MASTER UNSTRUCTURED DATA STORAGE

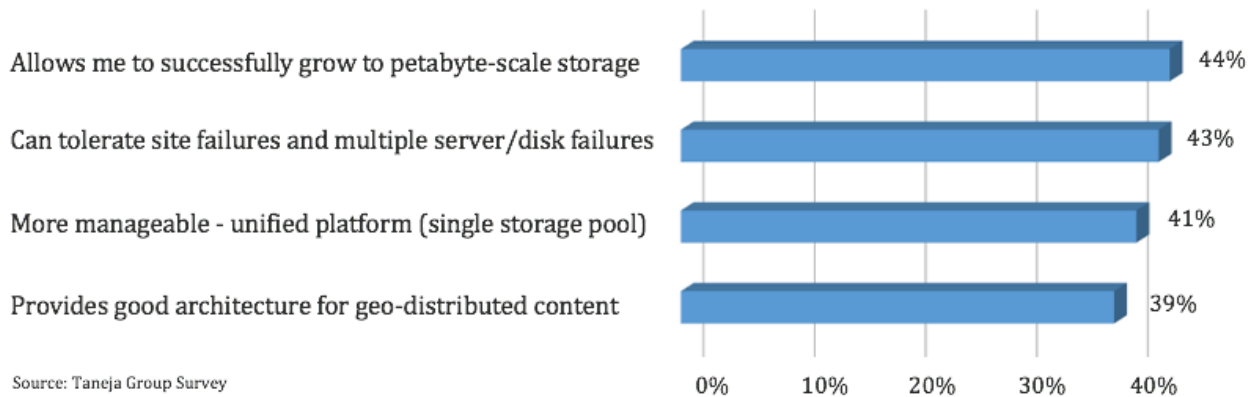## CISCO UCS S3260 – A VERSATILE SCALE-OUT STORAGE OFFERING

SEPTEMBER 2017

It may be time to change the way we store and mange unstructured data. Many companies are coming to this realization as they fully embrace the digital era. Organizations are using web-scale applications, content repositories, the Internet of Things (IoT) and big data analytics to improve customer insights, streamline processes and improve collaboration. However, as their digital transformation unfolds, companies often discover unstructured data is increasing at an alarming rate and the massive amount of unstructured data usually quickly strains legacy file storage systems.

Many companies are finding traditional file storage approaches simply aren't well-suited to manage vast amounts of unstructured data. The most common file storage limitations are highlighted in a recent Taneja Group IT survey that found lack of flexibility, poor storage utilization, inability to scale to petabyte levels and failure to support distributed data are the top challenges with traditional file storage solutions. These obstacles often lead to high storage costs, complex storage management and the inability to support a broad range of unstructured storage use cases.

As a result, companies are turning to scale-out object storage solutions that break through traditional storage scaling limitations. The Taneja Group Survey revealed that (see chart below) scale-out object

What Companies Like Most About Object Storage

Allows me to successfully grow to petabyte-scale storage	44%
Can tolerate site failures and multiple server/disk failures	43%
More manageable - unified platform (single storage pool)	41%
Provides good architecture for geo-distributed content	39%

Source: Taneja Group Survey

storage is ideal for managing huge amounts of unstructured data because it easily scales to hundreds of petabytes by simply adding storage nodes. Scale-out storage also provides efficient fault tolerance and simplifies storage management by eliminating the hierarchical structure used in file storage and placing everything into a flat address space or single storage pool. These object storage features are essential to cost-effectively manage large-scale storage environments. Add to this support for geographically distributed environments and it's easy to see why scale-out storage solutions are the preferred storage approach for multiple use cases, such as file backup, active archives, storage services, big data analytics, content repositories and web-scale applications.

So, what should companies look for in a scale-out object storage offering? As with any IT solution, it is essential to have the right architecture. For scale-out object storage this means having a modular design that supports multiple unstructured workloads, applications and use cases. This is an area where we believe Cisco's scale-out storage offering based on UCS S3260 really shines. Cisco UCS S3260 modularity enables multidimensional scalability and I/O flexibility. As a result, admins can quickly vary computing, network, and storage resources to accommodate IOPS-intensive, throughput-intensive and capacity-intensive workloads. Cisco UCS S3260 also includes Cisco Validated Designs (CVDs) for reliable implementation of data services from major object storage vendors, such as Scality, SwiftStack, IBM Cloud Object Storage (COS) and Red Hat Ceph. By combining the Cisco UCS S3260 modular storage server architecture with the flexibility enabled through Cisco CVDs, we believe Cisco has delivered a scale-out object storage solution that provides an exceptional opportunity to deploy versatile scale-out object storage that provides a true multipurpose advantage.
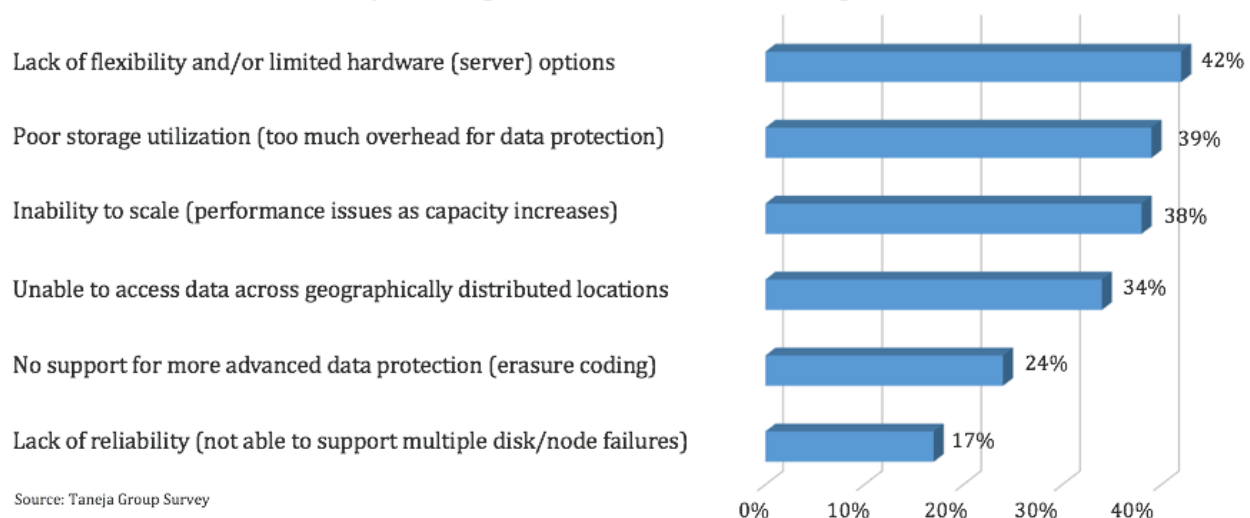
In this paper, we will look at the limitations of legacy file storage and the key elements of scale-out object storage platforms. We will also highlight capabilities and advantages of the Cisco UCS S3260 scale out offering and we will examine several scale-out storage use cases.

MANAGING UNSTRUCTURED DATA WITH LEGACY FILE STORAGE

For years, companies have used traditional file storage systems to store unstructured data. On the surface these systems appeared to be a viable long term option. Unfortunately, the turnkey nature of legacy file storage also makes it inflexible and not very scalable.

The most common traditional file storage shortcomings were highlighted in a recent Taneja Group survey, see the chart below. The survey found the top four challenges to be lack of flexibility, poor utilization, inability to scale and the inability to support distributed data.

Key Challenges with Traditional File Storage

Challenge	Percentage
Lack of flexibility and/or limited hardware (server) options	42%
Poor storage utilization (too much overhead for data protection)	39%
Inability to scale (performance issues as capacity increases)	38%
Unable to access data across geographically distributed locations	34%
No support for more advanced data protection (erasure coding)	24%
Lack of reliability (not able to support multiple disk/node failures)	17%

Source: Taneja Group Survey

Companies that have continually dealt with these issues have begun implementing software-defined scale-out storage solutions that address the deficiencies of traditional file storage. To find out when the shift to scale-out object storage occurs, Taneja Group recently surveyed companies to find out their preferred solution for various storage requirements. The Taneja Group survey results clearly showed that when storage requirements reach half a petabyte, companies start thinking seriously about object storage and once their storage needs reach a petabyte or more, object storage becomes the clear choice.

OBJECT STORAGE REDEFINED AS MULTIPURPOSE SCALE-OUT STORAGE

As object storage continues to mature, so too does the adoption of object storage solutions increase and evolve. Object storage started as proprietary back end infrastructure for storing massive amounts of unstructured data (photos, videos, etc.) at cloud focused companies like Facebook and Twitter. More recently, object storage solutions have gained significant traction in enterprise organizations for a range of unstructured data use cases, such as active archives, big data analytics, scalable backup storage and content repositories.

We believe this multipurpose theme will influence the next wave of object storage solutions as companies look for scale-out offers capable of supporting block, file and object based applications as well as IOPS-, throughput- and capacity-intensive workloads. Below are key elements of multipurpose scale-out solutions.

Category	Key Elements of Multipurpose Scale-Out Storage Solutions
Petabyte-level scalability	Scale-out architecture that easily expands to multiple petabytes by simply adding additional server nodes
	Performance that scales linearly across mixed workloads as the number of nodes in a cluster increases
Unified system management	Unified control plane for provisioning, monitoring and managing compute and storage resources and storage and network connections.
	Policy-based management with service profiles and workflow templates that automate and accelerate routine admin tasks and simplify operations
Extreme durability	High fault tolerance - ability to tolerate site failures and multiple server and disk failures
	Ability to automatically rebuild data with self-healing technology
Support for geo-distributed data	Global shared name space for universal data access, where file information is shared across all nodes in a multi-site cluster
	File sharing protocols designed for the internet (RESTful / HTTP protocols)
Cost-effective storage	Freedom to select industry standard servers with good price performance
	Efficient data protection with low overhead to enable high utilization
Versatile architecture	Ability to vary computing, network, and storage resources to accommodate IOPS-intensive, throughput-intensive and capacity-intensive workloads.
	Support for file, block and object protocols (NFS, SMB, iSCSI, FC, S3, etc.)

KEY OBJECT STORAGE USE CASES

The adoption of object storage was initially driven by companies building web-scale applications. Companies discovered that easy storage access (via RESTful APIs) and limitless scalability on demand via cloud services made object storage well-suited for their unstructured data storage requirements. The advent of object storage solutions has moved object storage beyond the realm of web-scale applications by providing a comprehensive set of data services, including rich metadata indexing and multi-protocol support (NFS, SMB, S3, etc.).

The adjacent chart highlights top scale-out object storage use cases identified in a Taneja Group Survey. We have also summarized these scale-out object storage use cases below:

File backup/archival: This was the most common use case sited in a recent Taneja Group survey. This is undoubtedly because there's significant demand for scalable file backup and online archival storage that facilitates timely document retrieval and cost-effective long-term data retention to ensure compliance.

Top Object Storage Use Cases

| Use Case	Percentage
File backup (archive)	57%
Storage-as-a-Service	44%
Big Data analytics	35%
Secure file sharing	35%
Content repositories	20%
Web-scale applications	13%

Source: Taneja Group Survey

Object storage is ideal for large-scale unstructured data storage because it easily scales to multiple petabytes by simply adding storage nodes and provides highly durable and incredibly efficient capacity-based storage. Support for multiple file protocols (NFS, SMB, POSIX, etc.) and traditional file service features, for instance file versioning and file locking, are also important considerations.

Storage as a service: Object storage gives service providers a cost-effective way to manage backups using a secure, highly scalable, multi-tenant architecture. Because storage is provided "as a service" companies save on personnel, hardware and data center facility costs. An IT administrator simply rents storage space on a cost-per-gigabyte-stored basis. These same benefits apply for private clouds where IT implements a portal/marketplace for providing Infrastructure-as-a-Service.

Big data analytics: Object storage is designed for large data sets, making it ideal for big data analytics. In a recent survey, Taneja Group found that about 30% of respondents aggregate 100 TBs of data or more for big data use cases, with the amount of data growing substantially every month. However, companies must tightly integrate object storage with low-latency storage and high-performance compute to support big data analytics and artificial intelligence (AI) for data mining, correlation and interpretation. A good example is IBM Spectrum Scale and UCS S3260. This is a high-performance solution for managing data at scale, which is ideal for big data analytics.

Media / Content repositories: Media and entertainment companies need to store and manage petabytes of images and videos and distribute digital content globally to users on demand. Healthcare providers have digital medical images in picture archiving and communications systems (PACS) and vendor neutral archives (VNAs) and patient care initiatives like personalized medicine and genomics require storing, sharing and analyzing massive amounts of unstructured data. Object storage's geo-dispersed data architecture and flexible data protection (replication and erasure coding) inherently facilitates efficient and secure file sharing across a multi-site cluster, which facilitates geographically distributed data access.
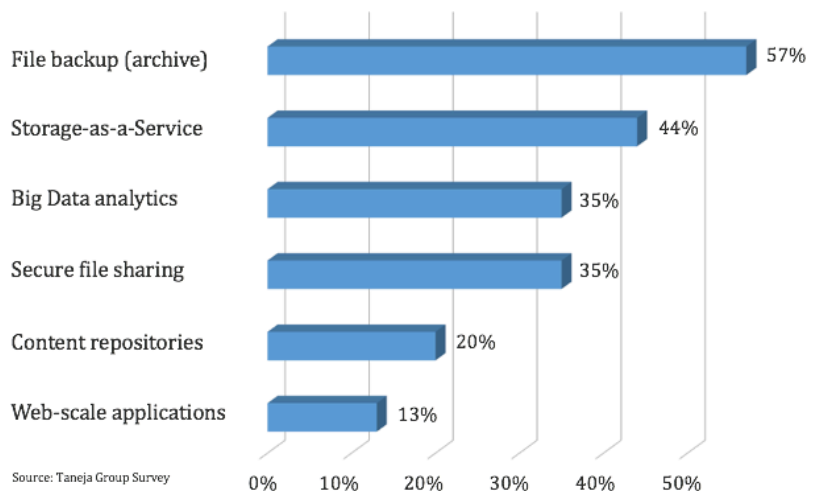
## CISCO UCS S3260 STORAGE SERVER OPTIMIZED FOR SCALE-OUT STORAGE

When it comes to storing and managing unstructured data, modern IT groups turn to storage solutions built for scale. They want storage solutions equipped to handle the massive amounts of data generated by the ever-expanding number of digital services. Therefore, IT seeks storage solutions that provide consistent performance at scale, as well as high fault tolerance and simplified management, i.e., capabilities essential for scalability, durability and cost-effectiveness.

Companies also want versatility. This is where Cisco's scale-out storage offering based on UCS S3260 separates from the pack. In addition to data services, such as rich metadata indexing, multi-protocol support and public cloud integration, Cisco UCS S3260 has gone a step further by offering a modular architecture that supports multiple unstructured workloads, applications and use cases. Below, we take a look at outstanding features and benefits of the Cisco UCS S3260 scale-out offering.

Petabyte-scale, according to a recent Taneja Group survey, the ability to successfully scale to petabyte levels is the capability companies like most about object storage. The Cisco UCS S3260 meets this requirement impressively through a scale-out architecture that easily expands to petabyte-scale by simply adding additional nodes. The Cisco UCS S3260 also offers massive capacity. With 10 terabyte (TB) disks, users can easily achieve a capacity of up to 600 TB with a single high-density server available in a compact 4-rack-unit (4RU) form factor.

Versatility is another hallmark of the Cisco UCS S3260 scale-out offering. The Cisco UCS S3260 achieves versatility by offering modularity throughout the Cisco UCS S3260 scale-out solution. This modularity, or composable architecture, enables multidimensional scalability and I/O flexibility. Admins can quickly vary computing, network, and storage resources to accommodate IOPS-intensive, throughput-intensive and capacity-intensive workloads. And admins can easily upgrade or replace computing nodes and network interfaces without the need to migrate data from one system to another, which reduces administration time and operating costs.

Simplified management was also highlighted by the Taneja Group survey as a key factor companies like most about object storage. Cisco UCS S3260 delivers ease of management with the embedded Cisco UCS Manager. UCS Manager provides an intuitive HTML 5 or Java user interface and a unified API-driven control plane for server provisioning, device discovery and configuration, diagnostics, auditing and statistics collection. The UCS Manager also updates, monitors, and manages computing and storage resources and storage and network connections. And last, but certainly not least, UCS Manager provides role- and policy-driven management using service templates and profiles (more details below).
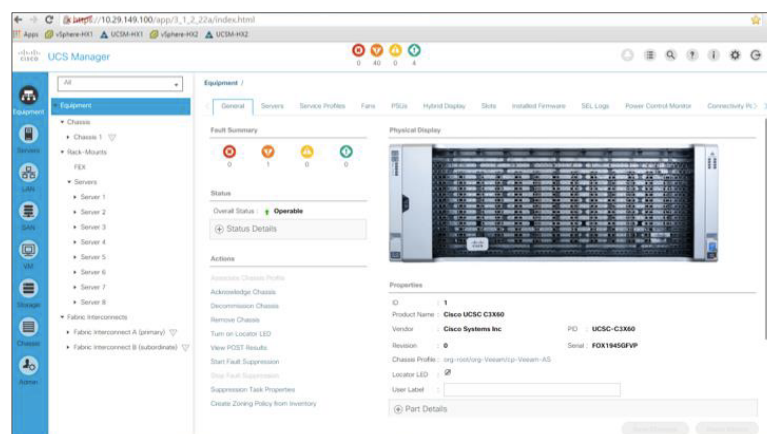
Storage automation is powered by UCS Manager and UCS Director. UCS Manager provides policy-driven management using service templates and service profiles. UCS service templates and profiles encapsulate server, storage and networking policies to enable standardized configurations and facilitate infrastructure automation. UCS Manager can be integrated with Cisco UCS Director to automate server, fabric and storage provisioning, along with a host of other functions. For example, a UCS Director workflow can automate the tasks required to expand a running Red Hat Ceph cluster.

UCS Director can also automate transformation to other object storage solutions. such as Scality, SwiftStack, IBM Cloud Object Storage (COS) and Red Hat Ceph.

Performance is a major requirement for throughput-intensive workloads, such as IoT, video on demand (VoD) and web-scale applications. To deliver consistent high performance, the Cisco UCS S3260 scale-out solution provides fast network access and easy access to additional CPU and memory. The Cisco UCS S3260 scale-out solution supports both 10 or 40 gigabit dual-port Ethernet network access and supports dual RAID controllers to double both read and write performance. Cisco UCS S3260 also supports Non-Volatile Memory Express (NVMe) flash memory and Fusion IoMemory. NVMe provides two or three times better flash performance as a result of the closer CPU architecture and direct connectivity to the PCIe bus. Fusion IoMemory, primary used with databases and Big Data applications, provides a scalable multi-queue block layer for the Linux kernel. This block layer uses parallelism, which enables higher performance.

Open Architecture is important to avoid technology lock-in. This is an area were the Cisco UCS S3260 scale-out solution really shines. Cisco UCS S3260 includes several Cisco Validated Designs (CVDs) for reliable implementation of data services from major object storage vendors, such as Scality, SwiftStack, IBM Cloud Object Storage (COS) and Red Hat Ceph. Through the CVDs, Cisco actively collaborates with these leading object storage solution providers to deliver valuable data services, like parallel data ingest, data replication and space efficient erasure coding for remarkable fault tolerance, shared global name space and rich metadata indexing for universal data access and multi-protocol (NFS, SMB, S3, etc.) support. Object storage solutions also provide public cloud (Amazon Web Services (AWS), Microsoft Azure, Google Cloud, IBM Bluemix) integration through RESTful APIs and S3 or Swift protocol support.

Single Support is vital to eliminate downtime and shorten time to resolution. Cisco offers a single support model for the Cisco UCS S3260 solution through the Cisco Technical Assistance Center (TAC). Cisco TAC provides customers with a primary point of contact and coordinates communication within Cisco and between key vendors which provides a no-hassle support experience from first call to resolution.

Flexible Financing is available for Cisco UCS Storage. Companies can choose to go with a traditional perpetual purchasing agreement paid upfront or select a standard lease that allows for even monthly payments. These options have advantages in the long run if companies are utilizing most or all of the Cisco UCS storage capacity. If a company's storage needs are less certain, they may prefer a flexible consumption model and go with the Cisco Storage as a Service subscription option where they only pay for the storage they use.

TANEJA GROUP OPINION

With companies facing massive unstructured data growth, IT groups are focused on transforming their unstructured data storage to improve scalability, lower storage costs and enable flexibility to support multiple unstructured data use cases. With these goals in place, companies invariably realize that true transformation requires embracing scale-out storage solutions.

Scale-out storage solutions support both scale-up and scale-out storage. Therefore, companies have the ability to maximize individual server performance and the ability to scale multiple petabytes by simply adding additional server nodes. Scale-out storage also provides linear performance as the number of nodes in a cluster increases. Moreover, scale-out object storage solutions provide simplified storage management across a multisite cluster, the ability to automatically rebuild data with self-healing technology, efficient data protection with low overhead to enable high utilization and hardware independence – core capabilities that are essential to cost-effectively manage large-scale unstructured storage environments.

Given all this, it's easy to see why scale-out object storage solutions play an essential role in unstructured data storage transformation. And when companies look at the other important dimension for transforming unstructured data environments, which is the flexibility to support multiple unstructured data use cases, they find that an object storage solution's versatility is key to enabling support for multiple unstructured data applications and use cases. This is an area where we believe Cisco's scale-out storage offering based on UCS S3260 really stands out. Cisco UCS S3260 modularity enables multidimensional scalability and I/O flexibility. As a result, admins can quickly vary computing, network, and storage resources to accommodate IOPS-intensive, throughput-intensive and capacity-intensive workloads.

Cisco UCS S3260 also includes Cisco Validated Designs (CVDs) for reliable implementation of data services from major object storage solutions, such as Scality, SwiftStack, IBM Cloud Object Storage (COS) and Red Hat Ceph. When you combine the Cisco UCS S3260 modular storage server architecture with the flexibility enabled through Cisco CVDs, we believe the Cisco UCS S3260 scale-out offering offers an exceptional opportunity to implement versatile scale-out object storage with a true multipurpose advantage.