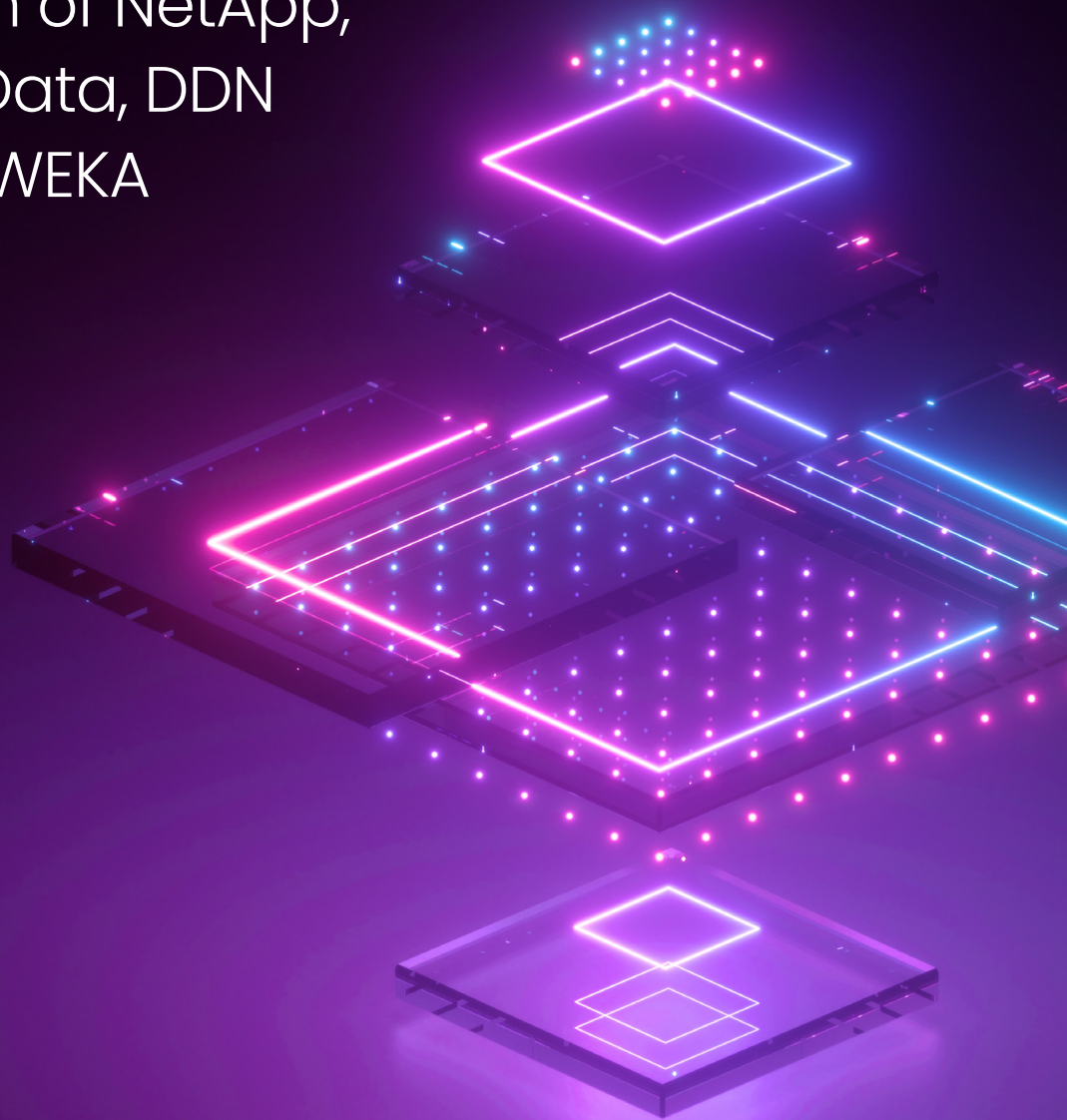




A Buyer's Guide to Modern Storage

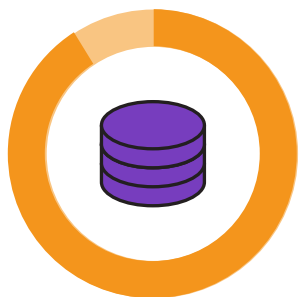
Comparison of NetApp,
Isilon, Vast Data, DDN
Lustre, and WEKA



INDEX

UNSTRUCTURED DATA GROWTH	2
HOW TO SPOT A LEGACY STORAGE VENDOR	2
WHAT ARE MODERN WORKLOADS?	3
THREE TENETS OF A MODERN STORAGE PLATFORM	3
26 KEY CONSIDERATIONS WHEN CHOOSING MODERN STORAGE	4
TECHNICAL CAPABILITIES OVERVIEW	5
WekaFS™ – MODERN STORAGE FOR MODERN WORKLOADS	6
ADDITIONAL RESOURCES	6

UNSTRUCTURED DATA GROWTH



ABOUT 90%
OF ALL DATA IS UNSTRUCTURED



GROWING AT 41%
PER YEAR

It may surprise you, but almost 90% of all data in the enterprise is unstructured and it is growing at 41% per year according to IDC.

Legacy vendors have built architectures to deal with each workload requirement separately. One for SAN, another for NAS, and another for Object storage. To top this, it is not uncommon to find separate architectures built for capacity and performance. These days most enterprises have a hybrid or multi-cloud strategy, which can introduce yet another layer of complexity, and this can lead to complicated management due to incompatible tools and processes and several different architectures. Also, as the data scale grows bigger it is very difficult to copy data and manage multiple copies of data, not to mention the cost of the same.

In the end, customers end up trading off between simplicity, speed, and scale. NAS is simple to use and manage, but not performant enough for modern applications. SAN or parallel file systems provide the performance but are complex and expensive. Object storage provides the scale, but is not performant for it to be really useful. There is a better way.

HOW TO SPOT A LEGACY STORAGE VENDOR

We have identified six aspects for identifying a legacy storage vendor, all of which make a vendor ineffective for modern data storage requirements:

- 1. Selling Systems Built with Proprietary Hardware:** If you are still using a storage solution that is available only in a customized proprietary hardware form factor from your storage vendor, it's a clear indication you're not using a solution based on current design principles, and you're buying legacy storage.
- 2. Hybrid Cloud is Limited:** If your current storage vendor does not support the same specifications for features, CLI, and performance, and for scaling both on-premises and in the cloud, these are clear indications that you're using a legacy vendor.
- 3. Limited Scale and Support for Mixed Workloads:** When considering your datacenter architecture decisions, if you must deploy more storage systems than functionally required by physical separation of resources, you're using a legacy storage vendor. If you must make sizing decisions that lock you in for the life of that system without the ability to expand or shrink as needed, you're also using a legacy storage vendor.
- 4. Limited Aggregate Performance & Single-Client Performance:** If the storage system you're using now has the same single-client performance limitations that existed about a decade ago, you're using a legacy storage vendor. If the system you're using now has an aggregate throughput\IOPS number that has not increased dramatically when compared with numbers a decade ago, you're using a legacy storage vendor.
- 5. Data Backup and DR Are Performed by Others—or They're Afterthoughts:** If your primary storage product, your backup product, and your cloud product are different entities, you are dealing with a legacy storage vendor. If your storage vendor forces you to treat backup and DR differently and store the data twice, you're using a legacy vendor. If your storage vendor forces you to go to a third-party solution in order to get sound backup or archive strategy, you're also using a legacy vendor. If the performance of the storage system drops significantly during rebuild, you're using a legacy vendor. If

you are still rebuilding blocks and not files, you are using a legacy vendor. If you don't have end-to-end data integrity protection for the client (each block has a checksum that is calculated at the client and verified on each step of the way to ensure no bit-rot), you're using a legacy storage vendor.

6. You Need to Make Tradeoffs: If your storage vendor has many products, each with slightly different tradeoffs, and you have to use a different mix of them as solutions to different projects, you're using a legacy vendor (for example, different utilities for taking snapshots, taking snapshots for backup, taking snapshots for bursting, etc.).

WHAT ARE MODERN WORKLOADS?

Technology has changed dramatically in the last 20 years. High speed connectivity has become a commodity, and compute now works at mind-blowing speeds. Just think about how quickly an autonomous vehicle processes and reacts to potential accident situations—it's done in milliseconds. Data is being generated all around us: street cameras, shopping centers, office buildings, your car, your phone, your watch, your family's home automation and security systems, and more.

Just as our personal and consumer data speeds and quantities have changed, so has the enterprise workload changed—from client-server technologies and relational databases to machine and deep learning processed by mini-edge processors or supercomputers in the core.

This leads to new workloads, use cases, and abilities to accomplish outcomes that were not possible before. Here are some examples of modern workloads and use cases:

Life Sciences

Ability to sequence genomes in record time, electron microscopy, AI in image processing

Research

Drug discovery and image processing in pharmaceutical development and research

Financial Services

New trading algorithms and modelling

Automotive

Autonomous driving and training, models that crunch enormous amounts of data

Media & Entertainment

Large file processing, rendering large files, 8k streaming video post production, etc.

Today's workloads demand a new class of storage that delivers the performance, manageability, and scalability required to obtain or sustain an organization's competitive advantage.

THREE TENETS OF A MODERN STORAGE PLATFORM

Simplicity

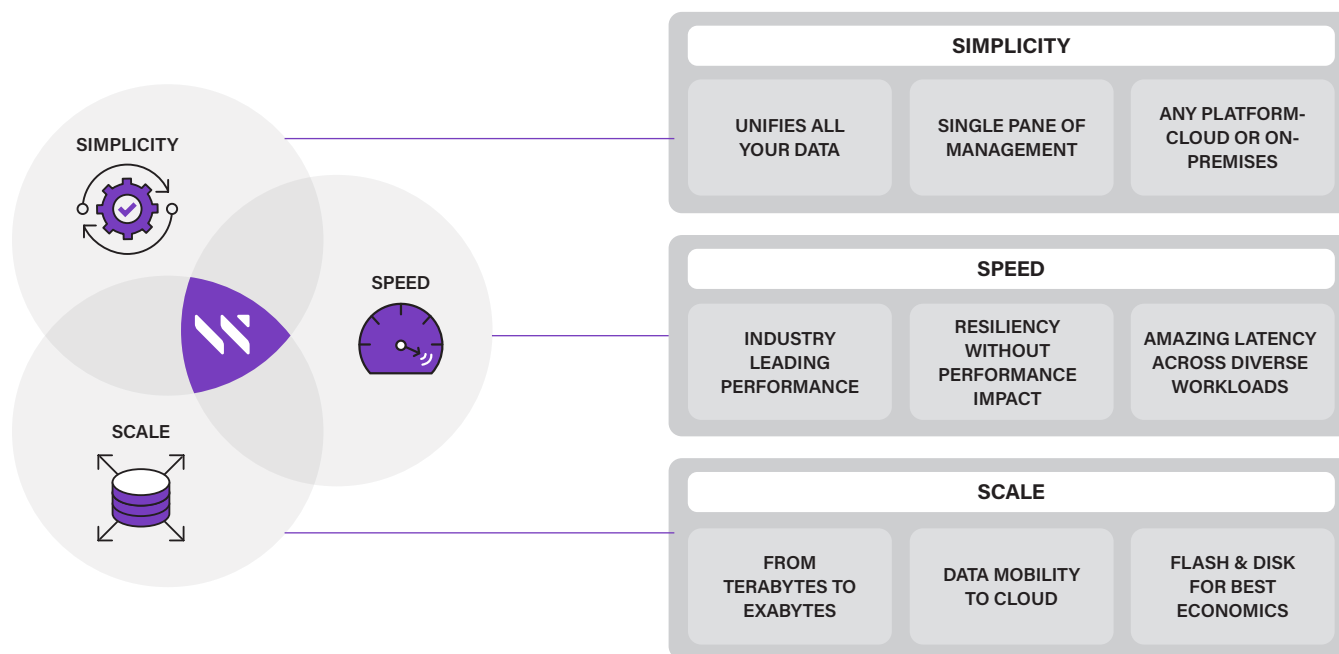
Consolidate all of your workloads on a single platform. No tuning required. For a long time it was believed that in order to get ultra-high performance you needed to sacrifice simplicity. Weka has cracked the code. Our modern file system, WekaFS, is different, combining top-notch performance with simple and reliable management. Weka makes it easy to deploy, configure, and manage storage on-premises or in the cloud.

Speed

Deliver the performance required to power your most demanding applications to dramatically accelerate time to insights. Weka's modern architecture is built from the ground up for flash and optimized for NVMe and cloud to power breakthrough innovation.

Scale

Weka enables high-performance computing at a massive scale without breaking the bank. Modern workloads and today's hybrid cloud environments have dramatically changed the needs of the data center. In addition to offering linear scaling with a scale-out file system, Weka redefines scalability in the cloud era and allows customers to scale in every dimension possible.



26 KEY CONSIDERATIONS WHEN CHOOSING MODERN STORAGE

SIMPLICITY

1. Ease of management
2. Hardware agnostic
3. Single namespace
4. Automated data optimization
5. Built-in data protection
6. Consolidation of data silos
7. Prometheus & Grafana support

SPEED

8. System performance IOPS/Bandwidth
9. Improved workload performance
10. Improving performance for customer's customer
11. Rapid deployment of new applications

SCALE

12. Ability to scale to exabyte
13. Extensibility to cloud
14. Scaling for Disaster Recovery
15. Scaling for capacity
16. Scaling for performance/compute
17. Scaling across mixed workloads
18. Scaling across availability zones
19. Deploying anywhere: on-premises, cloud, or hybrid

OTHERS

20. Full cloud-native capabilities
21. Container Support
22. Support multiple data types (file and object)
23. Multi-protocol support
24. Ecosystem of server vendors
25. Ecosystem of storage vendors
26. The ability of your storage environment to utilize new technologies as they are released. For example, new CPUs, NVMEs, servers, accelerators, etc.'

TECHNICAL CAPABILITIES OVERVIEW

This section provides a detailed comparison of NetApp, Isilon, Vast Data, DDN Lustre, and WekaIO.

COMPARISON OF SIMPLICITY, SPEED, AND SCALE

Criteria/Capabilities		Weka	NetApp	Isilon	Vast Data	DDN Lustre
SIMPLICITY	Easy to deploy and manage	Yes	Yes	Yes	Yes	No
SPEED	Mixed workloads	Yes	No - Performance fall-off underload	No - Performance fall-off underload	No - Read-only Workloads	Legacy metadata processing capabilities curtails small file I/O
SCALE	1000's of compute servers and storage with linear scaling	Yes	No	No	No	Yes

COMPARISON OF CAPABILITIES

	Criteria/Capabilities	Weka	NetApp	Isilon	Vast Data	DDN Lustre
DEPLOYMENT MODELS	On-premises	Yes	Yes	Yes	Yes	Yes
	Cloud	Yes	Yes	Partial	No	Yes
	Converged	Yes	Yes	Yes	No	Yes
ADMINISTRATION	GUI, CLI, API	Yes	Yes	Yes	Yes	Yes
DATA TYPES	File, Object	Yes	Yes	Yes	Yes	No
MIXED WORKLOADS	Efficiently perform large and small IOPS on the same filesystem*	Yes	No	No	"Read-only"	No
CLOUD BURSTING	Natively moving data to the cloud**	Yes	Yes	Yes	No	Yes
CLOUD EXTENSIBILITY	Runs in the cloud same as on-premises. (deployment, management and functionality)	Yes	Yes	Yes	No	Yes
	Ability to extend namespaces to Object Storage	Yes	No	No	No	No
MULTIPROTOCOL SUPPORT	POSIX, NFS, SMB S3	Yes	No	No	No	Yes
CONTAINER SUPPORT	Works with CSI interface	Yes	Yes	Yes	Yes	Yes
SECURITY	End-to-end encryption (in-flight and at rest)	Yes	No	No	No	No
	Multi-tenancy (Ability to create multiple and separate namespaces with separated management capabilities)	Yes	Yes	Yes	No	Yes
	Support for ACLs	Yes	Yes	Yes	Yes	Yes
HARDWARE	Running on off the shelf servers	Yes	No	No	No	No
SCALING LIMITS	Agnostic (Not tied to a specific hardware vendor)	Yes	No	No	No	No
	1000's of storage servers	Yes	No	No	No	No
	1000's of compute servers	Yes	Yes	Yes	Yes	Yes
	Scale capacity and performance independently	Yes	Yes	No	Yes	Yes
	Integrate different tiers such as flash and HDD	Yes	Yes	Yes	No	Yes

* Same file system can accommodate maximum throughput and IOPS without a configuration requirement

** Ability to move data between different on-premises and public cloud environments without the need of 3rd party data mover applications or back-up applications.

Companies are not optimizing only for bandwidth, which is what Lustre and GPFS have historically done. They're delivering on IOPS and metadata performance, the things that matter when you are doing anything that scales out that is not a single large-scale simulation.

WekaFS™ – MODERN STORAGE FOR MODERN WORKLOADS

Today's workloads demand a new class of storage that delivers the performance, manageability, and scalability required to obtain or sustain an organization's competitive advantage. The Weka Data Platform was designed and optimized for data-intensive modern workloads. It is ideally engineered to take storage performance and data availability to the next level as performance demands intensify in Artificial Intelligence (AI), Machine Learning (ML) and Deep Learning (DL). Its architecture and performance are designed to maximize your usage of GPUs across cloud, on-premises, or hybrid deployments, providing data management capabilities that can accelerate time to insight to EPOCH by as much as 80x.

ADDITIONAL RESOURCES

WekaFS™ Architecture Whitepaper

<https://www.weka.io/resources/wekaio-architectural-whitepaper/>

Weka for AWS

<https://www.weka.io/parallel-file-system/public-cloud/aws/>

Redefining Scale for Modern Storage

<https://www.weka.io/blog/scalable-storage/>

Schedule a Free Trial

<https://www.weka.io/get-started/>

Contact Us

<https://www.weka.io/contact-us/>



910 E Hamilton Avenue, Suite 430, Campbell, CA 95008 T: 408.335.0085 E: info@weka.io www.weka.io

©2019-2021 All rights reserved. WekaIO, WekaFS, WIN, Weka Innovation Network, the Weka brand mark, the Weka logo, and Radically Simple Storage are trademarks of WekaIO, Inc. and its affiliates in the United States and/or other countries. Other trademarks are the property of their respective companies. References in this publication to WekaIO's products, programs, or services do not imply that WekaIO intends to make these available in all countries in which it operates. Product specifications provided are sample specifications and do not constitute a warranty. Information is true as of the date of publication and is subject to change. Actual specifications for unique part numbers may vary.