All-Flash File & Object for the AI Era

VAST Data Platform

Tier-1 Performance.
Archive Economics.

Exabyte-Scale Namespace.

No More Tiers.
No More Compromises.

An End to Decades of Data Management Complexity

When it's possible to consolidate NAS and Object workloads to a single data platform with all-flash performance, cloud scalability, and archive-level economics the barriers to insight disappear. Today VAST is the foundation for innovation across every industry, with over 8 exabytes of data under management and an unprecedented growth trajectory.

VAST Data has a singular focus on solving the problems that matter the most to our customers. Starting with technology that breaks the long held trade-offs, delivered by a simple, transparent business model with support engineering that creates partnerships with customers, the VAST experience simplifies every aspect of data management.

VAST set out to address the decades old limitations of scale and state inherent in storage and database systems. Legacy platforms forced customers to choose between performance or scale and to accept fractured data pipelines. Explosive data growth fueled by advanced analytics and AI workloads created an opportunity for a new platform that could provide fast access to all data, over any protocol, while delivering economics such that customers could afford to consolidate all of their data into one simple pool.

Scale to Exabyte & Beyond
Independently scale performance & capacity

82x Faster AI Performance
Compared to standard NFS

80% less total cost
Compared to other scale-out and cloud offerings

NAS Simplicity
No tiers, no trade-offs, no "nerd knobs"

The Principles of VAST Data.

Unlock fast access to all data

New insights, game-changing breakthroughs, and competitive advantage are powered by data. VAST delivers scale and performance so that all your data is ready for machine learning. With one simple tier of flash data, silos are eliminated, workflows are simplified, and insights are unleashed.

Break the performance, cost trade off

VAST makes the all-flash data center an affordable reality. By implementing a set of data efficiencies that are only possible with a new architecture, VAST reduces the cost of flash to that of hybrid storage systems, while delivering superior performance and none of the complexity.

Simplify cloud-scale operations

VAST composability, multi-tenancy, and resiliency are the foundation for public and private flash clouds. Consolidating multiple workloads to a single data platform requires Quality of Service (QoS) to ensure consistent performance and uptime and durability that delivers greater resiliency as the platform grows.

Simplified data pipelines

Data at massive scale requires a new approach to managing data. The old ways of locking data to the protocol by which it was ingested and using tools designed for TBs of data can't keep up. VAST stores data as elements for multiprotocol access with equal performance and granular security policies.

A new business model for modern infrastructure

VAST combines the flexibility of software-only with the simplicity of an enterprise appliance. VAST Gemini lets you buy hardware at cost, and license just the capacity you need in multi-generation clusters where you scale performance and capacity with zero migrations.

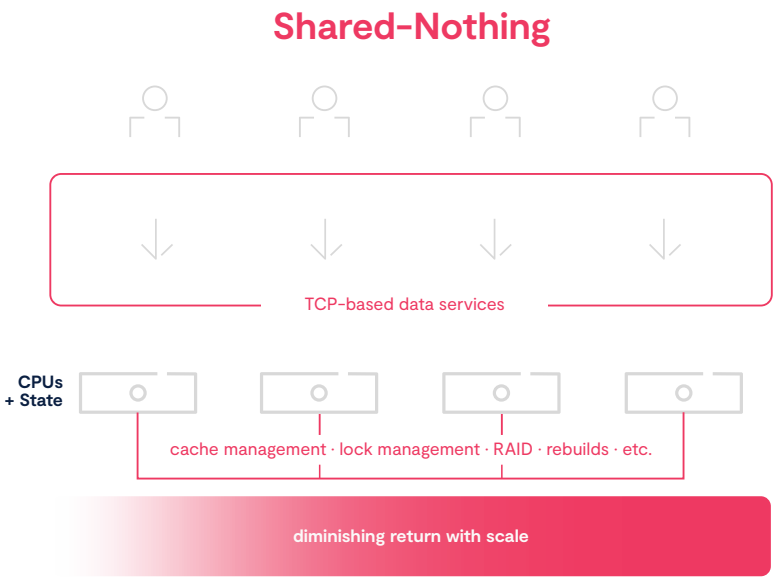
VAST | Gartner Challenger

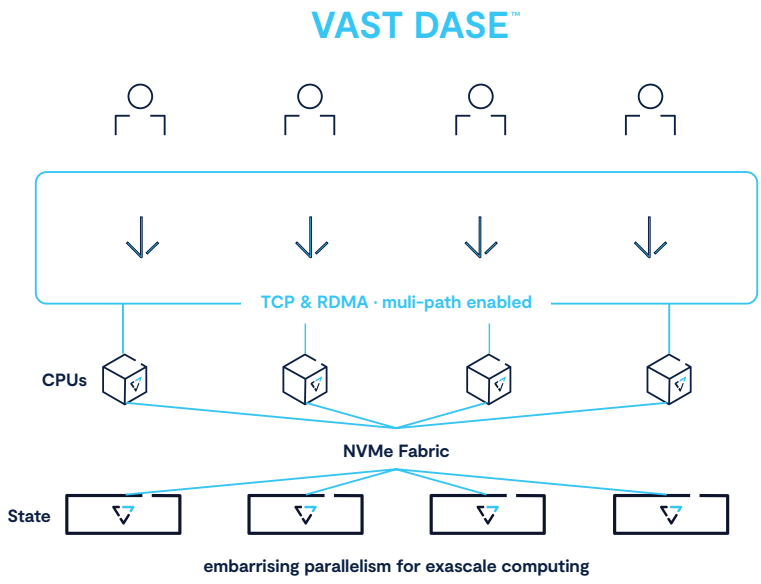
Highest Debut Ever on the **Magic Quadrant for Distributed File Systems and Object Storage.**

Technology Overview.

Disaggregated, Shared-Everything: A Modern Platform for All Data

VAST made the decision to break away from the Shared-Nothing architecture that had dominated the market for the last two decades. Shared-Nothing systems are built of nodes that bind together compute and storage which made independently expanding performance or capacity impossible. Moreover, as each node has exclusive access to its data, every I/O requires extensive communication between nodes such that as clusters grow, adding more resources produces diminishing returns.





While the rest of the industry layered on bolt-ons and work-arounds to their aging solutions VAST addressed the core problems with a new approach built exclusively for NVMe over Fabrics (NVMe-oF). Now its possible to serve client I/O from an unlimited number of stateless containers that have a global and consistent view of every storage device across a high-speed low-latency network. VAST calls this architecture appropriately Disaggregated, Shared-Everything (DASE).

In DASE, metadata, and data are stored on NVMe devices accessible from every stateless container eliminating the need for East-West traffic. Removing these limitations then allowed VAST to devise novel data protection and efficiency technologies to deliver economics on par with hybrid systems without sacrificing performance.

No Cache Coherency Challenges

Performance scales linearly when operations no longer need global coordination.

No Rebuilds During Server Failures

Statelessness eliminates rebuilds when servers fail, failover is instantaneous.

No Batteries Required

VAST cluster are 100% non-volatile. Never worry about a UPS or battery failure again.

Container-based Auto-Scaling

Containerized architecture lays the foundation for autonomous, adaptive cluster scaling.

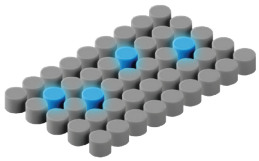
> *We require a high-performance big data solution to run our machine learning algorithms, that's also infinitely scalable to meet our future needs."*
>
> agoda

Break The Performance Cost Trade-Off

10-year flash longevity saves up to 85%

Global Wear Leveling & Write Amortization

Ultra-dense flash features the highest NAND capacities by fitting multiple bits per flash cell but at the cost of reduced endurance. To extend the life of flash for up to a decade VAST always writes in optimized blocks. VAST data intelligence groups data with similar expiration to further reduce eliminate the write amplification that occurs on traditional storage associated with deletes.



Ultra-efficient data protection saves up to 66%

A New Erasure Code To Break The Cost/Resilience Trade-off

VAST's locally decodable erasure codes provides data resilience with as little as 3% overhead. Built for flash VAST is able to create wider stripes for greater efficiency while only needing to read from 1/4th of the drives in a stripe. VAST's fail-in-place, declustered, approach to erasure coding brings all of the system resources to bear during a rebuild event in order to dramatically reduce the time to recovery and increase resilience.

Global data reduction saves 50-90%

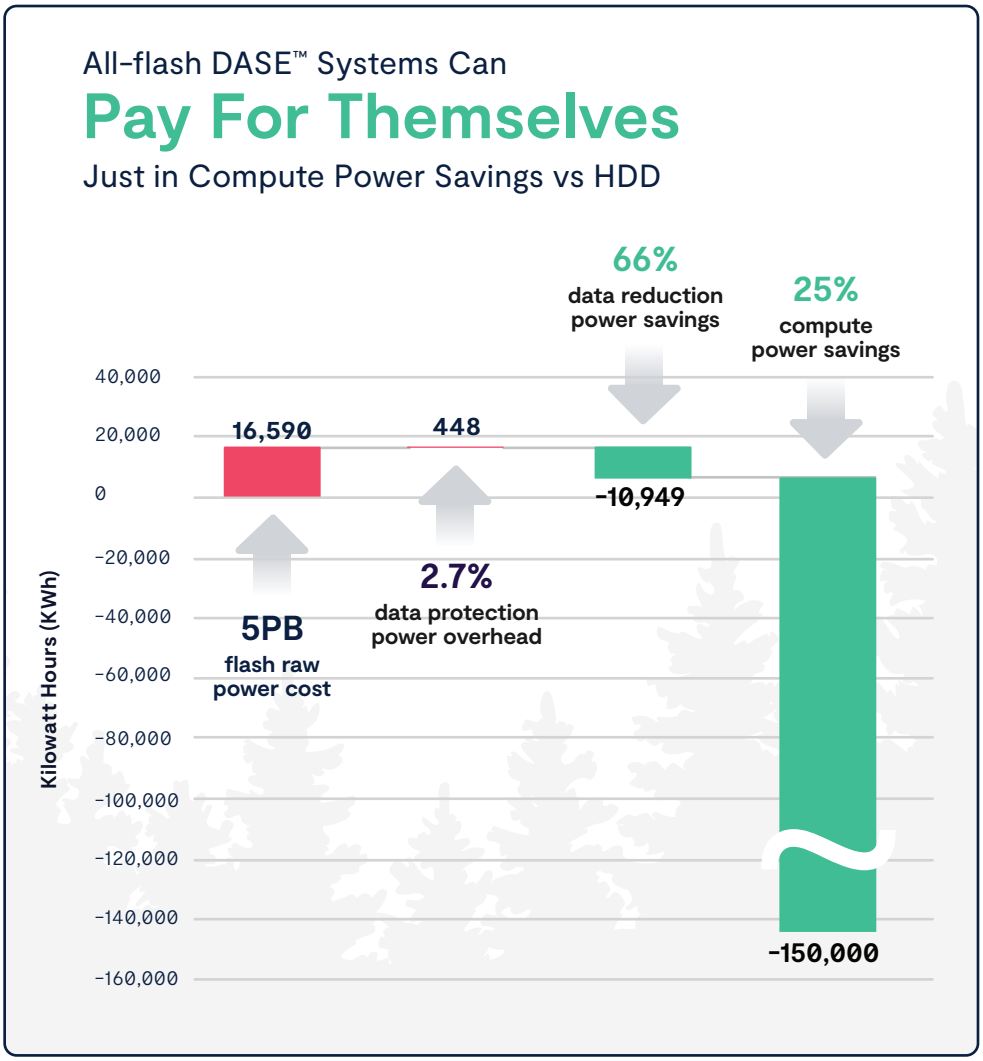
Similarity Based Data Reduction

Combining the global nature of deduplication with the fine-grained byte-granularity of pattern matching to achieve unprecedented levels of storage efficiency without compromising performance or endurance. Unlike legacy architectures where deduplication is limited to a volume or subset of the cluster, each VAST cluster is a single deduplication realm that finds more similarity as the system grows. VAST's data reduction technology can find additional savings even on already compressed and reduced data sets. As such, VAST is an ideal solution for all-flash data protection with rapid recovery.

> *VAST has changed the economics of flash storage. It's making flash affordable at large scale — and less expensive for all kinds of data and data sets."*
>
> jumptrading

Efficiency that Drives Sustainability

In environments such as high-performance computing (HPC), servers can account for as much as 80% of data center power consumption. VAST delivers all-flash performance enabling customers to run workloads in less time thereby consuming less energy on the compute layer.



This example synthesizes several VAST deployment experiences in large-scale computing centers. 5PB of effective storage capacity and 800 compute nodes (some with GPUs) in a large cluster environment accelerated by 33% vs. HDD.

> *"NIH deploys massive amounts of compute power to analyze massive datasets and attempt seemingly impossible projects. VAST infrastructure reduces processing time by up to 25% on NIH's analytical workloads, creating massive energy saving opportunities.*

NIH

Service Delivery at Cloud-Scale

To consolidate multiple workloads and applications to your own flash cloud you need to operate like a hyperscaler. Today VAST customers are supporting workloads together on VAST that would have required independent storage systems in the past to prevent “noisy neighbor” resource contention issues. VAST multi-tenancy and QoS features are key to enabling application consolidation onto one scalable all-flash platform. The VAST API allows dynamic provisioning to be automated, scaling resources up and down as needed.

Quality of Service Policies

Create I/O and bandwidth service levels

Encryption

External key manager for per tenant encryption

Quotas

Define capacity, file, and directory limits

Secure Multitenancy

Restrict tenants to authorized paths

Today, VAST is the foundation for both public and private cloud service providers and they rely on VAST to provide the performance and reliability to meet Service Level Agreements. Resilience is engineered into every aspect of the cluster:

- **No more file system repairs:** Storage class memory provides a high-performance, persistent write layer for zero volatility in the data path
- **No more late-night disk swaps:** A Fail-in-Place Cluster where the failure of any element from protocol servers to NVMe devices does not impact end-user workflows
- **No more “outage windows”:** Non-disruptive upgrades and expansions mean there is never a need to work the weekends.

“We can run 10s of petabytes of storage across three continents with no storage admin at all.”

dug

Data Management at Exabyte Scale

A True Multi-Protocol Platform

VAST was built from the ground up to be a data platform that in many ways defies classification. Most other storage solutions have a primary protocol with bolted on “multi-protocol” features that separate namespaces (no data interoperability) or limit performance and features of the secondary protocol.

Taking a new approach VAST stores data as elements and enables access by NFS, SMB, or S3 with granular security permissions. The VAST Element Store enables file and object-based methods of storing and accessing data without compromising on performance, data access, or efficiency.

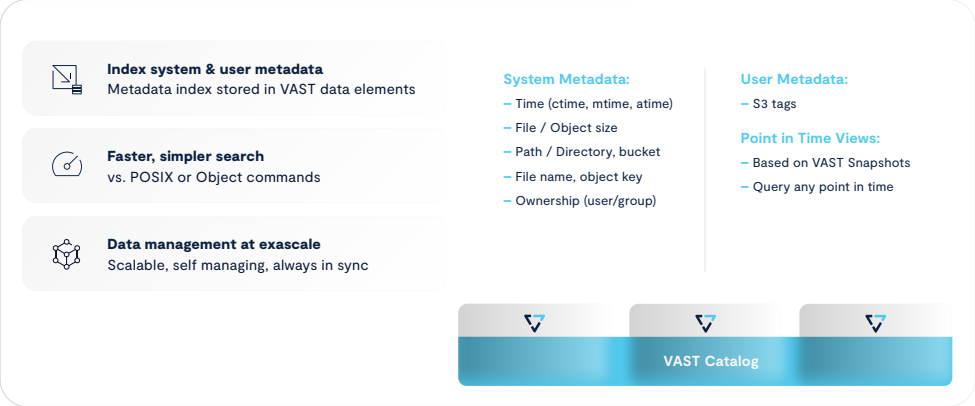
- **No protocol gateways**
- **VAST native NFS, SMB, and S3**
- **Unified permissions model**
- **Protocol auditing**

> *“We have been able to consolidate all of our storage to one platform – VAST is quite simply the last storage system you need to buy.*
>
> KRYSTAL
> Honest. Reliable. Personal.

Simplified Data Pipelines

VAST Catalog: Semantic Layer, Unified View of Data

The Element Store enables possibilities beyond just File and Object storage protocols. The first demonstration of this advanced capability introduces a built-in, high-speed index of all object and file metadata that we call the "VAST Catalog." Data sets that contain billions of files and objects are extremely difficult to manage with POSIX or Object utilities. VAST Catalog creates a high-speed, always-in-sync index, that lets you find data in seconds with a simple graphical interface or use the SQL-compatible query engine of your choice for advanced queries..



VAST Catalog: User Metadata examples

Search on key user metadata

Find all files where:
"ProjectName=GEMINI"

Combine ownership & filename

Find all objects where:
"group = HR" and where file name contains "contract"

Let users manage their space usage

Find all files where:
Path = my_homedir and file size is >10GB

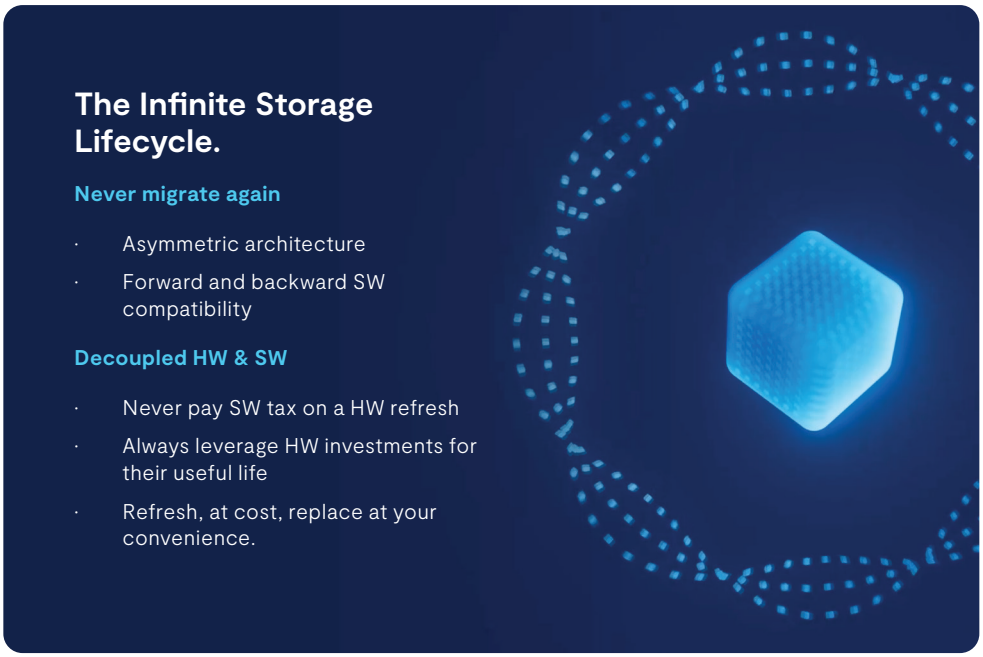
VAST

FAST COMPANY

Number 10 Most Innovative
for Data Science

A New Business Model for Infrastructure Sold as Software, Delivered & Supported as an Appliance

Meet Gemini – the flexibility of software-only storage with the simplicity and support of an enterprise appliance. With Gemini, customers purchase managed software on hardware purchased directly from VAST partners at cost.



Working with some of the world's leading technology manufacturers to specify resilient, scalable, and efficient equipment VAST builds an architecture that supports mix & match across generations of flash and compute. Because VAST sells hardware at cost there is no incentive to force hardware refreshes and customers decide when to expand or refresh to meet their needs.

Gemini provides an all-inclusive software license based on just the capacity needed, with full warranty support for hardware for up to 10 years for a seamless support experience.

Buy Hardware at Cost

VAST is not in the purchase path

Get full visibility to supply chain costs

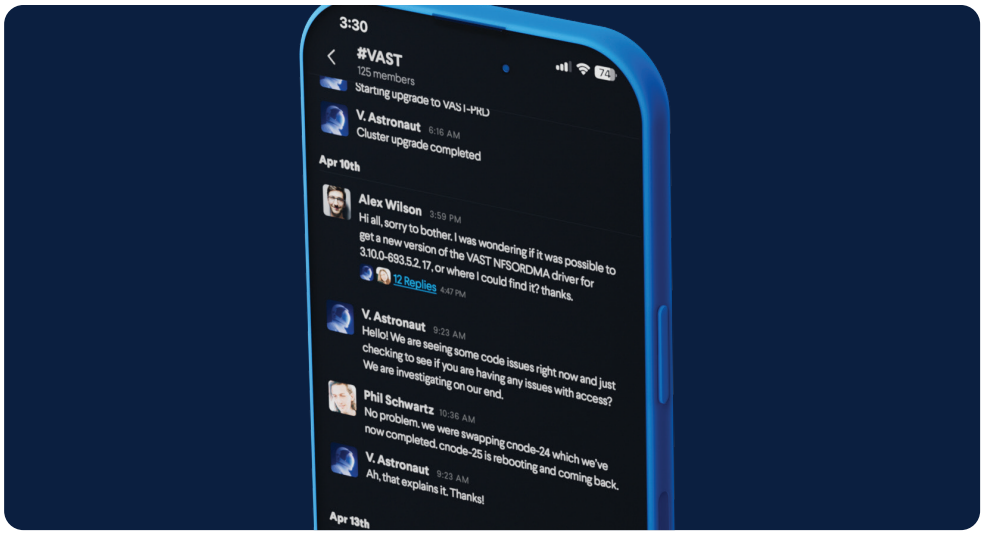
Maximize Buying Power

VAST manages the supply chain for you

A Simple, Turnkey Appliance

VAST telemetry ensures 24/7 monitoring

VAST processes all HW RMAs



Partnership, White-Glove, Jump the Queue

VAST eliminates the typical tiered support model, where customers are passed from engineer to engineer until a solution is found. VAST connects customers directly with L3 engineers, known as “Co-Pilots” for a white glove service experience and quick time to resolution. Co-pilots do far more than basic support, they handle the planning of deployments and upgrades, and are available to answer any questions about the system.

Gartner peerinsights™

Recommended by 100% of VAST Customers

★★★★★ **Evaluation & Contracting (4.8/5)**

★★★★★ **Integration & Deployment (4.8/5)**

★★★★★ **Product Capabilities (4.7/5)**

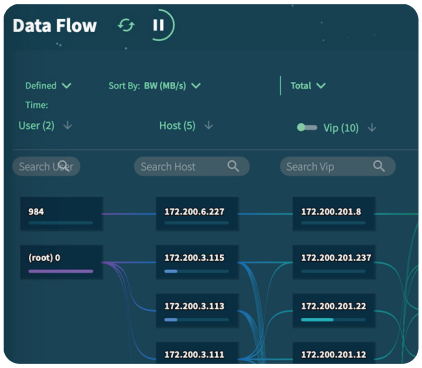
★★★★★ **Service & Support (4.9/5)**

> *“We’ve never gotten this level of support and near-instantaneous feedback from any vendor before. It’s amazing.*
>
> Albinaki A. Martinos Center for Biomedical Imaging

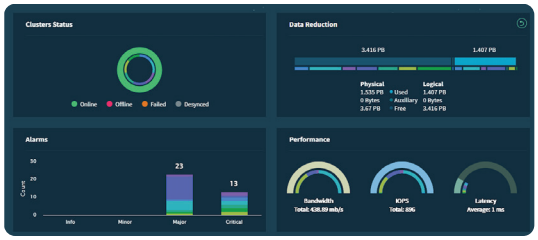
Simplicity Meets Power

The VAST Management System includes intuitive analytics tools such as Data Flow that gives admins a graphical view the IO load in real time as it is distributed across the cluster. Along with an elegant user interface, VAST utilizes an API-first approach to allow for seamless automation of any task.

You can easily manage multiple VAST clusters with Uplink, a cloud portal with active AI capabilities to help you proactively add resources and stay ahead of application demand.



Real time IO monitoring with Data Flows



Multi-cluster cloud portal with Uplink

VAST

Deloitte.

5th Fastest Growing
Technology Fast 500 North America

Enterprise Features

One Platform for All Data

Vertical Solutions	Backup & Recovery	Big Data & AI	Enterprise Infrastructure
Life Science	COMMVAULT, IBM	splunk>, elastic	vmware
Finance	Veeam, VERITAS	VERTICA, spark	Microsoft Hyper-V
Content	COHESITY, rubrik	trino, dremio	kubernetes
HPC	model	TensorFlow, PyTorch	openstack, Red Hat OpenShift

Protocols, Api

- NFSv3, oRDMA, Multipath
- NFSv4.1, oRDMA, Multipath
- SMB 2.1
- SMB 3, Multi-channel
- S3
- Container plugins: K8S, Openstack
- GPUDirect™
- IPv6 support

Data Services

- Exabyte scale namespace
- LDAP, Active Directory
- Multi-protocol ACL
- Metadata Index "VAST Catalog"
- Directory and User Quotas
- Snapshot and object immutability
- Similarity Data reduction: adaptive chunking, data-aware compression
- Capacity and Performance Analytics
- Distributed / Multiprotocol File Locking
- Multi-tenancy
- QoS
- S3 life cycle policy
- Access Audit: User Access (NFS, SMB), Admin Audit

Infrastructure Services

- Transactional Storage System
- Distributed, persistent metadata
- Global Hyperscale Flash Translation
- Encryption at rest and in flight
- Rapid rebuild erasure coding
- Fail-in-place architecture
- Tenant-based container pooling
- Asynchronous Replication w/automated failover
- Cloud/Object Replication Native replication (VAST to VAST)
- Multi-factor authentication

Cluster Management

GUI | CLI | REST APIs | Automation Plugins | Global Cloud Monitoring Dashboard (Uplink)