and get to the relevant data securely and easily. Fundamentally, this requires a new underlying data platform, which can meet the needs of the data scientist for accessing innovative new tools at a rapid rate, while also providing the industrial-grade reliability and security that IT organizations insist on. As a result, there is a 30-year replatforming process that is in progress.

The problem is that it is easy to fall into the trap of choosing many point solutions and not think about the foundational importance of a data platform. Experience makes it clear that history will repeat itself with limited point solutions and more silos, unless a deliberate decision is made to architect and build on the right foundation. The success rate of applying AI and analytics at production scale can be up to 5x better than alternatives in the market, but it depends greatly on a data platform foundation optimized for mission-critical capabilities, linear scalability, elasticity, and the ability to deploy seamlessly in a hybrid-cloud world, while also harnessing the power of Kubernetes containerization for elasticity. A data platform that can support the innovation to come in the years ahead and enable an enterprise-wide data fabric becomes of paramount importance.
MAPR VISION OF A NEW DATA PLATFORM

A unique and groundbreaking approach is needed, which combines essential new tool technologies, such as Hadoop, Spark, machine learning and AI tools, while optimizing for high scale, reliability, and elasticity through containerization. Also critical is global deployment flexibility by bridging seamlessly from on-premises to the edge, or to one or more clouds. MapR architected, designed, and implemented the new data platform, using a set of principles to meet essential customer criteria for making a thoughtful data platform selection:

1. **Supports a variety of data** from big to little, structured and unstructured, in tables, streams, or files, IoT and sensor data – essentially all data types from any data source, including a range of ingest mechanisms
2. **Supports diverse computational tools and frameworks**, such as Hadoop, Spark, machine learning, TensorFlow, and Caffe
3. **Does AI and analytics applications simultaneously**, without requiring multiple clusters or silos, which means faster time to market, less maintenance engineering, and more consistent results, due to the same data sets being used by data scientists and analysts
4. **Broad range of open APIs** for no-lock-in: POSIX, HDFS, S3, JSON, Hbase, Kafka, REST
5. **Pub-sub streaming and edge first** for all data-in-motion from any data source, including IoT sensors
6. **Trusted and secure** by design: security built-in, not bolted on
7. **Reliability, security, and scale** to operate in global, mission-critical production AI and analytic applications
8. **Easy data and application movement between on-premises and in-cloud** through stateful application support with Kubernetes
9. **Operates on every cloud**, a critical must-have so that a customer can enjoy cloud economics and no cloud lock-in across clouds, including Amazon, Google, Microsoft Azure, CenturyLink, country clouds, and private clouds within their own data center
10. **Enables a global data fabric** to simultaneously ingest, store, manage, process, apply, and analyze data

**SUMMARY**

MapR has capitalized on a paradigm shift with its unique vision of an innovative new data platform to deliver exceptional value for applying AI and analytics at production scale. Customer experience from many industries and use cases around the globe have proven the underlying platform innovations to speed time to value and achieve a higher success rate of going into production for AI, analytics, edge-first, IoT applications in hybrid and multi-cloud, containerized applications. This is possible because the underlying data platform provides unequaled scale, performance, and reliability to deliver a clear business value and competitive advantage.