



Storage Multi-Tenancy for Cloud Computing

Paul Feresten, NetApp;
SNIA Cloud Storage Initiative Member

March, 2010



Table of Contents

- Introduction..... 1**
- What is Multi-Tenancy?.....2**
 - Storage Multi-Tenancy2
- Enabling Cloud-Based Data Management — CDMI3**
- Virtual Storage Containers.....5**
 - Security.....6
 - Performance and Quality of Service7
 - Data Protection and Availability9
 - Manageability10
- End-to-End Multi-tenancy11**
- Conclusion12**
- About the CSI & CDMI.....12**
- About the SNIA12**

List of Figures

- Figure 1: Shared infrastructure1
- Figure 2: Cloud Data Management Interface (CDMI).....4
- Figure 3: Attributes of a virtual storage container6
- Figure 4: End-to-end multi-tenancy11



Introduction

Organizations of all types are struggling to control costs while facing increasing demands created by explosive data growth and ever-changing regulations. To address these challenges, storage industry professionals are turning to cloud computing and cloud storage solutions.

Cloud computing is not in itself a new technology; it is a new business model wrapped around a set of technologies—such as server virtualization—that reduce the cost of using information technology resources. Cloud computing takes advantage of Web based mechanisms that allow scalable, virtualized IT resources to be provided as a service over a network. The advantages of cloud storage and other cloud services include pay as you go, the perception of infinite capacity (elasticity), and the simplicity of use/management.

When virtualized storage is available on demand over a network, an organization is freed from the need to purchase—or often even to provision—storage capacity before storing data. Significant cost-savings result because organizations typically only pay for storage actually consumed.

Despite the potential advantages, however, many organizations hesitate to expose potentially sensitive data to cloud computing or commit such data to cloud storage because of concerns about security in cloud environments where infrastructure elements—servers, networks and storage—may be shared among many different organizations (Figure 1 below). The high utilization that results from sharing, however, is in large part what makes the economics of cloud computing compelling.

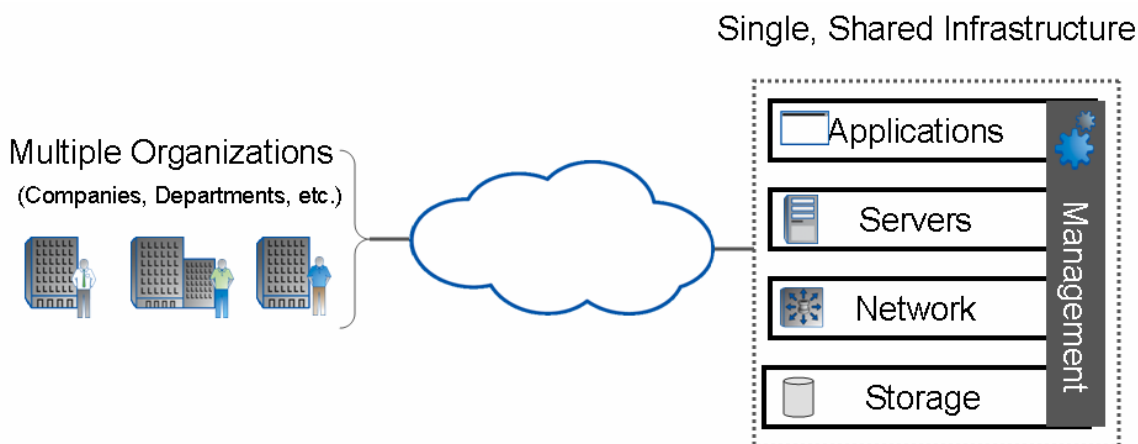


Figure 1: Shared infrastructure



This paper defines a set of requirements for storage in such “multi-tenant” cloud environments in four key areas of concern for cloud adopters: security; quality of service; data protection and availability; and manageability. Where appropriate, these requirements are defined in terms of the Cloud Data Management Interface (CDMI).

What is Multi-Tenancy?

The terms **multi-tenant** and **multi-tenancy** are not new; both have been used to describe application architectures designed to support multiple users or “tenants” for many years. With the advent of cloud computing, this terminology has simply been extended to include any cloud architecture—or infrastructure element within that architecture (application, server, network, storage)—that supports multiple tenants. Tenants could be separate companies, or departments within a company, or even just different applications.

To provide “secure” multi-tenancy and address the concerns of cloud skeptics, a mechanism to enforce separation at one or more layers within the infrastructure is required:

- **Application layer.** A specially written, multi-tenant application or multiple, separate instances of the same application can provide multi-tenancy at this level.
- **Server layer.** Server virtualization and operating systems provide a means of separating tenants and application instances on servers and controlling utilization of and access to server resources.
- **Network Layer.** Various mechanisms, including zoning and VLANs, can be used to enforce network separation. IP security (IPsec) also provides network encryption at the IP layer (application independent) for additional security.
- **Storage Layer.** Mechanisms such as LUN masking and SAN zoning can be used to control storage access. Physical storage partitions segregate and assign resources (CPU, memory, disks, interfaces, etc.) into fixed containers.

Achieving secure multi-tenancy may require the use of one or more mechanisms at each infrastructure layer.

Storage Multi-Tenancy

While mechanisms to support multi-tenancy and enforce separation exist at every infrastructure layer, this paper is primarily concerned with storage and the



requirements for secure and effective storage multi-tenancy in a cloud environment. To understand the full set of storage requirements, it is necessary to consider cloud storage from both the perspective of the tenant (user) and the provider of cloud services.

Cloud computing services can be broken down into a variety of types, ranging from Software as a Service (SaaS)—in which the provider delivers specific application services to each tenant—to Data storage as a Service (DaaS) —which is virtualized storage on demand over a network. Regardless of the type of cloud service, from a tenant perspective there will be specific requirements that apply directly or indirectly to data storage.

Tenant requirements are typically defined in terms of service level agreements (SLAs), which cover a variety of capabilities including:

- **Security**
- **Performance**
- **Data protection and availability**
- **Data management**

From the provider's perspective, multi-tenant storage should provide convenient mechanisms for satisfying these and other tenant SLAs as well as supporting additional capabilities such as:

- **Accounting.** The ability to monitor usage by each tenant for billing or other purposes.
- **Self service.** The ability to allow a tenant to perform a defined set of management tasks on their data and the storage they use, thereby offloading these functions from the provider.
- **Non-disruptive upgrades and repairs.** Downtime in multi-tenant environments may be difficult or impossible to schedule, so maintenance activities must be possible without incurring downtime from the point of view of the tenant.
- **Performance management.** The ability to balance cost and performance as the lifecycle requirements of data changes over time.

Enabling Cloud-Based Data Management — CDMI

Designed to enable multi-tenant storage offerings, the SNIA's Cloud Data Management Interface (CDMI) for cloud storage and data management integrates and is interoperable with various types of client applications. CDMI offers a standard



approach to data portability, compliance and security, as well as the ability to connect one cloud provider to another, enabling compatibility between cloud vendors.

Using this approach, a client will be able to discover the capabilities of cloud storage and use this interface to manage data containers and the data elements that are placed in them. CDMI makes extensive use of metadata to simplify application access and enable multiple levels of service as required by a diverse set of users. The model behind the Cloud Data Management Interface is shown in Figure 2 below.

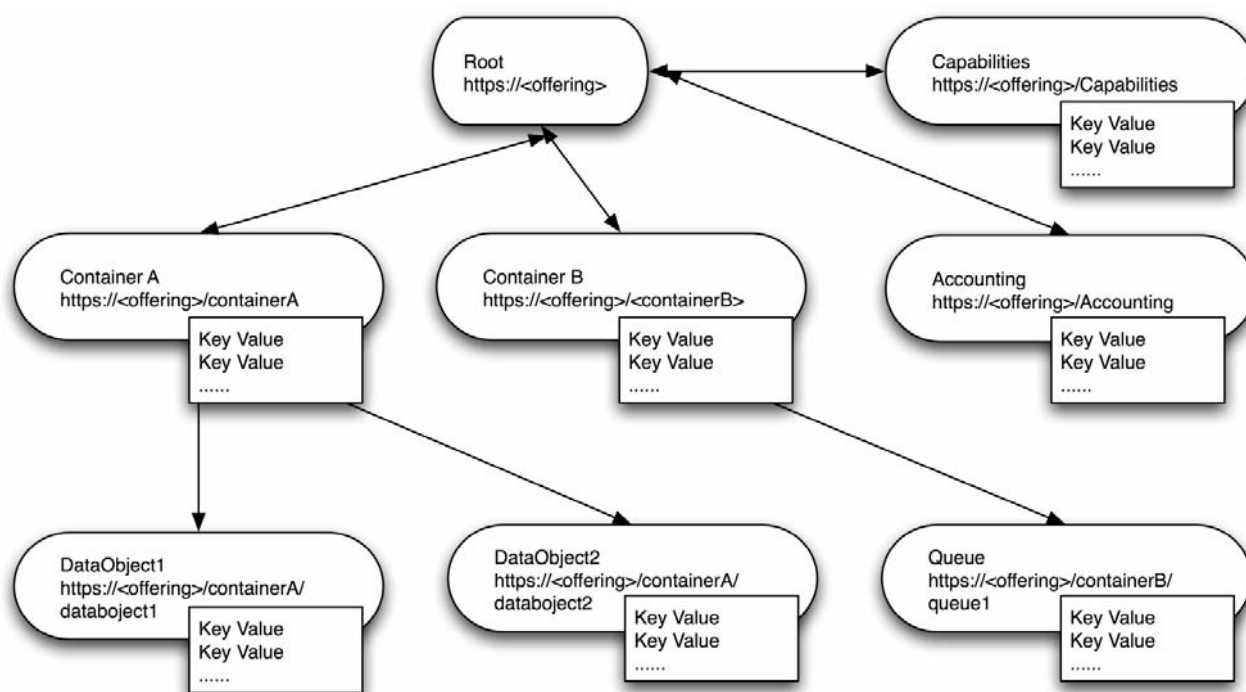


Figure 2: Cloud Data Management Interface (CDMI)

In the storage layer, the CDMI interface can simplify management since data system metadata can be applied to container hierarchies. For the functional data path interface for data storage, CDMI assigns each data object a separate URI (Uniform Resource Identifier). Since objects can be fetched using the standard HTTP protocol employing RESTful (REpresentational State Transfer) operations, each data element can be managed as a separate resource. In this way, it is possible to separate and classify data elements and containers for secure access as well as service levels. The result is a level of isolation suitable to tenant based, on-demand data access.



Virtual Storage Containers

The traditional mechanisms for enforcing storage separation mentioned above—LUN masking, SAN zoning and physical storage partitioning—do not adequately address all the requirements of multi-tenant storage in a cloud environment. These methods are too rigid to deliver the flexibility and high utilization required.

Several storage vendors have independently arrived at the idea of a “virtual storage container” as a way of delivering storage multi-tenancy. Naturally, each vendor uses its own descriptive language, so the term virtual storage container in this context is used as a generic term to allow discussion of the concept in a general way and should not be taken to imply any vendor’s particular implementation.

A virtual storage container is a contained management domain that grants the tenant some or all of the management capabilities of the overlying storage system. (Restricted to the storage available to the tenant, of course) In effect, virtual storage containers provide each tenant one or more “virtual storage arrays”. From the perspective of the storage provider, a virtual storage container is a discrete entity. Virtual storage containers can use CDMI to ensure that metadata is correctly applied in the data hierarchy thus providing a simple and predictable interface for applications and individual tenants. CDMI’s use of metadata can extend down into individual data elements or can apply to the entire virtual storage container. Thus, any data placed into a container essentially inherits the metadata of the container into which it was placed. The extension of metadata to managing virtual storage containers enables a reduction in the number of paradigms for managing the components of storage—a significant cost savings. Providing metadata in a cloud storage interface standard and prescribing how the storage and data system metadata are interpreted to meet the requirements of the data, delivers the simplicity required by the cloud storage paradigm, while still addressing the requirements of enterprise applications and data.

A service provider should be able to perform a variety of management actions on a virtual storage container as a whole including the ability to create and destroy containers or apply specific policies on behalf of the tenant.

A virtual storage container is in many ways analogous to the virtual machines of server virtualization, in that resources can be dynamically shifted between virtual storage containers. This paper is more concerned with the capabilities that multi-tenant



storage and virtual storage containers should deliver—for both tenant and provider—than it is with implementation specifics.

Virtual storage containers are aligned with individual tenants through unique identifiers (object identifiers) and are required to also encapsulate security attributes that prevent unauthorized access. Figure 3 illustrates these basic attributes.

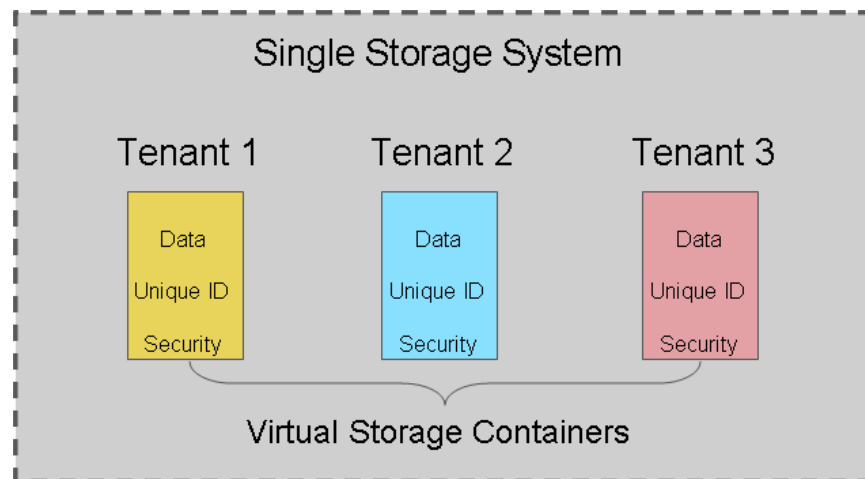


Figure 3: Attributes of a virtual storage container

Security

As we saw in the introduction, the first requirement for multi-tenant storage is to ensure the security of tenant data. A virtual storage container must protect the tenant data it contains from multiple classes of threats, including:

- **Snooping.** One tenant should not be able to gain unauthorized access to another tenant's data. A tenant must be restricted to their own virtual storage containers such that symbolic links or other possible mechanisms for connecting to storage outside the tenant's domain are secured.
- **Unauthorized Discovery.** Virtual storage containers should be invisible to everyone except their owners. Only authorized tenants should be aware of the existence of their associated virtual storage containers.
- **Spoofing.** Authentication mechanisms must ensure that no one can assume a tenant's identity to gain data access.
- **Deletion.** (Accidental or malicious.) No action external to the virtual storage container should cause tenant data within the container to be deleted or corrupted.



- **Denial of service.** Tenant data access must not be disrupted by direct denial of service attacks against the virtual storage container, the normal activities of other tenants, or abnormal tenant activities such as an application error that creates an I/O storm in another tenant's virtual storage container. (This is discussed further in the following section on Performance and Quality of Service.)

Multi-tenant security is achieved by isolating one tenant's virtual storage containers from another tenant's. This can be achieved in part by providing separate IP addresses for each storage container and binding a tenant's data containers to VLANs to ensure separation of network traffic. Careful tenant authentication is necessary to ensure security against possible malicious attacks. Encryption of data as it is stored on the underlying media may also be provided as an option to meet the security concerns of the most sensitive tenants.

In each of the above categories, the CDMI interface is able to provide a standard and interoperable suite of protective measures (e.g. user authentication, access control, data retention, encryption) that can be implemented as optional components of the overall security solution. In this case, the mandatory transport mechanism will be HTTP (TLS) and each CDMI implementation will function as the authentication vehicle.

In addition, CDMI supports the use of secure storage protocols to be used in a cloud-computing infrastructure. If using an iSCSI protocol, for instance, CHAP and IPsec can be used to secure the storage connections from the server. For Fibre Channel protocols, DH-CHAP, FC-SP, LUN masking and secure fabric zoning can accomplish similar security features. For file protocols such as NFS, Kerberos can be used to secure the storage network. CDMI thus allows secure access to storage no matter the data path used to access that storage.

Performance and Quality of Service

The second biggest concern with cloud storage after security is quality of service. Concerns about performance and performance consistency in multi-tenant environments may cause many potential purchasers of cloud services to hesitate.

From a storage perspective, a cloud service provider must be able to ensure that storage I/O doesn't become a bottleneck, preventing the provider from meeting tenant SLAs. To do this, a service provider must be able to offer different classes of service and be able to ensure that the storage infrastructure supports delivery of each



class of service; lower classes of service must not interfere with delivery of higher classes of service.

For example, a cloud storage service provider might offer four classes of service:

- Platinum: highest storage performance
- Gold: high storage performance
- Silver: intermediate storage performance
- Bronze: low-performance, high-capacity storage for archival

To accomplish this, the provider would have to ensure that a Platinum tenant received a Platinum virtual storage container capable of delivering the highest storage performance. This probably implies a virtual storage container on a fast controller, utilizing high performance disks (incl. solid-state storage) and an adequate number of spindles.

More important, the service provider needs a mechanism to ensure that I/O traffic to and from Gold, Silver, and Bronze virtual storage containers does not interfere with Platinum-level traffic. This might be accomplished in two ways:

- Assign more resources (memory, cache, CPU and interface bandwidth) to higher-priority virtual storage containers.
- Provide a mechanism to set the priority of the I/O transactions of each individual virtual storage container.

Accommodating Tenants That Require Multiple Classes of Service. Some tenants will require multiple classes of storage service to meet different needs. For instance, a tenant might need high performance storage for production applications and high-capacity storage for archiving. This could also be accommodated in one of two ways:

- A tenant's single virtual storage container provides multiple classes of service.
- A tenant receives multiple virtual storage containers, each delivering one class of service. In this case, a mechanism may be needed to federate multiple virtual storage containers into a single management view.

CDMI is able to simplify the provisioning of service class based on its ability to extend metadata to virtual storage containers as well as individual data elements. Once metadata settings are established for a specific container, for example, service-level parameters are automatically extended to any file, LUN, or object placed in the container, thus ensuring consistent tenant-level performance. CDMI allows the cloud storage provider to advertise different types of containers with corresponding



metadata values in an interoperable manner. A client can compare the offerings of Platinum containers between different cloud providers.

Scaling Performance. For some tenants, performance and capacity may need to scale rapidly. Accommodating the needs of such tenants suggests two additional requirements for virtual storage containers:

- **Non-disruptive migration.** Virtual storage containers may need to be moved from one storage system to another to accommodate a tenant's growing need for performance, capacity or both.
- **Scale out.** The ability to spread the access to a single data object across multiple physical storage systems would make it simpler to meet performance needs, load balancing access across the systems, even geographically.

Data Protection and Availability

Data protection and availability are also naturally a concern for anyone considering cloud service adoption. Well-publicized outages for public cloud services, such as Google's Gmail, have heightened concerns about service availability.

In cloud environments, mechanisms to protect data, ensure data availability, and provide disaster recovery must be closely integrated with storage, such that data is never overlooked and left unprotected. The provisioning process for virtual storage containers should ensure that some default level of data protection is applied to all data within the container.

A virtual storage container should provide convenient mechanisms for either the tenant or the provider (or both) to exercise additional control over data protection and availability functions. For instance, in our previous example of classes of service, each class would include a specified level of availability and a specified level of data protection via the data system metadata on each. Platinum service might include hourly backups plus offsite replication for disaster recovery, while lower classes offer just backup at some specified schedule. Tenants may wish to override the defaults in certain situations to increase the backup frequency or add additional replication after a critical event.

Cloud storage providers can use CDMI's capability tree to define their various levels of data protection and availability in a standard and interoperable manner.



Manageability

Manageability is the final, but certainly not the least, of the considerations for multi-tenant storage. The fear of losing control over data management is certainly among the reasons for hesitancy about cloud services.

Flexible management options. Certain classes of tenant will desire a cloud service where they can manage and monitor data more or less as they would in their own data centers, while others will prefer an environment where data is managed to specified SLAs with little or no tenant involvement.

From the perspective of cloud providers, virtual storage containers should be flexible enough to allow for different levels of management control by tenants. With some cloud services, tenants will want or need no direct control or very limited control while others will want or need full control within the confines created by the virtual container.

Self service. Allowing tenants to perform self-service of ad hoc tasks such as provisioning, data protection and replication, can significantly reduce management overhead for a cloud provider. If the provider's environment is built from multiple vendors' equipment, SMI-S can be utilized under the cloud layer to remove the need to deploy multiple adapters for this self-service management.

Storage efficiency. The ability to utilize storage efficiency technologies such as thin provisioning and deduplication can significantly increase storage utilization. From the provider's perspective, a more efficient service is cheaper to provide and, therefore, more competitive. From the tenant's perspective, these technologies reduce the amount of storage they consume and thus may lower their overall storage bill.

Storage protocol selection. For some cloud services, particularly those offering infrastructure as a service (IaaS), there is a need to offer tenants different storage protocol options including both file-based options (NFS and CIFS) and block-based options (iSCSI or FC SAN). (This is also related to the ability to provide multiple classes of service as discussed above in the section, Performance and Quality of Service.) This could be accommodated either in a single virtual storage container or via multiple virtual storage containers of different types.

CDMI supports all standard block, file and object storage protocols for use as the Functional interface (Data Path). CDMI also works with Cloud Computing to make containers available for use by the virtual machines in that environment.



Upgrades and maintenance. From the perspective of a cloud provider, upgrades and maintenance in a multi-tenant environment become difficult or impossible to accomplish using traditional means because scheduling downtime is impossible in an environment with multiple tenants spread across multiple geographies, and all with different operating schedules. Therefore, it must be possible to perform all upgrade, repair, and maintenance activities in a non-disruptive fashion. It may be necessary to be able to non-disruptively migrate all virtual storage containers off of a particular storage system, so such work can be performed.

CDMI's provides a rich and flexible container model that covers common management aspects such as allocation and monitoring of storage. This enables Cloud storage providers to advertise their adherence to standard levels of efficiency and manageability, while supporting custom vendor extensions that can be used for differentiation or specialization.

End-to-End Multi-tenancy

It should be clear that provisions made at any infrastructure layer for security, quality of service, availability, or manageability cannot ensure those attributes in other layers. For example, to provide complete security in a cloud providing infrastructure as a service (IaaS), data must not only be protected from inappropriate access of storage as described above, but must also be protected as it traverses storage networks to servers, as it resides in server memory, and as it traverses the network to the tenant. A secure end-to-end "lane" from user to data that offers secure multi-tenancy at each layer is needed. This is illustrated in Figure 4. Similarly, appropriate measures are needed at each layer to ensure quality of service, availability, and manageability. While these measures are beyond the scope of this paper, it may be necessary or desirable for each layer to communicate with every other layer to ensure appropriate levels of service, particularly with regard to security and quality of service.

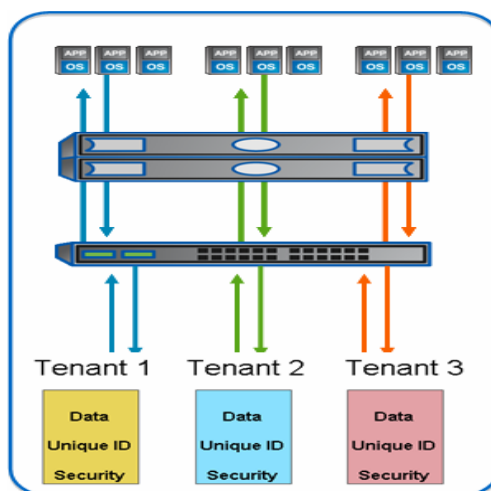


Figure 4: End-to-end multi-tenancy



Conclusion

The cloud creates unique requirements for data in terms of security, performance, data protection, availability, and manageability. To dispel the concerns of many potential cloud adopters, these requirements must be addressed in a systematic way, and the concept of a virtual storage container provides a useful construct for thinking about how to meet these requirements.

CDMI now provides an approved storage industry standard to richly define the properties and capabilities of such a virtual storage container. CDMI also defines management interfaces to efficiently operate it across one or more Cloud storage providers.

By addressing the requirements, a storage vendor or cloud provider will be able to create a multi-tenant storage infrastructure that is secure, flexible, highly functional and interoperable.

About the CSI & CDMI

The SNIA Cloud Storage Initiative (CSI) was created to foster the growth and success of the market for cloud storage. Members of the SNIA CSI work together to educate the vendor and user communities about cloud storage, perform market outreach that highlights the virtues of cloud storage, collaborate with other industry associations on cloud storage technical work, and coordinate with SNIA Regional Affiliates to ensure that the results of CSI activities are felt worldwide. The CSI, along with 140 individuals from more than 30 organizations, promotes the adoption of standardization through the Cloud Data Management Interface (CDMI) standard specification. For more information or to get involved, visit the SNIA CSI website at www.snia.org/cloud.

About the SNIA

The Storage Networking Industry Association (SNIA) is a not-for-profit global organization, made up of some 400 member companies spanning virtually the entire storage industry. SNIA's mission is to lead the storage industry worldwide in developing and promoting standards, technologies, and educational services to empower organizations in the management of information. To this end, the SNIA is uniquely committed to delivering standards, education, and services that will propel open storage networking solutions into the broader market. For additional information, visit the SNIA web site at www.snia.org.