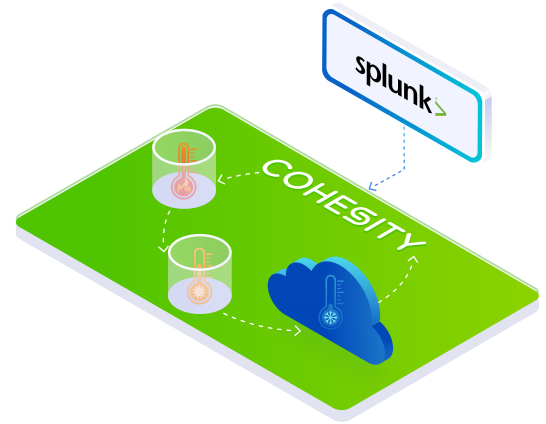


Cohesity SmartFiles: Simplify Splunk Data Management



Key Benefits

- Cohesity SmartFiles with Splunk reduce data protection costs with a space and cost-efficient storage tier
- Scale-out platform with unlimited scalability, nondisruptive upgrades, pay-as-you grow scalability
- Multiprotocol access supporting NFSv3, CIFS, SMB2.x, SMB 3.0, S3 and OpenStack Swift APIs
- Global deduplication and compression, with Google-like search across all file and object metadata
- Public cloud integration—Amazon Web Services, Microsoft Azure, and Google Cloud Platform—for archival, tiering, and replication

A number of sources, including machine data and artificial intelligence, are driving unstructured data growth. As data volumes continue to grow, knowing what data exists and where it resides is critical. So is how to search, analyze, and visualize the unstructured and machine-generated data generated by websites, applications, and IoT devices, which can be analyzed for more meaningful insights at different stages of the data lifecycle. That’s why Cohesity and Splunk are so valuable.

Cohesity and Splunk

Cohesity SmartFiles simplifies the management of data and apps at web-scale, on a single software-defined platform. Combined with Splunk, they create a powerful, purpose-built platform to make machine data accessible, usable, and valuable at scale to businesses, and to simplify the management of massive amounts of data and help enterprises achieve their data retention goals.

The Splunk Data Lifecycle

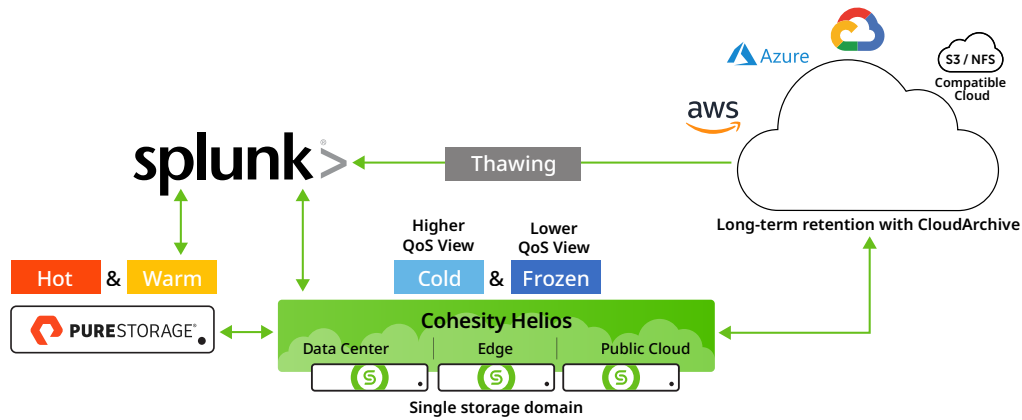
Splunk helps administrators gain operational intelligence from machine data to support better customer service. In Splunk, events are stored as raw data in a compressed format and indexes are added that point to the raw data. Together, these files constitute the Splunk Enterprise Index, also known as the indexer, with data now transformed into searchable events.

The Splunk Enterprise index typically consists of many buckets, organized by age of the data. Some directories contain newly indexed data, while others contain previously indexed data. The number of directories can become large, depending on how much data is being indexed. As data ages, Splunk’s indexer gracefully handles and moves data through the following five stages in a transparent, policy-driven mode—where policies are set as attributes in the index configuration until by default at a preset time, it removes old data from the system:

Bucket Stages	Description	Searchable?
Hot	Contains newly indexed data. Open for writing. One or more hot buckets for each index.	Yes
Warm	Data rolled from hot. There are many warm buckets. Data is not actively written to warm buckets.	Yes
Cold	Data rolled from warm. There are many cold buckets.	Yes
Frozen	Data rolled from cold. The indexer deletes frozen data by default, but enterprises can choose to archive it instead. Archived data can later be thawed.	No
Thawed	Data restored from an archive. If enterprises archive frozen data, they can later return it to the index by thawing it.	Yes

Ever increasing regulatory and industry requirements surrounding data retention mean that enterprises indexing large amounts of data cannot simply allow default processes to proceed as scheduled. Instead, enterprises must carefully plan and orchestrate specific aging policies for data





depending on the bucket stage of the data in the Splunk platform. However, customization of the indexer flow presents IT teams relying only on Splunk for data storage with a new challenge: where to store data for varying lengths of time and how to quickly retrieve it when needed?

Storing Cold and Frozen Splunk Buckets in SmartFiles

SmartFiles overcomes variable-length data storage obstacles by providing enterprises with a scale-out platform designed to converge—data silos, including backup and recovery infrastructure and target storage. Because the paths for the hot, warm, cold, and thawed directories are configurable, to achieve optimal access and streamlined compliance enterprises can store buckets in one stage in a separate location from buckets in other stages.

SmartFiles is the ideal solution to store cold and frozen buckets of Splunk, while providing complete visibility to the Splunk administrator via the Splunk interface. Concurrently, hot and warm data can be stored in primary storage arrays, such as Pure Storage, and accessed via Splunk. As data becomes cold, it will automatically move to a Cohesity View with high quality of service (QoS) settings and be accessible via NFS endpoints.

SmartFiles simplifies cold and frozen data storage by providing globally distributed NFS, SMB, OpenStack Swift and S3 object storage on Cohesity Helios, a revolutionary web-scale platform. NFS, SMB, and S3 volumes are provisioned as Views on SmartFiles and leverage all of the Helios capabilities including scale-out storage, global variable-length deduplication and compression, and unlimited snapshots and clones.

Cohesity SmartFiles capabilities for storing Splunk buckets include the following:

Capacity optimization	Globally, variable-length deduplicated data is distributed across all nodes
Global search	Powerful indexing of all files and object metadata to enable global Google-like search
Quotas	Volumes, file shares, and object buckets
Security	Software-based encryption
Cloud Integration	Native support for Amazon Web Services, Microsoft Azure, and Google Cloud Platform for policy-based archival, tiering, and replication

Finally, as cold data programmatically moves in the Splunk index to the frozen bucket, the data will move to a Cohesity View with low QoS settings, which can still be accessed via NFS endpoints. However, data in a frozen bucket cannot be searched and requires manual intervention to be either deleted or archived. Enterprises opting for archival can take advantage of Cohesity CloudArchive for long-term retention and archival to reduce reliance on tape and lower TCO, while also gaining an easy way to retrieve data back on-premises or recover data to a different site.

By combining Cohesity SmartFiles and Splunk, enterprises can uncover more value from all their data while achieving data retention goals and meeting regulatory requirements. If your enterprise is already using Splunk, SmartFiles is the ideal complementary storage solution.

Learn more at www.cohesity.com



© 2021 Cohesity, Inc. All rights reserved.

Cohesity, the Cohesity logo, SnapTree, SpanFS, DataPlatform, DataProtect, Helios, and other Cohesity marks are trademarks or registered trademarks of Cohesity, Inc. in the US and/or internationally. Other company and product names may be trademarks of the respective companies with which they are associated. This material (a) is intended to provide you information about Cohesity and our business and products; (b) was believed to be true and accurate at the time it was written, but is subject to change without notice; and (c) is provided on an "AS IS" basis. Cohesity disclaims all express or implied conditions, representations, warranties of any kind.