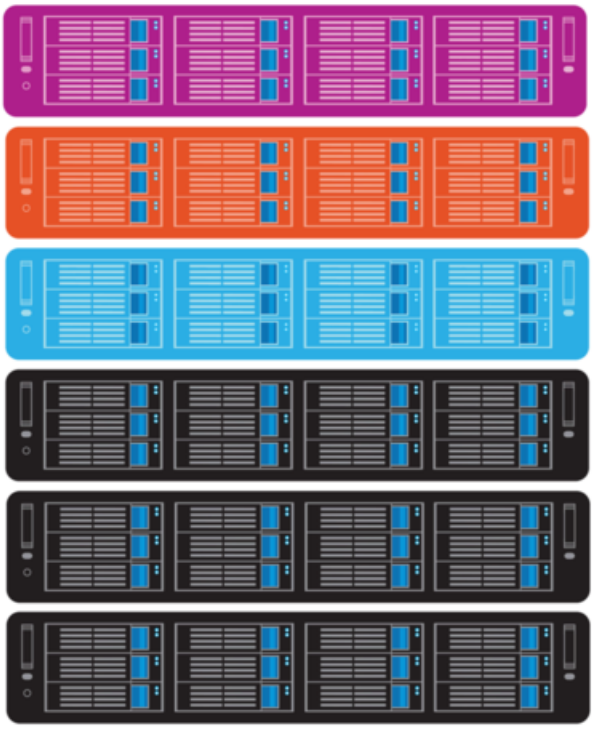
SCALA STORAGE

White Paper

Scala Storage Scale-Out Clustered Storage White Paper



www.scalastorage.com | info@scalastorage.com | phone 800 980 1988 | fax 626 322 3878

Chapter 1 Introduction 3
Capacity - Explosive Growth of Unstructured Data 3
Performance - Cluster Computing 3

Chapter 2 Current Storage solutions 4
DAS 4
NAS 4
SAN 4

Chapter 3 Clustered Storage System 5

Chapter 4 Scala Storage Scale-Out Solution – Cluster Storage Systems 6
Scala Storage Scale-Out Solution at glance 7
Scala Storage Scale-Out Solution Software Components 8
ScalaFS - Scala file System 8
ScalaManager - Scala Management 8

Chapter 5 Functions and Features of Scala Storage Scale-Out Solution 9
Scala Storage Files Systems Features: 9
Management of Massive Data 9
Fast filename lookup 9
Performance 9
Global Namespace 10
Large Directory Support 10
Scalability 10
Load Balancing 11
Data Security 11
Data Redundancy 11
Preventive Failure detection and Auto Recovery 11
System Management 12
Fast deployment and Easy Management 12
Compatibility with Current Environments 12
Cost Effective Scale-Out Solutions 12
Low System Cost-of-ownership 12

Chapter 6 Summary 14

Chapter 1 Introduction

In today's computing environment, there are two major challenges for high performance storage systems: capacity and performance. Below are few examples of how these two requirements are important in different environments

Capacity - Explosive Growth of Unstructured Data

Increasing usage of computers for different industries creates a high demand for data storage. Data storage is growing at a rapid pace. Data can be well structured or unstructured. Traditional data such as databases, files, and emails are structured or semi-structured usually at the MB or GB level, but unstructured data is different. Both the size of every file and total amount of unstructured data is growing. For example, audio or video is enlarging gradually as the definition gets higher and a single file can reach a dozen GB. In scientific computing, seismic data could reach hundreds of GB. As a result, the total data storage capacity could easily amount to hundreds of TB. Mass unstructured data is in dire need of being conveniently stored and managed. Companies need a storage platform that is large enough to hold that data and at the same time be flexible enough to grow with the highest performance possible.

Performance - Cluster Computing

With fine expansibility and high cost performance of cluster architecture, cluster technique has been widely applied into various industries. In many companies, the relative services have switched the use of traditional expansive and specialized "large computer" and "minicomputer" to cluster architecture. As a result, the capacity for computing and processing has been enormously increased and can meet the demands brought by the development of services.

With flexible computing and processing capacity, cluster architecture also poses new challenges for companies. The powerful computing and processing capacity of cluster architecture has to operate on data and when data is accessed concurrently. Storage systems that immediately respond to cluster computing requests could lower performance and waste numerous computing and processing resources. The development of cluster architectures will need a large-scale storage platform that possesses flexible expansibility and efficient data processing capacity, in order to support the powerful cluster computing.

Chapter 2 Current Storage solutions

There are three kinds of popular storage solutions: DAS (Direct Attached Storage), NAS (Network Attached Storage) and SAN (Storage Area Network).

DAS – Direct Attached Storage

DAS (Direct Attached Storage) devices are designed to provide direct access to small-shared data pool for the IT environment of department level. The storage device is directly attached to servers through a cable. For small data centers, this might be an acceptable solution but as storage scales enlarge, companies will have to add DAS devices one by one. These systems cannot communicate to each other and so becomes an isolated information island, which results in managing nightmare in storage system and hampers the efficiency. Moreover, once server breaks down, information resources will also be lost with the failed server and cannot be accessed. DAS storage solution in datacenter lost its advantage in coping to work with higher requirements in many application.

NAS – Network Attached Storage

NAS (Network Attached Storage) is a relatively simple solution for data sharing. It provides the global file sharing through NFS/CIFS/AFP protocol to front end so that data input and output can be coordinated among more than one application server. Traditional NAS provides storage service through a single server whose performance is limited to single server configuration. When access becomes more frequent, NAS will soon turn to the bottleneck of performance.

There is still other limits to NAS such as the maximum storage capacity, storage capacity of file systems and amount of files. With the continuous increase of data, these limits will cause trouble for the unified management of data.

SAN – Storage Area Network

SAN (Storage Area Network) provides access interface through fibre channels in 8 GB/s and 16 GB/s or in Gigabit or 10G Ethernet network in iSCSI in IP SAN. The operating systems will connect to the SAN hard drives and format, as if they are local hard drives. SAN works as if they are DAS structures but are more flexible and scalable. This structure does not support data sharing capabilities. The complexity of operation and maintenance from fibre channel networking and iSCSI performance is a concern for large storage solutions.

Chapter 3 Clustered Storage System

DAS, SAN and NAS are three most common storage architectures. As unstructured data and digital content is growing rapidly, the storage scale of unstructured content is enlarging. The traditional storage systems are not enough to serve today needs in high performance, capacity, flexibility and yet easy to manage. The traditional storage architectures are designed mainly for structured data containing small files and advanced service such as database and email servers. However, unstructured data has unique features that are beyond the management of traditional storage architecture abilities. Large file size and data quantity, higher throughout requirement, read intensified access and concurrent access by large amounts of computers are few example that DAS, NAS and SAN are not designed for. For internet cloud based storage, streaming media, high-performance computing (HPC) and energy industry, they require data usage in highly intensified and accessed in concurrently. Expansion of capacity and performance while there are no downtimes is something traditional storage systems can never achieve.

Cluster storage experienced fast development and has now been widely accepted as new generation enterprise-level storage architecture. It has been years since cluster technique was applied to the server and achieved remarkable effects. Cluster storage, as the development and extension of cluster technique, has flexible expandability in term of storage capacity and performance. Through the cluster structure, clustered storage systems are constituted by several storage systems, and can easily expand performance and capacity seamlessly. It has effectively improved the efficiency of file sharing.

Cluster storage integrates the storage space of more than one storage device into a storage pool. Each storage device provides access for application servers so that all the disks on the storage devices are making full use of the performance. Data will be written and read from more than one storage devices according to preset policies to achieve a higher concurrent accessing performance.

Beside performance, the advantages of cluster storage improves the whole performance of concurrent or partition I/0 in cluster server environments, especially in workflows that read-intense files or accessing large files.

Energy industries, broadcast television, scientific computing and internet on demand have all become customers of cluster storage systems. They have gained high returns on their investments into cluster storage systems due to the nature of scalability in capacity and performance.

Chapter 4 Scala Storage Scale-Out Solution – Cluster Storage Systems

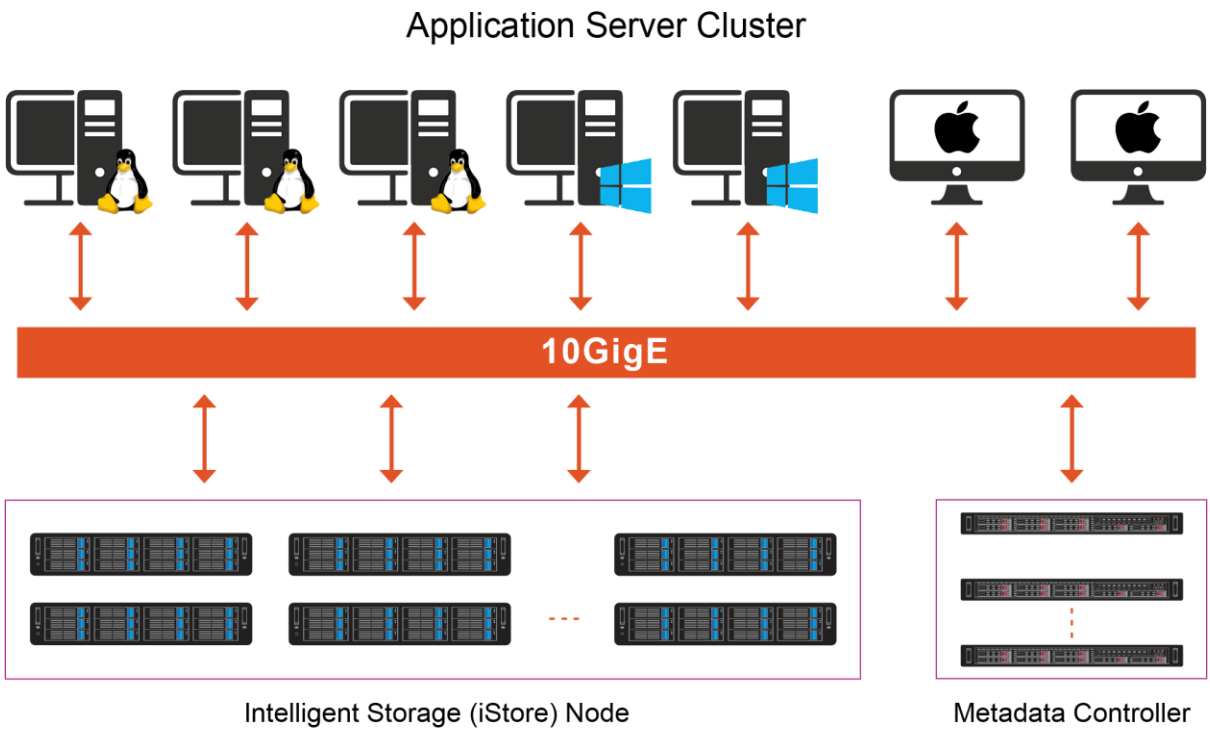
Scala Storage Scale-Out Solutions is a clustered storage system designed for large-scale, data-intensive applications. With many years of hardware and software experience, Scala storage aggregates computing powers and storages into a storage cluster system with interface to user applications is full POSIX-compliant.

Scala Storage provides a storage capacity in hundreds of PB levels and can dynamically expand online as required, based on the needs of application and usage. The dedicated metadata server cluster in Scala Storage Scale-Out storage has eliminated the limits of the current storage systems concerning the file limitation under a single directory and processing speed of small files. These dedicated servers provide an unlimited number of file storage and extremely high performance and response times for filename lookups. Scala Storage clustered storage architecture efficiently support tens of millions of large directories (tens of millions of files can be easily built under single directory and random searches for files can be done at highest performance speed available). Meanwhile, Scala Storage Scale-Out Solutions is designed to be highly flexible with high redundancy in mind. The system can Scale-Out with performance and capacity, while completely eliminating single point failures. Built-in preventative failure detection and immediate auto failure recovery, ensures continuous and stable operation for mission-critical environments.

While Scala Storage provides many features related to today never-ending storage requirements, it can be manage and review performance in a single management interface, allowing IT administrate with minimum learning curve to the storage systems.

Different successful cases in oil and energy, media entertainment, high performance computing, life science research and cloud based on demand application proven Scala Storage Scale-Out Storage is capable for today storage demands.

Scala Storage Scale-Out Solution at glance



Scala Storage Scale-Out cluster storage systems consists of three parts: application server cluster (Windows/Linux/Mac server or workstation), metadata server controller (MDS), and intelligent storage node (iStore).

- Application Server Cluster: The cluster runs user applications such as scientific computing, information processing and web services, etc. By installing Scala Storage client software, these application servers can access the shared storage pool. Scala Storage client supports Windows, Linux and Mac operation systems.

- Metadata Controller (MDS): Scala Storage Scale-Out Solution can support up to 128 metadata servers depending on the number of files in the storage systems. Metadata controllers has three responsibilities. First, it manages the metadata of file system such as directory tree structure, time of creation, owner ID, access permission settings, etc. Second, it is responsible for managing global namespace, allowing each application server to access the files and coordinates the traffic between application server and storage server. Finally, metadata server provide the management interface for the systems. Metadata is distributed across multiple metadata servers.

- Intelligent Storage Node (iStore): iStore nodes stores data and provide I/O services to application servers. Data is replicated and distributed across multiple iStore servers.

Metadata controller are only involved when there is a metadata related operation is required, such as file creation or file open. When application servers access data, they only communicate to metadata controller once, and access the requested data from iStore node without further interaction with metadata servers. By separating data and metadata, Scala Storage Scale-Out Solution is able to break performance bottlenecks and provide better cache utilization, performance and scalability.

Scala Storage Scale-Out Solution Software Components

- **ScalaFS - Scala File System**

  ScalaFS is a completely POSIX-compliant cluster file system. It replicates and distributes files across different iStore nodes.

  Key Functions of ScalaFS are as following:

  - Virtualize storage across all available storage servers into a single unified storage pool and provide a global namespace.
  - Metadata controller support metadata-intensive applications and large amount of file concurrent access.

- **ScalaManager – Web Based Management**

  Scala Storage Scale-Out Solution includes a centralized management GUI, ScalaManager, to configure, manage and monitor the storage systems. One administrator can simple manages ~hundreds of PB of storage. ScalaManager is available for Windows, Mac and Linux via Web based interface.

Chapter 5 Functions and Features of Scala Storage Scale-Out Solution

Scala Storage Files Systems Features:

- **Management of Massive Data**

  The massive amount of data generated by graphics, audio and video, combined with fast growing Web 2.0 applications creating a critical demand for innovative file storage products that can economically and efficiently manage massive data. Scala Storage Scale-Out Solution supports PBs of storage per volume, tens of billions of files per namespace and tens of millions of files per directory.

- **Fast filename lookup**

  One of ScalaFS's shining points is fast filename lookup. When the amount of files reaches tens of millions, filename lookup will create huge pressure to file system. The situation will be even worse if there are a lot of small file IO or any workload that generates many non-sequential metadata operations.

  Scala Storage world class distributed metadata techniques of multivariable and high efficient mass file search. While Scala Storage Scale-Out Solution hold tens of billions of files, the metadata controllers and its technologies ensure high filename lookup efficiency. The systems could manage tens of millions of files under one directory and enable searching rate of tens of thousands of files per second. Scala Storage systems is designed for not only read or write-intensive applications, but also metadata-intensive applications.

- **Performance**

  IO pressure increase when application access concurrently to storage system. Scala Storage Scale-Out Solution outperforms SAN and NAS solutions when there are large amount of concurrent access. Scala Storage Scale-Out Solution stripes files across multiple iStore nodes, so that both read and write operations from/to multiple iStore nodes is a parallel process and utilize aggregate bandwidth of multiple iStore nodes. When an application reads one file, it access different stripes of that file concurrently from multiple iStore nodes improving read

performance. This parallel read/write process eliminates the bottleneck formed by single data path in systems like NAS, and meets the bandwidth requirement of concurrent access from multiple application servers.

Scala Storage Scale-Out Solution is also well tuned for small files I/O and large files access. Scala Storage Scale-Out Solution realizes efficient and comprehensive load balancing through data striping distribution within the systems, fully uses the performance of hardware and network, and ensures the highest IO throughput possible.

- **Global Namespace**

Global namespace allow all application server access the same directories and folders. Scala Storage Scale-Out Solution provides a global namespace for all the application servers for easy management. The unified management of data management simplify application development and provides effortless management.

- **Large Directory Support**

The global names space allow all data stores into the same locations, but traditional storage solution can only host limited number of files under single namespace and single directory. When the number of files in the storage system reaches a certain limit, the file system performance will drop dramatically.

ScalaFS large directory support software breaks the limit on number of files under single directory. It outperforms other competing file system, especially when dealing with massive amount of small files. In one of our production clusters, there are tens of millions of photos under one directory, but the file name lookup is still fast enough to meet the requirement of concurrent access.

- **Scalability**

Scala Storage Scale-Out Solution supports dynamic expansion of storage capacity without interrupting operation or downtime. Storage systems expansion is expanded in just a few clicks to include a new iStore nodes into a running systems, while application server is running. It is completely transparent to application servers and there is no downtime involves. The aggregated

throughput of the whole cluster will increase linearly. Metadata controller can be expanded in the same way. The current system limit is 128 metadata controllers and no set limited on iStore node.

- **Load Balancing**

  Most of the current storage architecture do not have the capability to expand storage capacity dynamically and utilize all the storage devices evenly. Some of them can add nodes online but administrators have to copy data manually to get balanced disk utilization and specify how much load each server should handle. Unbalanced disk utilization will be a performance bottleneck and a management burden for administrators.

  ScalaFS load balancing process is specially designed for systems growing over time. After adding new iStore nodes to a running cluster, administrator could use ScalaManager to schedule a time load balance all iStore nodes. The load balance process will automatically redistribute data across all iStore nodes until disk utilization on each iStore nodes is relatively equal. By doing so, the Scala Storage systems can utilize all the possible resources and aggregate throughput, allowing growing linearly when iStore nodes are added.

Data Security

- **Data Redundancy**

  Redundancy is an efficient way to protect data. Users can set number of replicas (1-4) based on file, directory. By replicating data across multiple iStore nodes, it prevents failures of disk, network and server. In addition, availability of multiple file copies improves the performance of concurrent access.

- **Preventive Failure detection and Auto Recovery**

  Preventive failure detecting and self-healing is crucial to a large storage systems. Scala Storage Scale-Out Solution includes an automated failure detecting and intelligent data recovery to provide pre-emptive care of the system. The system can automatically detect failures of disk, server and network. Once failure is detected, the system will automatically start recovery process to create new copies of data on other servers so as to ensure data availability. The whole

recovery process is transparent to applications and will not interrupt the user operations. Moreover, the recovery itself is a parallel process and based on actual data on disk (compare to RAID which has to recover the whole disk). Recovering 1TB of hard drives takes 1/5 of RAID traditional RAID controller.

System Management

- **Fast deployment and Easy Management**

  Scala Storage Scale-Out Solution is easy to deploy and does not requires a lot of complex configurations like SAN. ScalaManager can guide system engineers to configure configuration, start services to application server, add additional storage, all in a single interface. Deploying 200 TB of storage and thousands of application only take hours instead of days.

  Daily management of the Scala Storage Scale-Out Storage is mange via the web based GUI interface. ScalaManager provide a dashboard for cluster monitoring and a smart alert system. Daily performance and statistic are recorded and monitor in the dashboard. Information about the system health can be export to different format for review. When failure occurs, the systems will go into a self-healing process automatically and send alert to administrators at the same time. Administrators can replace the failed hardware later. Only one storage administrator is required to manage over hundreds of PBs of data.

- **Compatibility with Current Environments**

  Scala Storage Scale-Out Solution is designed to fit most computers environments, using 10Gb Ethernet network as data transferring channel to ensure the full utilization the performance of iStore nodes and Metadata controllers. Application servers is support in Gigabit Ethernet or 10Gb Ethernet network. Linux, Mac and Windows is supported concurrently with single namespace. There is no need to make any modifications to operating systems and can seamlessly work with the current software and hardware environment.

Cost Effective Scale-Out Solutions

- **Low System Cost-of-ownership**

Scala Storage Scale-Out Solution can largely reduce the cost of purchasing, operating and maintaining system.

- Scale-out Structures:
  Scala Storage pay as you growth allow you to start as small as 144TB and purchase additional iStore node when your capacity and performance requirements change.

- Always Online Storage:
  Scala Storage is a high availability storage solutions without single point of failure. Preventive failure detection and auto self-healing to maximized the uptime. Online expansion to increase storage and scheduled load-balance maintains the best performance with the systems.

Chapter 6 Summary

The explosive growth of unstructured data and concurrent performance has posed higher request to storage system including IO bandwidth, PB level storage capacity, tens of billions of files, tens of millions of directories, transparent access and data sharing, high security, intelligent management, fine expansibility, so on and so forth. All these demands can be barely met by the traditional storage technology.

Scala Storage Scale-Out Solution provide Scale-Out storage solutions, which provides high IO performance and expansibility, in an easy to management storage systems. Easy operation, uninterrupted service and maintain the systems at highest performance can be provided with Scala Storage Scale-Out solution. While satisfying users' demand, the system effectively reduces the possessing cost and managing cost, enabling users to build information infrastructure economically and effectively improve your core competency.

If you want to learn more about Scala Storage and our solutions, please contact us by phone or email.