SCALE COMPUTING

Storage Buyers Guide

Featuring:

ICOS TECHNOLOGY

Table of Contents

Shared Storage Buyer's Guide
for the
Small to Medium-Sized Enterprise

Introduction

An organization's dependence on data cannot be separated from its dependence on storage. For this reason, the need for shared storage may be evident, but the right storage can still be difficult to sort out. The cost of buying, owning and maintaining storage concerns businesses of any size, and in particular small to medium-sized enterprises (SMEs) operating on tight budgets. Quite simply, the right storage for your company meets your present needs, is scalable for the future and includes enough features to support the continuous access, availability, and protection of critical business information.

This guide describes the risks of keeping data isolated on individual servers or maintaining too many storage systems and covers such topics as the types of disks available, SAN and NAS differences, and storage system features.

Disadvantages of Direct Attached Storage and Multiple Storage Systems

Direct attached storage (DAS) is sometimes referred to as "islands" of information, and that description alone may be enough to understand the types of limitations encountered as the amount of data grows. Although initially it can be simple and inexpensive, storage isolated on multiple servers or storage systems has limitations that you'll want to consider as your environment grows.

- **Over-Purchasing:** Some drives may be full and others practically empty but there's no efficient way for an administrator to share capacity between servers. Low overall utilization means you pay more overall for storage because you can't fully utilize the resources you already own.

- **Availability:** Data attached directly to the host creates a dependency between the two that increases the complexity of making the application and data available. If the server fails, the data is trapped on a dead server until it can be moved or recovered from backup to a working server.

- **High administration costs:** Storage for each server must be administered separately. Efforts to keep data available on each server are multiplied out across the environment. Other tasks like indexing, archiving, replicating, or analyzing data have to be performed across multiple locations, driving up management costs and efforts.

The Case for Shared Storage

Purchasing a storage system may appear to be more expensive than simply placing disk drives in servers; but in reality, the benefits of a shared storage system outweigh the initial costs.

- **Higher utilization:** Hard drive capacities on shared storage systems are aggregated together into one big storage pool that can be provisioned out as needed. With the addition of features like thin provisioning that allow provisioning of capacity without allocation of unused blocks, a shared storage system can reach a much higher utilization percentage before additional capacity must be purchased.

  In the example below, a server in figure "A" requires additional capacity; therefore a new hard drive must be attached, driving the overall storage utilization down to 40 percent. With a shared storage system, the additional capacity can be

The Case for Shared Storage - (Cont.)

provisioned from the remaining pool for an overall higher utilization of resources.

Fig.A

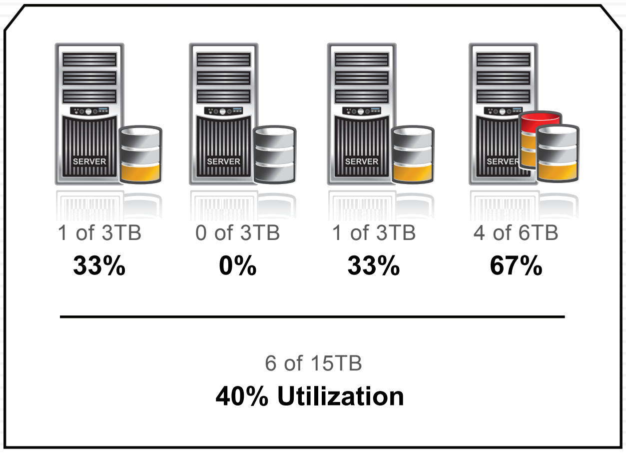
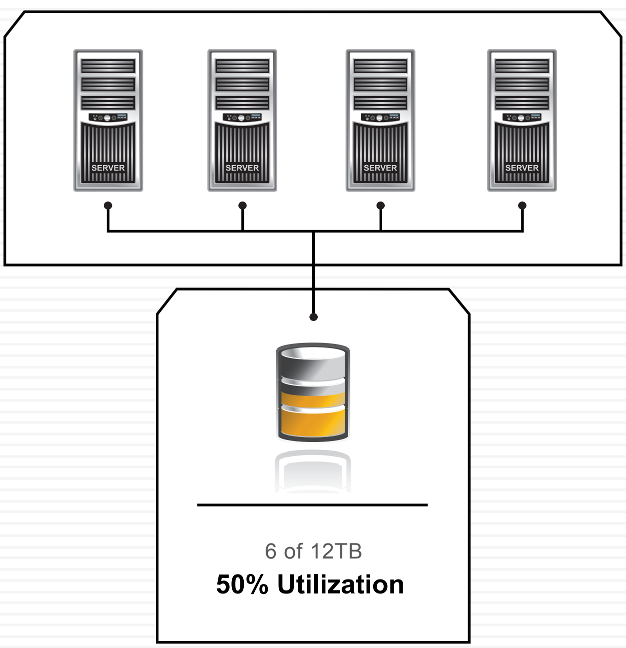


Fig.B



Disk utilization in a 6TB environment

- **High availability:** Shared storage systems integrate data redundancy to ensure data is highly available, reducing the cost impact of downtime due to failed disks. Scale-out storage systems even provide redundancy across physical storage chassis for the highest level of redundancy.

- **Easier disaster recovery:** Having all data in a single storage pool makes identification and management of protecting that data offsite for disaster recovery purposes easier.

- **Streamlined management:** Shared storage allows you to provision manage and monitor storage from a single point, lowering administrative costs and frustrations.

- **Access to advanced server cluster and virtualization features:** In addition to reduced power and cooling consumption in virtual environments, virtualization platforms offer high availability and disaster recovery features that require a compatible shared storage system.

Types of Storage: SATA, SAS and SSD Storage

While magnetic tape mediums may still be preferred for long-term offline storage, the vast majority of data is stored on hard disk drives (HDDs) that offer large capacities and fast random access to data. Not every HDD has identical characteristics. In addition to capacity variations, different methods of interfacing HDDs with computers have evolved to solve slightly different problems. Storage systems can support one or multiple types of disk storage, so understand what type of HDD fits your need and ensure the storage system of your choice supports it.

Serial Attached SCSI (SAS) drives offer higher platter rotational speed (15,000 rpm) and generally last longer than competing Serial Advanced Technology Attachment (SATA) drives; however, SATA drives offer larger capacities at a lower cost than SAS drives. SATA drives were once only used in the consumer market while SAS drives were applied to more demanding 24/7 business applications; but, when combined with RAID schemes to create redundant copies of data, SATA drives with

Types of Storage: SATA, SAS and SSD Storage - (Cont.)

enhanced reliability referred to as "Enterprise SATA" are now often used in storage systems designed for small to medium enterprises at lower prices than SAS-based systems. SAS drives are also available in slower 7200 rpm formats referred to as near-line SAS (NL-SAS) drives that offer capacity per dollar ratios closer to enterprise SATA drives.

Some storage systems now also support SSD (Solid-State Drive) devices, which, as the name implies, use solid-state memory rather than spinning magnetic disks to store data. SSDs support the same SATA or SAS protocols and match the same physical dimensions of HDDs, allowing storage systems to integrate SSDs without extensive hardware or software modifications.

SSDs retain data in flash memory. Because there are no moving parts they are also quieter and less susceptible to damage than HDDs. For random access I/O, SSDs are significantly faster than HDDs, however for large, highly sequential I/O, SSDs lose their significant speed advantage. While SSDs offer many benefits, they are also more expensive per GB and not yet available in the larger capacity choices of HDDs. SSD technology continues to evolve quickly, so prices will continue to decrease while capacities rise.

Connecting Hosts to Storage

Network Attached Storage (NAS) Devices

A storage system that makes files available to users over the network using a file sharing protocol is called Network Attached Storage (NAS). File sharing protocols were created to allow one or more computers to access files on demand from another computer. File sharing protocols such as NFS for Linux, CIFS for Windows and AFP for Apple exist to support file sharing on any operating system. The practice of assigning a general-purpose computer the sole duty of storing and sharing files for everyone made it easier for people to find information and for administrators to manage it. Eventually NAS devices evolved to serve that role more efficiently than a general-purpose computer.

The Storage Area Network (SAN)

Most enterprise applications do not support storing transactional application data on network file shares. File sharing protocols do not offer the same level of granular block-level access to files that is possible with storage attached to the server. A Storage Area Network (SAN) is a storage solution that allows consolidated data to be accessed by remote servers as if the data were local to the server. With a SAN, applications running on the server can access data over the SAN as if it was a physical hard disk drive installed directly on the server.

The applications for SAN and NAS devices are usually distinct. NAS devices generally host files for end users on shares they can browse over the network. SANs deliver storage to servers that are often used to store data from enterprise applications like e-mail. This too is changing. Enterprise virtualization platforms like VMware vSphere support both SAN and NAS protocols to store virtual disk files.

SANs normally leverage a dedicated network infrastructure that is separate from the public network. This network may use equipment and protocols different than the public TCP/IP network, as is the case with Fibre Channel SANs. A SAN may also use cabling, NICs, and switches identical to the public TCP/IP network, as in the case of iSCSI SANs that support the SCSI protocol over Ethernet.

Fibre Channel networks have continually increased in speeds and are now capable of greater than 8 Gb/s bandwidth, but only large enterprises are normally able to afford the more expensive and dedicated network infrastructure required for a Fibre Channel SAN.

For many enterprises and almost all small to medium-sized ones, iSCSI is now the dominant SAN architecture. Over 30% of mid-market and enterprise companies already use the iSCSI storage protocol in their environment. Increasing support for 10GbE network hardware is also opening the door to even greater iSCSI storage system speeds.

The advantages of using an Ethernet network are clearly

The Storage Area Network (SAN) - (Cont.)

preferred by the market and have prompted the introduction of other storage protocols over Ethernet. Fibre Channel over Ethernet (FCoE) and ATA over Ethernet (AoE) both use Ethernet cable, but differ from iSCSI in that they do not use IP, and therefore cannot use IP routing for remote configurations.

Features of storage systems

Scalability

Eventually a storage system will run out of capacity or reach a performance limit. What happens after that depends on whether the storage system scales up or scales out. When a storage system scales up capacity can normally be added to the device, and in some cases, components of performance—such as extra I/O ports— can be added. But, once the expansion ports are full, that system cannot be expanded any further. A larger system can be purchased and data migrated to the new system in a process typically described as a forklift upgrade. Alternately, a second identical system with a separate pool of storage must be managed, compromising the original benefits of storage consolidation.

A storage system that scales out can manage a single pool of storage across multiple devices. This avoids the large financial impact of a forklift upgrade, and in the case of some scale-out storage systems, adding more devices can increase the I/O performance of the overall storage system.

RAID Data Protection

HDDs are comprised of spinning magnetic disks with moving heads that face eventual mechanical failure. Even SSDs with no moving parts have a lifespan measured by the number of times data can be written to the microchip.

RAID, or Redundant Array of Independent Disks, is a scheme to avoid data loss due to disk failure by storing data redundantly within a set of drives. Because RAID can seamlessly continue delivering data during failures, it also enables high availability in disk storage systems.

RAID arrays combine multiple disk drives into one set that can be accessed by the operating system so data can be distributed across the set using a particular RAID implementation. RAID has different implementations defined as "levels", each offering a different level of fault tolerance and performance.

In the event of a failure, the RAID controller will take recovery steps to rebuild the data from the failed drive onto a new drive that was defined as a 'hot standby' or was manually identified by the administrator. Not every storage system supports every RAID level, so determine the right level of data redundancy and fault tolerance needed when considering a storage system.

Summary information on the most common RAID levels is provided below.

- **Mirroring (RAID 1)**
  Data on any given drive is fully mirrored block-for-block to a second drive. As long as drives containing both copies do not fail at the same time, no data is lost.

- **Striping with distributed parity (RAID 5)**
  RAID 10, or RAID 1+0, is referred to as a "stripe of mirrors." Multiple mirrored sets of disks are configured and then data is striped across the sets. Since data is striped across multiple disks and can be read back using multiple disks simultaneously, it is faster than RAID 0 in reads and writes. There is less performance impact during disk failure events than RAID 5 or 6 because data is read from a mirror copy rather than being reconstructed from parity files.

Hardware Redundancy Features

Hardware redundancy features prevent downtime caused by hardware failures. Imminent failures are detected before or immediately after the event and operations transition to standby hardware. Storage systems that provide hardware fault tolerance generally define their implementation as active-active or active-passive. Active-active configurations mean that both pieces of hardware are being used, but one can

[1] *ESG Research Brief: iSCSI SAN Adoption Update 2010, Enterprise Strategy Group*

Hardware Redundancy Features - (Cont.)

still failover for and assume the duties of the other. Active-passive configurations keep standby hardware idle until the active hardware fails. Each configuration is a trade-off. Active-active systems have higher utilization of resources, but may deliver reduced performance in a failure scenario as the remaining hardware takes on substantial additional load. Active-passive systems maintain performance in a failure scenario, but during normal operations the standby hardware is not utilized, essentially resulting in double the hardware costs for an insurance policy.

Another alternative to active-active or active-passive configurations are scale-out storage systems designed in clustered architectures that provide Active-active benefits across a number of storage devices. This type of architecture fully leverages the hardware in the cluster for maximum performance, but also removes the risk of overloading a single standby piece of hardware in the event of a failure by distributing the work of the failed node across the entire cluster.

Thin Provisioning

Thin provisioning is a critical feature in increasing the utilization of disk capacity in a storage system. Thin provisioning lets the administrator provision more storage to clients than actually exists on the system. For example, while the client may have an entire 2TB of capacity provisioned, if the client is only storing 1TB of data, the unused 1TB of underlying space is available for any other client to use. As various clients actually need more of their provisioned space, the administrator can add more capacity without the need to create new storage allotments. Thin provisioning optimizes available resources and reduces storage management effort by allowing you to buy less storage capacity up front and then fully utilize that capacity.

Snapshot Support

A storage system with RAID data redundancy can protect against disk drive failure, but it cannot protect against human error. Snapshot features built into the storage can store previous points in time for easy recovery in the case of data corruption by human or software error.

There are multiple methods for taking snapshots of data. Copy-on-write methodologies move the old version of data to a snapshot location when new data is written to the original location. The redirect-on-write method of snapshots writes new changes to the snapshot location. Copy-on-write snapshots must perform two write operations for every new data write (once to move the old data, and once to write the new data). Redirect-on-write snapshots do not incur any performance penalty during data writes, but must merge snapshot data into the production data as older snapshots are pruned. Copy-on-write snapshots are more easily managed and can be pruned by simply deleting the oldest snapshot data.

A storage system likely only supports one type of snapshot creation, so understand what type of snapshot is supported by the system and plan for the extra consumed capacity of snapshots, which may start out small, but grow large over time.

Replication

Even when data is protected with snapshots on a highly available and redundant storage system, that system is still at risk from disaster on a geographical scale. A replication feature on the storage system can simplify the task of protecting large amounts of data to a geographically distant location.

Replication must work over limited amounts of bandwidth, so carefully analyze how much data must be replicated, how much bandwidth is available, and any capabilities of the replication that can be used to conform the replication load to the available bandwidth. Replicating only the changed data at the block or byte level will reduce the amount of data that must be transmitted. Transmission scheduling can also restrict replication traffic to off-peak network times. For challenging bandwidth scenarios, compression or deduplication features may be required to successfully replicate. While these features can be helpful when bundled into the storage, they are more beneficial when implemented outside the storage for all WAN traffic as part of a WAN optimization solution that allows all consumers of the bandwidth to benefit.

Data Tiering

The value of data, measured by how frequently it is accessed, changes over time. Storage systems that offer differing levels of performance within the system may provide features that allow the movement of data between those performance levels, either automatically or manually.

This movement of data allows the storage system to address data hot spots and load balance overall performance by placing popular data on the fastest tier of storage.

For storage systems that distribute data fully across all disks in the system and thus only have one tier of performance, this type of load balancing for performance is not necessary.

Deduplication

Data deduplication eliminates physical instances of redundant data while maintaining logical representation of individual instances. This improves storage utilization in a similar way to compression at the expense of extra resources to identify and catalog duplicate content.

Multiple methods of deduplication exist. Sub-file deduplication can identify parts of files that are identical, even if they are in different locations within the file. Sub-file deduplication realizes a greater reduction in data size at the expense of extra overhead to identify duplicate content.

Single file instancing looks for identical files. If some blocks of data differ, they will not be processed. This method of deduplication cannot find as much duplicate data, but may require less overhead.

Management

Storage designed for small to medium enterprises must be manageable without specialized training and should provide an intuitive graphical interface to configure and manage the system. The system should allow for simple installation and storage configuration. It should also provide monitoring and management tools that have the ability to identify and correct inefficient or incorrect configurations.

Summary

Small and medium enterprises (SMEs) have smaller storage needs than large enterprises but still require key features for availability and protection of their critical business information. Stripped down systems from large enterprise vendors may not represent the needs of the SME and are often missing essential features that cost extra to add back into the system. The best storage system for SMEs is one where everything is in the box and the storage is supported by an organization experienced with SME environments.

For a SME that doesn't maintain a storage administrator, your storage vendor's technical support immediately becomes a trusted (or mistrusted) member of your storage team. They should have a working knowledge of your business and your storage needs, be available and responsive and have a high customer service satisfaction rating.