

2022-23 DCIG TOP 5



RISING VENDORS IN STORAGE FOR HEALTHCARE SOLUTIONS

By Mike Matchett

Table of Contents

3	Rising Vendors
3	Storage for Healthcare Challenges
3	Benefits of an Effective Storage for Healthcare Solution
4	Distinguishing Features of Storage for Healthcare
5	Similarities Among the DCIG TOP 5 Rising Vendors in Storage for Healthcare
5	Differences between the DCIG TOP 5 Rising Vendors in Storage for Healthcare
6	DCIG TOP 5 Rising Vendors in Storage for Healthcare Solution Profiles
6	Qumulo
7	Nasuni
7	Pavilion HyperParallel Data Platform
8	VAST Data Universal Storage
9	WekaIO
10	Inclusion and Evaluation Criteria for Storage for Healthcare Solutions
11	DCIG Disclosures

Rising Vendors in Storage for Healthcare Solutions



Qumulo File Data Platform

Nasuni

Pavilion HyperParallel Data Platform

VAST Data Universal Storage

WekaIO

**Licensing vendor is listed first;
others listed in alphabetical order*

SOLUTIONS EVALUATED:

- DDN EXAScaler
- Dell PowerScale F900
- Hitachi Vantara HCP
- HPE ClusterStor E1000
- Huawei OceanStor 9000
- IBM Spectrum Scale
- Nasuni
- Netapp HPSS (E-series)
- Panasas ActiveStor Ultra
- Pavilion HyperParallel Data Platform
- Pure Storage FlashBlade
- Quantum StorNext
- Qumulo File Data Platform
- Quobyte
- VAST Data Universal Storage
- WekaIO WekaFS

SOLUTION FEATURES EVALUATED:

- *Deployment Capabilities*
- *Data Protection*
- *Product and Performance Management*
- *Documentation Support*
- *Technical Support*
- *Licensing and Pricing*

Rising Vendors

DCIG recently conducted research on the Top Storage for Healthcare solutions. This report focuses on the top solutions by vendors of size less than \$200M whom we consider to be “rising” in the market. The use cases are the same, but rising companies tend to be younger companies bringing newer technologies and disruptive solutions to market. Those organizations looking for modern approaches and potential competitive advantage may want to consider working with a rising vendor.

Storage for Healthcare Challenges

Healthcare IT must tackle significantly more challenging data handling requirements than most enterprises. In this environment, data storage is perhaps the most difficult architectural component to structure and operate both effectively for the mission and correctly for security. Not only is ultra-high availability, remote collaboration and fast data access critical to treating patients and achieving good patient outcomes, but data protection needs to be paramount. And both compliance and data privacy regulations ensure that healthcare IT has a lot of demanding work also cut out for them in managing data access, integrity/retention and disaster recovery.

Because of these challenges, healthcare is lagging behind many industries in IT modernization efforts. But there is incredible opportunity. Many healthcare organizations have now adopted key transformation initiatives in providing connected/remote care, conduct clinical and population health analysis, coordinate patient care across locations, improve patient engagement and foster more personalized/precision medicine.

Specific healthcare workloads requiring aligned modern data storage can include PACs/VNA (medical imaging and image archives), telemedicine/telehealth distance sharing and collaboration (VDI etc.), EMR/EHR (patient records) and big data powered research analytics. These applications require local performant access to large media files (both images and video streams), secure sharing and collaboration across distances, and large on-line archives.

Imagine a common radiology scenario in which a successful patient outcome might depend on accessing and sharing a critical patient medical image in real-time between a hospital and an off-site pathologist. Large image data needs to be available on-demand, the image application needs fast streaming quality access, and online sharing needs to be supported and secured—all can be challenges for healthcare IT stuck with weak data storage solutions. In another common scenario, patient health records may be hard to access even internally (much less transferred to the point of patient care) due to overburdened or legacy storage systems that are still in use because of the fear of meeting legal compliance or security concerns (i.e. protecting data in a cloud or cloud-like distributed system).

Benefits of an Effective Storage for Healthcare Solution

The modern storage solutions evaluated in this report all have features that address IT modernization challenges and concerns. From a healthcare business perspective, they support and enable many of the desired healthcare IT initiatives mentioned above. For example, these storage solutions can be used to modernize and consolidate unstructured file data storage that is often found locked inside aging NAS arrays local to each

Rising Vendors in Storage for Healthcare Solutions

Data storage can become a significant expense and useful medical data is growing at an exponential rate.

hospital data center. By aggregating such data into a more global, distributed “file system”, data management policies can be applied immediately and universally while ensuring wide secure access for collaboration and sharing. Many of the solutions address performance with “edge” devices that can provide flash-speed streaming media access to local applications.

It's hard to imagine truly proving compliance and really protecting critical data without the support of a modern healthcare storage solution that can natively capture immutable copies, audit access, enforce retention/deletion, and recover on-demand. Storage with native compliance features and policy-based management directly address the data concerns found in HIPAA, HITECH, GDPR, PIPEDA and other regulations. In addition, data with provable integrity proves a better foundation to ensure healthcare insurance and billing are correct, defensible and complete.

Cost is a major concern in the healthcare industry, data storage can become a significant expense and useful medical data is growing at an exponential rate. The best storage solutions can help address growing storage cost/capacity concerns with features like deduplication, copy reduction, inherent snapshots or built-in backup/archive. Some will tier from flash all the way down to tape, while most enable integration to cloud storage to take advantage of utility scale and pricing. And top storage solutions will make storage management easier on staff with many consolidating end-to-end storage management into a single pane-of-glass.

Finally, data protection requires security from both hackers and insider threats. Encryption features are critical. And ransomware attacks can be a life or death issue in healthcare. Top storage solutions will offer native features to help protect and/or recover quickly from ransomware attacks like immutable snapshots, WORM options, and quick restore to fine-grained recovery points.

Distinguishing Features of Storage for Healthcare

In addition to broad feature sets mentioned above, all of the storage solutions evaluated in this report have a few things in common that distinguish them for the healthcare use case from the wider market of IT data storage solutions.

Scalable capacity. They can scale capacity (to many PB) cost-effectively, enabling global NAS consolidation and the creation of a single file system architecture.

Performance at scale. They have the ability to deliver low latency and high throughput streams to support the most demanding data-intensive applications even as they scale.

Interoperability and access. Multiprotocol support and a wide range of connectivity to meet a variety of workload requirements while delivering IO to a large set of simultaneous, distributed users.

Architectural flexibility. They generally can adopt or include a variety of underlying storage media, storage server node types and cloud extensions as storage tiering needs grow and expand with a proven track record of refresh and updating to new technologies.

Enterprise management. They offer solid, mature management approaches and policy-based management for data protection, security, replication/synchronization, access and audit, monitoring and lifecycle data governance.

DCIG TOP 5 solutions are all based on cost-effective and performant scalable capacity architectures, and offer cloud integrations for additional capacity and functionality.

Similarities Among the DCIG TOP 5 Rising Vendors in Storage for Healthcare

In addition to the distinguishing features above that all evaluated storage solutions share, the select DCIG Rising Vendors in Storage for Healthcare solutions all have the following traits in common:

Application files focused. These storage solutions focus on delivering files to optimize consuming healthcare application (e.g. medical imaging, records retrieval) performance at large scales. They all serve large-scale big data/data lake/AI and enterprise datacenter applications too.

Designed for data-intensive verticals/industries. These solutions have been adopted across many data-intensive verticals like finance, media & entertainment and oil & gas because of the benefits of having a globally accessible namespace, creating an enterprise-wide single data and storage management solution, and generally accelerating aggressive IT modernization initiatives.

Cost/capacity focused. These solutions are all based on cost-effective and performant scalable capacity architectures. Additionally they all offer cloud integrations (e.g. to AWS S3 or Google Cloud) for additional capacity and functionality.

The Rising Vendors in Storage for Healthcare solutions also all deliver the following product features:

- Space efficient snapshots
- Asynchronous replication
- Pre-integrated appliance options
- Array-based secure encryption

Differences between the DCIG TOP 5 Rising Vendors in Storage for Healthcare

However, not all storage solutions are the same. Many of these solutions have evolved from differing original design implementations. The DCIG Rising Vendors in Storage for Healthcare solutions can differ from one another in the following ways:

Integration with or deployment into cloud. Cloud integration features for hybrid storage can vary. Most storage these days has been recast as software-defined even if mainly purchased via fixed appliances, but some of our top solutions are also certified to run centrally or in part in cloud environments with varying levels of functionality.

Unified and/or multi-protocol storage. Some storage designs provide file and object protocol access, some are tiering files over object storage, and others are unifying file and object storage in the array.

Space efficiency features. Our top solutions all offer capacity optimizing features, but not all offer the same efficiencies or approaches. The details of deduplication, compression, incremental snapshots, versioning, and other features as well as implementation design and policy-manageable processes can vary greatly.

Rising Vendors in Storage for Healthcare Solutions

Qumulo is especially suited for the capacity storage of petabytes of media files (images, video) while simultaneously servicing demanding video processing or research workloads with cloud-elastic or bursty IO patterns.

Encryption in-flight or in-cache. Data protection features like encryption and immutability can be critical for healthcare storage. It's important to note how, when and where each solution supports data encryption and key management to make sure that the storage architecture will support complete end-to-end privacy and compliance constraints.

Levels and types of storage media tiering. It's normal for various designs of storage solutions to support different levels and types of tiering. When evaluating a large capacity solution, one may need to dive into the details of the various quantities and types of media available or required, the ability to expand or upgrade tiers over time, and the relative intelligence of any auto-tiering or policy-based tiering provided.

DCIG TOP 5 Rising Vendors in Storage for Healthcare Solution Profiles

Each of the DCIG Rising Vendors in Storage for Healthcare Solution Profiles highlights three notable solution features that make the product attractive for this market.

Qumulo

Qumulo is a cloud-native, "single-tier" global file system designed for efficient, extreme-scale unstructured data capacities and HPC-class file performance. Qumulo is especially suited for the capacity storage of petabytes of media files (images, video) while simultaneously servicing demanding video processing or research workloads with cloud-elastic or bursty IO patterns.

Qumulo applies native intelligent file-level analysis, predictive pre-fetch and caching best utilize assigned NVMe resources to maximize performance, while delivering linear scalability through auto-tiering across diverse hybrid cloud and active-archive storage capacities. Qumulo clusters can be deployed across all major public cloud providers, HPE, Dell, Pure, Fujitsu and other underlying storage arrays or are available through Qumulo pre-integrated hardware.

Three of the key features that earned Qumulo a spot among DCIG TOP 5 Rising Vendors in Storage for Healthcare solutions include:

Simple global hybrid cloud. Qumulo delivers a single namespace with multiprotocol access (SMB, NFS, FTP, REST) across all assigned storage. No application migration or transformation is required to access files in any environment, including in the cloud, easing both storage consolidation and data collaboration tasks.

Consistent workflow integration. Rich API's from Qumulo enable the tight workflow integration of storage with complex data pipelines and critical application workflows found in key use cases like genomics, molecular simulations, image processing and PACS.

Real-time analytics. Qumulo provides instant data and usage visibility across billions of files. The resulting intelligence enables the "full utilization" of storage resources and powers the Qumulo platform's automatic predictive file-level pre-fetch and caching for high performance.

Rising Vendors in Storage for Healthcare Solutions

One of the main architectural advantages to Nasuni is the ability to collect regional NAS consolidations into a global storage domain based on back-end cloud storage.

The Pavilion HyperParallel Data Platform provides massive IO parallelism through its unique architectural approach that basically hyper-converges a high-speed network switch with a unified flash storage array.

Nasuni

Nasuni is enterprise file storage delivered through SaaS. Nasuni cleverly combines the best of distributed file and cloud object storage together to bring their customers both the unlimited on-demand capacity scaling and utility economics of cloud and integrated storage simplicity from having a single management console covering global file system provisioning and management. The single SaaS Nasuni management console covers all storage, edge appliances, volumes, snapshots, protocols in use, and shares in the system, across all cloud and on-premise deployments.

Three of the key features that earned Nasuni a spot among DCIG TOP 5 Rising Vendors in Storage for Healthcare solutions include:

Distributed access with performant local latency. One of the main architectural advantages to Nasuni is the ability to collect regional NAS consolidations into a global storage domain based on back-end cloud storage. Nasuni then fosters sharing and collaboration of that cohesive data set through global distributed access using both multi-file synchronization and global file locking. For performance, the Nasuni Edge Appliances forward cache and deliver fast local data access.

Natively integrates storage, backup, DR and file sync. At the edge, Nasuni UniFS de-dupes, encrypts and cloud-stores incremental changes (configurable). This obviates the need for separate backup (or DR) processes or systems. On the backend, Nasuni stores all files, metadata, and nodes as cloud storage objects, with the benefits of inherent data protection and direct recovery of past versions. Data sets can be directly restored or cloned in the cloud for DR or on-demand analytics. UniFS ensures compliance with end-end encryption and policy driven lifecycle management.

Cloud-based scaling, capacity and economics. Nasuni can utilize cloud-side storage across all three major cloud platforms (AWS, MS Azure, and GCP) letting users decide source provision, easing migrations, and avoiding lock-in. Nasuni Edge Appliances can be run as VMs on virtual or hyperconverged infrastructure, on pre-configured hardware, or in the cloud.

Pavilion HyperParallel Data Platform

The Pavilion HyperParallel Data Platform provides massive IO parallelism through its unique architectural approach that basically hyper-converges a high-speed network switch with a unified flash storage array. The base 4RU system can hold over 2PB and up to 20 controllers, while serving high-performance block, file and object I/O as every controller can talk directly to any disk. Pavilion can cluster an unlimited number of these base arrays with linear (or even increasing) performance.

While the Pavilion HyperOS and its inherent Pavilion HyperParallel File System provide a combined NFS/S3 global namespace across an unlimited number of array units and assure both performance and availability at scale, Pavilion's flash arrays can also host other high-performance file systems (e.g. IBM Spectrum Scale) if desired.

Three of the key features that earned Pavilion HyperParallel Data Platform a spot among DCIG TOP 5 Rising Vendors in Storage for Healthcare solutions include:

Unlimited linear performance (and capacity) scaling. A remarkable storage density enables Pavilion HyperParallel platform users to start fairly small but grow both scaling up (to 2PB per unit) and scaling out (unlimited) cost-efficiently as necessary. Because of the integrated network switch design, performance across clusters can be linear with growth.

Built for serious workflows. The Pavilion HyperOS and HyperParallel File System provide a native multi-protocol access to enable a single, unified storage repository to power both very high-speed data ingest and parallel client consumption of fast data in disparate forms. Protocols include NVMe-of(RDMA, IB or TCP), iSCSI, NFS and S3 with native support on any combination of controllers across any number of arrays.

Integrated enterprise data services. Tiering, replication, snapshots, clones, security, encryption and application plugins (popular S3-based apps, ML, Big Data, Spark, Tensor-Flow, Kafka, Splunk, Teradata and more) are all included in the core platform. Hardware upgrades and cluster expansion can be made non-disruptively, and the platform design eliminates single points of failure common to other arrays.

VAST Data Universal Storage

VAST's all-flash scale-out storage is designed to take advantage of several newer technologies—NVMe-of, a “hyperscale” flash architecture and storage class memory. The VAST Universal Storage solution attempts to converge primary and secondary storage into one tier for all data, enabling deep capacities of data to reside essentially inside a single Exabyte-scalable globally name-spaced system.

VAST front-end servers are essentially fungible, stateless points of access that can be scaled up to 10,000 nodes (VAST appliances or containers). Each server can access every storage node over an NVME fabric. Every one of up to 1000 storage nodes can support over 2PB/2RU (VAST's own storage appliances, nodes can also be run on other certified hardware).

Three of the key features that earned VAST a spot among DCIG TOP 5 Rising Vendors in Storage for Healthcare solutions include:

Performance in and out. VAST's leverage of large amounts of Storage Class Memory buffers writes at memory speed. This memory staging then enables background deduplication and destaging out to flash-optimized wide stripes across cost-efficiently driven flash storage distributed across the cluster. Keeping all data in flash (and cached in memory) ensures maximum parallel flash read performance for all data—great for random access, metadata intensive operations, and AI/ML applications.

Storage simplicity. As a single Exabyte-scalable tier with multiprotocol access to all data (NFS, NFS over RDMA, S3, SMB and containers), management is simplified and usage is completely “democratized”. Storage is protected with a proprietary Global Erasure Code that declusters error correction to provide near instant recovery of data due to failed devices.

Capacity optimization. Incoming data is cleverly deduplicated using a proprietary algorithm that leverages a measure of similarity to existing blocks to create a “local” and more performant encoding/decoding capability within the larger scale-out cluster. Decompression is fast (out of flash). Overall, this capability makes the investment in their cost-efficient flash architecture go even farther.

VAST leverages of large amounts of Storage Class Memory to buffer writes at memory speed. Flash-optimized wide-striping across cost-efficient flash storage ensures maximum parallel flash read performance for all data.

Weka can integrate and internally auto-tier with a third-party object store to create a “single-tier”, single-namespace solution with combined high-performance and scalable capacity storage.

WekaIO

Weka is a POSIX-compliant high-performance clustered parallel file system designed to run natively over NVMe drives. It can integrate and internally auto-tier with a third-party object store to create a “single-tier”, single-namespace solution with combined high-performance and scalable capacity storage.

WekaFS is software that runs on commodity hardware and is deployable across a range of environments from virtual servers, containers and on-premise x86 Linux servers to cloud instances as a dedicated appliance, converged, or cloud-native solution.

Three of the key features that earned WekaIO a spot among DCIG TOP 5 Rising Vendors in Storage for Life Sciences solutions include:

Parallel flash performance. WekaFS optimizes the use of local NVMe drives on each storage node and converts local (Linux) node caching into “adaptive caching”, a type of cluster-coherent shared caching. Weka can be leveraged simply to accelerate the performance of slower object storage implementations.

Resiliency at scale. WekaFS provides enterprise NAS management features like storage native snaps, clones and a “snap to object” which can be subsequently rehydrated on non-mirror systems for a wide variety of agility-enhancing use cases. Weka is resilient to failures with a unique node striping approach and advanced erasure coding that enables rapid drive rebuilds.

Cloud-like economics. In addition to essentially pooling industry-standard server and storage for performance, WekaFS also integrates tiering to offload colder data to cloud and/or on-premise object storage through S3 and Swift protocols. File-based apps can run on Weka in the cloud without modification, and data can be served out of archive object stores with high performance.

Inclusion and Evaluation Criteria for Storage for Healthcare Solutions

In this report, DCIG specifically focused on Storage for Healthcare solutions possessing the following characteristics. DCIG identified sixteen different solutions meeting these inclusion criteria:

- Commercially available on December 1, 2021.
- Sufficient, publicly available information available for DCIG to make an informed decision.
- Clear vendor intention to support life sciences and healthcare use cases as evidenced by publicly available solution-focused content.
- Demonstrated business in these industries.
- Capable of supporting global initiatives, meeting demanding application file performance requirements, and scalable to PB's.

DCIG evaluated each of these solutions in the following areas:

- 1. Deployment capabilities.** Evaluate the capabilities concerning on-premise deployment options, cloud provider deployment options, cloud provider targets supported, storage protocols supported, virtual environments supported, and certifications with equipment, operating systems, and applications.
- 2. Data protection capabilities.** Evaluate solution capabilities supporting availability, encryption, replication, and snapshot features.
- 3. Product and performance management features.** Evaluate options to manage the underlying hardware and optimize it for performance. Examples include dashboard views, predictive analytics, storage optimization, quality of service features, auto-tiering capabilities, and directory service integration.
- 4. Documentation support.** Evaluate the breadth and depth of documentation the provider makes available to customers. Examples include whitepapers, knowledge bases, online manuals, as well as community forums.
- 5. Technical support.** Evaluate the availability and technical support options of the solution provider. Examples include support availability, response time commitments, options to open cases, escalation support, and proactive problem resolution.
- 6. Licensing and pricing.** Evaluate the relative ease of doing business through flexibility and simplicity in contract lengths, pricing elements, and bundled pricing options.

DCIG Disclosures

Vendors of some of the solutions covered in this DCIG report are or have been DCIG clients. There are some important facts to keep in mind when considering the information contained in this report and its merit.

- No vendor paid DCIG any fee to research this topic or arrive at predetermined conclusions.
- DCIG did not guarantee any vendor that its solution would be included in this DCIG TOP 5 report.
- DCIG did not imply or guarantee that a specific solution would receive a DCIG TOP 5 designation.
- All research is based upon publicly available information, information provided by the vendor, and/or the expertise of those evaluating the information.
- DCIG conducted no hands-on testing to validate how or if the features worked as described.
- No negative inferences should be drawn against any vendor or solution not covered in this TOP 5 report.
- It is a misuse of this DCIG TOP 5 report to compare solutions included in this report against solutions not included in it.

DCIG wants to emphasize that no vendor was privy to how DCIG weighted individual features. In every case the vendor only found out the rankings of its solution after the analysis was complete. To arrive at the DCIG TOP 5 solutions included in this report, DCIG went through a seven-step process to come to the most objective conclusions possible.

1. DCIG established which features would be evaluated.
2. The features were grouped into six general categories.
3. A DCIG analyst internally examined the feature data for each solution and completed a survey for it based upon the analyst's own knowledge of the solution and publicly available information.
4. DCIG identified solutions that met DCIG's definition for a Storage for Healthcare solution.
5. DCIG weighted each feature to establish a scoring rubric.
6. DCIG evaluated each solution based on information gathered in its survey.
7. Solutions were ranked using standard scoring techniques. ■

About DCIG

The Data Center Intelligence Group (DCIG) empowers the IT industry with actionable analysis. DCIG analysts provide informed third-party analysis of various cloud, data protection, and data storage technologies. DCIG independently develops licensed content in the form of TOP 5 Reports and Solution Profiles. Please visit www.dcig.com.



DCIG, LLC // 7511 MADISON STREET // OMAHA NE 68127 // 844.324.4552

dcig.com

© 2022 DCIG, LLC. All rights reserved. Other trademarks appearing in this document are the property of their respective owners. This DCIG report is a product of DCIG, LLC. All other brands or products are trademarks or registered trademarks of their respective holders and should be treated as such. Product information was compiled from both publicly-available and vendor-provided resources. While DCIG has attempted to verify that product information is correct and complete, feature support can change and is subject to interpretation. All features represent the opinion of DCIG. DCIG cannot be held responsible for any errors that may appear.