

Building High Performance Splunk SmartStores with MinIO

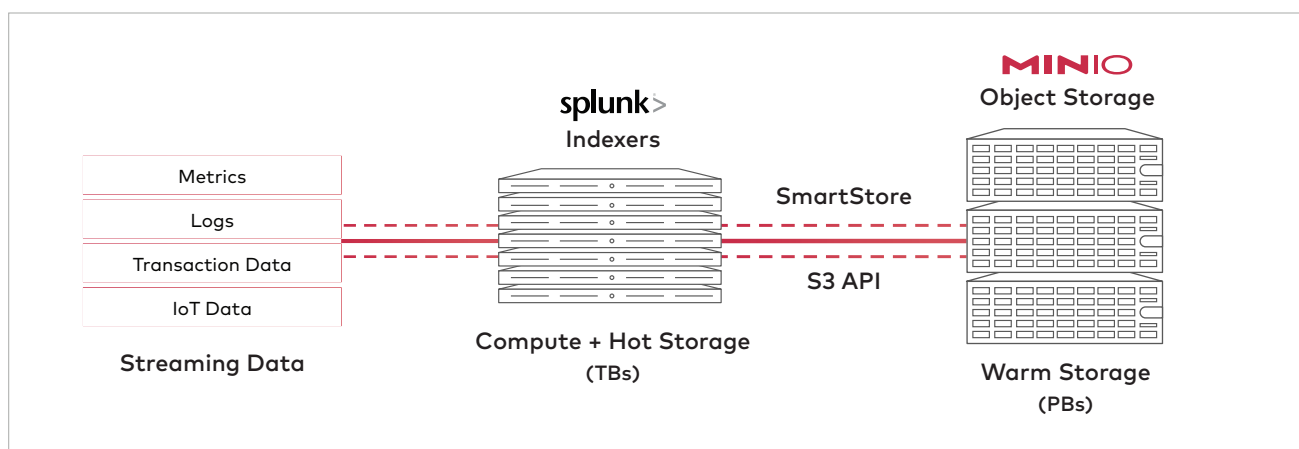
SmartStores built with MinIO object storage are performant, cloud-native, secure and resilient - all the reasons Splunk uses MinIO in its own products.

Splunk's massive success as an analytics platform created distinct challenges for the enterprise. As organizations elected to source more data and to keep that data longer - both from an insight generation and compliance perspective - they created material economic challenges under Splunk's traditional architecture.

Splunk responded by introducing SmartStore, a new, disaggregated architecture featuring a distributed scale out model that allows compute (Splunk Indexers) and storage to scale independently. By intelligently tiering data, SmartStore can deliver vastly superior economics without sacrificing performance.

The efficiency of the SmartStore model depends on the scalability and availability of the remote data. These requirements make MinIO an ideal SmartStore endpoint. Designed for high-performance, peta-scale architectures, the open-source, S3-compatible object store is specifically targeted at the private-cloud.

- Fastest SmartStore solution with 55 GB/s+ read speeds
- The most Kubernetes friendly SmartStore offering
- Superior encryption in flight and at rest
- Inline erasure coding and bitrot protection offer superior resiliency
- Designed for seamless multi-PB scale



MinIO object storage is performant, cloud-native, secure and resilient while offering superior economics. Of all the SmartStore alternatives, only one, MinIO, is deployed inside Splunk's own product suite*. The following features outline why MinIO is the object store of choice for Splunk SmartStores.

Erasure Coding

MinIO protects data with per-object, inline erasure coding which is written in assembly code to deliver the highest performance possible. MinIO's implementation ensures that objects can be read or new objects written even if up to half the drives are lost or unavailable - delivering resiliency in a fraction of the disk space that traditional replication would require. MinIO's erasure code implementation delivers economic, performance and resiliency benefits over both traditional approaches and other SmartStore solutions.

Bitrot Protection

Silent data corruption, or bitrot, is a serious problem faced by the corruption of disk drives without the user's knowledge. MinIO's optimized implementation of the HighwayHash algorithm ensures that Splunk's Indexers will never read corrupted data - it captures and heals corrupted objects on the fly. This capability is critical for the types of investigative use cases that Splunk generally runs.

Encryption

It is one thing to encrypt data in flight, it is another to protect data at rest. MinIO supports multiple, sophisticated server-side encryption schemes to protect data - wherever it may be. Server side and client side encryption are supported using AES-256-GCM, ChaCha20-Poly1305 and AES-CBC. Encrypted objects are tamper-proofed with AEAD server side encryption. Given the exceptionally low overhead, auto-encryption can be turned on for every application and instance.

Continuous Replication & Lambda Compute

To protect against data center failures, Splunk can be configured for multi-site indexer clustering. Multi-site indexer clustering with SmartStore requires support for cross-site replication between physical object stores. MinIO's continuous replication designed for large scale, cross-data center deployments makes it the perfect choice for Splunk remote object store. By leveraging Lambda compute notifications and object metadata, MinIO can compute the delta efficiently and quickly, keeping the data loss to a bare minimum should a failure occur - even in the face of Splunk's highly dynamic datasets.

Scaling with Global Federation

MinIO allows disparate instances to be combined to form a unified global namespace. Any number of MinIO servers can be combined into a Distributed Mode set and multiple Distributed Mode sets can be combined into a MinIO Server Federation. The result is a SmartStore that can scale massively for large, geographically distributed enterprises.

Multi-Cloud Gateway

Making SmartStore data available, wherever it may reside, is one of the key challenges that MinIO addresses. For SmartStore users this can be important if they choose to back up to the public cloud. MinIO's compatibility with AWS S3 have made it the most used S3 gateway and can make any cloud - even those without S3 capabilities look like an S3 bucket.

Performance

MinIO is designed for high performance workloads and is a natural fit for the Splunk SmartStore. With NVMe speeds in excess of 55GB/s on an 8 node configuration and HDD speeds in excess of 18GB/s on a 16 node configuration, MinIO is the fastest object store in existence. This means less waiting and more learning for the teams using Splunk.

Simplicity

MinIO can be installed and configured within minutes simply by downloading and executing a single binary. The amount of configuration options and variations are kept to a minimum resulting in near-zero system administration tasks and fewer paths to failures. Upgrading MinIO is done with a single command which is non-disruptive and incurs zero downtime - lowering total cost of ownership.

Software Defined

MinIO is software-defined and runs on a broad range of standard hardware meaning the enterprise can optimize their cost/performance tradeoff at the hardware level. In addition, MinIO is 100% open source under the Apache V2 license. This means that MinIO's customers are free from lock in, free to inspect, free to innovate, free to modify and free to redistribute.

Cloud Native

Built from scratch over the last four years, MinIO offers the highest level of interoperability with modern, cloud-native technologies such as Docker

(280M+ pulls), Kubernetes and other microservices. This cloud orientation, coupled with MinIO's software-defined approach and open source licensing provide enterprise-class certainty that is measured in decades.

MinIO and Splunk: Better Together

Deploying a Splunk SmartStore with MinIO has multiple benefits. First, with 80-90% of the data moving out to the SmartStore, the enterprise can save 60-75% in infrastructure costs. Second, by selecting MinIO, those cost savings come without any performance degradation. Third, the scalability and economics of MinIO allows enterprises to collect more data and hold that data longer, increasing the likelihood of unique insights - enhancing the value of the overall Splunk investment. Finally, because MinIO is open source and software-defined, enterprises can tailor their SmartStore to their performance and economic requirements.

*MinIO is built into Splunk's Digital Stream Processor product.

World's Fastest Growing Object Storage System



18.9K+

GITHUB STARS

281.9M+

DOCKER PULLS

6006

SLACK MEMBERS

486

CONTRIBUTORS