STARFISH

November 2017

WHITE PAPER Take the *meh* out of metadata

Summary

Starfish Storage is a software solution that:

- Provides visibility (both current and historic) into file systems, directories and files
- Orchestrates the data management lifecycle from file detection to data migration among file systems and object stores
- Provides data movement capabilities such as archive, backup and restore
- Keeps going long after applications, scripts and ad hoc management methods exceed their limits

With Starfish, previously opaque file system metadata is made accessible and actionable. Quickly locate files and groups of files, get immediate information on file counts, sizes, duplication and ages, disk utilization (`du`), and change rates on billions of files and petabytes of data. Make data-informed decisions and automate actions (move, migrate, backup/restore, archive/restore) from the results of those queries even at very large scale.

Starfish creates and maintains a unified database of file metadata, file history, and customizable data such as ad hoc tags or metadata extracted from the files. The file metadata can be gathered using file system event integration and high-performance scanning. The Starfish database can be queried and updated using our sophisticated CLI, HTML 5 GUI, and API interfaces.

Starfish includes a job manager that provides automation and data management orchestration, and includes default operations such as archive and backup. The job manager is a flexible framework that also allows integration with a variety of custom workflows, and scales jobs easily across any number of agents. Job results can be referenced or acted upon by querying the Starfish database.

Table of Contents

Overview

Over the past five years Big Data has made a big splash, with alarms and warnings of the upcoming Big Data deluge. Countless talks have been given, articles written, and research presented on the three, four, five V's of Big Data and the challenges it presents. It makes for great copy. And – it's true. As the amount of data has increased, challenges that used to be limited to the very top tier of organizations like the Library of Congress, NIH, Facebook, and Google, are now faced by small companies, individual departments, and even single research projects. Owing to the growth of research instruments generating large amounts of data, the ease of creating new image and video content, the sharing of research data, and the reduction in the price of storage, more organizations have petabytes of data, in billions of files, on-site and remote.

End-to-End Lifecycle Management
Billions of Objects
Petabyte-scale File Collections
Extremely Rapid Metadata Ingest

Managing this ensemble of data is time consuming, risky, expensive and usually tedious. What is simple to do with one million files gets complex as it grows through billions. Some sites try to ignore the problem, letting files accumulate or buying more storage, which is inefficient and gets very expensive very quickly. Others try to build their own tools for monitoring and management of storage. However, such development often pulls resources away from other projects and in many instances the homegrown software eventually becomes abandonware. Others attempt to use open source or commercial tools to try to solve the problem, but many tools are unable to function efficiently in large-scale environments. Many of the tools are inflexible and designed for one role rather than all of the stakeholders in the enterprise's storage efforts. Also, these tools are often appropriate to one type of file system or storage device when, in practice, many sites have many different storage targets, including the cloud.

Some Use Cases for Starfish

- Locate all of an employee, group, project, or arbitrarily tagged files
- Facilitate management of files when an employee leaves
- Chargeback/Showback reports
- Automatically update old versions of files (for example to new codecs)
- Automatically compress files of a certain type
- Add metadata to files and directory trees, ad hoc or based on criteria
- Manage file provenance
- Automate file lifecycles according to grant specifications
- Move files into the cloud or object storage (Oracle and Amazon)
- Rapidly migrate data using Starfish parallelized data mover
- Ensure file integrity with fixity checks
- Find all files impacted by a virus (CryptoLocker) or ransomware
- Quickly generate reports about files and data storage on large/multiple file systems
- More rapid legal discovery request response
- Determine how much space a subdirectory is using (ability to do du again via Starfish `sfdu`)
- Integrate with BA Insight for broad searching across data stores, including Microsoft SharePoint and clouds
- Integrate with RSpace to keep files accessible in lab books, even if they are moved
- Extract the liquid produced when threatened and use it to treat inflammatory illnesses. (A feature of the echinoderm not the software.)

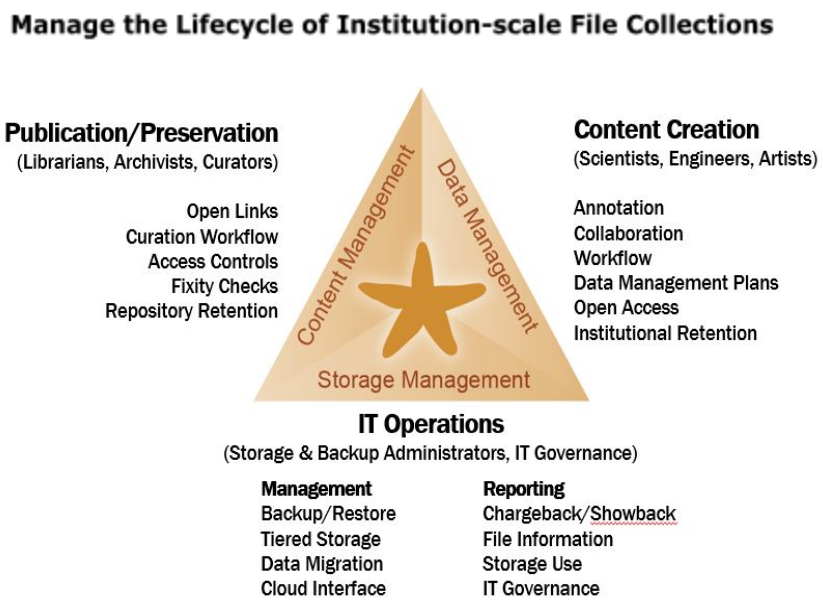
See our other white papers for more detailed use cases.

In conversations across a broad range of industries - scientific research, media and entertainment, oil and gas, government and education - the same problem is revealed: there is an abundance of data, and there is a desire to manage this data by some meta concept, such as a project, grant, or contract. Once finished with the data, the site wants to move it all off to an archive, but it doesn't want to manage the 50,000 files in the project. Rather, it wants to manage all data in aggregate. Those activities must be lightweight and highly efficient. The activities

should include local and remote file, block, and cloud storage.

Roles that Need Starfish

Most storage management and content management systems are geared toward one type of user, the IT Operator, the Content Creator, or the Publisher and Preservationist. The people managing the storage have different drivers than the Content Creators (scientists, engineers, and artists), who also have different incentives and concerns than those involved with Publication and Preservation (librarians, archivists, curators). While some tools meet the needs for one type of user, they can leave others out in the cold, with no way to access or control the tools. That can leave an organization using multiple tools, or empower one type of user while not enabling others to do their jobs efficiently. Starfish was designed and implemented to be useful for all three of those roles.



If there were a way to annotate data at creation and have that metadata follow the work regardless of where it is stored or moved, the challenges associated with data management for all of these three roles would be greatly reduced.

The Problem – Current Technologies are Incomplete

Current technology and purchasing processes have resulted in large organizations having multiple storage products, including file systems, storage devices, types of storage, and storage locations. No matter whether this is due to procurement requirements, vendor viability, or changing needs, many organizations have more than one storage system. Thus, it is highly unlikely that a single hardware vendor's integrated solution will be able to meet all requirements.

Some current systems support transparent tiering (Isilon pools, GPFS AFM, etc.) to allow for migration of data to less expensive storage or tape. Other systems support archiving management (SGI DMF or HPSS). Many of these have some type of lock-in to their technology and proprietary formats. Few tools work with multiple storage systems and allow multiple types of data movement.

Similarly, there are systems for data management (SRB, iRods, and Datadobi). These tend to lack scalability and functionality, requiring administrative effort and expensive servers to maintain. This reduces effectiveness. On the other side of the coin are virtualization solutions such as MarFs from Los Alamos which uses GPFS to front-end a wide range of object stores. These tools provide solutions for specific use cases (data migration, data federation, data stewardship, or large-scale data storage) but don't solve the general problem of large-scale metadata management and actions based on metadata queries.

As the complexity of environments grew, the paucity of useful tools became obvious.

The Solution – Metadata Management with Starfish

For a solution to succeed, it must meet the requirements of the IT Operators, Content Creators, and Publication and Preservationists. Metadata is the key to managing data from a higher concept level. As an example – NCBI's "bioproject" includes experiments, samples and analyses. Over time, it may be optimal to archive the raw data and keep the analyses. At that point, it would be optimal to have high speed `find` of all the data according to an attribute, and migration of all the raw data to archive storage as a single batch operation. Tracking the old and new

locations via metadata would allow contiguous operation; the view of the file system via its metadata would stay the same while the raw data moved out of the way.

Such a metadata database can be used in many ways. If it is updated periodically (based on site need) with all file changes, it can provide a history of all files. For example, if data is copied info an archive, Starfish, with a simple query based upon the creation time, can show the system before, during or after that copy.

What if the data was moved to a new storage system? How does this affect users? Currently, this is usually handled via folklore – that is, a note is posted that file system `PRJ24` has been moved to `Isl-south:/2015`. Here Data replaces Lore. All files and paths mentioned in workflows are now broken and need to be fixed. By tracking the history of all data in an environment, folklore is no longer needed. The metadata now reveals where a set of data was and where is has moved to, allowing tools such as electronic lab notebooks to still find the files they linked to but are now stored in new locations.

Driving activities based on metadata liberates the IT Operators to focus on data at the appropriate level of abstraction, rather than worrying about files, file systems, tracking spreadsheets, myriad tools, and ad hoc file movers like rsync. It becomes simple to ask questions such as how many files or how much space is taken by:

- duplicates
- files unaccessed in the year
- files owned by a given user
- files owned by a given group
- files market with an ad hoc tag
- files matching specific criteria.

Also easy to find are:

- details about whether files have changed (fixity)
- how many files have been added/removed/changed
- space allocation over time
- what files have caused free space to decrease
- and so on.

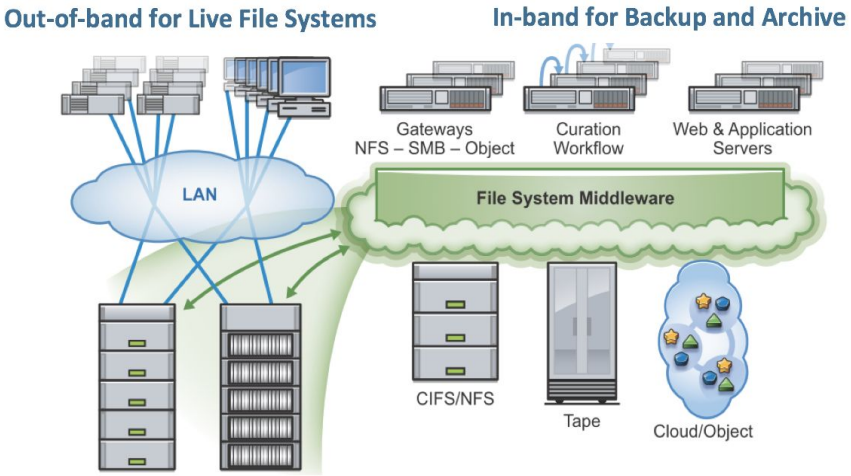
Once the complete metadata of all file systems and object stores, including history, is in a fast, up-to-date database, potentially useful questions are endless and answers become key to data and site management. Of course, the ability to take actions based on the metadata is useful for all three roles; consider project owners archiving a specific file type, storage managers migrating unused files to lower cost storage or the cloud, and archivists tracking files via ownership tags.

STARFISH

This backdrop lays out the vision and mission for Starfish - to add intelligence to the management of large data sets, keeping the points of view of the IT Operators, Content Creators, and Publication and Preservationists in mind. As in the classic illustration of blind men describing an elephant, each of these groups view the world of data through different lenses. Starfish enables all three roles to be more effective and more efficient, accomplishing activities difficult or impossible before. Starfish meets all of their requirements by harnessing metadata from the file system and objects, metadata from the file contents, and provenance information and metadata from processing pipelines. It assists in the management of data. It works in heterogeneous environments, both on premise and in-cloud storage.

Product Vision

Starfish came about due to the inability of POSIX file systems to carry additional information and to easily manage large numbers of objects. In POSIX, a file is labeled by its path; move the file and you have a new path. It is the same object, but for applications, there is no record of this change or ability to track the file. Files themselves contain inherent information - metadata - that if easily accessible can permit more intelligent operations to be performed on the unstructured data.



As a broad generalization, IT Operations cares about the bits and bytes, data location, backup/archive, compliance, and purchasing and implementation of new hardware systems. Moving petabytes of data around becomes a chore and 99% correct is not quite good enough. That leads to a great deal of time being spent checking and double checking during data movement operations. Attributing data use to projects is a large task, and managing and maintaining the total allocations to a project is often so daunting it is only done on an as-needed basis. Starfish provides IT Operations with an easy way to search and address any set of data and gain visibility into their storage. Starfish also goes one step further - allowing IT Operations to quickly and efficiently take action on any queried set of data (*a predicate*) by applying a command against that data set. Alternatively, IT Operators can also create polices to perform repeated actions on any data which fits a given set of criteria.

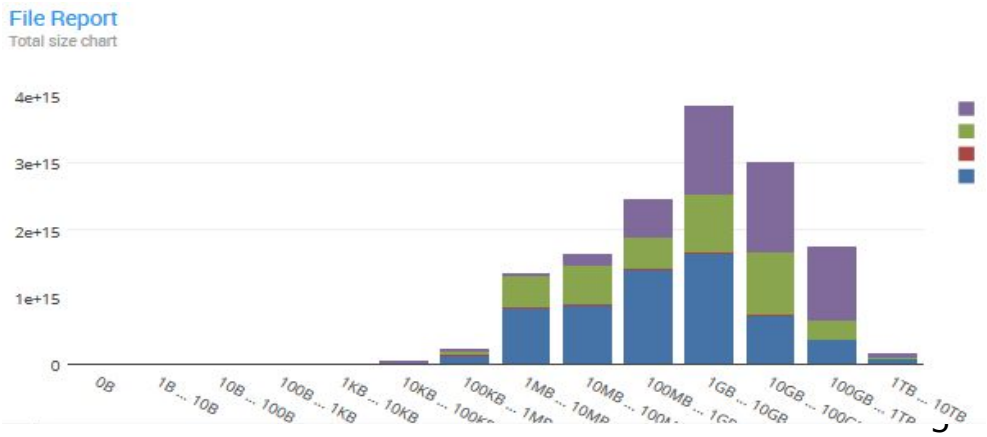
Content Creators are interested in their data and spend a large amount of time managing the experiments, projects and workflow of that data. If they can be freed up from managing thousands, millions or even billions of files and can instead manage a single project, with different states of data (for example "ingest", "qa", "scratch", "derived", and "publication"), it allows them much more time for research, engineering, video production, and other activities that are more important to their jobs. If researchers at the end of a grant can press an "archive" button, it eliminates the time spent by Content Creators and IT Operations who previously had to initiate the archiving and tagging themselves; all appropriate files move to an archive with the correct data retention flag associated. Auditing burdens are lessened because the process is automated. Starfish can assist Content Creators in automating workflows, movement, and actions taken on files. Starfish can also be used to apply metadata tags such as project names, grant specifications, or retention requirements on files that have certain attributes.

For the Publication/Preservationists, which includes librarians and curators, Starfish provides a wide variety of solutions. It can provide full identification of the types and contents of files via its suite of interfaces. Tags can be added (for example to track authorship), subsequently queried, and actions performed. In a future version, metadata namespaces will be added to track all aspects of file authorship and curation. This then opens the door to integrate curator-specific metadata management systems by isolating their metadata schemas from other metadata aspects.

What does Starfish do?

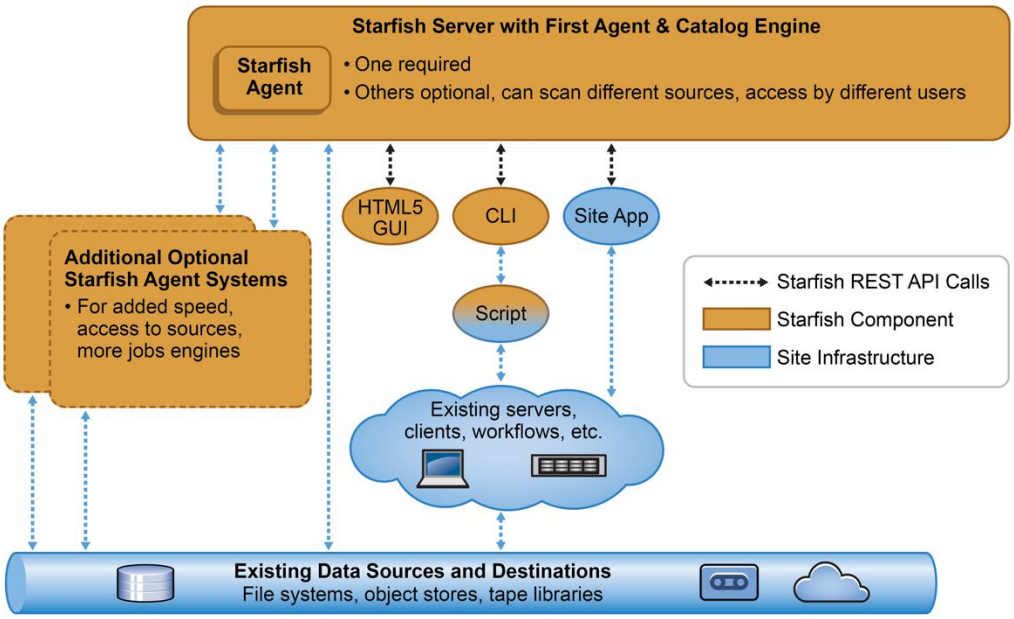
- Starfish is an application designed to manage the lifecycle of institution-scale file collections.
- Starfish associates metadata with files and directories on conventional POSIX file systems and NFS-mounted devices (with CIFS, cloud storage services, and object stores forthcoming), and stores the derived data in full.
- Starfish extracts raw and derived properties from the contents of identified volumes into a Postgres database.
- Starfish can execute actions based upon query predicates, such as archival based upon tags.
- Starfish keeps version histories of all scanned files and objects, allowing data provenance and change tracking.
- Starfish can store checksums of all objects and alert you when policies are violated, such as copying data off of access-restricted storage or onto a restricted target.
- Starfish can even be used to help mitigate ransomware events through historic data and metadata queries.

Metadata are leveraged for search, classification, reporting, graphing, and driving a rules engine that enforces storage management policies. Storage management policies include things like backup, archiving, disposition, data integrity checking, and deduplication. Policies are programmable, not limited to any built-in functionality.



Starfish metadata collection and management sits outside of the data path, such that it is transparent to users and does not introduce latency or points of integration or failure.

Starfish is enterprise-class storage management software. It is designed to handle billions of files, provide version histories, and provide a unified namespace across a diversity of storage devices. Starfish enables both common workflows (such as backups and archiving) as well as customized ones using policies that are refined by extensible metadata.



Starfish creates an ecosystem that incorporates both the file system metadata (names, create times, permissions) and enables storage of file metadata (author, DICOM header, PDF version, etc.). It creates a catalog that spans file systems and storage subsystems, unifying them into a common infrastructure. Starfish maintains this information via out-of-band scanning or event monitoring from the file systems to keep this catalog current.

Starfish is built with a modular architecture that has three main sets of components:

- The Catalog Server – the metadata catalog database for all of the files and objects in your enterprise or stored within a selected group of devices. It maintains history of those items over time, is fed by the Scanning Engine and is used by the Jobs Engine. It can be explored via a API, CLI and a GUI that includes summary charts as well as a file/directory browser.
- The Scan Engine – acquires file and object metadata and updates the catalog. It includes very sophisticated tools for synchronizing the metadata contents of your storage devices with the database, either via direct integration to event logs, via syslog, or via POSIX file system scanning. Event log monitoring currently includes GPFS (DMAPI), Lustre (changelog), with Isilon and NetApp following soon. These file systems generate events in near-real-time, allowing Starfish an up-to-date view of the file system.
- The Jobs Engine – runs actions against your files. Jobs are initiated by a scheduler. A job consists of a query to the Catalog Server metadata database and an associated process. The results of the database query go into a central queue. Starfish agents pull work out of the queue and run the process. Agents can be run in parallel (scale out, across multiple systems) to handle large volumes of work. Agents can be metadata managers or data movers. Agents then return results to the catalog. Jobs can be defined as policies and continuously run against any data that meets a set of query criteria.

The Catalog Server

Once you have a catalog, there are many opportunities to derive value. The first example is reporting. Once the catalog is in place, querying for "all files belonging to user X" becomes straightforward, even though there may be millions of files across dozens of stores (file system and object). Aggregate analyses become possible. For example, reporting on all files belonging to each user, binned by age of access, or similarly by group.

Since metadata can be extracted from files, reports listing the cost per department of storage are enabled for showback or chargeback. Often, all that is missing for better data management is better communication of costs and utilization to the appropriate individuals.

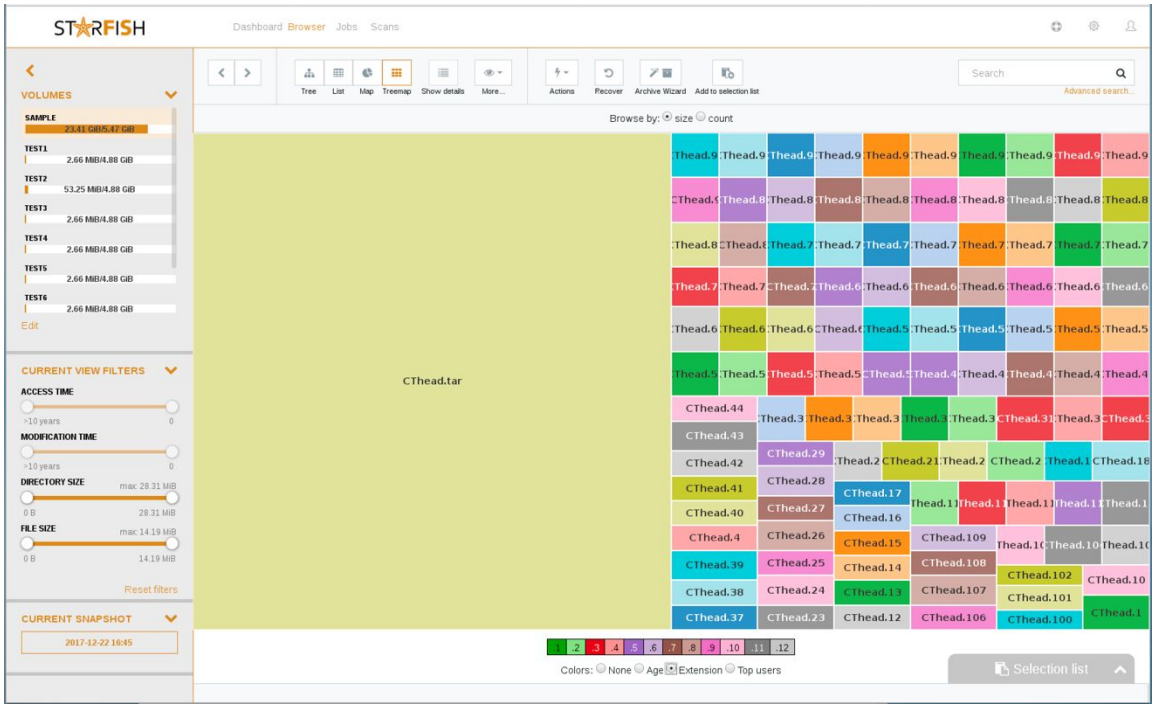
The catalog permits you to create subsets of data. It's a database - so database queries can reduce the datasets to useful components. For example, copying data to an archive can be accomplished by requesting a copy of all files not yet migrated to the archive, or that request can be filtered to only migrate all files not migrated with an age greater than two years and an extension of ".MOV". Creation of rules and filters enable flexibility in processing data. Starfish allows action rules to act upon data sets.

Starfish supports cloud and object stores, allowing archive/recovery, backup/restore, and migration to S3 services (including on-site S3 like Amplidata, GPFS, NooBaa, Symphony), and to Openstack/Swift object stores (including Oracle Cloud).

The Scan Engine

The Starfish Scan Engine ingests file and object metadata into the catalog. Starfish scales out to support larger workloads by parallelizing the scan effort across all the cores permitted, to break the work effort down into smaller tasks. In a large environment with multiple file systems and multiple starfish agents, the work is distributed across the agents, with information fed back to the starfish Catalog server.

Starfish uses REST interfaces to interact among its component services, and these services are available for integration with external systems as well. As an example, the **storage service** provides views into directory structures; it returns file and directory metadata, as well as derived aggregates (tree summaries, totals of extensions, etc.) in order to provide rapid responses to queries.



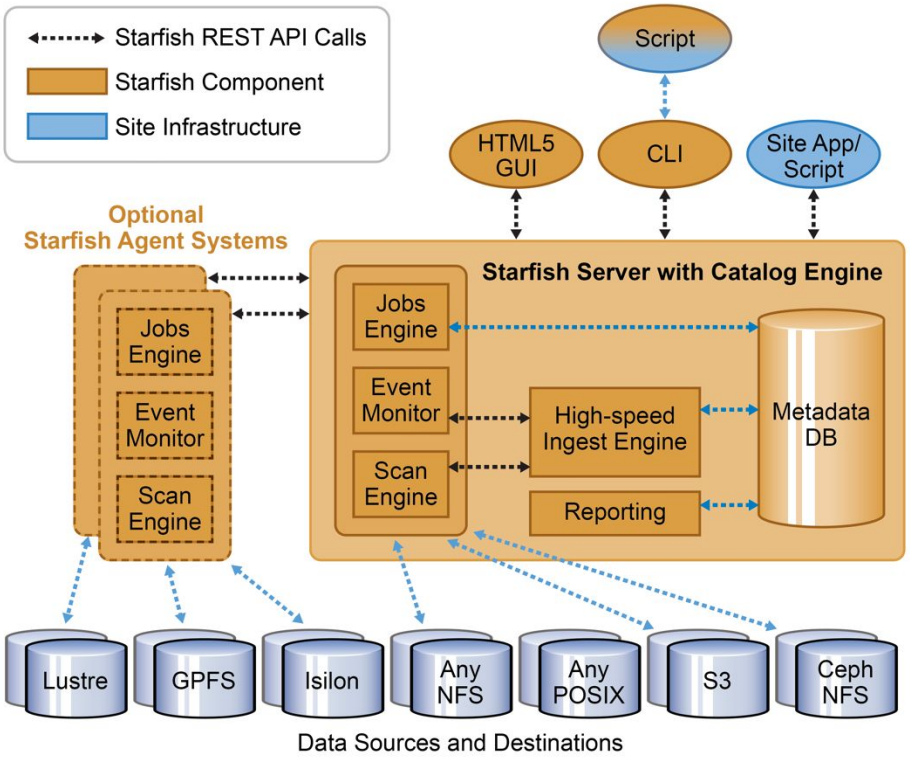
Other services provide administration - the **volume service** manages information about a Starfish-monitored storage unit - which could be a subdirectory, share or a complete file system. The **scan service** similarly manages the execution and reporting of information about file systems scans and monitors.

The Jobs Engine

Actions on files and objects are managed by the Starfish Jobs Engine. The Jobs Engine essentially provides a small compute cluster scheduler that analyzes queued requests and intelligently distributes that work to servers capable of meeting the need, dynamically balancing for load, data access, and priority. This jobs engine accepts both predefined

and custom tasks. Pre-defined tasks such as moving data, archiving, copying to cloud, or hashing of data for validation come standard. Custom tasks are created by the Starfish administrator or via professional services, and are simple Python scripts that extend the functionality of Starfish. As an example, one site was interested in extracting several fields from MRI files to drive their workflow. Using open source ImageMagick and a simple filter, these fields were extracted, added to the catalog and are now available for further queries and dataset operations.

Jobs that extract metadata place their results as JSON back into the catalog, where they are available as structured metatdata for future queries and jobs. Starfish also supports a more freeform tag structure, which permits inherited and non-inherited tags to be applied to files and directories. As an example, a directory could be tagged by a grant and owner. Subsequent jobs may execute an archive or move task including those tag values as parameters.



In Starfish, a job is a task that is run based on a schedule and against a collection of data. The collection is defined as a set of files or directories with records retrieved from the catalog. The collection is specified by several methods: matching tag, file type, age, modification time, size, extension, and regex on file names or paths. If this is not sufficient, an extended query can be defined which utilizes the full power of the database behind the catalog to specify the collections.

Once a job has been created, it collects its dataset and pushes that onto a jobs queue. The queue manager identifies eligible worker processes across the job cluster and distributes tasks across the cluster. Jobs respond with status and results. These results are, in turn, updated back into the catalog and tagged back to the original objects.

The Ensemble

Together, these components provide a fast, complete set of tools for monitoring and managing data sets.

Starfish resource use scales from minimal, for a small installation, to moderate, for facilities with large data needs. All tasks and cataloging can be scaled out across multiple servers and agents to distribute and parallelize the workload. Scaling may not be needed though, because of the efficiency of Starfish. For example, one site has 20PB of data in 9 billion files; one Starfish instance on a single server monitors and manages that storage. Current Starfish-managed sites scale to greater than 50PB and 17 billion files. The rate of scanning of files for tracking in the metadata database has reached 2.8 billion objects per day at one site.

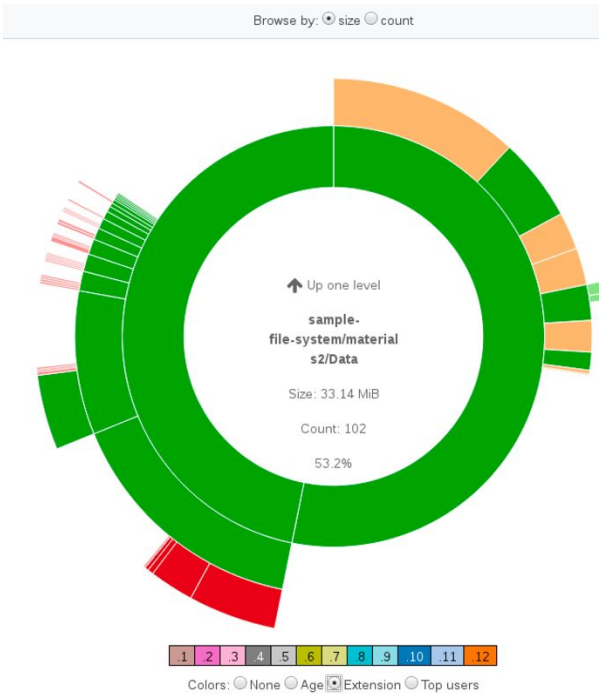
Estimate Duplicates

DUPLICATES COUNT	TOTAL COUNT	DUPLICATES COUNT PERCENTAGE	DUPLICATES SIZE (GB)	TOTAL SIZE (GB)	DUPLICATES SIZE PERCENTAGE	MIN FILE SIZE (BYTES)
87,605,935.00	7,622,594,708.00	1.10	8,424,678.53	17,616,849.86	47.80	10,000,000

Some sites appreciate Starfish for its fast, informative reporting features. Others appreciate action. Actions are based on sophisticated metadata queries. Imagine using Starfish to tag a subset of billions of files as low priority when they are created. This "low priority subset" then goes unaccessed for three months which triggers Starfish to move the files to archive, a lower-cost tier, or the cloud. Built-in actions such as archive/restore can be augmented by arbitrary actions implemented via programs or scripts and run via the job scheduler.

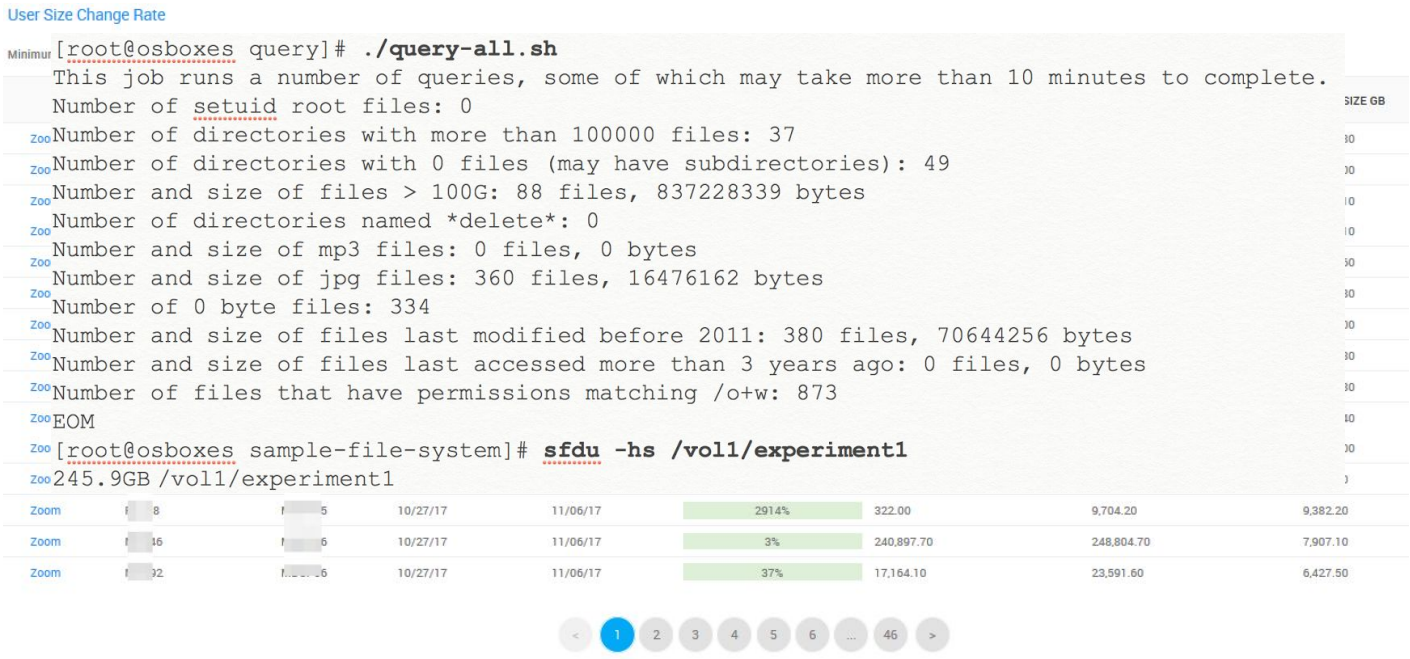
Starfish includes a CLI, APIs, and a GUI. All three components provide access to the catalog and the ability to create and manage Jobs. The GUI is especially useful for visualizing the state and history of volumes, files, and summary information.

By adding intelligence to the data, Starfish enhances and enables data management at petabyte scale.



What Can STARFISH Do for You?

- Starfish installs in as little as 15 minutes, and can provide you with useful, actionable information about all of your files within hours.
- It works with heterogeneous storage, including on premises file stores, object stores, and cloud stores.
- It captures file metadata, and tracks file history over time, and you can add programmatic or ad hoc metadata.
- You can then run pre-built or custom queries against the metadata, via CLI, GUI, or API, to explore the data, or create jobs to take actions based on the metadata, such as migrate, archive/restore, and backup/restore, or any action that is scriptable.
- Starfish can integrate with third-party tools such as ElasticSearch, BA Insight, and RSpace. Others coming based on customer demand.
- With all of this file metadata in our database, adding even more insightful queries is fast and easy. The next set of features include advanced search in the GUI and time-delta reports (such as largest directory-size changes in the past 30 days).



Take the *meh* out of your metadata and use it for data-informed decisions and actions.

To learn more about how Starfish Storage can help solve your business and IT challenges, contact (mailto:info@starfishstorage.com) us or an authorized reseller or visit http://www.starfishstorage.com/.