Qumulo

Qumulo Core Overview

June 2017

Contents

Summary . . . 3

Background . . . 3

Some possibilities . . . 4

Qumulo Core . . . 5

Real-time visibility and control at scale . . . 7

Cluster activity . . . 7

Capacity analysis . . . 8

Capacity trends . . . 8

Performance analysis . . . 9

Client performance . . . 9

Throughput distribution . . . 10

Directory-based capacity quotas . . . 10

File system snapshots . . . 11

Performance considerations . . . 11

Qumulo subscriptions . . . 12

Customer success . . . 12

Conclusion . . . 12

Summary

Qumulo Core is a modern scale-out storage system designed from the ground up for the new era of multi-petabyte data scale on premises and in the cloud.

Qumulo Core stores tens of billions of files with scalable throughput and is the only product that provides real-time visibility and control for file systems at petabyte scale. Storage administrators can instantly see usage, activity and throughput at any level of the unified directory structure, no matter how many files are in the file system. They can pinpoint problems and effectively manage how storage is used.

Qumulo Core provides reliable and efficient use of raw capacity at scale. Unlike legacy NAS, Qumulo Core can rebuild failed drives in hours not days.

Qumulo Core's software-based approach is cost effective and provides a path to continuous innovation. The system uses no custom device drivers or expensive custom hardware. Instead, it works with a variety of standard hardware platforms. Giving customers a choice of platforms lowers costs and helps ensure that they have access to the latest technology. For example, Qumulo Core runs on HPE Apollo 4200 servers, which have some the densest storage arrays available in the standard 2U rack format.

Background

The current explosive growth of digital data is well known and shows no signs of slowing. Approximately 90% of this growth is for file and object storage. IDC predicts that the amount of data will reach 40 zetabytes (a zetabyte is a billion terabytes) by 2020, and that there will be more than 163 zetabytes by 2025. This is 10 times the data generated in 2016.

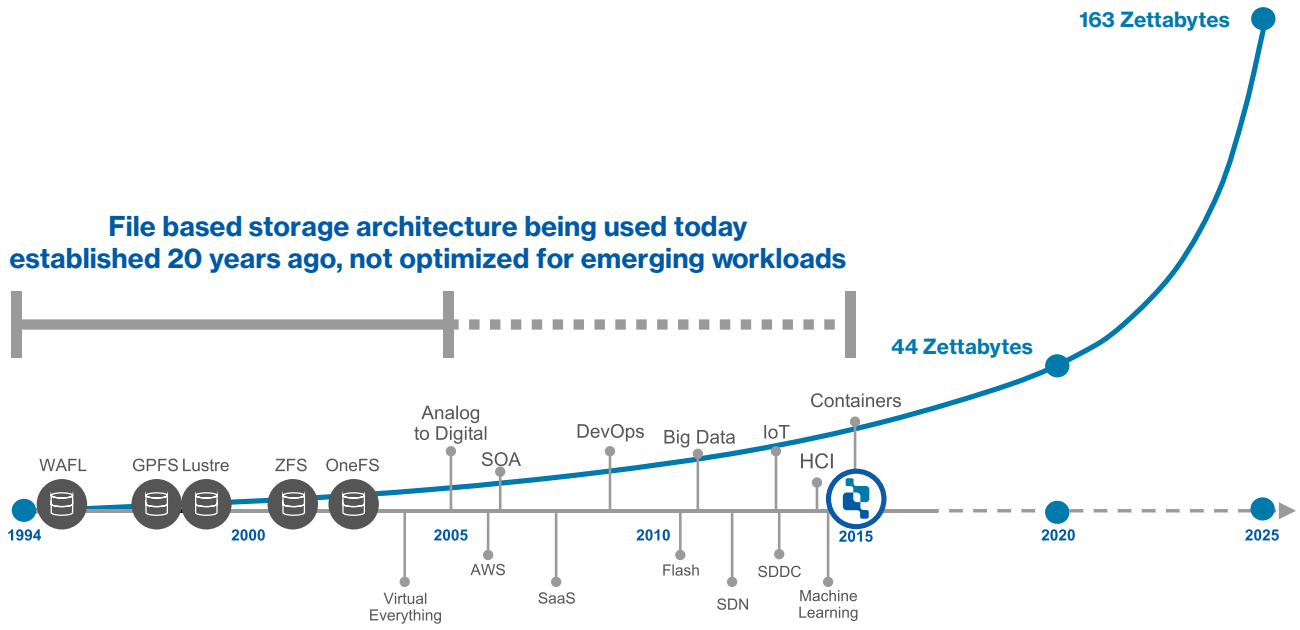
There are many technological reasons for this growth. The public cloud, service-oriented architectures, new media types and the Internet of Things are just a few of them.

Distributed computing, where pools of individual computers act together, is the only way to handle the emerging scale requirements. In highly distributed scale-out file storage, each additional node increases the network throughput of the entire cluster while transparently growing the total storage size of the cluster. It should be no surprise that scale-out storage is set to dominate the file and object category going forward. Gartner estimates that by 2021 more than 80% of enterprise data will be in scale-out storage systems in enterprise and cloud data centers, up from 30% today.[1]

Unfortunately, legacy scale-out has reached the limits of its architecture-- the scale at which data is created and the challenges of managing data at today's scale are fundamentally different than what they were 20 years ago, when the most commonly used data storage systems were developed. For example, when the number of files stored becomes large, fundamentally different techniques are required to manage file directories and other metadata. Systems with 20-year-old designs have no answer for the problem of "big metadata" in data-intensive file-based workflows.

1 Gartner, Magic Quadrant for for Distributed File Systems and Object Storage, by Julia Palmer, Arun Chandrasekaran, Raj Bala, October 20, 2016

This diagram shows the temporal relationship between the growing data footprint and the technologies used to handle it.



Some of the major file storage products were developed as early as 1994. The principal open-source file systems were built in the 1990s. The most popular scale-out file system was designed in 2002.

These legacy architectures are simply not capable of handling the ever-increasing amounts of data that enterprises will generate. Not only will there be petabytes of data, but the number of digital assets stored in a single system will grow into the billions. Traditional scale-out networked attached storage (NAS) has difficulties at that scale, for the type of data we're describing, in terms of performance, management and availability. Additionally, storage lifecycle expectations become a concern, as does affordability.

Some possibilities

As they try to cope, companies may feel they're limited to three choices, all of them with significant drawbacks. These choices are to continue buying legacy appliances, use open source and modify it, or move to object storage.

Legacy storage appliances. Staying with legacy storage products means further investment in 20-year-old technology that is tied to proprietary and expensive hardware. Companies who take this route face the severe performance issues that arise when administrative tasks rely on walking a tree or scanning the entire file system. Companies also have to cope with data protection schemes that do not scale to large numbers of files. In many cases, inefficient compromises such as splitting the file system into many volumes are required. The cost of managing such systems is high.

Homegrown solutions. Some companies are deciding to build their own storage solutions using open source software that the company adapts to suit its existing hardware. However, just as with proprietary storage appliances, open source counterparts are also not designed for file systems at current and expected scale, so

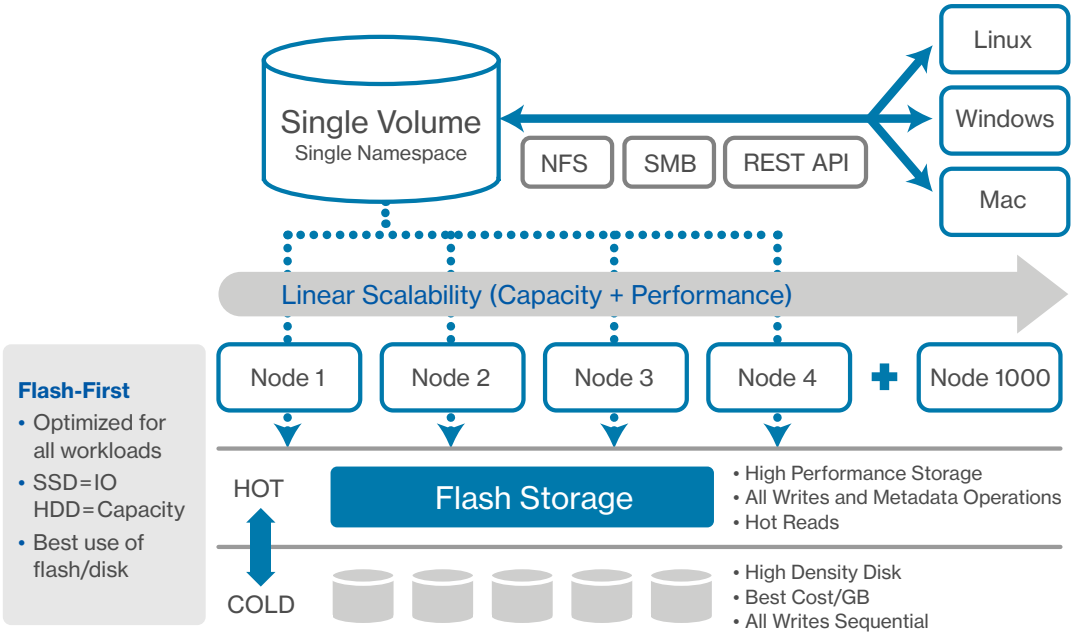
the effort is already based on an outmoded approach. Also, you lose the advantages and support that comes with commercial systems and become completely responsible for maintaining the end-to-end solution. The cost of managing home grown systems is extraordinarily high.

Rewriting for object stores. As an alternative to file-based NAS, some enterprises consider switching to object storage. Unfortunately, in addition to the time and money needed to rewrite existing applications for object stores, it turns out that object stores don't offer the benefits of a file system for organizing data. A surprising amount of valuable business logic is encoded in the directory structure of enterprise file systems.

Qumulo Core

There is a better way. Qumulo Core is a modern scale-out storage system designed from the ground up to handle multiple petabytes of data both on premises and in the cloud. Qumulo Core uses advanced software-based techniques that give storage administrators up-to-the-minute file system analytics, intelligence, and control. With a block-based data protection mechanism, Qumulo Core uses raw capacity far more efficiently than other systems, especially as the number of files becomes large. Since Qumulo Core is software, it runs in a variety of standard environments and delivers innovation quickly with frequent software updates.

Here is a diagram of the Qumulo Core architecture.



Flash-first design. Qumulo Core has a flash-first hybrid design that uses both solid-state drives (SSD) and high-density hard disk drives (HDD). This balanced design optimizes capacity, performance and cost for a wide range of workloads. For speed, all writes land on the SSD first. Reads typically access recently written data, so the SSDs can be thought of as a cache, as well. As the SSDs fill up, less frequently accessed data is pushed to the HDDs. All writes to HDDs are sequential for speed and are optimized for best storage layout.

Unlimited numbers of files. The file system itself has been designed for massive scale. There is no preset limit to the number of files or their sizes. Unlike other systems, Qumulo Core can accommodate, with real-time management, billions of files and directories. Qumulo Core uses techniques similar to those used for large, distributed databases to manage file metadata. Using a proprietary distributed algorithm, it also maintains totals, updated dynamically, of file sizes and counts at each directory level. The aggregated metadata enables many of the advanced sampling techniques used in the Qumulo Core dashboard displays.

Efficiency and reliability at scale. Qumulo Core uses the raw storage very efficiently. Even with enough redundancy to handle multiple concurrent disk drive failures, Qumulo Core offers as much as 85% of raw capacity as usable space.

With Qumulo Core, rebuild times to replace a failed disk are measured in hours, not days, regardless of cluster size. Qumulo Core implements data protection (using state-of-the-art erasure coding) at the block level instead of the file level, where a block is a 4096-byte file segment.

Qumulo Core's block-level protection improves scalability by eliminating time-consuming tree walks—when a disk fails the system knows which blocks are affected without consulting the file system. Block level data protection also allows highly optimized use of cluster nodes to restore data from failed disks or nodes, with the effect that rebuild times decrease as the cluster grows. Further, Qumulo Core uses write layering with independent locking schemes, so rebuild operations don't contend with normal use of the file system. Qumulo Core does not require any user disk space to be left available for its internal use. User files can occupy 100% of provisioned capacity, while legacy scale-out recommends only using 80%.

Real-time capacity quotas. Quotas in Qumulo Core are easy to manage, no matter how large the file system. Unlike some other systems, Qumulo quotas do not require breaking the file system into volumes. Also, with Qumulo Core, moving a quota, or moving directories across quota domains, involves no lengthy tree walks or cumbersome, multi-step processes. Quotas in Qumulo Core can be applied to any directory, even nested directories. If an allocation has more than one quota due to nested directories, all quotas must be satisfied in order to allocate the requested space.

Quotas in Qumulo Core include features that customers need to provision storage to their users and projects. Qumulo quotas can be set on any directory. They limit visible capacity to the size of the quota domain (as opposed to the size of the cluster). They also provide notifications when quotas are approaching full via a pub/sub API which admins can subscribe to, and can be created/modified/deleted via a management UI (or a programmable API).

Snapshots. Snapshots in Qumulo Core solve the usability and scaling problems of snapshots found in legacy NAS. Snapshots in Qumulo Core are based on a fundamentally smarter file system technology-- snapshots can be taken instantly, and there is no practical limit to the number that can be taken.

Qumulo Core's easy-to-use scheduling system allows storage administrators to control the frequency of snapshots. Snapshots can be set to expire after a designated interval, such as 30 days.

A storage administrator who wants to reclaim space by deleting snapshots before they expire can use Qumulo Core's innovative snapshot management feature to calculate how much storage would actually be reclaimed by deleting a set of snapshots. Qumulo Core allows the administrator to select any number of snapshots in any order and then reports how much space would be reclaimed by deleting that set.

Replication. Qumulo Core includes asynchronous data replication at scale. In addition to replicating additions, deletions and changes to files, Qumulo Core replicates configuration data such as users, shares/exports and permissions to the secondary cluster.

API. Qumulo Core includes a REST API that makes it completely programmable.

Real-time visibility and control at scale

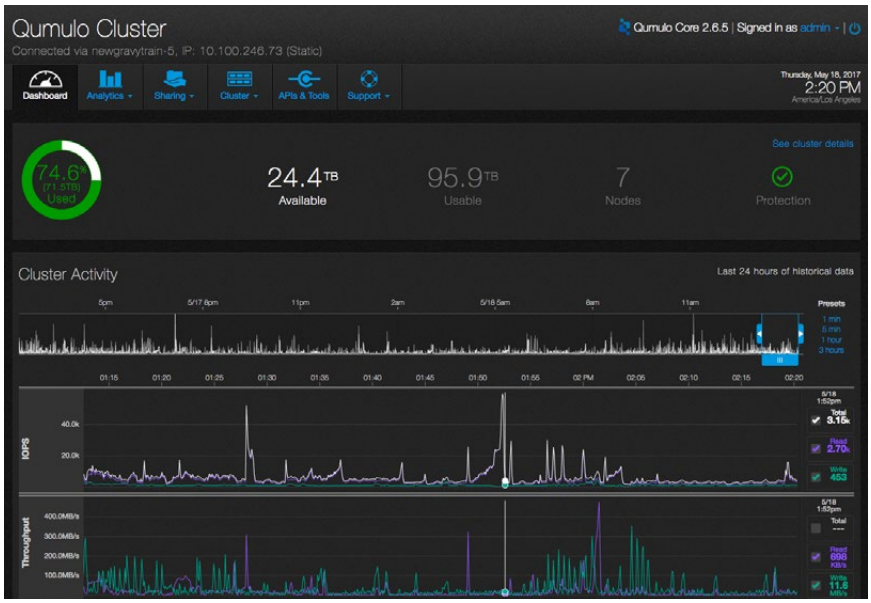
Storing large amounts of data isn't the only problem, you also need to be able to understand and manage it. Storage administrators must be able to manage billions of files within a single file system and to analyze what they see. The Qumulo Core dashboard lets you drill down into all aspects of your file system, instantly.

The inability to manage large file systems is the Achilles heel of legacy systems. With legacy NAS, a simple directory query can take months to run, which means even standard management tasks become impossible. In contrast, Qumulo Core provides immediate access to the information needed to diagnose problems, and its management features operate in real time, regardless of the number of files. For example, Qumulo Core's directory-based capacity quotas take instant effect. Rather than having to wait hours or days, administrators have immediate control over storage allocation.

Here are some examples from the Qumulo Core dashboard that illustrate common management and analysis tasks.

Cluster activity

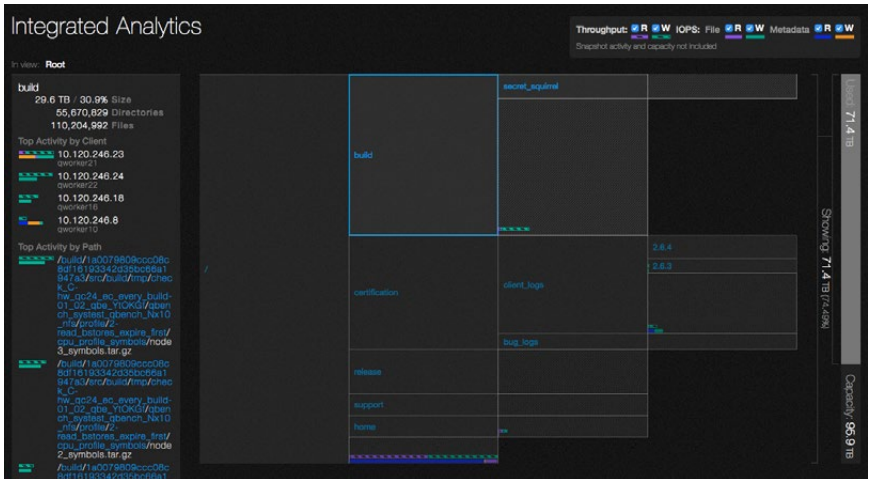
What is the status of the cluster?
What's recent activity of the cluster?



The Qumulo Core dashboard is an overview of the cluster's status and activity.

Capacity analysis

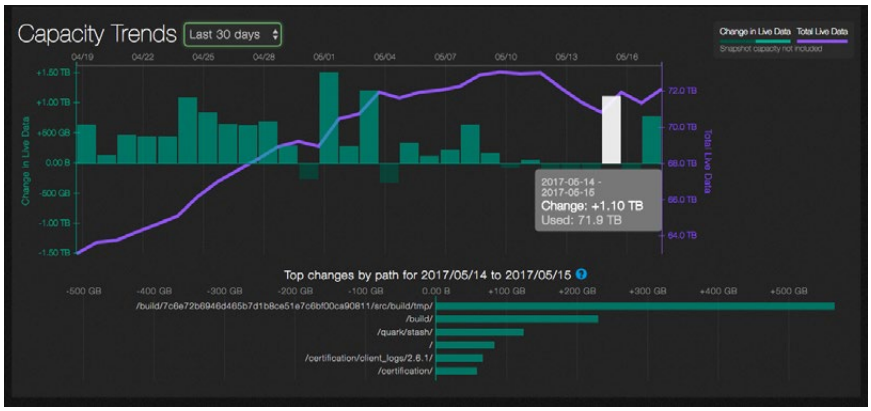
How much capacity is there, and how is it divided among directories?
How much performance is getting consumed on a per-file basis?



The right panel in this view shows how much capacity is available and how it is divided among directories.

Capacity trends

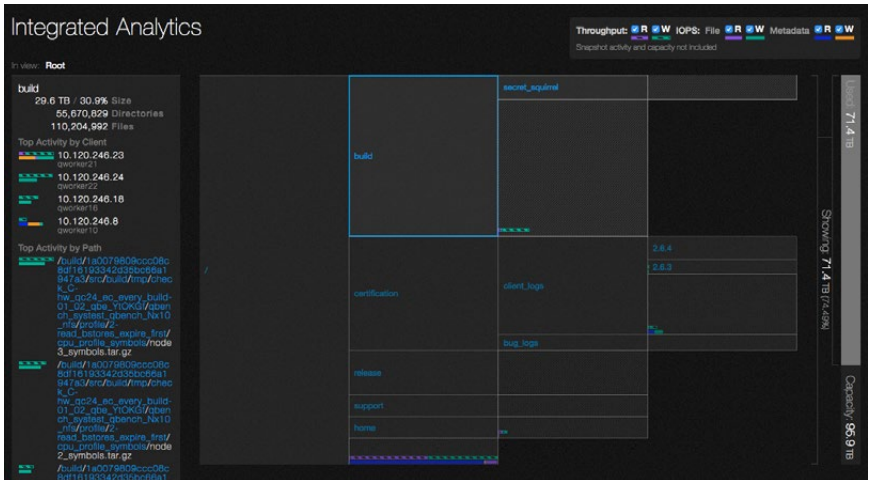
What file data is being added to the file system and using the system capacity?



This view shows what file data has been added to the file system over time. You can identify usage patterns and plan for future capacity needs.

Performance analysis

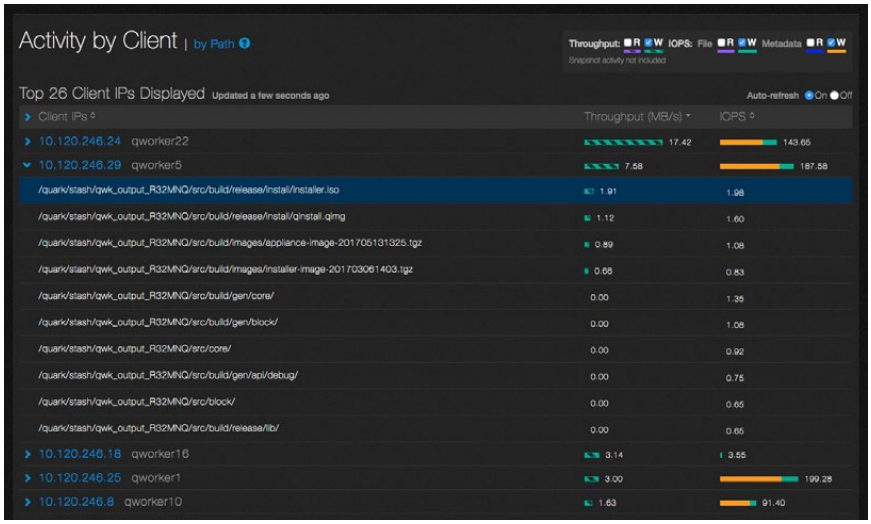
How much performance is getting consumed on a per-file basis?



The left column of this view shows how much performance is being consumed on a per-file basis.

Client performance

Performance has slowed. Which individual files and directories are getting accessed?

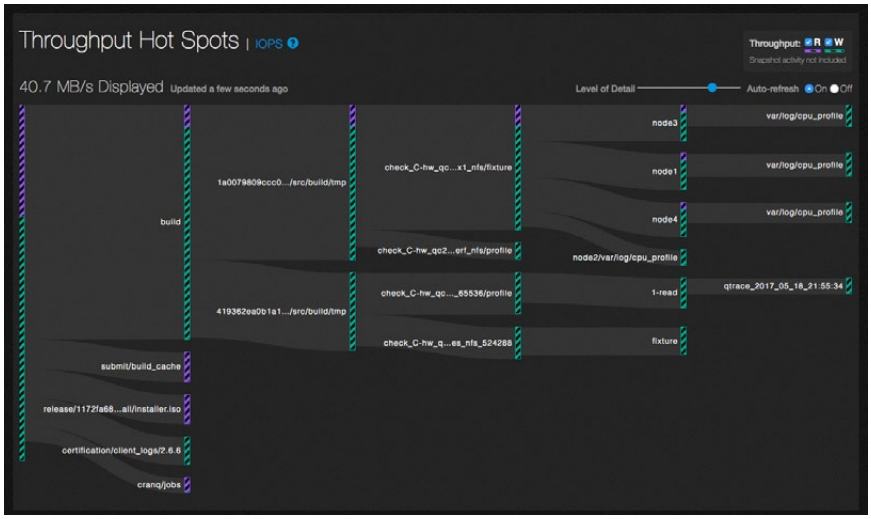


This view shows which individual files and directories are being accessed.

Qumulo also provides up-to-the-minute analytics that allow administrators to pinpoint problems and effectively control how storage is used. For example, you can instantly see how many IOPS and how much throughput you're consuming.

Throughput distribution

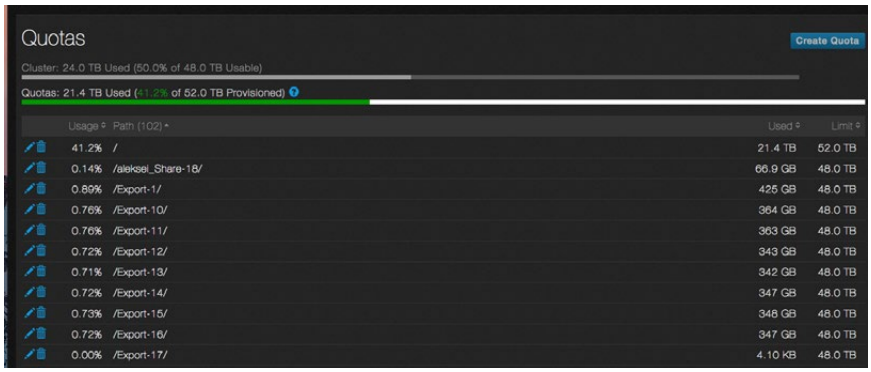
How is throughput being distributed across the file-system right now?



This view shows, in real time, how throughput is distributed across the file-system.

Directory-based capacity quotas

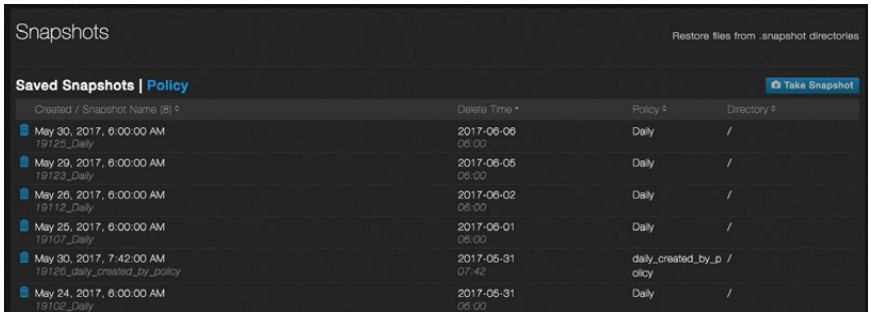
What directory levels quotas have been set and what is the usage of each?



This view shows the directory levels quotas that have been set and how they're being used. Administrators receive instant notification when quotas are approaching capacity.

File system snapshots

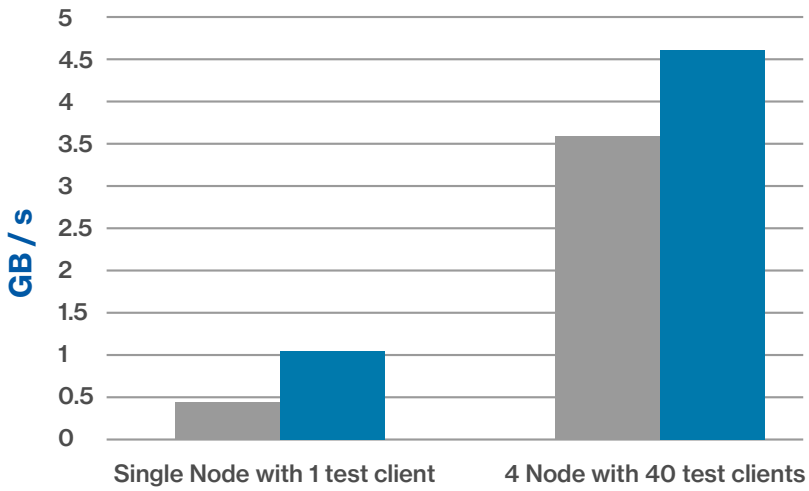
How many snapshots have been taken, and when?



This view shows how many snapshots have been taken, and when.

Performance considerations

Qumulo Core gives balanced performance for a variety of data-intensive workloads. For illustration, here are some Iozone test results on an HPE Apollo 4200 4-node cluster with 720TB total raw capacity.



All tests were run using Qumulo Core 2.6.4 over NFSv3. Speed is linear no matter the number of test clients. SPEC SFS® 2008 was measured as 4,870 per node on the Apollo 4200 server.

Qumulo regularly improves performance with software updates.

Qumulo subscriptions

Qumulo uses a subscription-based licensing model. That subscription is transferrable to future platforms, and you can license more capacity as you grow. The subscription also entitles you to be involved with the Qumulo roadmap and to receive new features as they are developed. With Qumulo, you do not pay extra for features such as capacity quotas and snapshots.

Customer success

A Qumulo subscription includes enrollment in the Qumulo Customer Care program. A customer success manager is assigned to your account and will be with you from the beginning. You can expect help with onboarding and regularly scheduled calls after that to see how you're doing.

Qumulo also monitors the health of your cluster. The data is aggregated and analyzed, and failures are detected in real time. For customers running Qumulo Core on Apollo 4200 servers, hardware issues will be reported to HPE's support organization.

The Qumulo online community is the central location for software release notes, new Qumulo Core builds, knowledge base articles, tutorials and open source scripts and utilities. The Qumulo Core open source community is on GitHub and you will find applications that Qumulo customers have built with the Qumulo Core API.

Qumulo customers give the company high marks. In the standard measure of customer satisfaction called "net promoter score" (NPS), Qumulo consistently scores well above the mean for enterprise software companies.

Conclusion

When the NAS category was invented more than 20 years ago, businesses saw the benefits of having a centralized, multi-user file system that was stored efficiently in hot-swappable disks and was accessible by anyone over the network. Data sharing improved, and administrators were able to easily scale up capacity by adding disks to their storage server. A few years later, the industry saw the emergence of scale-out NAS which, because of its distributed nature, was more fault tolerant and could scale storage capacity and I/O throughput at the same time. Scale-out NAS was a huge success.

Ironically, the immense success of scale-out NAS is a contributor to the current storage crisis. The amount of data stored by a typical scale-out NAS cluster has grown by a factor between 1,000 and 10,000 over 15 years. Storage capacities today are often measured in petabytes, and file counts numbering hundreds of millions or even tens of billions of individual files are common. Exabyte storage is on the horizon, representing a 20 million fold increase in scale since the birth of scale-out NAS.

While Qumulo Core could be thought of as scale-out NAS, its unique scale characteristics means it has little in common with legacy approaches to scale out. Because it can effortlessly store and manage billions of digital assets, Qumulo Core defines an entirely new category of modern scale-out storage for the new era of multi-petabyte scale.

Q150 0617

Qumulo, Inc. 1501 4th Ave. Suite 1600 Seattle WA 98101 1-855-478-6856 | sales@qumulo.com | qumulo.com