GIGAOM

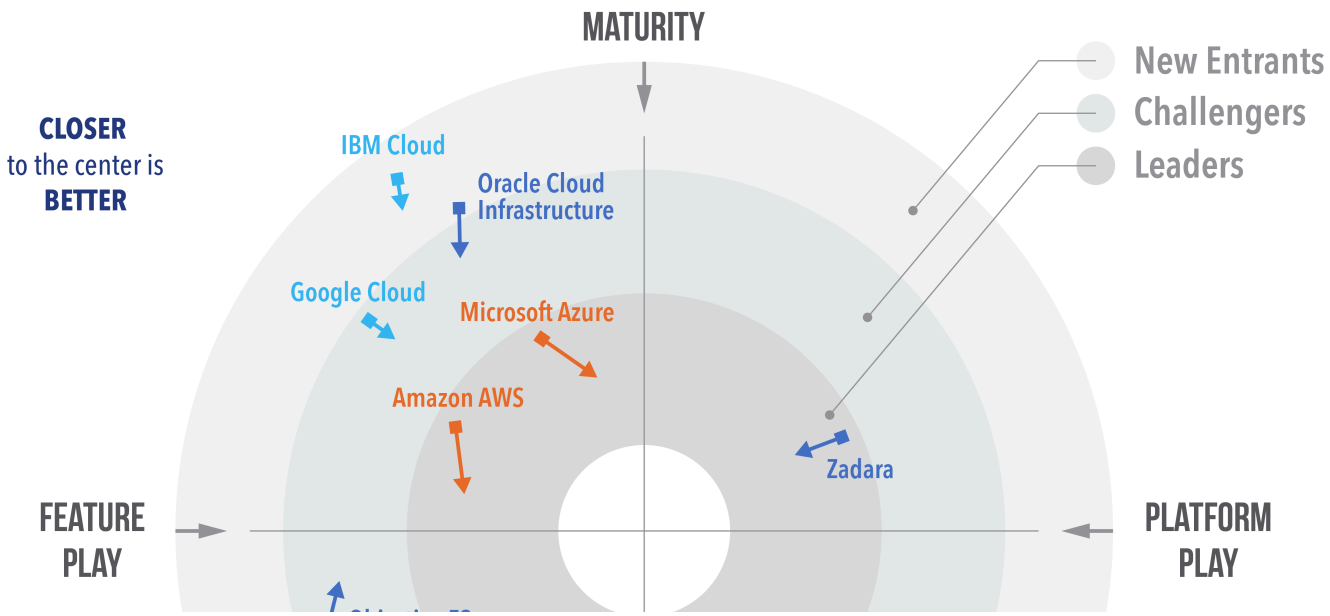
MARKET RADAR

GigaOm Radar for File-Based Cloud Storage v1.0

Key Criteria for Evaluating File-Based Cloud Storage

ENRICO SIGNORETTI

TOPICS: CLOUD INFRASTRUCTURE DATA STORAGE



GIGAOM

GigaOm Radar for File-Based Cloud Storage

Key Criteria for Evaluating File-Based Cloud Storage

TABLE OF CONTENTS

1. Summary

FileFile storage is a critical component of every hybrid cloud strategy, and enterprises prefer this type of storage to block and object in several cases.

Cloud providers started to offer this type of service very late when compared to block and object, and this created the opportunity for several storage vendors to provide products and services to fill the gap. Especially now, with the increasing need for data mobility and more and more workloads that are moving across on-premises and cloud infrastructures, file storage is better, easier to use and more accessible than other forms of storage.

Lift-and-Shift migrations to the cloud are one of the most common scenarios where file storage is a key factor to keep the environment as identical as possible to the original one, but, again, simplicity and performance are important as well.

In particular regard to performance, file systems still provide the best combination of performance, usability, and scalability for many workloads. It is still the primary interface for the majority of big data, artificial intelligence/machine learning (AI/ML), and High-Performing Computing (HPC) applications, and usually offers data services such as snapshots to improve data management operations.

Collaboration is another area where file storage excels. Now more than ever, with hybrid cloud infrastructures and distributed organizations across the world, having the possibility to access data from everywhere, seamlessly, simplifies teamwork while keeping data under control.

In recent years file systems have also become more cloud-friendly, showing better integrations with object storage, thus offering better scalability, a better combination of speed and $/GB, and advanced features for data migrations and disaster recovery.

Both traditional storage vendors and cloud providers now offer different types of file services or solutions that can run both on-premises and in the cloud, trying to cover the largest number of use cases. Approaches are different, though, and it can be very difficult to find a solution that solves today's needs while being ready to evolve and face future challenges. Even more so, most cloud providers offer the best integration across the entire stack, but also more risk of lock-in and services are not always the best in class. On the other hand, storage vendors offer solutions that provide more flexibility, performance, and scalability, but can be less efficient or lack the level of integration offered by an end-to-end solution.

2. About the GigaOm Radar

HOW TO READ THIS REPORT

The Market Roundup/Landscape/radar profile is part of a series of documents aimed at giving the reader the tools to better understand a technology, evaluate it, and explore the market to find the best solutions for his organization.

In this context, and in order to get a complete view of the state of the solutions available in the market, the reader should consider the following documents:

Key Criteria to Evaluate File-Based Cloud Storage, is an introduction to the technology, defines the necessary evaluation metrics, the key criteria to evaluate new solutions, and the impact of the latter on the former. It is dedicated to those end users that are approaching a new technology for the first time or want an update on the latest evolution.

GigaOm Radar for File-Based Cloud Storage, offers a brief 360° view of the market including market and technical positioning of most notable vendors, a short introduction of their solutions and differentiations, including a high level graphic comparison of the vendors.

3. Market Categories and Deployment Types

For a better understanding of the market and vendor positioning (table 1), we categorized solutions for Kubernetes data storage by the target market segment (small-medium or large enterprise) and their architecture (enterprise storage systems with a CSI plug-in, software-defined solutions with optimizations for containers, or cloud-native solution):

- **Small-medium enterprise:** In this category, we find solutions that meet the needs of very small businesses that can grow up to address those of medium-sized infrastructures. They can also be solutions adopted by large enterprises for departmental use cases without a very rich feature set, with limited data mobility and management capabilities, but easy to use and deploy.
- **Large enterprise:** Usually adopted for larger and business-critical projects. Solutions in this category have a strong focus on flexibility, performance, data services, and features to improve security and data protection. Scalability is another big differentiator, and the possibility of having the same service in different environments is another important characteristic.
- **Specialized:** The Solution is designed for specific workloads and use cases, such as big data analytics and HPC, for example.
- **Cloud-only solutions:** Available only in the cloud. Often designed, deployed, and managed by the service provider, they are available only from that specific provider. The big advantage of this type of solution is the integration with other services offered by the cloud service provider (functions, for example) and its simplicity.
- **Hybrid and multi-cloud solutions:** These solutions are meant to be installed both on-premises and in the cloud, allowing them to build hybrid or multi-cloud storage infrastructures. The integration with the single cloud provider could be limited compared to the other option and more complex to deploy and manage. On the other hand, they are more flexible, and the user usually has more control over the entire stack about resource allocation and tuning. These solutions can be deployed in the form of virtual appliances, like a traditional Network-Attached Storage (NAS) filer but in the cloud, or a software component that can be installed on a Linux VM (i.e., a file system).

Table 1: Vendor Positioning

	MARKET SEGMENT			DEPLOYMENT MODEL	
	Small/Medium Enterprise	Large Enterprise	Specialized	Cloud-only	Hybrid
Amazon AWS	+++	+++	+++	+++	+
Google	+++	++	+	+++	-
IBM Cloud	+	++	+	+++	+
Lucid Link	++	++	++	+	+++
Microsoft Azure	+++	+++	++	+++	++
Nasuni	++	+++	+	+	+++
NetApp Volumes ONTAP	+++	+++	++	+++	+++
ObjectiveFS	++	++	++	++	+++
Oracle Cloud Infrastructure	+	++	+	+++	+
Panzura	++	+++	++	+	+++
Qumulo	+	+++	++	++	+++
Weka	+	++	+++	++	+++
Zadara Storage	+++	+++	+	+++	+++

+++: strong focus and perfect fit of the solution
++: The solution is good in this area, but there is still room for improvement
+: The solution has limitations and a narrow set of use cases
-: Not applicable or absent.

Source: Gigaom 2020

4. Key Criteria Comparison

Following the general indications introduced with the Key Criteria for Evaluating File-Based Cloud Storage, table 2 quickly summarizes how each vendor included in this research performs in the areas that we consider differentiating and critical for modern data protection. The objective is to give the reader a snapshot of the technical capabilities of different solutions and define the perimeter of the market landscape.

Table 2: Key Criteria & Evaluation Metrics Comparison

	KEY CRITERIA					EVALUATION METRICS					
	Multi-zone and region availability	hybrid and multi-cloud	Integration with object storage	Serverless Integration	Analytics	Architecture	Scalability	Flexibility	Efficiency	Performance	Manageability and ease of use
Amazon AWS FSx for Lustre	++	-	+++	+	+	++	+++	++	+	+++	++
Amazon AWS FSx for Windows File Servers	++	-	+++	-	++	++	++	++	++	++	+++
Amazon AWS Elastic File Service (EFS)	+++	-	-	+	+	++	++	++	++	++	+++
Google Cloud FileStore	+++	-	-	-	+	++	++	++	++	++	+++
IBM Cloud	++	-	-	-	+	++	+	++	++	+	++
LucidLink	+++	+++	+++	-	+	+++	++	+++	+++	++	+++
Microsoft Azure Avere	++	++	+++	-	+	+++	+++	++	+++	++	++
Microsoft Azure File Storage	+++	-	-	-	+	++	++	++	++	++	+++
Microsoft Azure NetApp Files	+++	+++	-	-	+	+++	++	+++	+++	+++	+++
Nasuni	+++	+++	+++	++	++	+++	++	+++	+++	++	+++
NetApp Volumes ONTAP	+++	+++	++	-	++	++	++	++	+++	++	+++
NetApp Cloud Volumes Services (including Azure NetApp Files)	+++	+++	-	-	+	+++	++	+++	+++	+++	+++
ObjectiveFS	+++	+++	+++	-	-	++	++	+++	+	++	+++
Oracle Cloud Infrastructure File Storage	+++	-	-	-	+	++	+++	++	++	++	++
Panzura	+++	+++	+++	++	++	+++	++	+++	+++	++	+++
Qumulo	+++	+++	++	-	+++	+++	+++	+++	+++	++	+++
Weka	+++	+++	++	-	+	+++	+++	++	+++	+++	++
Zadara	+++	+++	+++	++	++	+++	++	+++	+++	++	+++

+++: strong focus and perfect fit of the solution
++: The solution is good in this area, but there is still room for improvement
+: The solution has limitations and a narrow set of use cases
-: Not applicable or absent.

Source: Gigaom 2020

5. GigaOm Radar

All the Key Criteria and the critical feature impact analysis are consolidated in the following graphic representation: The GigaOm Radar (figure 1). This vector-based graphic gives an overall perspective on all the vendors included in this research in terms of technical capabilities and features (table 2), execution on the vision, regardless of their market share or segment (table 1).

The GigaOm Radar focuses on technical capabilities, roadmap, execution, and the ability to innovate. In this graphic, for each vendor representation of the market, there are three fundamental informational points:

- **The current position** in the chart gives information on the state of the solutions provided by the vendor
- **The direction** indicates where the vendor is going with the product strategy and development (potentially where the vendor will be in 12/18 months from now)
- **The module** signifies how fast it is executing on vision and strategy

The metrics on which the GigaOm Radar focuses consists of the following:

- **Maturity:** Expresses the maturity and solidity of the product, user acceptance of the solution, and overall ecosystem sustainability.
- **Innovation:** The differentiation of the solution, innovative technical aspects, and overall vendor approach to the market.
- **Feature Play:** Gives indications about the focus of the solution in terms of single differentiating features and technical aspects of the product. It usually includes new vendors that are defining new product categories, niche players, and point solutions.
- **Platform Play:** Contrary to “Feature Play,” solutions positioned in this area provide a horizontal platform that can face a broader range of challenges, with a comprehensive feature set and an extensive ecosystem.

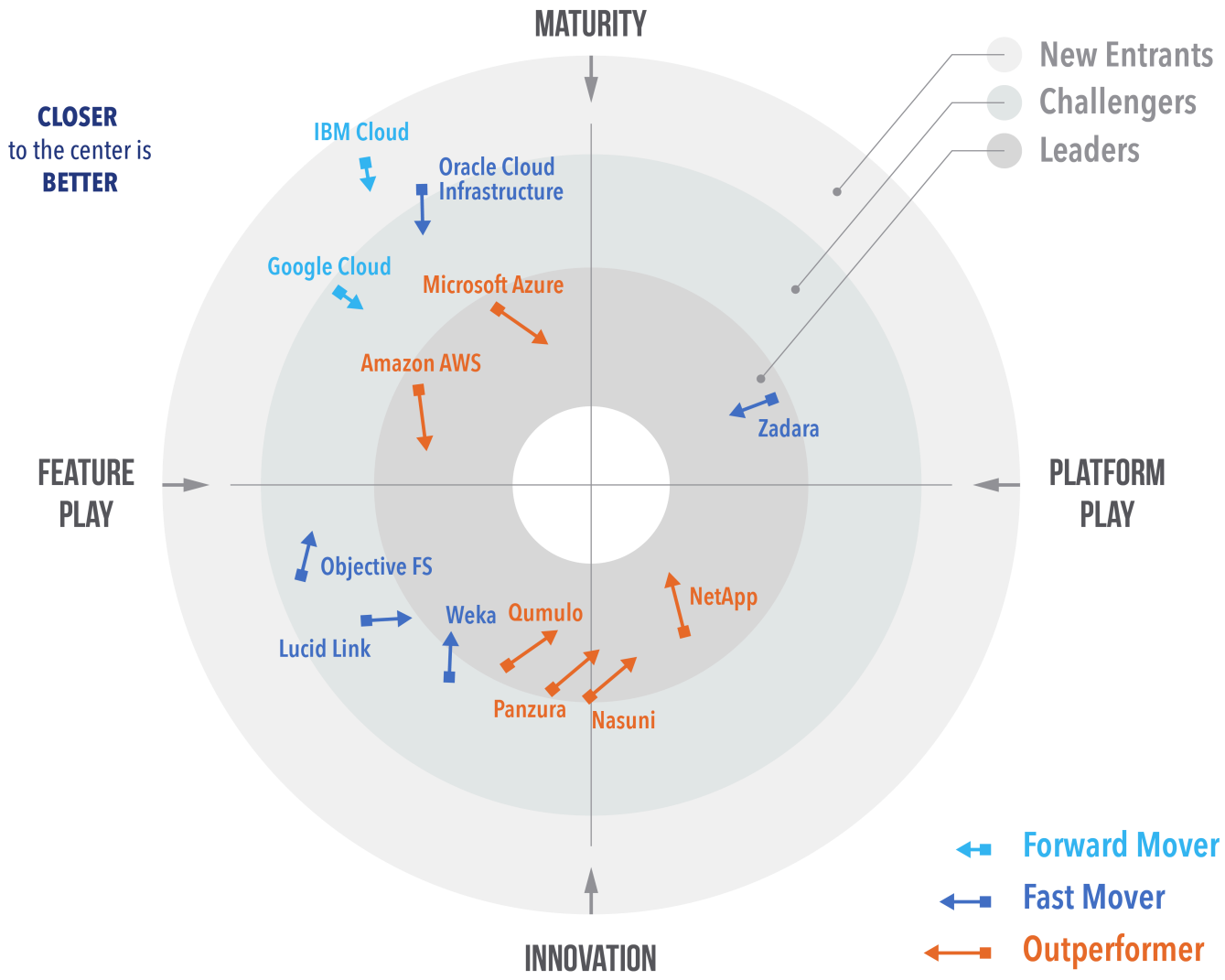
The GigaOm Radar is organized in three concentric areas and, the closer to the center, the better:

- **Leaders:** Vendors that are competing for market leadership for the metrics described above, even if they are competing in different market segments.
- **Challengers:** Vendors, with the potential to become a leader soon, as well as niche or traditional players with an established market, and other companies that have interesting solutions who still have to prove their real differentiation.

- **New entrants:** Usually companies with a limited feature set and too young to be included in inner circles, but with potential to move there soon.
- The center-most circle of the GigaOm Radar is usually empty. It is reserved for **extremely mature and consolidated markets** with very few vendors remaining and solutions that are mature, comparable, and without much space for further innovation.

The market share metric is excluded from the GigaOm Radar to give a more clear view of the potential of the vendors. It is a forward-looking market landscape representation that leaves more room for innovation and differentiation instead of weighting established positions.

Figure 1: GigaOm Radar for File-Based Cloud Storage



Source: GigaOm 2020

6. Vendor Roundup/Overview

Amazon AWS

Amazon AWS offers several file-based services for its cloud platform target at different workloads and use cases, including general purpose file sharing options (EFS, FSx for Windows) and Big Data/HPC/AI (FSx for Lustre). The cloud provider can also take advantage of its solution ecosystem to simplify data migrations and data movements from on-premises data centers to the cloud.

Strengths: Long list of services and options that can take advantage of a comprehensive ecosystem which enable the user to configure tailored solutions that can respond to demanding application and business requirements. Continuous effort to innovate.

Weaknesses: Limited hybrid cloud options, with most of the solutions designed only to move data to the cloud. Some services could be improved especially when it comes to integration with Lambda functions.

Google Cloud

Google Cloud Filestore is the only file storage option directly available from Google, with additional services provided by third parties on its marketplace. Google acquired Elastifile last year in order to enhance the Filestore offering.

Strengths: Filestore is available in two performance tiers, easy to use and configure, with support for Network File System (NFS). It offers performance up to 1.2 GB/s of sequential read, and 60,000 IOPS.

Weaknesses: The service is limited both in terms of capacity (maximum volume size is 64 TB) and write throughput is limited to 350MB/s. File options for HPC, big data and AI/ML workloads are not yet available, though we expect that to change once Elastifile is integrated.

IBM Cloud

IBM cloud file service is limited, but IBM can count on several partners to provide file services to its users. An option available for IBM customers is to use bare metal cloud servers to build a storage infrastructure tailored for their needs.

Strengths: Even though the burden on the user can be high and the flexibility very limited, there is the possibility to build a storage infrastructure starting from bare metal servers and software that the customer can purchase separately.

Weaknesses: The service is limited both in terms of capacity (12TB max per volume) and IOPS (48K

max per volume). It offers an NFS protocol only.

LucidLink

Simple, yet effective and secure, LucidLink is a SaaS-based storage solution for data sharing that integrates an object storage backend with clients available on major operating systems. LucidLink is compatible with major cloud providers, object storage vendors and operating systems, enabling the user to build a NAS-like experience for distributed organizations across the globe.

Strengths: Simple and scalable, elastic, compatible with most object storage services and systems, with good performance at the client level. Excellent solution for small organizations and workgroups that want to replace traditional NAS systems with cloud solutions. Can be deployed both on premises and in the cloud.

Weaknesses: The feature set is still limited compared to other solutions in this space, but the company is actively working to close the gap.

Microsoft Azure

Microsoft Azure has built a complete storage service line up now able to respond to most disparate requests from low-end Server Message Block protocol (SMB) and NFS file services up to HPC workloads. Many of these services are also integrated with data analytics tools from the Azure catalog. Hybrid cloud options for high demand workloads are also available thanks to Avere virtual and physical appliances.

Strengths: Microsoft Azure provides several options for file services, and many of them have good integrations with other services available from this cloud provider. Microsoft's friendly partnership policy is helping to expand the storage ecosystem as well.

Weaknesses: Even though Avere is a good solution for hybrid environments and Azure NetApp Files shows very good performance, there isn't a native scale-out option for HPC and big data workloads with multi-petabyte capacity per volume in Azure. In this regard, Microsoft is currently working with its partners to close the gap with other providers.

Nasuni

Nasuni's Cloud File Services Platfrom solution enables file sharing, data protection, and synchronization across multiple clouds and on-premises infrastructures using standard network protocols such as NFS and SMB. Thanks to the object-storage backend, the solution is very scalable while offering advanced integrated data protection features. Edge and cloud appliances are usually deployed as virtual machines. Nasuni has partnerships with major cloud service providers and object storage vendors.

Strengths: It is a solid and mature architecture, easy to use, elastic and scalable with a large number of features aimed at improving file serving and data protection for distributed organizations across the globe. Now with the possibility to take advantage of third-party analytics tools thanks to a dedicated connector.

Weaknesses: The solution has been designed for large scale and multi-cloud environments and not for small enterprises. Not originally designed for cloud-only deployments, a use case that is becoming more common though.

NetApp

NetApp has one of the most comprehensive product and service file storage portfolio for hybrid and multi cloud, providing both a seamless user experience for those customers that want to maintain a traditional approach (Cloud Volumes ONTAP) and those that need a cloud-native storage solution (Cloud Volume Services and Azure NetApp Files available from major public clouds). Additionally, storage is part of a broader strategy that includes several solutions for application and data management. At the same time, the recently acquired Talon Storage will help NetApp to address emerging hybrid cloud use cases such as remote collaboration and Cloud-NAS consolidation.

Strengths: Simple to deploy and easy to use, with good performance and scalability, compatible with the latest NFS and SMB protocols, and available on multiple clouds with tools that allow you to replicate and migrate data quickly across different clouds and object stores. Integrated with NetApp monitoring tools such as Cloud Insights and Compliance.

Weaknesses: Automated tiering for Cloud Volumes Service is not available yet, remote data replication capabilities need improvements to satisfy requirements of high demanding use cases.

ObjectiveFS

A straightforward, scalable, and effective solution for organizations of all sizes that need a shared file system for distributed environments. ObjectiveFS is a log-structured POSIX-compliant FUSE-based file system that leverages an object store backend for data persistence. Compatible with Linux and Mac OS.

Strengths: Very simple and straightforward solution that can be used both on-premises and in the cloud at a reasonable cost and with good performance for many workloads.

Weaknesses: It needs a client installed in every single computer or VM instance using it. No analytics and remote replication tools for data migrations to a different cloud.

Oracle Cloud Infrastructure

Oracle provides a very good, flexible and scalable, NFS service for its users and a storage gateway to help its customers to manage data migrations to the cloud or hybrid cloud scenarios.

Strengths: Easy to use and manage, flexible, very scalable, and highly available, this storage service can be deployed for several use cases.

Weaknesses: Even though the NFS service is good, with the possibility of starting with very small volumes and growing from there, the file storage ecosystem is limited to this option, lacking SMB protocol support and solutions specifically designed to manage high-performance workloads.

Panzura

Panzura Freedom is a cloud based NAS solution for unstructured file data with an object storage backend, either public or private. Freedom can be deployed either in-cloud or on premise, providing software defined file services for globally distributed environments. The Solution is now even more interesting thanks to Vizion.ai, a service aimed at unstructured data analytics and management.

Strengths: Simple to use, scalable, available for on-premises and in-cloud installations, provides local performance with a cloud backend. Compatible with both cloud object storage (such as Amazon AWS S3 and others) as well as on-premises object storage systems.

Weaknesses: It could be expensive for cloud-only workloads where data is shared only locally. Even if Panzura Freedom has good performance in most use cases, it is designed for unstructured data not structured data like databases and HPC workloads.

Qumulo

Qumulo has recently launched the latest iteration of its scale-out file storage solution, now focused on hybrid cloud infrastructures. Based on an innovative, efficient file system with embedded technology for deep analytics, the product is available for on-premises and cloud deployments enabling users to migrate data across different environments for high-performance seamlessly and for large-capacity workloads that range from media rendering, big data analytics, commercial HPC, AI/ML, archiving, and more.

Strengths: High performance, scale-out, easy-to-use, and flexible storage solution with an enterprise feature set that can run on multiple clouds, with good object storage integration for data tiering.

Weaknesses: Despite its good feature set, the solution remains expensive for small installations due to its scale-out nature and the necessary resources for entry-level configurations.

Weka

A parallel scale-out file system solution capable of providing top-notch performance when compared to other file systems, perfect for all types of HPC, AI/ML, and big data workloads both in the cloud and on-premises. Designed to take advantage of the latest flash memory technology and with good object storage integration to reduce the cost for capacity and protection.

Strengths: Good user interface that allows hiding the complexity of the scale-out architecture. Available on the AWS marketplace. NFS and SMB gateways integrated into the product.

Weaknesses: Even though the solution can also be deployed for less demanding workloads, it provides limited analytics and reporting functionalities, and an absence of data replication features.

Zadara Storage

Zadara offers its Storage-as-a-Service platform on all major cloud providers as well as on-premises. The solution is very flexible, secure, highly configurable and provides block, file and object storage service in a single system. Advanced data services and other innovative functionalities enable users to address several use cases.

Strengths: The solution is mature and solid with services that allow users to build hybrid cloud infrastructures and move data across different environments for several types of workloads and applications. Zadara has also developed an interesting serverless solution, with containers that can run directly into its virtual controllers. Easy to use, scalable, and flexible.

Weaknesses: Zadara focuses on enterprise workloads and the solution is not primarily aimed at serving very high (multi-petabyte) capacity and performance demanding applications even though more than 15% of Zadara customers use the solution also for big data analytics.

7. Conclusion

The market is moving very fast, with several players offering file services for different use cases. 'One size fits all' is still hard to reach, especially when complexity and cost come into play.

Most cloud service providers still offer a limited set of file services to their user base, imposing constraints and making even lift-and-shift migrations more complicated than expected by users. In this group, Amazon AWS leads in the number of options, granularity, integration with the ecosystem, and innovation, while Microsoft Azure is the only one in the position to compete while also having a better hybrid cloud strategy and approach.

NetApp has the lead among storage vendors at the moment. It is building a good solution ecosystem that enables its users to build file-based cloud-native multi-cloud infrastructures with a seamless experience while keeping an eye to existing ONTAP customers. In the same category, another compelling solution is Zadara, which offers storage as a service that can be consumed both on-premises and in the cloud with the same flexibility, granularity, features, and user experience. Both of them are designed for enterprise workloads.

For high performance and large capacity, scale-out NAS workloads, Weka and Qumulo have the most interesting solutions in the market. The first leads in terms of absolute performance, while the latter has a more robust feature set for enterprise use cases while still providing very high performance compared to other solutions. They are both working to integrate object storage in their architecture to combine performance with capacity at a decent cost while enabling better data protection and use cases.

When it comes to file storage solutions with an object storage back end, we find two leaders and two runners-up. Nasuni and Panzura have tremendous solutions when it comes to sharing files across different clouds and on-premises locations, and they are very familiar with enterprise environments and the necessary feature sets to support their needs. On the other hand, LucidLink and ObjectiveFS offer simpler solutions with very good performance, lower cost, and enough flexibility to satisfy small and medium enterprise or vertical use cases.

With more enterprises moving to the cloud, in the next couple of years we will see an increasing interest in file and cloud storage, with hybrid and multi-cloud solutions that will be more and more focused at data management, analytics, and integration with microservices such as Kubernetes and functions-as-a-service (for example, AWS Lambda).

8. About Enrico Signoretti

Enrico has 25+ years of industry experience in technical product strategy and management roles. He has advised mid-market and large enterprises across numerous industries and software companies ranging from small ISVs to large providers.

Enrico is an internationally renowned visionary author, blogger, and speaker on the topic of data storage. He has tracked the changes in the storage industry as a Gigaom Research Analyst, Independent Analyst and contributor to the Register.

9. About GigaOm

GigaOm provides technical, operational, and business advice for IT's strategic digital enterprise and business initiatives. Enterprise business leaders, CIOs, and technology organizations partner with GigaOm for practical, actionable, strategic, and visionary advice for modernizing and transforming their business. GigaOm's advice empowers enterprises to successfully compete in an increasingly complicated business atmosphere that requires a solid understanding of constantly changing customer demands.

GigaOm works directly with enterprises both inside and outside of the IT organization to apply proven research and methodologies designed to avoid pitfalls and roadblocks while balancing risk and innovation. Research methodologies include but are not limited to adoption and benchmarking surveys, use cases, interviews, ROI/TCO, market landscapes, strategic trends, and technical benchmarks. Our analysts possess 20+ years of experience advising a spectrum of clients from early adopters to mainstream enterprises.

GigaOm's perspective is that of the unbiased enterprise practitioner. Through this perspective, GigaOm connects with engaged and loyal subscribers on a deep and meaningful level.

10. Copyright