DATA CENTER FILE SYSTEM

Architecture White Paper

Quobyte®

Scalable, Unified Storage

Abstract:

This document offers a thorough exploration of the Quobyte architecture, and provides the reader with a complete picture of how our modern, scale-out software delivers robust performance without limitations.

Table of Contents

1.0. Introduction / Background

Quobyte storage software has its roots in an open source project that was created by several company founders as part of their PhD work. A fundamental aspect of this early research was a focus on distributed systems. Only after solving the challenges of building a fault-tolerant, fully distributed system—one that scales linearly and remains manageable even with thousands of nodes—did they turn their attention towards creating a powerful and equally scalable file system. This unique product evolution sets Quobyte software apart from nearly all other software storage solutions in the marketplace, and it's why Quobyte customers can be confident they will never outgrow the product.

After leaving university two of the company's founders went to work for Google where they learned what "management at scale" truly means. The founders applied the lessons learned to their own company, the result being Quobyte engineers understands large, complex systems. They recognize the need for storage products that can deliver fault tolerance, automatically handle disk and node failures, and unburden administrators tasked with managing such installations. Quobyte software delivers remarkable scale-out performance—a hallmark of the product—without driving up headcount or maintenance budgets.

In all, there are more than 10 years of development behind the product, making it truly one of the most mature software storage products shipping today.

Key Features

- Software-only solution (zero hardware or kernel dependencies maximize compatibility)
- Self-monitoring and self-healing (manage large clusters with minimal staff)
- High IOPs, low-latency client (also provides seamless failover)
- Policy-based data management and placement
- Volume Mirroring (perfect for disaster recovery, or as a backup source)
- Runtime reconfigurable (eliminates downtime, changes on the fly)
- Simple, scripted cluster installation (up and running in less than an hour)
- Compatible with any modern Linux distro
- Supports POSIX, NFS, S3, SMB, and Hadoop file access methods (one platform for all workloads)
- Works with HDDs, SSDs, and NVMe devices (no need to buy expensive flash when it's not needed)

2.0. Core Strengths

Three key aspects of Quobyte software distinguish it from other offerings. No. 1 is its linear scaling. Doubling the node count doubles the performance, it's as simple as that. No. 2 is unified storage, which allows multiple users to simultaneously work on the same file regardless of the access protocol being used. No. 3 is the system's ability to monitor, maintain, and heal itself, permitting small teams to start with a few terabytes and grow to manage tens to hundreds of petabytes while continuing to deliver data despite hardware failures.

2.1. Scalable Performance

Scale-up storage systems of the past invariably hit performance roadblocks, but so did early scale-out architectures. That's why Quobyte engineers spent years designing a system free of scaling constraints. The company's deep distributed systems knowledge led to the creation of a non-blocking design that enables Quobyte's highly optimized storage to deliver high throughput with very low latency (see figure 1). The architecture excels at handling metadata-heavy workloads like those found in EDA, media & entertainment, the life sciences, and financial modeling.

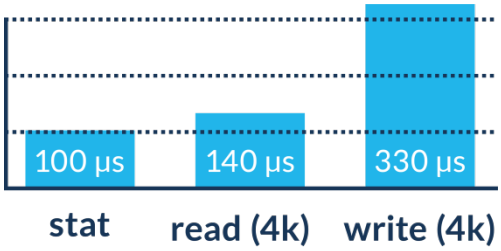


Figure 1

When accessing the Quobyte software distributed parallel file system, the native client can issue parallel reads and writes to all nodes of the storage cluster (see figure 2), a key factor in delivering stellar performance. In contrast, a typical NFS client requires an extra network hop and only connects to a single NFS service or controller, greatly limiting its performance envelope.

This "n-to-m" parallel communication is done via Quobyte's lightweight RPCs, and clients take advantage of all storage devices in the cluster unless restricted by a placement policy. As more nodes are added to the cluster, the performance increases linearly. Quobyte customers need not worry about "hot-filer" problems.

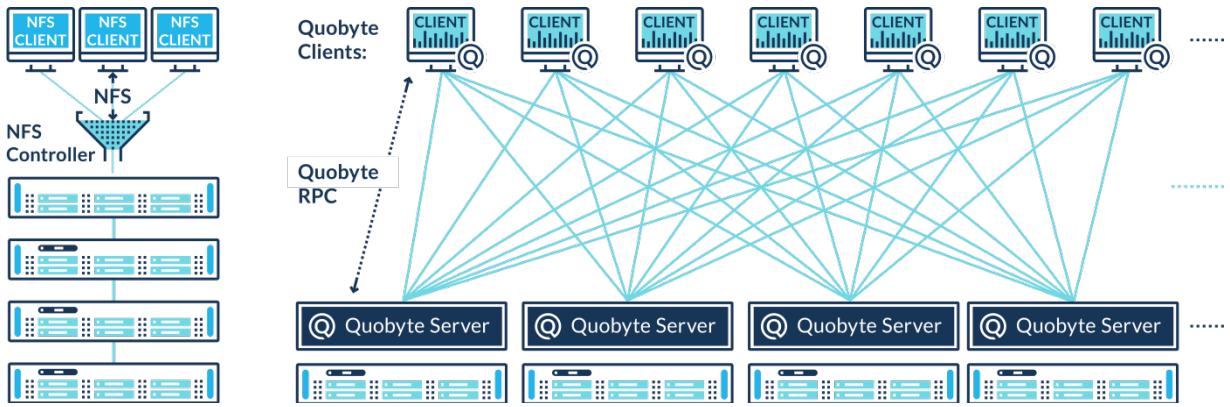


Figure 2

2.2. Unified Storage

Traditional storage systems have led to storage server sprawl problems. These aging, monolithic system architectures lead to IT infrastructures built around data silos, preventing easy accessibility between users and applications. Quobyte's unified approach to storage empowers admins to deliver performance, capacity, and scalability from a single storage infrastructure.

As depicted in figure 3, all clients have access to the same file, simultaneously, regardless of their access protocol. Locking (byte-range or file-level) is maintained across all protocols that support locks. Internally, the system maps NFSv4-style ACLs to each client, maintaining consistent file permissions regardless of their access method. In media & entertainment, a Windows user can be editing a video sequence while a Mac client is reading the same file – no need to copy or move the file to another system. HPC users can easily share their data across the globe via S3.

Quobyte software allows IT departments to provide a single consolidated storage platform, thus removing data silos imposed due to differing storage interface protocols. Virtually any environment where data needs to be accessible to and transferred between Linux, Windows, or Mac systems, via NFS, SMB or S3, can benefit from Quobyte storage.

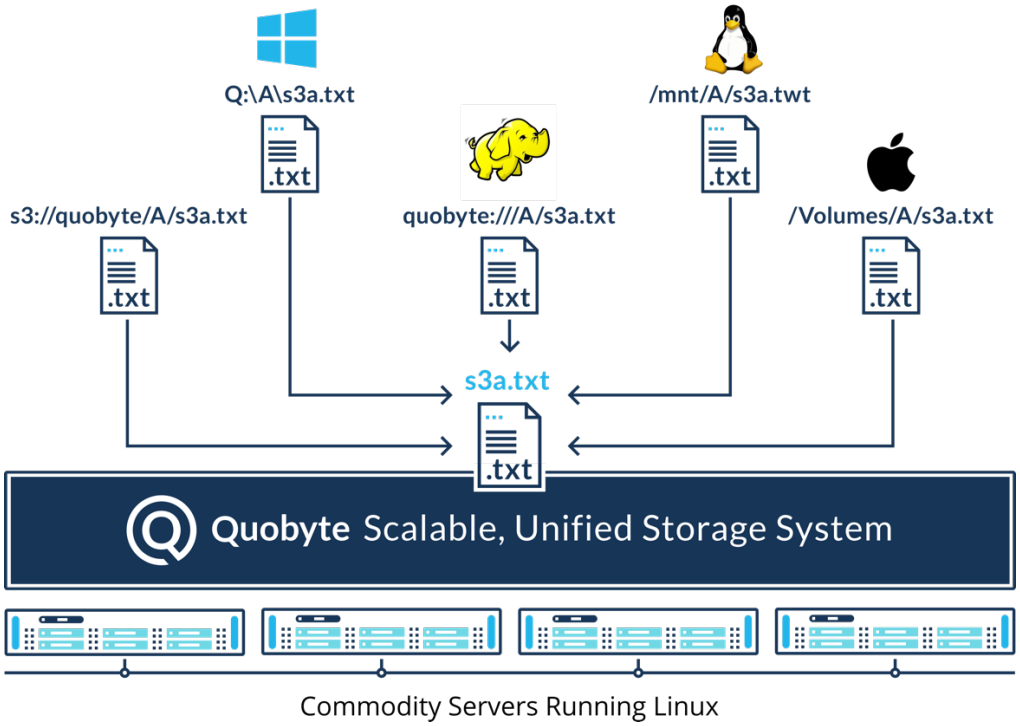


Figure 3

2.3. Scalable Operations

The Google Lesson

Back in the late 90's when Google built its eponymous search engine, the company quickly realized that the amount of data they needed to store and analyze would overwhelm existing storage systems of the day. And even if they could have purchased the requisite hardware, the manpower needed to operate and maintain it would have been prohibitively expensive.

Faced with these challenges, Google engineers capitalized on their distributed systems background and the new, at the time, Paxos research to develop a massively scalable, self-healing storage infrastructure. The ground-breaking system they developed easily scaled to keep up with the web's explosive growth and was fully manageable by just a handful of people.

Quobyte's founders have much in common with the Google infrastructure pioneers: from distributed systems PhD work to experience in HPC and hyperscaler environments. You could say that Quobyte shares much of the same DNA.

Applying the Knowledge

Typically, as the storage footprint grows, staff must be added to manage the larger environment. Quobyte storage software has removed much of the complexity that goes with managing modern storage systems. A Quobyte system can perform de-clustered rebuilds and takes care of replacing lost resources without requiring human intervention. Isolated islands of storage can be replaced with a single system that scales. Quobyte software's fault tolerance, built-in automation, hardware monitoring, and simple management interface greatly increase the productivity of any storage team.

Scalable Fault Tolerance, Scalable Operations

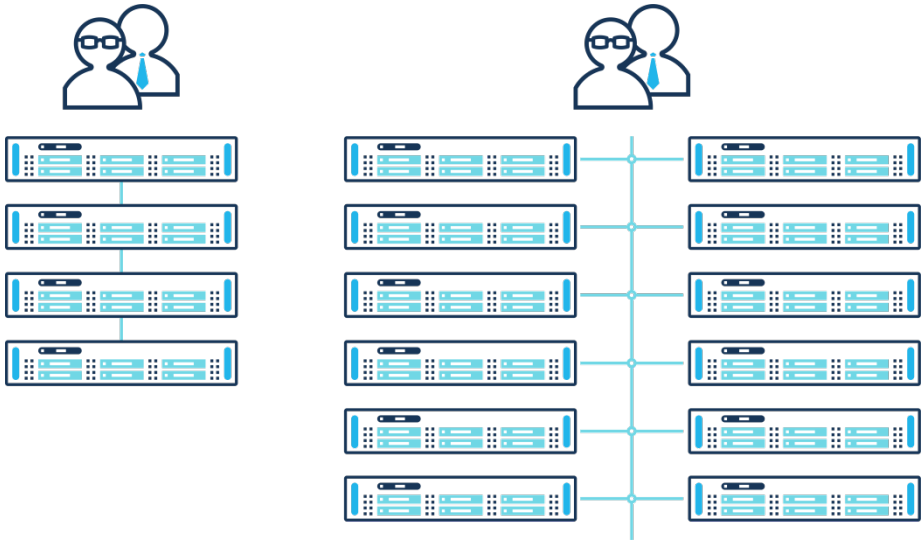


Figure 4

3.0. Software Architecture

3.1. Fundamental Objectives

Performance and scalability were the driving factors when design work commenced on Quobyte's storage software, with the foremost goal to achieve a near-limitless ability to scale-out. Other factors also came into play. For example, it had to be fast and reliable, with no single points of failure. During the development process the need for "low-touch" maintenance in large-scale installations was also factored in, and guided the product's automation specifications. The team then set about determining the best way to deliver peak performance without compromising on these additional design elements.

The Case Against Block Storage

Being keen to avoid the pitfalls that plague existing products, Quobyte engineers evaluated and rejected a number of designs, starting with the block-based approach that grew out of Storage Area Networks (SANs). SANs work well when only a handful of machines need to share a common resource, but this design fails miserably for scale-out.

Tracking millions or billions of storage blocks adds complexity, hampers resizing operations, requires a large memory footprint, and increases inter-node cache synchronization headaches. Quobyte software doesn't suffer from these problems because each storage node's local file system handles block mapping duties.

Object Storage Doesn't Cut It

Since object storage is all about handling billions of entities, this seemed to make sense in the world of hyperscale. Unfortunately, object storage was never designed for high IOPs performance, especially with small-file workloads. Object storage also has a completely different set of access semantics, meaning many applications must be rewritten to take advantage of it. Either that, or a performance-robbing gateway is placed in front of it to provide file-like semantics. Lastly, the object storage eventual consistency model is adequate for archival use but doesn't meet the needs of primary storage workloads that expect a POSIX-compliant file system – the Quobyte software file system, on the other hand, delivers strong consistency and POSIX-compliance.

File System for the Win

The product design path chosen by Quobyte engineers was really rather simple; provide a high-performance, low-latency, and POSIX-compliant file system that offers block-like speed with file system flexibility and ease of use.

Quobyte then went even further and developed its own key value store (KVS) rather than relying on existing products. This database was designed and optimized for file system use, and automated, no-touch, replica set management is deeply integrated into the underlying architecture. Because metadata operations are not part of the I/O path and disk-blocks aren't tracked globally, Quobyte storage delivers a scale-out system that block and object-based storage products cannot equal. And dynamic (per file) primary elections during data placement is completely decentralized, enabling near-infinite scaling.

3.2. Quobyte Services

The Quobyte software architecture consists of several services that all run inside Linux "userland." Not only does this design eliminate kernel dependencies and recompile issues, it also means Quobyte storage supports the widest range of Linux distributions of any software storage product. Binaries are supplied in .deb/.rpm package formats. Services have no local configuration, making swapping out a failed system incredibly simple. All client-to-server and inter-server communication utilize Quobyte RPCs, and the RPCs, as well as all data blocks, are Cyclic Redundancy Check (CRC)-protected to ensure data integrity.

Registry Service

The registry service stores the location and configuration information for the cluster. Clients use the registry as their roadmap to find all other services within the storage network. In addition, the registry service performs all of the cluster's monitoring and automation tasks. The registry is not a centralized locking service, nor does it determine node availability. To guarantee these functions are always available, the registry service runs on a minimum of four nodes.

Metadata Service

Next, is the metadata service, which performs the bulk of the work during file metadata operations. This is where everything except the file data itself lives. All attributes, ACLs, permissions, and a file's location are stored here. The metadata service is also responsible for determining the devices on which files are stored. With so many important functions, the metadata service is always deployed on at least four nodes and can scale to hundreds of servers. It can run on all nodes, if desired, even in very large clusters. If the lowest possible metadata latency is the goal, Quobyte recommends using SSDs or NVMe devices to store the metadata.

Data Service

The data service is where actual file I/O occurs. One data service instance runs per node and can manage one to hundreds of devices. In practical terms, the number of data service nodes can be in the hundreds or more. Lightly loaded clusters can run all three Quobyte storage services on the same nodes if desired. Customers who require maximum performance will likely run dedicated nodes for metadata and data services.

4.0. Policies

4.1. File Layout

Quobyte software gives administrators the flexibility to use differing protection schemes for different use cases. The setting is assigned as part of the volume configuration but can also be applied at the individual file level based on the file's naming pattern, file age, or file extension. Layouts can be created for specific application profiles (e.g. Cassandra), and the layout can be changed after a volume has been in use to provide maximum flexibility.

4.2. Data Placement

Device Tags

Device tags are simple text strings that administrators can assign. The system automatically assigns HDD and SSD tags based on a device's type. But administrators can add additional tags to enable very sophisticated placement strategies.

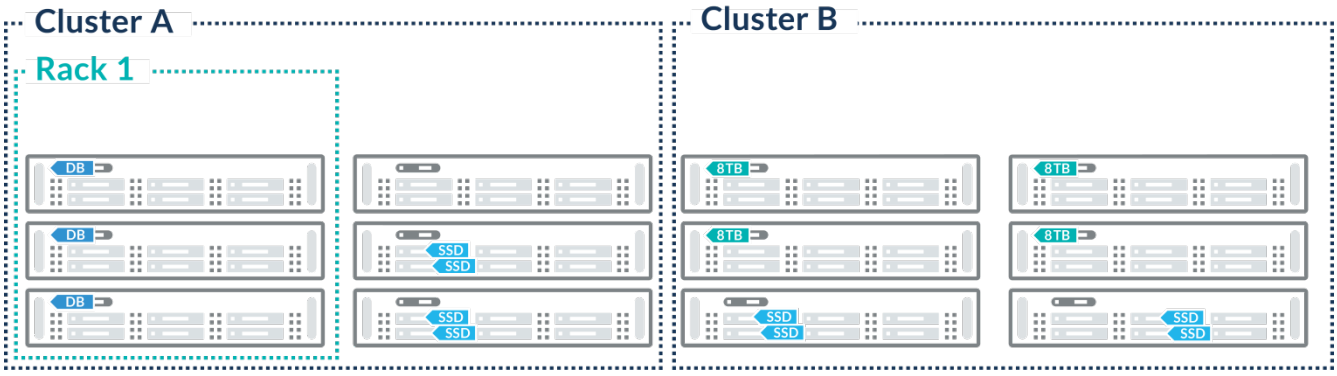


Figure 5

Partitioning

Using the aforementioned tags, admins can partition the hardware to service different workloads with a specific set of devices. For example, tenant X's user data can land on nodes 1-10 while tenant Y's users are serviced by nodes 11-20. Administrators can take it a step further and isolate user or application workloads on demand so that the workload gets guaranteed performance from dedicated hardware. All of this flexibility is fully configurable at runtime, giving a level of agility not widely available. If the partitioning layout is changed, the software will move the data to the right location.

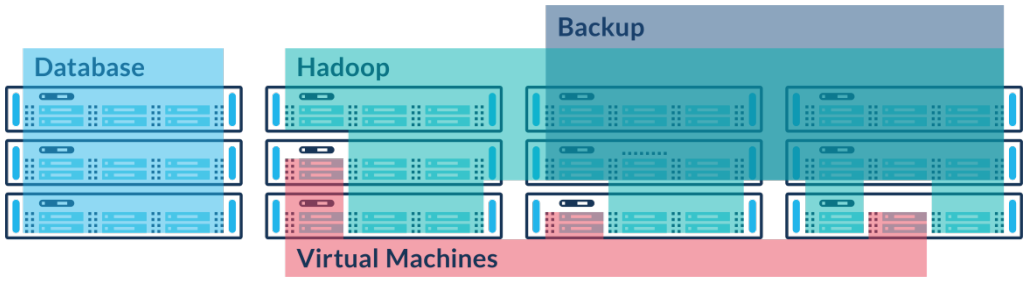


Figure 6

Automated Tiering

Because all devices are automatically tagged with their device type, admins can define a simple age-based placement rule that tiers older data from SSD to HDD automatically.

Figure 7

Placement Policies

Quobyte software's flexible data placement options extend beyond simple tiering. For example, database admins could specify that files with the .log extension should always go to SSD, ensuring that database logs always get the best performance.

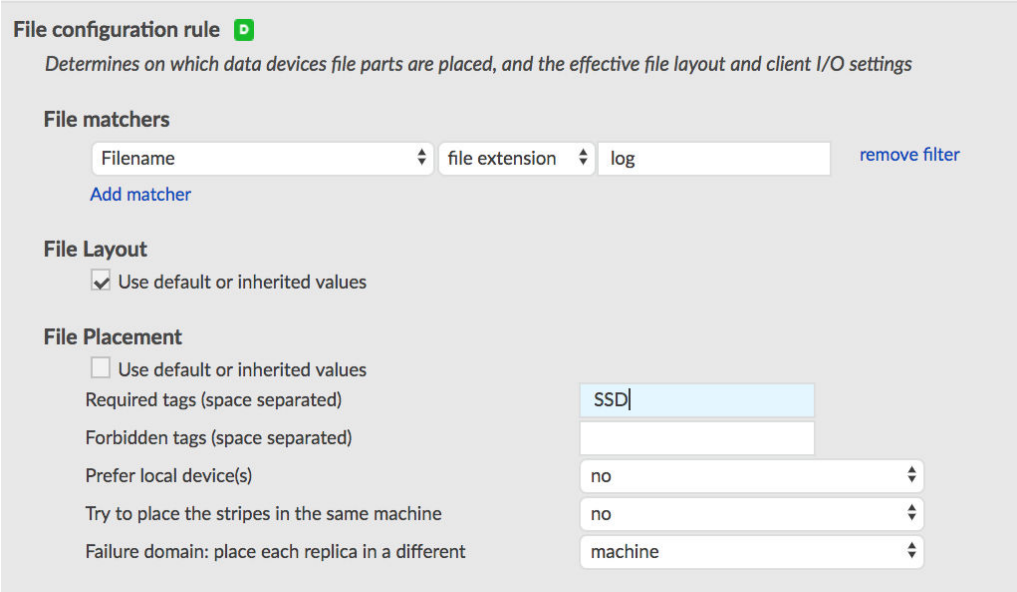
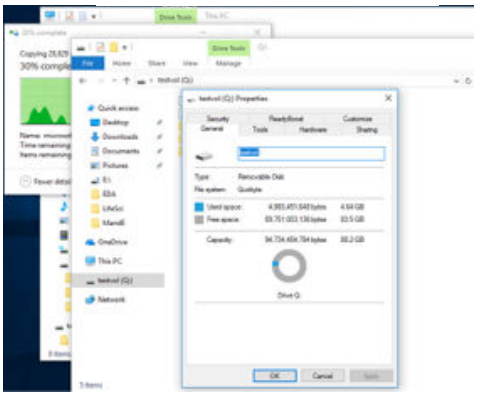


Figure 8

Client Software

Quobyte client software is available for Mac, Windows, and most Linux variants. When accessing Quobyte storage — a distributed parallel file system — the client can communicate with ALL nodes in the cluster, a key factor in delivering stellar performance. In contrast, a typical NFS client requires an extra network hop and only connects to a single NFS service or controller, greatly limiting its performance envelope. Extensive client-side performance and monitoring metrics are also collected by the client and can be viewed via the web console. The client requires no local configuration, making rollout out to thousands of nodes easy.

Windows mount example:



Linux / Mac mount example:

```
/quobyte/volumeA/foo.mp4
```

Figure 9

FUSE for Linux and Mac

Both the Linux and Mac clients deliver a POSIX-compliant file system that mounts just the same as an NFS share. Linux or macOS applications can run on Quobyte storage immediately, without any need for modifications. Based on FUSE, the client doesn't require any kernel changes or custom modules, as all processes live in "user-land."

5.0. Reliability Features

Keeping data safe is the single most important task for any storage system. Furthermore, data needs to remain safe *and* available even if there are hardware or software failures. This has been the primary focus of Quobyte research, and the engineers never forgot it as they developed a product from the ground up that addresses data safety and integrity with multiple layers of protection.

5.1. Replication

Customers seeking the highest possible performance and lowest latency for random IO and small file workloads should deploy them on Quobyte storage replicated volumes. Data is stored on a primary device and simultaneously replicated using a 4k (user-configurable) block size to two backup copies. Since Quobyte clients can read from multiple replicas, the combined bandwidth of all three servers is delivered and the "hot file/filer" issue is mitigated.

5.2. Erasure Coding

Erasure coding (EC) is offered as an alternative to replication and is user-configurable on a per-volume or per-file basis. Besides being more space efficient, erasure coding is ideal for large-block sequential workloads, like those found with streaming media, surveillance, analytics, or transcoding. EC is also ideal for long-term archiving, where space efficiency is key.

5.3. Failure Domains

There are several means by which a Quobyte storage administrator can further protect the cluster from unplanned outages. Failure domains tell the software that a set of physical nodes or individual devices should be treated as a unit. Simply put, a failure domain is a physical or logical boundary in a network or storage system that becomes unavailable through some sort of fault. The distinction can be based on which nodes connect to the same power feed, which machines are part of the same physical rack, etc. When data is written, logic code guides the placement of the data copies in such a way that all file replicas land in different failure domains.

5.4. Checksums and Hardware Monitoring

Quobyte software employs end-to-end checksums that are generated on the client and are saved on disk along with every data block. This feature prevents possible data corruption that can occur due to faulty network cards, misbehaving switches, and even disk controller anomalies. In addition, the RPCs themselves are also CRC protected, eliminating network-induced error. Furthermore, watchdog timers work to ensure IO operations don't get stalled by faulty disk drives.

This high degree of protection must be baked-in to the core architecture from the start, not as an afterthought. It's why Quobyte customers can safely deploy on low-cost SATA drives. Other solutions force you to run more expensive "enterprise drives" or even specific firmware versions.

5.5. Split-Brain Safe

Quobyte storage is fully protected against network anomalies as well. Partial outages can result in "split-brain" scenarios where it's unclear which version of the data is correct. In these circumstances, Quobyte storage will *never* deliver data it cannot be sure is consistent and verified. The software protects against real-world issues like broken or misbehaving switches, faulty NICs, or bad "heartbeat" cables that can result in packet loss or corruption. In fact, there is no single point of failure *anywhere* in a Quobyte cluster.

5.6. Health Manager

Delivering high-performance storage while reducing the administrative workload is a key design goal for Quobyte. Health Manager is a set of fully automated tasks the system performs to keep the storage cluster operating properly, maximizing resource utilization, and eliminating potential failures.

Periodic rebalancing, data scrubbing, and enforcement of placement policies are some of the tasks that the Health Manager executes. Individual tasks can be further refined as to the actions they take, when they will run, and how many tasks will execute simultaneously. Tasks are always run at low priority to minimize their impact on the production workload. Administrators can designate maintenance windows to assure the tasks don't interfere with production schedules.

5.7. Automated Monitoring

Numerous aspects of the Quobyte storage cluster are continuously monitored for proper behavior. If something drifts beyond reasonable presets, an automatic alert is sent to the web console, and an optional email alert can be dispatched.

5.8. Maintenance on Your Schedule

Quobyte storage fully supports rolling updates (both forwards or backwards) and running mixed minor releases is not an issue. Tasks like hardware break/fix, BIOS updates, etc. can take place while the rest of the cluster continues to service users. In fact, a live server can be pulled with no deleterious effects to the system. Hardware that dies can be repaired or replaced when it's convenient, so say goodbye to "fire drills." Likewise, kernel or security patches can be applied when deemed necessary. Quobyte likes to say storage teams can take back their weekends!

6.0. Management Flexibility

6.1. Single-Script Installation

Deploying a new cluster or adding/removing nodes on an existing cluster has been simplified by Quobyte's easy-to-run

installation script. Just by answering a few prompts a brand-new cluster can be installed and operational in less than 45 minutes. With no local configuration to worry about, hundreds of servers can be deployed using a tool like Ansible.

6.2. Runtime Reconfigurable

Unlike the old monolithic storage hardware of the past, almost any setting can be modified while the Quobyte storage system is active. Need to change where the data lands? No problem. Have some very important data that needs extra protection? Just change the replication factor and it's good to go. Temporarily move some devices or whole nodes to another failure domain with a few mouse clicks. Downtime is expensive, so why not eliminate the need for it in the first place?

6.3. Composable Quotas

Quobyte storage fully supports several types of quotas. Admins can set a limit on the number of files, the raw or logical space being consumed, or the number of volumes. These quotas can then be applied to tenants, users, or groups.

6.4. Selective Prefetching

This Quobyte-exclusive feature is aimed at administrators needing to support frame-per-file video applications. Content creation tasks like VFX and color grading place huge demands on both storage and the network, especially when dealing with 4K raw resolutions and above. After defining a file naming/numbering sequence, the Quobyte client will automatically read in additional files matching that sequence. Prefetching is adaptive based on the performance of the cluster's drives and network. The performance difference can be substantial, and delivers jitter-free playback.

6.5. Elastic

Quobyte software was built to be flexible in all regards. Nowhere is this more evident than when adding or removing storage. Unlike some products, the system doesn’t force a painful rebalance operation whenever adding or removing nodes.

Added performance and capacity are useable immediately; individual drives or entire new servers are auto-detected and made available right away. And if a storage resource must be taken offline it can be drained with one click. Any replicas it hosted are recreated elsewhere in the cluster.

6.6. Thin Provisioning

Naturally, Quobyte storage supports volumes, and yes, they are thinly provisioned. Each volume behaves like a separate file system, and the system supports hundreds of thousands of volumes. By default, every storage device of every storage node is utilized for maximum performance. Administrators can alter this behavior using placement policies.

7.0. Security

Along with its strong data availability features, Quobyte software provides high security for critical data as well. ACLs are maintained and enforced across all interfaces, and X.509 certificates can be used to verify that machines and users are—in fact—trusted. IP filters can also be employed to assure access only comes from trusted networks. Access to the management APIs or web console can be validated against external LDAP, Active Directory, or OpenStack Keystone data, or the system's own internal database.

7.1. Multi-Tenancy

Quobyte software storage fully supports multi-tenancy, letting customers define isolated namespaces for their storage. Tenant group members can only see the volumes assigned to them, while tenant admins can manage resources of their own group but are prevented from seeing the entire cluster.

Quotas can be shared by tenant members, and administrators can further isolate tenants by controlling which physical hardware they have access to.

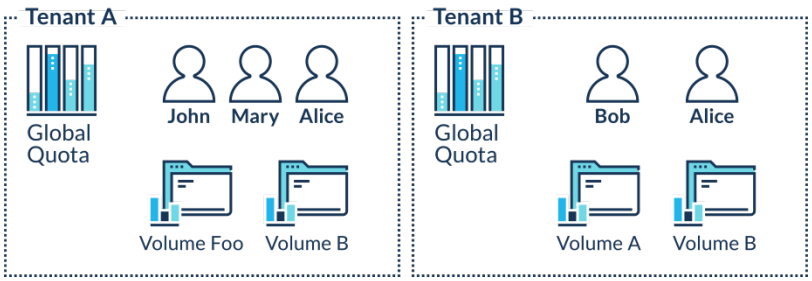


Figure 10

8.0. Performance Monitoring

Understanding and predicting storage needs is difficult, especially when there are many workloads and users hitting the system. Quobyte software provides tools to better understand what is going on in the system, to find "troublemakers" and to identify potential and future performance and capacity bottlenecks before they turn into a problem.

The software offer two methods for keeping tabs on the cluster to make sure it's running as intended. Built-in real-time analytics let administrators pinpoint the top IOPs, throughput, and metadata consumers. Long-term trends can be studied by exporting data for external analysis.

8.1. Real-Time Analytics

The web-based Quobyte management console provides a "right-now" assessment of system performance. IOPs, throughput, and metadata operation charts give an instantaneous view of what a volume's load is.

Rich client-specific information, such as username, application or process name, and file name are all displayed for the top consumers, making it easy to track down who or what is slowing down the system.

IOPs, latency, and throughput graphs are displayed for each individual storage device.

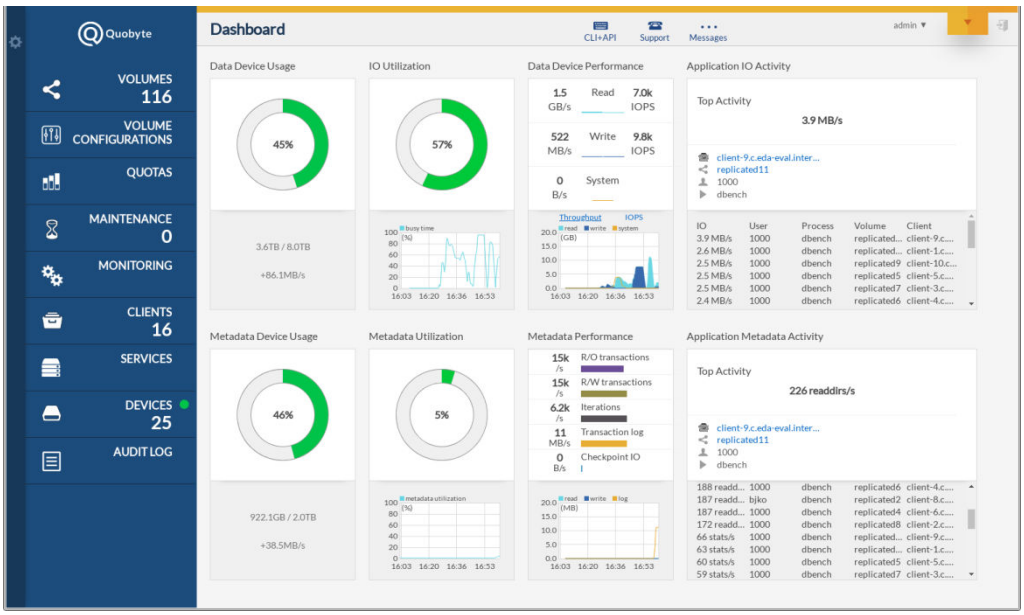


Figure 11

8.2. Long-Term Statistics

Long-term trending analysis can take advantage of Quobyte software's Prometheus-compatible output. Collecting and parsing data over extended time periods can then be done using one of the many open-source tools available, such as Grafana. In contrast, several current storage products rely on off-premise analysis of your data, which often violates corporate security policies.

9.0. Disaster Recovery

Being able to withstand a catastrophic outage is something companies today must factor into their planning. Whether from natural causes or deliberate malfeasance, business continuance is always a worry.

9.1. Volume Mirroring

The Quobyte software volume mirroring feature gives administrators the ability to remotely mirror a volume from one cluster to another, regardless of location. The asynchronous replication is done on a per-file basis and operates continuously. Source volumes can even have different encoding than the targets, e.g., replicated volumes can mirror to erasure-coded volumes, resulting in greater space efficiency.

Users have read-only access to the mirror target during normal operations. If an administrator disconnects the mirror, the target volume then becomes read-write. Source volumes can be mirrored to multiple target clusters for even greater levels of redundancy.

10.0. Management Console

Quobyte software greatly reduces the need for highly-trained storage administrators, beginning with the management interface. Starting with a clean, simple-to-understand dashboard, the web console provides "at-a-glance" status information.

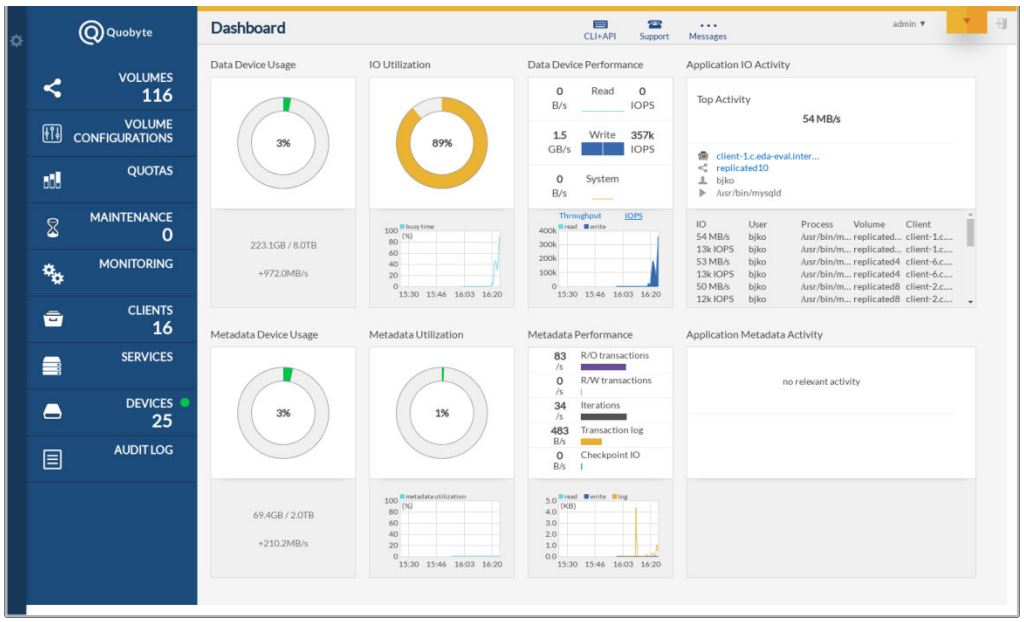


Figure 12

Many screen elements can be expanded to display operational and performance data, such as this view of the storage devices (see figure 13). Administrators can drill down even further for comprehensive information about each of the Quobyte software services, as well as client-side operations.

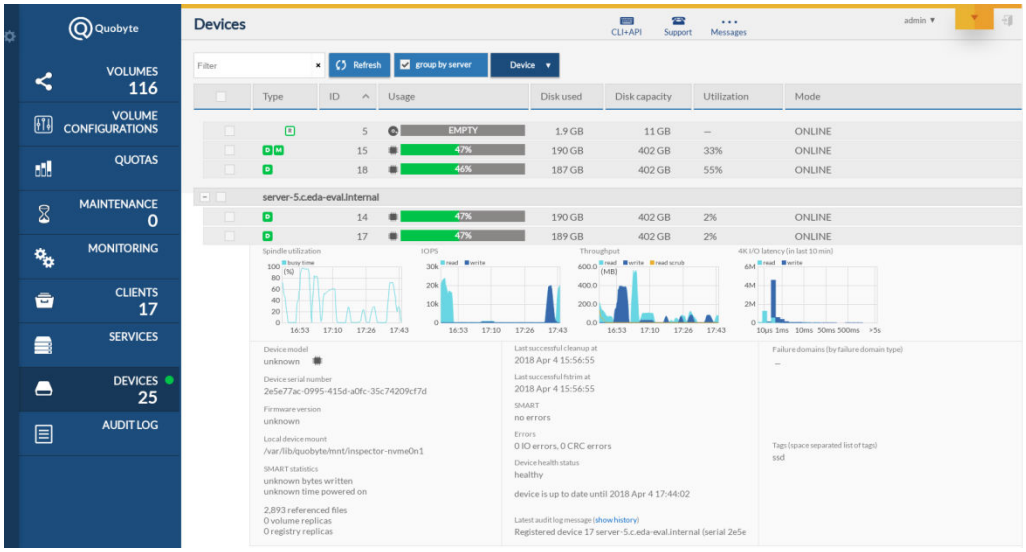


Figure 13

11.0. Quobyte API

Customer preference today is for open systems, and Quobyte software fully support this via the product's JSON-RPC APIs. Administrators can use these APIs to deliver things like self-service provisioning portals for users, or integration with their internal management platforms. In fact, Quobyte's OpenStack and Kubernetes drivers make use of these APIs.

12.0. Integration

Today’s storage needs to work in today’s ecosystems. Quobyte software provides broad support for everything from OpenStack to bare metal. Regardless of where a company is on the technology adoption scale, Quobyte storage can support it. Proxies are available to support NFSv3/v4 and S3 protocols, as is a drop-in HDFS driver replacement for Hadoop. The software also supports using Samba proxies for SMB (Windows) clients.

12.1. NFS Proxy

Customers who must support NFS for legacy clients can deploy the Quobyte NFS proxy software. The service can run concurrently with other Quobyte services on the cluster nodes or as a stand-alone. Multiple proxies can be utilized for higher performance. The NFS proxy employs Quobyte client connectivity to the storage cluster for peak performance and automated failover.

12.2. Scalable S3

Quobyte's S3 software service differs from most competitors' products in that it doesn't force customers to use erasure coding when storing their data via the S3 protocol. Because S3 objects are really just files in the Quobyte file system, the proxy merely translates S3 operations into Quobyte RPC file calls.

Admins can pick either replication or erasure coding for the storage method, resulting in greater performance and efficiency.

Any volume can be exported via S3, and an object's ACLs are the same as for file system access. SSL and multi-part uploads are supported, and users can be authenticated against LDAP, Keystone (OpenStack) or internal databases. The Quobyte S3 proxy software service is stateless, so customers looking for maximum performance can run multiple S3 proxies in parallel behind a load balancer.

12.3. Docker

People utilize Docker because it delivers a responsive infrastructure on which to deploy applications. Quobyte's Docker plugin software only requires a single mount point while providing access to multiple volumes. By exposing each volume as a subdirectory under the mount point, it saves time as well as system resources. With Docker version 1.10 or above, it's possible to mount Quobyte storage on a host while running the client inside a container. Application containers can then access this shared mount point using only one Quobyte client per host.

With Quobyte software’s implicit locking feature, admins can make standard Linux applications highly available by taking advantage of the failure handling of Quobyte software in a container infrastructure.

When a host fails, the application container is rescheduled on a new machine and can recover its previous state. Quobyte software ensures that two instances cannot accidently access and potentially corrupt the same data.

12.4. Kubernetes

Having deployed some form of containerized infrastructure, many IT departments are embracing Kubernetes (K8s) as their orchestration platform of choice. Quobyte was one of the early providers of an "in-tree" storage plugin for Kubernetes.

External Quobyte Cluster Providing Storage Resources to Kubernetes Nodes

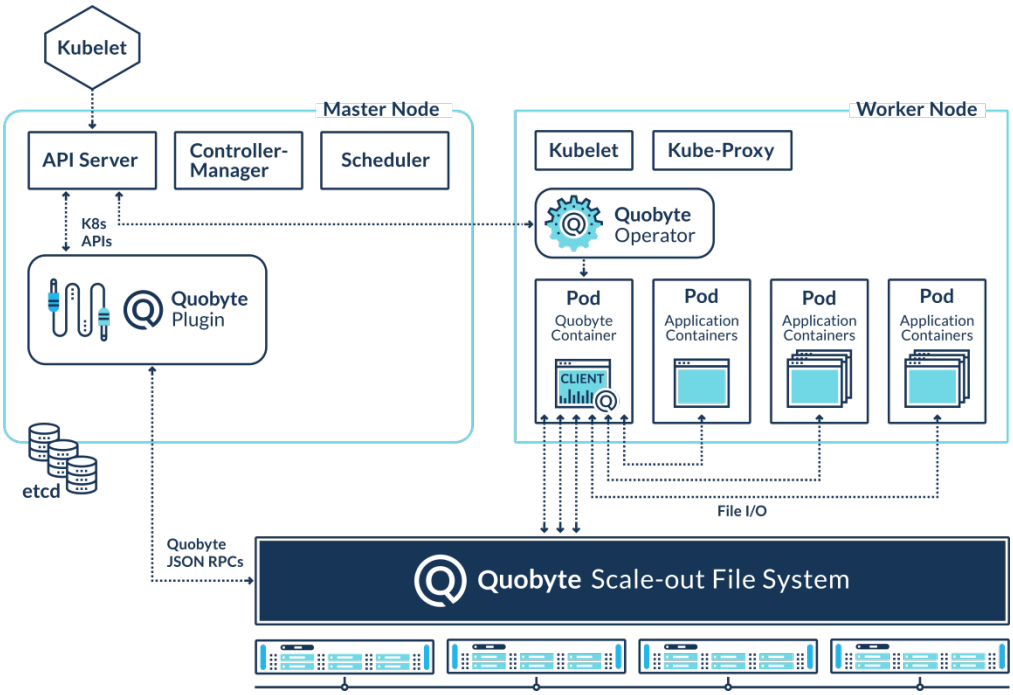


Figure 14

Quobyte's software plugin lets administrators share volumes to a Kubernetes cluster in several ways. First, it can be used as a Volume Source, although this methodology has since been superseded. Existing Quobyte volumes can also be exposed using a PersistentVolume (PV) definition. Once the PV is defined, containers acquire access to them via a PersistentVolumeClaim (PVC). Quobyte software also supports Kubernetes' dynamic provisioning feature through its StorageClass driver. Using this approach, administrators no longer need to predefine volumes ahead of time. Instead, when the PVC makes a request and specifies the StorageClass, the driver creates the desired volume automatically.

Leveraging the May 2018 release of Kubernetes' Operator SDK, containerized client and server services can now be configured using the Quobyte software operator. Besides simplifying initial setup and ongoing maintenance of any sized cluster, the operator helps assure requested services are running properly, and can execute rolling upgrades. Additional functionality will be delivered in future releases.

Fully Containerized Quobyte Services Providing Storage Resources to Kubernetes Nodes

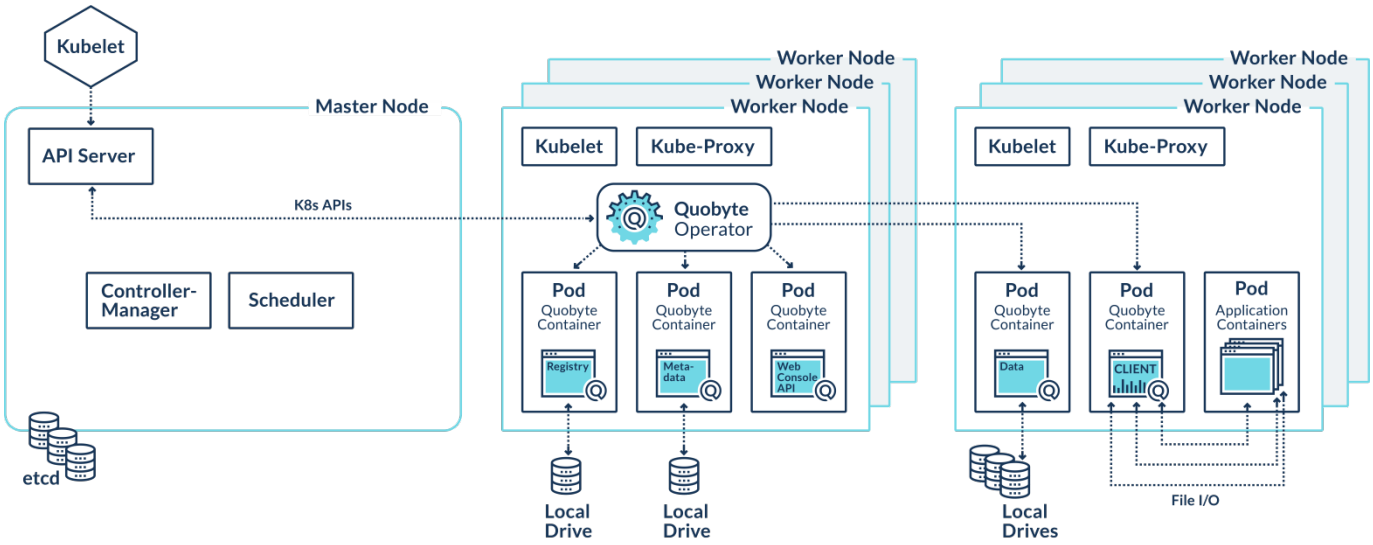


Figure 15

Kubernetes is moving to replace its plugin architecture with the Container Storage Interface (CSI) specification. Quobyte already supports CSI, and is actively developing enhanced management features based on this standard.

12.5. Hadoop Driver

By running existing big data workloads on Quobyte software using the drop-in HDFS driver, customers can replace HDFS with efficient, erasure coded storage. The driver works by translating HDFS file operations into Quobyte native RPCs. Administrators running an HCI configuration can enable a volume option to provide data-aware placement and scheduling for Map-Reduce jobs.

12.6. OpenStack

Ever since the “Kilo” release, support for Quobyte storage has been built into OpenStack. The software supports block-based storage under Cinder, shared files under Manilla, and Glance for images. OpenStack domains also integrate with the Quobyte software's multi-tenancy. Objects are handled via the S3 proxy, and authentication via Keystone integration makes Quobyte software one of the most complete OpenStack storage products around.

13.0. Hardware Requirements

When it comes to hardware, Quobyte storage has modest requirements. Obviously, performance requires horsepower, so running with fast SSD or NVMe data devices will necessitate a powerful CPU. A reasonable starting point is an eight-core Intel E5-series CPU (or equivalent), and 32GB of RAM.

For networking Quobyte recommends 10Gbps or faster Ethernet-based IP networks. Costly RAID controllers aren't needed because fault tolerance is built into the software. Customers can extend the useful life of their servers because Quobyte software allows mixed hardware generations—even from multiple vendors—to exist within the same cluster.

13.1. Basic Node Configuration

- Reasonable CPU and RAM requirements
- IP network (TCP, UDP, no multicast, no storage backbone)
- No need for RAID controller, NVRAM or journaling devices
- Heterogeneous hardware (mix generations, models)
- Almost any modern Linux distribution

Typical 2U, 12 bay storage node

Figure 16

14.0. Summary

Quobyte software provides non-stop, high-performance storage for the most-demanding workloads using economical, commodity hardware. Engineered to be a complete solution, the product does away with operational complexity, replacing it with scalable operations that require fewer human resources. Linear scaling, run-time configuration flexibility, and real-time performance monitoring means responding to changing storage requirements is nearly instantaneous. Support for rolling upgrades and non-disruptive cluster expansion removes the burden of planned outages.

Quobyte storage offers the freedom to choose how data is stored, and delivers that data to clients over the broadest set of access protocols, all while maintaining consistent access control. Customers are able to combine the benefits of flash with the storage efficiency of hard drives, the product never forces the use one when the other will do. Quobyte delivers storage independence.

Quobyte Inc.
4633 Old Ironsides Drive · Suite 150
Santa Clara, CA 95054 USA
(650) 564-3111

Quobyte GmbH
Hardenbergplatz 2 · 15th floor
10623 Berlin, Germany
49 30-814 591 800