

# Optimize MongoDB with Dell EMC PowerStore

MongoDB® is a modern NoSQL database that uses a document-based data model to store both structured and unstructured data. It is designed to be highly scalable and is capable of processing massive amounts of data efficiently. A MongoDB database can scale up to hundreds of systems with petabytes of data distributed across them. With a modern database architecture comes the need for modern storage and application-driven infrastructure engineered to optimize and consolidate existing and new business use cases of the future. The Dell EMC™ PowerStore™ storage platform, together with the latest capabilities of MongoDB, introduces AppsOn and a new era of onboard application support.

## Dell EMC PowerStore X model



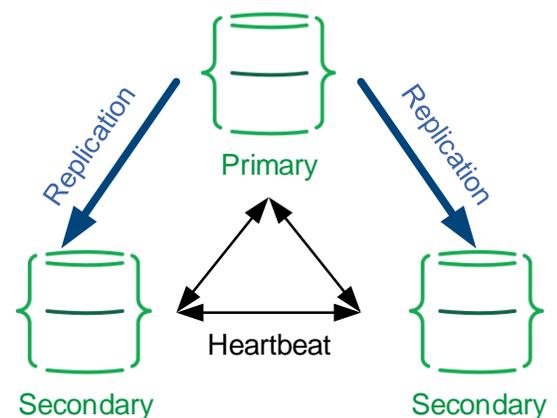
### Highlights

- Predictive analytics with built-in machine learning engine
- Adaptable architecture for operational simplicity and agility
- Performance for modern database architecture
- Data-centric design to optimize and consolidate mixed workloads with ease
- Flexible scaling for distributed database workloads
- Adaptable with next-generation media and non-disruptive hardware upgrades
- Co-locate applications directly on PowerStore bringing apps closer to the data
- Outstanding data efficiency with thin provisioning, snapshots, and up to 4:1 advanced data reduction technologies
- Mission-Critical availability for high value business analytics deployments
- Enterprise-class data encryption using 256-bit AES
- Tight integration with industry-leading VMware® ESXi™ software - embedded in PowerStore X model appliance

## A modern database with high availability and data protection

MongoDB is engineered with replica sets to increase data availability and fault tolerance of the MongoDB servers. Full copies of the data are replicated to multiple secondary members. A single replica set supports up to 50 members. Using a larger number of replicas increases data availability and protection. It also provides automatic failover of the primary member during planned or unplanned events such as server updates, server failures, rack failures, data center failures, or network partitions. Replicating the data to a different server in a different data center further increases data availability and data locality for distributed clients.

Deploying a high number of replicas is typical in an environment where commodity hardware is used and is vulnerable to hardware failures and where service interruptions can occur. Too many replicas can lead to lower storage efficiencies, higher network bandwidth usage, and increased management complexities.



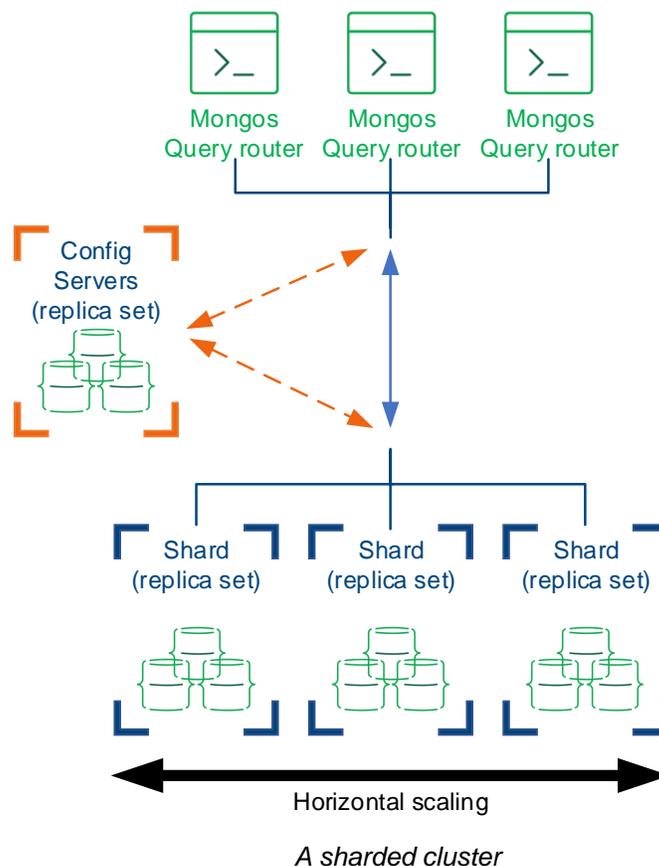
*A three-member replica set*

The primary member is responsible for the write and read operations for the replica set by default. The clients can specify a read preference to send read operations to the closest secondaries as well. A data-bearing member can also be configured to be hidden to serve as a backup copy if needed.

As the workload grows, the primary or secondaries need to be able to scale their processing capacity by adding additional CPU, memory, or storage. In a read-oriented workload, additional secondaries with a full copy of data might be needed. To increase the data durability and to avoid data to be rolled back when a primary member fails over, a write majority concern can be specified and enable journaling on all voting members. This guarantees data is replicated to most of the replica set members before returning acknowledgment to the client.

### Building a flexible scale-out distributed database architecture

With large data sets and high throughput environments, MongoDB uses a process called sharding to distribute data across multiple systems to increase storage capacity, throughput and improve performance. A sharded cluster consists of three components: shard, mongos, and config servers. The shards hold a subset of the data and are deployed as a replica set. The mongos process communicates with the config servers and routes the client requests to the appropriate shards. The config servers store the metadata for the cluster configuration settings. The config servers are deployed as a replica set. In a non-sharded database, there is only one primary member in a replica set that is responsible for write operations. However, in a sharded cluster, each shard can perform write operations respective to its dataset.



### Modern pluggable storage platform engines

MongoDB supports a wide variety of traditional and business-critical workloads including both operational and real-time analytics workloads. MongoDB's pluggable storage architecture extends new capabilities to the storage platform depending on the different workloads. These storage engines are responsible for storing the data and on how the data is stored. Starting with version 4.2, MongoDB supports various storage engines including the WiredTiger storage engine, the in-memory engine, and the Encrypted Storage Engine. The MMAPv1 storage engine has been deprecated in version 4.2.

**WiredTiger storage engine** is the default and preferred storage engine for most workloads. It persists data on disk and provides features such as document-level concurrent model, journaling, checkpoints and compression.

**In-memory storage engine** stores the dataset in the memory to reduce data access latency but does not persist data on disk. It is available only in the MongoDB Enterprise Edition.

**Encrypted storage engine** is the native encryption option for the WiredTiger storage engine. It provides encryption at rest and is only available in the MongoDB Enterprise Edition.

It is possible to mix the different engines based on use case in the same replica set to optimize and meet the needs of specific application requirements and most benefits the specific engines. For example, combine the in-memory engine for ultra-low latency operations with the WiredTiger engine for on-disk persistence.

## The advantages of MongoDB on Dell EMC PowerStore

PowerStore achieves new levels of operational simplicity and agility. It uses a container-based microservices architecture, advanced storage technologies, and integrated machine learning to unlock the power of your data. PowerStore is a versatile platform with a performance-centric design that delivers multidimensional scale, always-on data reduction, and support for next-generation media.

PowerStore brings the simplicity of public cloud to on-premises infrastructure, streamlining operations with an integrated machine-learning engine and seamless automation. It also offers predictive analytics to easily monitor, analyze, and troubleshoot the environment. PowerStore is highly adaptable, providing the flexibility to host specialized workloads directly on the appliance and modernize infrastructure without disruption. It also offers investment protection through flexible payment solutions and data-in-place upgrades.

## AppsOn brings MongoDB closer to the infrastructure and storage

Bringing applications closer to the data increases density and greatly simplifies infrastructure operations. AppsOn integration with VMware vSphere® results in streamlined management where storage resources plug directly into the virtualization layer. The use of VMware as the on-board application environment results in unmatched simplicity, since support is inherently available for any standard VM-based applications. When a new PowerStore is deployed in the cluster, the VASA provider is automatically registered and the data store created, eliminating manual steps and saving time. PowerStore seamlessly integrates the VMware ESXi software into the same hardware. Two ESXi nodes are embedded inside the appliance which has direct access to the same storage resources. This close integration allows applications like MongoDB to take full advantage of server and storage virtualization with simplified deployment and management. Running MongoDB directly on a PowerStore model appliance brings the applications closer to the data and reduces latency and network bottlenecks.

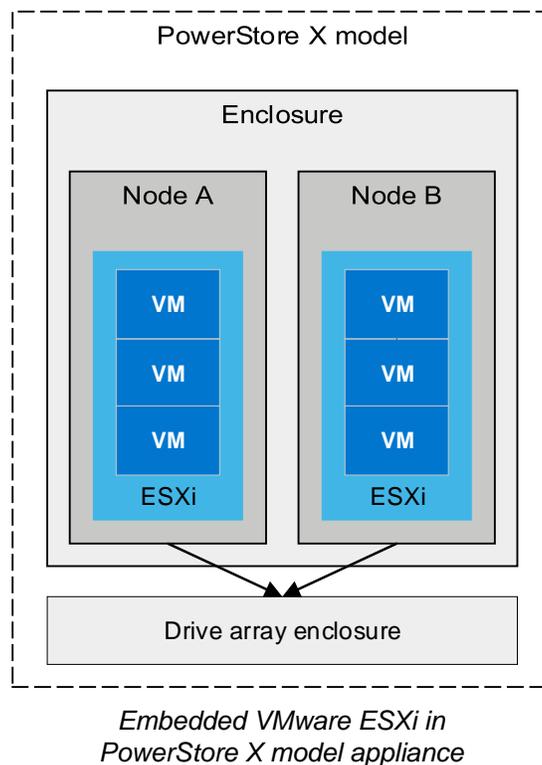
## Agile infrastructure, flexible scaling on a high performing storage and compute platform

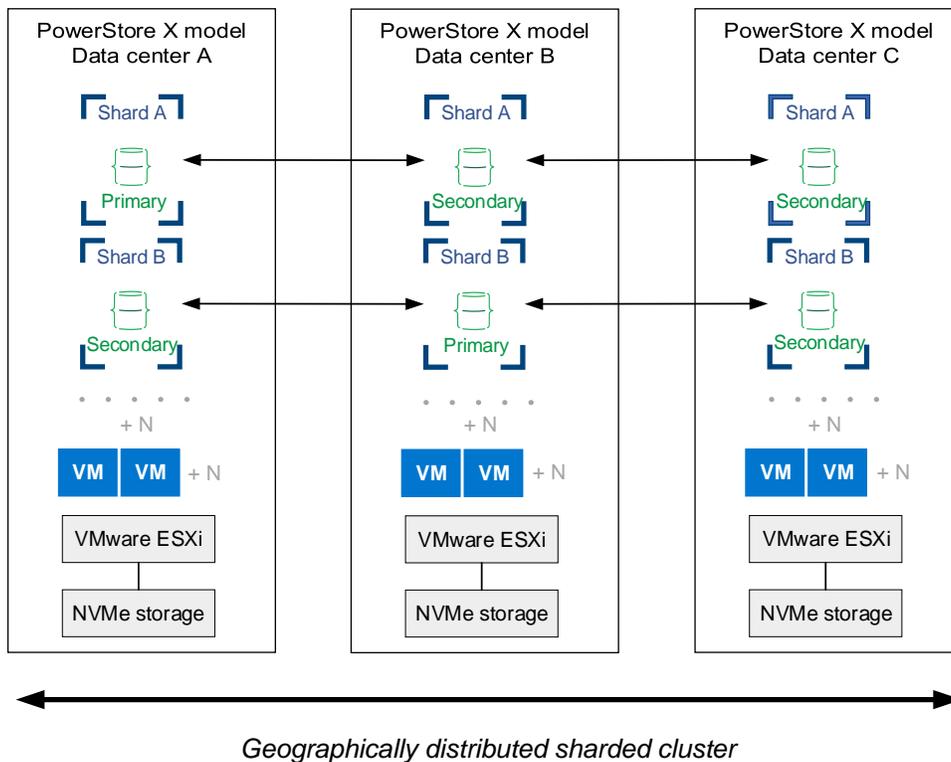
PowerStore provides flexible scaling with ease of management that greatly complements the MongoDB scale-up and scale-out distribution model. The built-in hypervisor dynamically scales up the replica set members when the workload requires it, while new replica sets, or shards, can be provisioned rapidly on the same or on additional appliances in a different location.

A single PowerStore appliance can scale up to 112 vCPUs, 2.5 TB of memory, and close to 900 TB raw storage capacity. The NVMe architecture is designed for the next-generation NVMe-based storage and low overhead NVRAM cache. PowerStore is powerfully engineered to handle the most demanding MongoDB mixed workloads.

## Mission-critical high availability and fault tolerant MongoDB platform

At the hardware level, PowerStore is designed to be highly available and fault tolerant. It monitors the storage devices continuously and will automatically relocate data from failing devices to avoid data loss. The PowerStore X model appliance includes two ESXi nodes and redundant hardware components. The non-disruptive upgrade (NDU) feature further maximizes overall PowerStore availability. NDU supports PowerStore software, hotfixes, hardware, and firmware.





To support high-value business workloads and the service requirements on the application level, it is essential to protect and ensure the availability of the primary member of a replica set. When the primary member of a replica set becomes inaccessible, the replica set cannot process any write operations until the primary member recovers or a new primary is elected. Furthermore, the election requires the majority of the members to be available.

With standard VMware vSphere High Availability (HA) integrated into PowerStore, the embedded VMware ESXi automatically restarts or migrates failed MongoDB servers to a different ESXi

node to resume operations. This helps to restore the MongoDB to full operation capacity and minimizes the chance of the database going offline or read-only.

To achieve an even higher level of redundancy and application availability, the MongoDB replica set and sharded cluster can be deployed across multiple PowerStore appliances in different data centers. PowerStore improves MongoDB availability and provides unparalleled flexibility and mobility to relocate and move across data centers and appliances.

### PowerStore inline data reduction reduces storage consumption and cost

As business data continues to grow, big data has become a critical component in the business analytics world. A tremendous amount of data is pulled from all kinds of sources continuously and run through cloud-scale applications like MongoDB to gain insights into customers and businesses. When putting MongoDB replica sets on PowerStore, the always-on inline data reduction feature greatly reduces the actual storage used but still maintains the application data availability and protection expected from MongoDB.

### Efficient and convenient snapshot data backup

Dell EMC PowerStore provides MongoDB additional data protection with the capability of array-based snapshots. A PowerStore snapshot is a point-in-time copy of the data. They are space efficient and take seconds to create. Snapshot data are exact copies of the target data and can be used for application testing, backup, or DevOps. Because of the tight integration with VMware vSphere, PowerStore is capable of taking VMware vSphere Virtual Volumes™ (vVols) VM snapshots directly from PowerStore Manager using a protection policy schedule or on-demand. The VM snapshot information can be viewed in PowerStore and vCenter. .

## Secure data protection with ease of mind

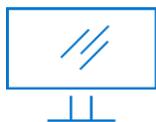
With high-value data driving business applications, data security is a top concern for all businesses. Lost or stolen data can seriously damage business reputation and result in huge financial costs and loss of customer trust. Dell Technologies™ engineered PowerStore with Data at Rest Encryption (D@RE) which utilizes self-encrypting drives and supports array-based self-managed keys. Once activated, data is encrypted as it is written to disk using the 256-bit Advanced Encryption Standard (AES). PowerStore D@RE provides this data security benefit to MongoDB while eliminating application overhead, performance penalties, and administrative overhead typically associated with software based solutions.

## Unified infrastructure and services management

PowerStore provides deep integration with VMware management tools and services with Virtual Storage Integrator (VSI), vRealize® Operations Manager (vROps), vRealize Orchestrator (vRO), and Storage Replication Adapter (SRA). Customers can easily incorporate PowerStore X ESXi into their existing vCenter and manage all VMware infrastructure and services from a unified management platform.

## MongoDB value and future expansion

New business analytics applications like MongoDB are fundamentally changing the way data is used to support the business. The blend of massive amounts of data and technology innovation provides the opportunity for businesses to transform. As the value and scale of this data grows, the need to provide a future-proof platform that can easily be managed, provides technical innovation for future growth, and is engineered to support the application architecture is critical. MongoDB on Dell EMC PowerStore combine to bring IT organizations the ability to be agile, efficient, and responsive to business demands.



[Learn more](#) about Dell  
EMC PowerStore  
solutions



[Contact](#) a Dell  
Technologies expert