DataCore Swarm

The DataCore Swarm system is a mature object storage system for private and public cloud usage and as a target archive system. Originally called CAStor, Swarm was developed by Caringo, acquired by DataCore in 2021. Swarm is delivered as software for x86 servers, or on Swarm Server hardware appliances. The scale out architecture allows the configuration to dynamically scale very large with commensurate performance. Swarm has implemented a symmetric parallel architecture that operates to manage the nodes as a distributed cluster. Nodes are automatically discovered and rebalancing occurs after booting from a network. All executing software is maintained in RAM on each node to enhance performance.

The system includes SwarmFS to provide file access to the underlying object storage using NFSv4 or SMB. FileFly is used to migrate data to Swarm from Windows and Linux file servers and NetApp ONTAP systems over CIFS/SMB or NFS.

Direct object access is accomplished by use of the S3 API, and a superset of S3 called Simple Content Storage Protocol (SCSP), which conforms to HTTP1.1 and an HDFS connector. Geographic data distribution and protection is standard through sub-clusters on the same latency network. For remote clusters, Feeds (routing function) automates the routing of objects for data distribution and data protection. Feeds are latency tolerant and take advantage of the full cluster-to-cluster capability of the system.

Key competitors include IBM Cloud Object Storage, Scality Ring, Cloudian HyperStore, Dell EMC Elastic Cloud Storage, and NetApp StorageGRID WebScale.

Evaluator Group Coverage

Related coverage on the Evaluator Series Research website includes:

Object Storage matrix
Object Storage Evaluation Guides
DataCore Swarm Product Brief
Competitive Product Analysis and Product Briefs

Highlights

- Characteristics
  - Swarm is delivered as software to install on X86 servers and can operate with various hardware generations. Also available as a series of Swarm Server appliances.
  - Additional functions are delivered with software systems: the SwarmNFS for support of file access over NFSv4 and FileFly for migration over S3 for files in Windows and Linux file servers and NetApp ONTAP NAS systems.
  - The maximum size of an object is 4TB.
  - S3 multipart writes are supported.
- Applications
  - Applications that access objects using S3 can use Swarm directly. Other applications that use file access require the gateways or SwarmNFS for use with Swarm. Object access is supported through a superset of S3 called Simple Content Storage Protocol (SCSP), which conforms to HTTP 1.1.
  - A number of applications are available with native access to Swarm.
- System environments
  - Access is through S3, SCSP, and HTTP.
  - With the use of SwarmFS or FileFly, Swarm may be used for file access.
- Deployment and Administration
  - Installation is a download of the software to a node and a network boot. The system auto-discovers the environment and joins in the Swarm distributed environment automatically.
    - Administration is through a web interface to nodes in the configuration.

Overview

Swarm is an object storage system that is sold as software for commodity servers and storage, usually integrated by resellers, and additionally as hardware appliances. Swarm is a scaling system with multiple nodes in a cluster where a node is a server executing Swarm software with captive storage attached.

The Swarm system has evolved from a records management / file archive system to an object storage system with uses not only in the compliance archiving area but also as a content repository and a target for retained backups. The advanced features with Swarm are both an indicator of the evolution of the system and the longevity. With the wealthy feature set, Swarm has a broad applicability in many different vertical markets including traditional enterprises.

Competitive functions such as forward error correction with erasure codes to protect from failures, geo-distribution of data, remote replication, transparent technology updating, and seamless scaling of capacity with dynamic capacity rebalancing are included in the Swarm system. Both S3 and an advanced custom API over HTTP v1.1 are used for access to Swarm.

The Swarm Server hardware appliances consist of 4 models: the SSA, the s3000, the hd5000, and the m1000. The SSA appliance is a single server appliance that contains all software and management features running on VM's in a 1U appliance. The s3000 and hd5000 are the standard and high-density models which may be deployed in clusters of 3 more. The m1000 appliance is a management node to handle Swarm services and search functionality for bare metal.

For file access, Swarm offers two solutions in addition to working with other vendor file gateways. SwarmNFS is a gateway for NFSv4 that maps files access to object on Swarm over S3. FileFly is a migration solution for files from Windows and Linux file servers and NetApp ONTAP NAS systems. The migration will leave a stub or symbolic link behind to provide transparent access to files stored as object on Swarm.

The Swarm Hybrid Cloud for Azure feature enables tiering of objects to Azure as blob storage.

Product Architecture – DataCore Swarm

Swarm is software that executes on nodes to provide the storage management for object data. Nodes are servers with storage devices attached. Objects are broken into chunks called segments and stored on the nodes in the system. The top level of the structure of the object storage is a **Cluster**. Within a Cluster, there are Domains. A **Domain** is a separate tenant with isolation of access and administration. Within a Domain, there can be multiple Buckets. **Buckets** are containers for Objects. Buckets can be assigned for different ownership and controls for access.

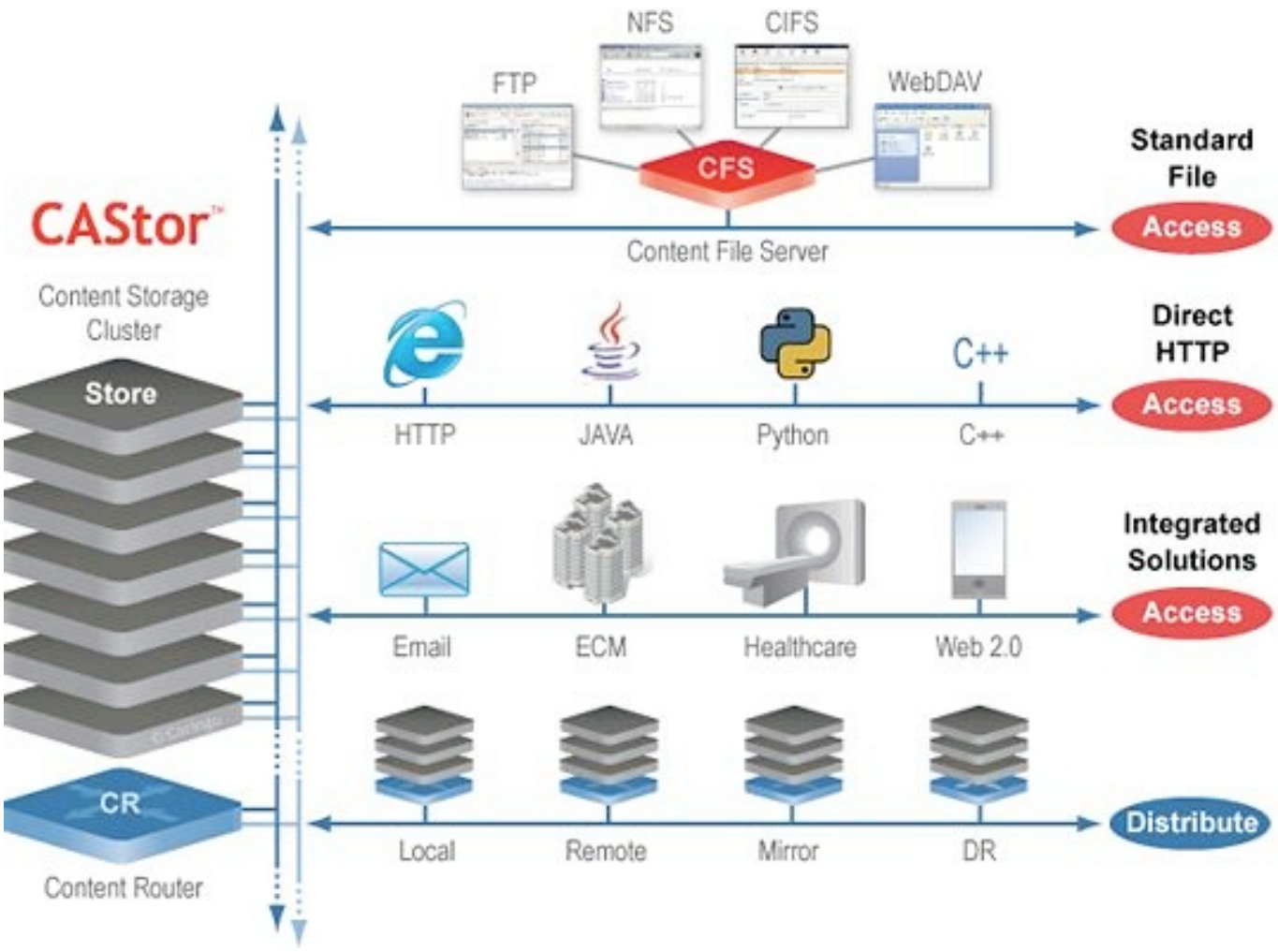

Figure 1: Swarm Architecture (source: Caringo)

Software Architecture

Swarm is storage software based on a decentralized architecture where individual nodes, executing the Swarm software make decisions about storage and protecting data based on the control settings established and assigned for buckets, domains, and clusters. The Swarm software (previously called CAStor) uses a set of rules to make decisions about routing data to nodes, making data protection copies or erasure coded segmentation, and distribution to geographic sites.

Additional functions are applied to the data through Swarm software on the nodes.

Figure 2: Swarm Overview (Source: Caringo)

Objects

Swarm deals with objects as either immutable or mutable. Mutable objects may be changed. Immutable objects may not be altered but may be deleted. The following describes the object concepts used in Swarm:

- Named Objects – Objects where the identifier are assigned by the client have a unique name within the bucket. The Swarm software must do a name lookup to access a named object, which does incur additional overhead. A named object can have its content replaced (mutable) but could also be set in WORM mode.

- Unnamed Object – The object has its identifier, the UUID, assigned by Swarm using a hash algorithm. The handling is optimized for performance. An unnamed object is immutable, WORM by default.

- Alias Object – An Alias Object has as permanent UUID assigned but the content is replaceable. To store an object as an Alias Object requires a specific write parameter.

A quota management feature is available through the GUI or REST API to control the number objects that can be stored in a bucket or domain.

A unique feature of Swarm is Object Rename, which is the ability to rename an object without have to read and rewrite the object. This is very useful for files stored as objects when managing file structures.

Starting with Swarm 11, Swarm allows for partial restore of objects.

Content Cache

A Content Cache is allocated in DRAM in storage nodes to store frequently accessed objects for performance acceleration. The objects stored in the Content Cache would be considered small objects and not large objects, however.

Routing

Routing in Swarm is called “Feeds”. Feeds is the distribution of data to the target storage cluster for replication and ElasticSearch search server for object metadata search.

Consistency

Swarm has eventual consistency where the read of data after a write depends on the replica or erasure coded fragments completing the storage operation.

Data Protection and Security

Device and Node Failure Protection – Elastic Content Protection

A device or node failure uses either the multiple copies of data written or the correcting erasure codes depending on the system setting for protection. There is no packing or containerization of small objects.

Replication

Replication will make multiple copies of data on different nodes to protect from a node or device failure. Replication only occurs for the first copy written and write complete is sent immediately as the first copy is written. Remote copies are treated as an extended local copy.

The setting Replicate on Write is a performance acceleration optimization where the write complete is not given until all copies are made.

Erasure Coding Forward Error Correction

Erasure coding, forward error correction using erasure codes, is selectable and used to optimize space used for data protection. Multiple encoding patterns such as 5:2 and 7:3 are available; provide some variation in the cost of data protection.

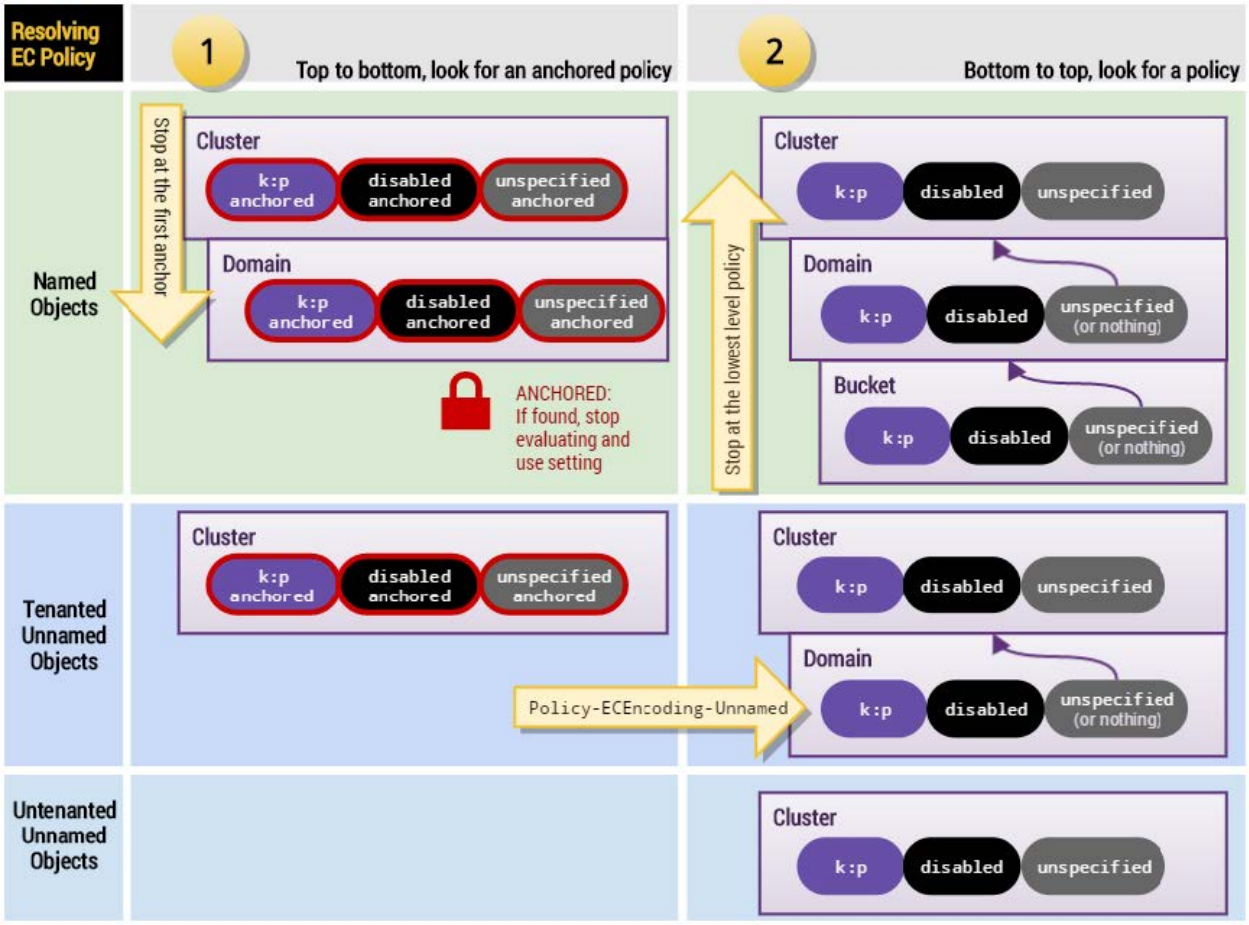

Figure 3: Swarm Erasure Coded Protection (source: Caringo)

Remote Protection

Geographic dispersion is supported as well as asynchronous remote replication. These are effected as a setting where nodes for a configuration may not be local but the same data protection mechanisms are used.

Data Integrity - Content Integrity Seal

A health processor function is use to assure data integrity. The health processor function monitors the data continuously, performing the integrity checks using the added unique hash code and object ID to the data and assuring the number of replicas or erasure-coded fragments are available. Any problem found is corrected by making additional copies, replacing bad copies, or recreating the fragments. The health processor also checks the metadata header.

Multi-Tenancy

Domains are the multi-tenancy unit within a Swarm system. The administrator can assign separate administrative privileges and system settings by domain.

Encryption

Encryption is done as Swarm writes data to devices and decryption is performed as data is read. Encryption is not performed for data during transfers. Swarm does the encryption in software and states that a 10-30% performance impact may occur when encryption is activated. Encryption is set at the disk volume level.

Advanced Features

The advanced features of Swarm extend the usage to meet additional customer requirements. With the history of regulatory compliance usage, Swam has many capabilities that may be critical for certain types of data.

Versioning

Versioning in Swarm is enabled at the domain or bucket level with cluster policy settings and supports an unlimited number of versions of an object. Versioning also works with search and replication. All versions will have a unique UUID.

Versioning with Swarm supports the S3 functionality and adds some additional capabilities such as the disabling versioning on a bucket level with automatic cleanup and simplified ACL management.

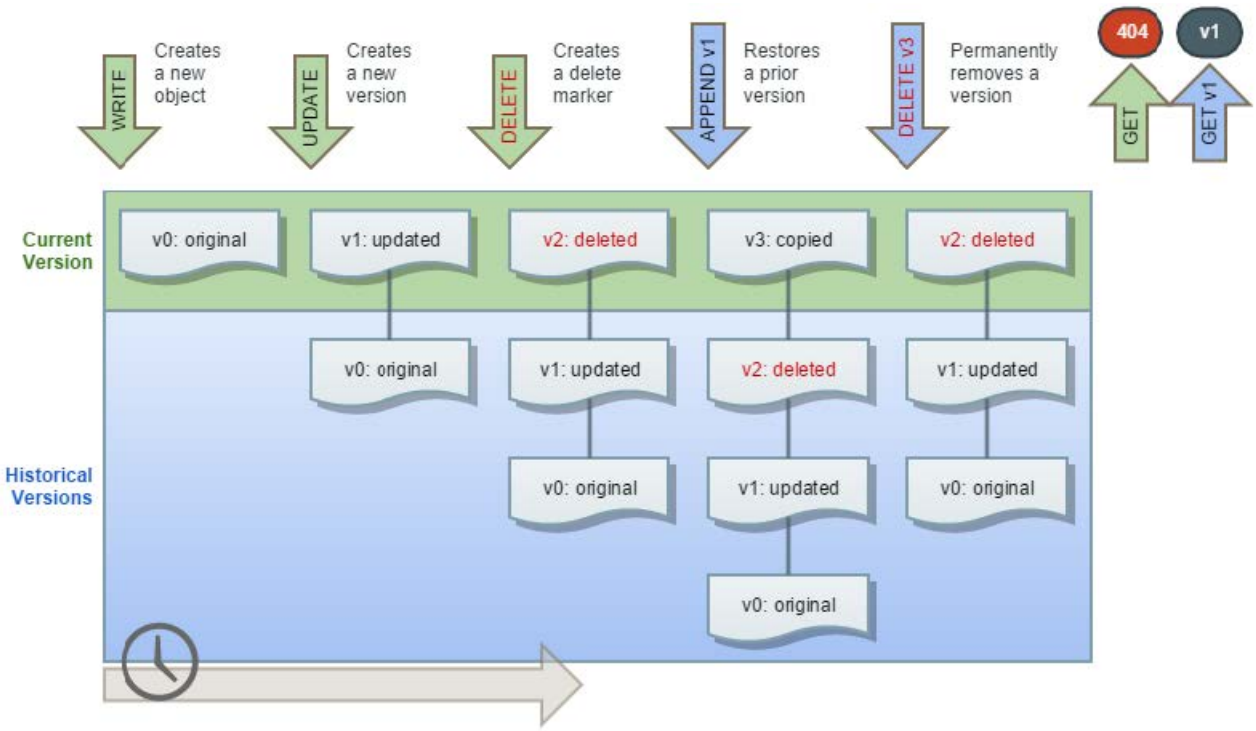

Figure 4: Swarm Versioning (source: Caringo)

Compliance

A number of features are available in Swarm to enable use in regulatory compliance environments. In addition to the WORM setting at the object or bucket level, legal hold can be set to hold off any

automated or manual deleting actions. As noted earlier, encryption is supported for data at rest. An audit trail of action is maintained to assure no alterations of data integrity or veracity.

Lifecycle Management

The lifecycle management function for objects is called Lifepoints. Lifepoints is a policy engine that runs in the health processor software and controls managed deletion and retention. The rules for the policy engine are defined by the administrator.

Data Reduction

Compression and deduplication are not currently available in Swarm.

Reports

Multiple pre-defined reports are available through the administrative GUI as well as ad hoc reports. An API is provided for software access to do reporting or data analysis.

Disk Spin Down - Darkive

Swarm has the capability to spin down devices on nodes that have not been active for a period of time. Access to data for a node that is spun down will be delayed while it spins up for access.

Search - ElasticSearch

The search engine ElasticSearch is used to search metadata – both system and user metadata on a Swarm system. Ad hoc queries can be used for the search. The search engine also has metrics and metering to track information in Swarm. Search can be isolated to buckets or domains and has assignable permissions along with LDAP and AD security when users perform searches.

There is a separate server used for ElasticSearch and it can be in a high availability configuration.

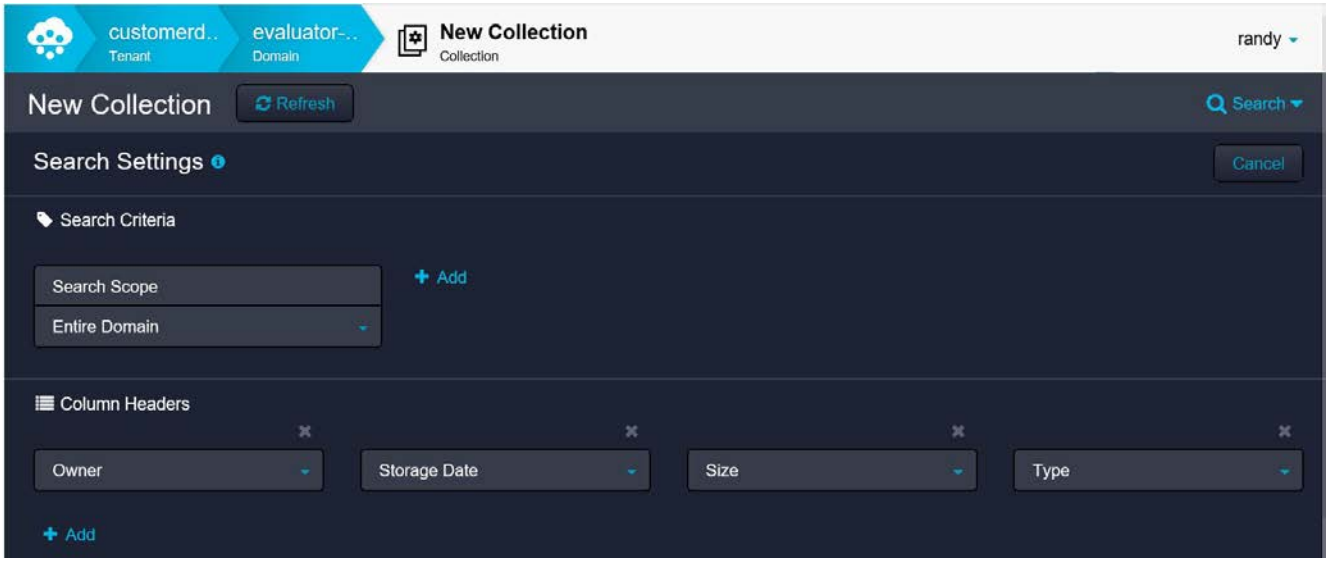

Figure 5: Search Example

File Access

Swarm has two current solutions for file access. The first is SwarmFS that is a gateway supporting NFSv4 and SMB access to Swarm. The second is FileFly, which is software that is either independent or delivered on a Windows 2016 server to provide file migration to and from Swarm with CIFS/SMB and NFS access and NetApp systems or Windows file servers. Other object storage gateways may also be used.

SwarmNFS

SwarmNFS is gateway software and is delivered on a server running Linux, as an appliance, or as a virtual machine. Files stored on the SwarmFS mount points are put into buckets within a domain with each file being an object. With the possible isolation of a filesystem to a domain, SwarmFS support multi-tenancy.

High availability configurations of the hardware with SwarmFS can be deployed for continuous availability. A global namespace across Swarm can be configured for access over NFS and SMB. Data stored in Swarm has universal access from protocols NFSv4, SMB, S3, SCSP/HTTP, and HDFS.

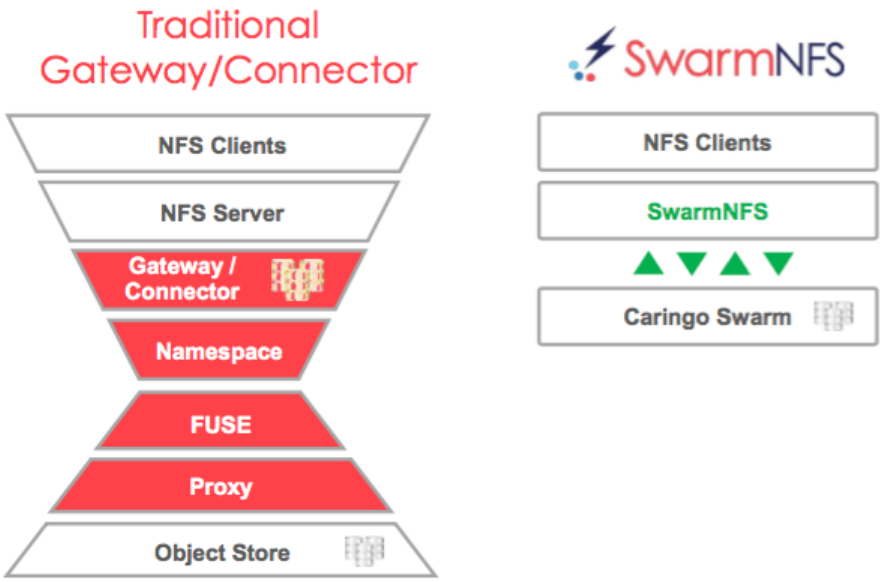

Figure 6: SwarmNFS (source: Caringo)

FileFly

FileFly is software that has a policy engine and data mover to migrate files from Windows servers, Linux filesystems, and NetApp filers and leave a symbolic link/stub behind. The link/stub provides means for transparent access to the file by causing a retrieval from Swarm. The NetApp file stubbing uses the FPolicy function of NetApp. CIFS/SMB 1, 2, 3, and NFSv3 are supported for the migration connection.

Active Directory support is included along with DFSR and FSRM (Filer Server Resource Manager) integration. FileFly has an option to add file metadata to the object.

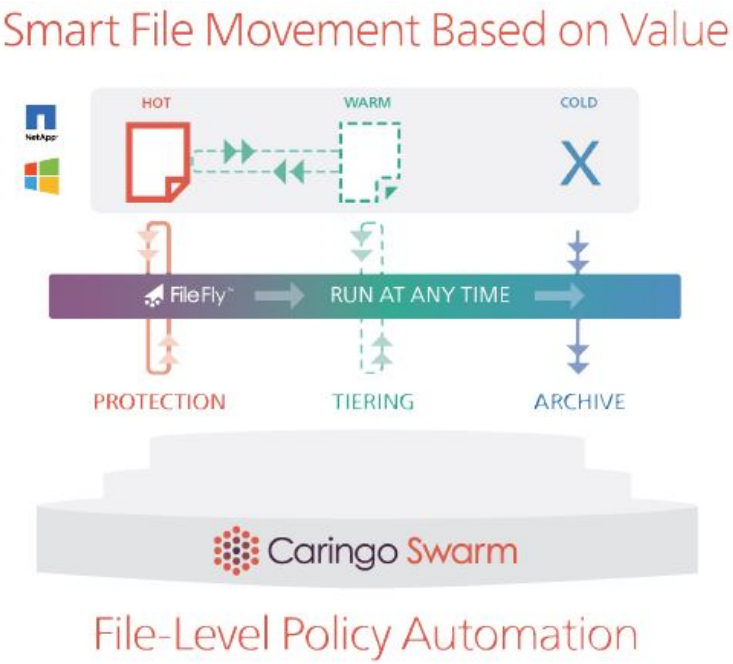

Figure 7: FileFly Overview (source: Caringo)

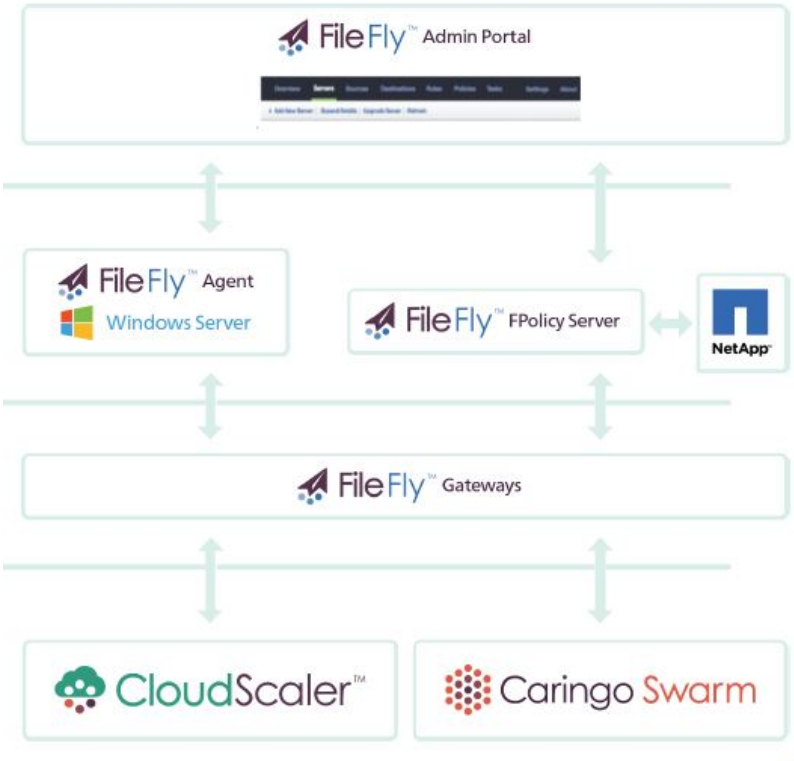

Figure 8: FileFly Deployment Options (source: Caringo)

Reliability, Availability, Serviceability Features

Reliability and availability are addressed with the scale-out architecture and the ability to access data after a node or device failure. Availability can be summarized as:

- Continuous integrity check of data and automatic replacement or rebuild.
- Information Dispersal Algorithms or multiple copies of data used to protect from node or device failures.
- Online node updates
- Ability to retire nodes for technology updates
- Automatic redistribution of data when a node is added / replaced
- Redundant data paths to all nodes
- REST management APIs

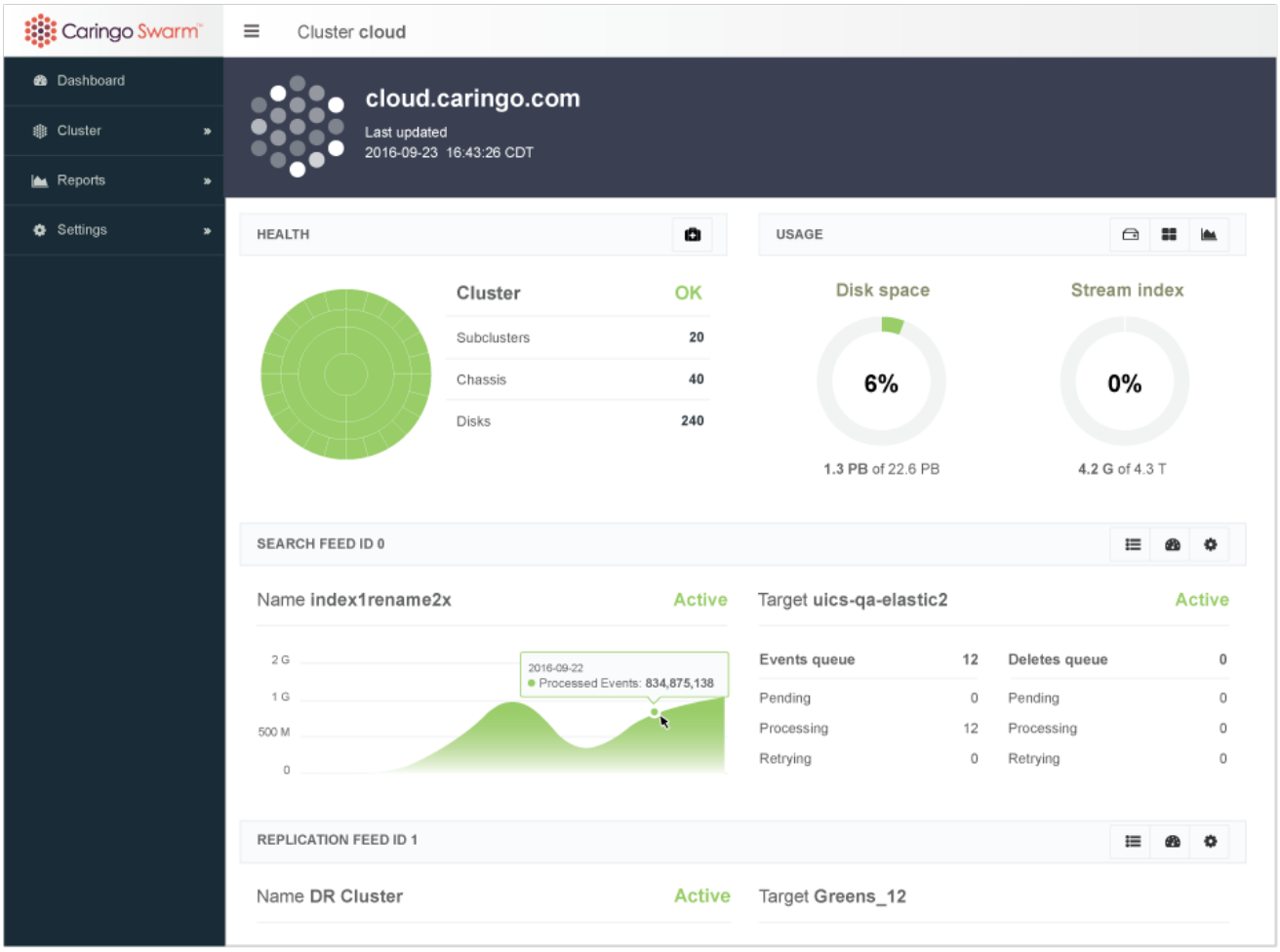

Figure 9: Swarm GUI Dashboard Example

Performance

No standard performance benchmarks for Object Storage are available at this time. One, IOMARK for Object Storage, is in process and will provide comparative performance data when available. Meanwhile, only vendor reported performance data is available.

Evaluator Group Comments

The Swarm object storage system has been available for a number of years primarily through an OEM to Dell. Swarm has been sold through resellers and with some direct sales, and the product continues to receive new capabilities over time. One challenge for Swarm under Caringo was the size of the relatively small company, but Swarm is a mature product with many advanced capabilities and provides a competitive object offering for DataCore's storage portfolio.

Strengths:

- ***Advanced features including compliance support for most regulatory areas***
- ***File access with NFSv4 through the SwarmFS gateway***
- ***File migration from Windows and Linux file servers and from NetApp filers with support for CIFS/SMB or NFSv3 access with FileFly***
- ***Versioning capabilities***
- ***S3 support***
- ***Geo-distribution of data and remote replication***
- ***Partial file restore and clipping***

Perceived Challenges:

- ***Overall performance information is limited. More information about performance under load would be useful.***
- ***A native file interface would be much easier to deploy and manage rather than the two solutions currently available.***
- ***Versioning controls for files would expand usage.***
- ***Data reduction, primarily in the form of compression, would be a good addition.***
- ***More advanced hardware encryption functionality would remove performance impacts.***

More detailed information is available at **http://evaluatorgroup.com**