

Nominee: MapR

Nomination title: MapR-XD

We are in the midst of a once in a 30-year infrastructure shift. Organisations must be able to combine legacy operational data and analytical data to create intelligent applications. Yesterday's storage and data management technologies weren't designed to take advantage of distributed computing environments, cloud infrastructures, containers and virtualisation, and IoT.

Additionally, the exponential growth of data volumes and rigid infrastructures make it difficult to move data and integrate analytics with operational processes - effectively creating data silos. These silos make it challenging to derive meaning and intelligence from the data and can lead to high costs of processing and storing data, which only increases when data volumes grow.

MapR-XD is the result of years of technical innovation and collaboration with our customers to develop a mission-critical modern data platform. MapR-XD Cloud-Scale Data Store is a high-scale, reliable, globally distributed data store that creates a data fabric for managing files and containers.

MapR-XD supports the most stringent speed, scale, and reliability requirements across multiple edge, on-premises, and cloud environments. MapR-XD makes it easy to store any data at exabyte scale and supports trillions of files, provides enterprise-grade features to be the system of record for large global enterprises, and uniquely combines analytics and operations into a single platform enabling intelligent application development.

MapR-XD is software for building intelligent applications with the MapR Converged Data Platform -an industry-first platform that integrates Hadoop, Spark, and Apache Drill with real-time database capabilities, global event streaming, and scalable enterprise storage to power a new generation of big data applications.

The new MapR-XD Cloud-scale Data Store includes:

- **Files, container support** - MapR-XD eliminates data silos and simplifies management across files and containers. MapR-XD provides unified security, data protection and high availability across diverse data types. The same underlying data can be accessed through a wide range of industry standard APIs including NFS, POSIX and HDFS to simplify development, administration and eliminate data sprawl.

- **Global exabyte scale - MapR-XD easily scales to support trillions of files, exabytes of data, on thousands of commodity servers or cloud instances, all accessible through a single global namespace. Additionally it reduces operational complexity and provides a single, scalable view of resources, simplifying access for users, applications and containers.**
- **Cloud-grade reliability - MapR-XD delivers high availability, data protection and disaster recovery with no single points of failure, fully distributed metadata, point-in-time snapshots and high-performance, distributed mirroring.**
- **Speed at scale with flash - MapR-XD utilises the full power of network interconnects and takes advantage of the available performance of underlying commodity hardware, such as disk and flash to meet the demands of graphics processing unit (GPU)-based architectures. Automated capabilities, such as logical partitioning, parallel processing for disparate workloads, and bottleneck avoidance with I/O shaping and optimisations, ensure maximum performance across a cluster. MapR-XD includes an extremely high-performance POSIX Client that provides up to 10x the performance of a typical Network File System (NFS) gateway.**
- **Stateful persistence for containerised applications - MapR-XD includes a secure, optimised container client for providing containers with access to persisted data. The client supports both legacy and new containerised event-based microservices applications; multiple data types of files, containers, database and event streams; works with multiple schedulers such as kubernetes, mesos and docker swarm; and across any infrastructure such as on-premises, multiple clouds and edge.**
- **Flexibility to leverage multiple infrastructures - MapR-XD supports edge, on-premises and cloud environments with the same platform. It enables multi-temperature capabilities across flash, disk and cloud tiers with support for containers and automated data movement to address performance, cost and compliance concerns.**
- **IoT Edge made easy - MapR-XD for the edge provides the ability to deploy processing and storage capabilities close to an IoT data source, such as in a car, medical device or jet engine. MapR-XD can store and process machine or sensor-generated data for seamless integration with a centralised Converged Data Platform where global aggregation and analysis would be performed.**

- **Extensible architecture - MapR-XD is a powerful component of the Converged Data Platform enabling customers to easily and seamlessly leverage additional capabilities including database, stream processing and integrated analytics on the same platform.**

MapR-XD includes the MapR multi-temperature Global Namespace and data management in the form of security, compression, snapshots, multi-tenancy, and self-healing. MapR-XD is delivered via either flash or disk.

One example of a successful deployment of MapR-XD is SAP Digital Interconnect, who chose MapR-XD as the underlying storage layer with SAP IQ because of its scalability, flexibility, speed, and cost.

SAP Digital Interconnect provides anonymised data on 50 billion events, roughly 40-50 terabytes of uncompressed data, each day. That's a yearly data volume of 1.3PB at a 1:10 compression ratio. This data needs to be cleaned up and normalised in order to provide customers with meaningful insights. SAP Digital Interconnect needed a cost effective storage solution that met its high SLAs.

MapR-XD provided essential scalability and flexibility, as MapR-XD can be scaled up or down as necessary, which SAP has taken advantage of over the last year. In addition, other system components can be upgraded or retrofitted without bringing the system down, so when additional storage is needed, more nodes are simply added.

MapR-XD also delivered on speed, with Steven Garcia, Head of Engineering and Operations for Cloud Solution Services, SAP reporting that "Even when searching against one trillion records, ten minutes is too long. With MapR, reports are happening in a fraction of that time." Finally, SAP Digital Interconnect estimates MapR-XD costs one-tenth that of commodity arrays.

Headquartered in San Jose, Calif., MapR provides the industry's only Converged Data Platform that enables customers to harness the power of big data by combining analytics in real-time to operational applications to improve business outcomes. With MapR's Converged Data Platform and wider product portfolio, enterprises have a data management platform for undertaking digital transformation initiatives to achieve competitive edge. Amazon, Cisco, Google, Microsoft, SAP, and other leading businesses are part of the global MapR partner ecosystem.

MapR-XD has been available globally since June 6th 2017.

Why nominee should win

- **MapR provides architectural innovations that provide critical success to a business**
- **MapR-XD Cloud-Scale Data Store is a high-scale, reliable, globally distributed data store that creates a data fabric for managing files and containers**
- **MapR-XD supports the most stringent speed, scale, and reliability requirements across multiple edge, on-premises, and cloud environments**
- **MapR-XD is software for building intelligent applications with the MapR Converged Data Platform, the proven solution for delivering business value in data-driven companies**
- **SAP Digital Interconnect was able to gain speed and scalability, while saving money, with MapR-SD**