

NetApp Plus Mosaic Application-Centric Data Protection for DataStax and Apache Cassandra

KEY BENEFITS

Minimize Application Downtime

- Near-zero recovery point and recovery time objectives (RPOs and RTOs)
- Single-click orchestrated recovery
- Granular table-level recovery

Increase Operational Efficiency

- Flexible policy-based management
- Recovery to different topology database clusters (testing and development)
- Automated failure handling

Increase Storage Efficiency

- Industry-first semantic deduplication
- Advanced support for compacted database tables

Simplify Deployment

- On-premises or public cloud use
- Compatibility with NetApp® ONTAP® 9 software
- Performance and availability benefits of NetApp arrays

THE DATA PROTECTION CHALLENGE

Rapid proliferation in social, mobile, cloud, and Internet of Things is driving enterprises to deploy hyperscale, distributed, data-centric applications such as customer analytics, e-commerce, security, surveillance, and business intelligence. To handle the data requirements of these high-volume, high-ingestion rate, and real-time applications, enterprises are rapidly adopting massively scalable and nonrelational databases such as DataStax Enterprise (DSE) or Apache Cassandra.

Like any business-critical application, these databases require data protection, including application-consistent backup; near-zero RPO and RTO; granular, repair-free recovery; failure handling; and backup storage efficiency. However, given the hyperscale, distributed nature of these databases, traditional backup and recovery products don't support these requirements, leaving a critical data protection gap.

THE SOLUTION: APPLICATION-CENTRIC DATA PROTECTION FROM RUBRIK MOSAIC

NetApp and Rubrik have partnered to extend the NetApp Data Fabric by providing application-centric data protection for next-generation applications. Enterprises can now leverage the performance, reliability, and flexibility of ONTAP 9 with the continuous data protection capability of Rubrik Mosaic. With this solution, enterprises can scale business-critical applications on a Cassandra database and be confident in the recoverability of data and the ability to maintain high application uptime. Mosaic is an industry-first scale-out, software-only data protection solution that is purpose-built for next-generation applications that are deployed on DSE or the Apache Cassandra database.

Scale-out architecture

Rubrik Mosaic is founded upon Consistent Orchestrated Distributed Recovery (CODR), Rubrik Mosaic's cloud-first, scale-out data management architecture that enables customers to harness the cloud for both next-generation data protection and management.

CODR uses elastic compute services that can be auto-scaled with load and removes the dependency on media servers. CODR also transfers data in parallel to and from file-based and object-based secondary storage for multiple workloads, including data protection and testing and development. CODR offers:

- High availability. Deploying Mosaic in a cluster (three-node) configuration prevents any software process or external hardware (node) failures from compromising backup and recovery operations.
- Enhanced backup performance. With the scale-out architecture of DSE or the Apache Cassandra database, users can easily scale their database according to application growth. Similarly, the scale-out capability of Mosaic with no reliance on media servers helps customers increase backup and recovery throughput, leading to lower RPO.

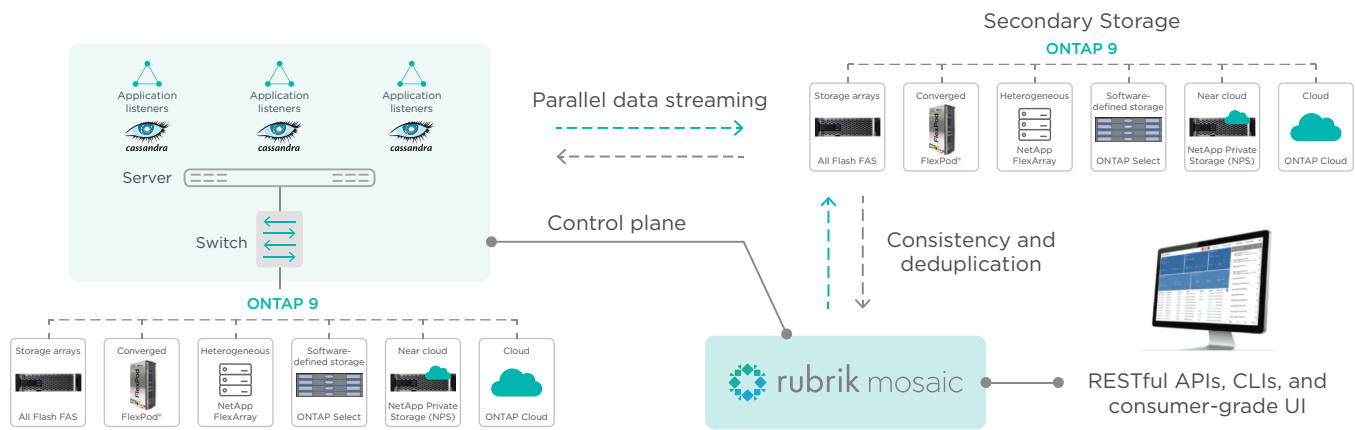


Figure 1 - Reference deployment of MongoDB with NetApp and Rubrik Mosaic.

Scalable versioning

By using native application intelligence, Mosaic creates an application-consistent point-in-time backup of the Apache Cassandra database at user-specified intervals, a concept that we call cluster-consistent versioning. A cluster-consistent version contains all the records that have achieved user-specified consistency. As a result, no repairs are needed when a version is restored back to the cluster, thus minimizing RTO. The backup operation is also highly parallel in its nature; Mosaic acts only as a control plane that orchestrates data movement from the data source cluster to version (secondary) storage. This approach allows Mosaic to handle large Cassandra databases and application workloads.

By allowing administrators to generate database backups at any user-specified time interval and at any granularity (table-level or entire database), Mosaic also simplifies operational use. Overall, versioning helps reduce data loss risk and minimizes capital and operating expenditures for an enterprise.

Fully orchestrated and granular recovery

Rubrik Mosaic provides single-click, fully orchestrated, any-point-in-time recovery. With Mosaic, customers can recover data directly back into the same database cluster (operational recovery) or to a different cluster (testing-and- development refresh) with a different topology (number of database nodes). This option reduces the operational burden of refreshing testing and development clusters for continuous development DevOps environments. Further, the recovery process deals only with the logical data, making it three times faster than with traditional approaches. During recovery, the data is directly (with no reliance on media servers) transferred from secondary storage into target databases, resulting in a very low RTO.

Industry-first semantic deduplication

Semantic deduplication is an industry-first capability that Rubrik has developed to reduce the cost of storing backups of nonrelational databases over their retention period. Most nonrelational databases keep multiple copies of the primary data, called replicas. As part of versioning, Rubrik Mosaic deduplicates all the replicas of a primary dataset, thus greatly increasing storage savings without losing native formats. For example, if the database uses a replication factor of 3 (RF = 3), Mosaic saves up to about 70% in secondary storage costs. And by using its application awareness, Mosaic optimizes the backup of compacted SSTables, resulting in significant additional secondary storage savings.

Why NetApp Plus Mosaic

NetApp big data solutions deliver an open, scalable storage system for building big data applications. Customers gain business insights—and value—more quickly. The NetApp All Flash FAS system offers predictable and high performance with consistent and low latency, resulting in a very fast response time to the most demanding applications that are deployed on Cassandra. And to deliver maximum uptime and high availability, it also offers non-disruptive operations and integrated data protection across applications, virtual infrastructures, and cloud architectures. Together with Rubrik, NetApp offers an industry-leading enterprise data management solution for Apache Cassandra that is based on ONTAP 9 and Rubrik Mosaic.

Rubrik Mosaic is a software-only data protection product that can be deployed on a physical server, a virtual machine, or any cloud compute instance (for example, Amazon EC2). It communicates with the Apache Cassandra database through a Secure Shell (SSH) connection that forms a control plane to orchestrate data movement. The data can be backed up to a secondary FAS array

or to a NetApp E-Series system. In addition to CLIs and RESTful APIs, customers can use the Mosaic consumer-grade UI to manage their data protection environment.

About NetApp

Leading organizations worldwide count on NetApp for software, systems and services to manage and store their data. Customers value NetApp's teamwork, expertise and passion for helping them succeed now and into the future.

www.netapp.com



Global HQ

1001 Page Mill Rd., Building 2
Palo Alto, CA 94304
United States

1-844-4RUBRIK
inquiries@rubrik.com
www.rubrik.com

Rubrik delivers a single platform to manage and protect data in the cloud, at the edge, and on-premises. Enterprises choose Rubrik's Cloud Data Management software to simplify backup and recovery, accelerate cloud adoption, and enable automation at scale. As organizations of all sizes adopt cloud-first policies, they rely on Rubrik's Polaris SaaS platform to unify data for security, governance, and compliance. For more information, visit www.rubrik.com and follow [@rubrikInc](https://twitter.com/rubrikInc) on Twitter.

20190614_v1