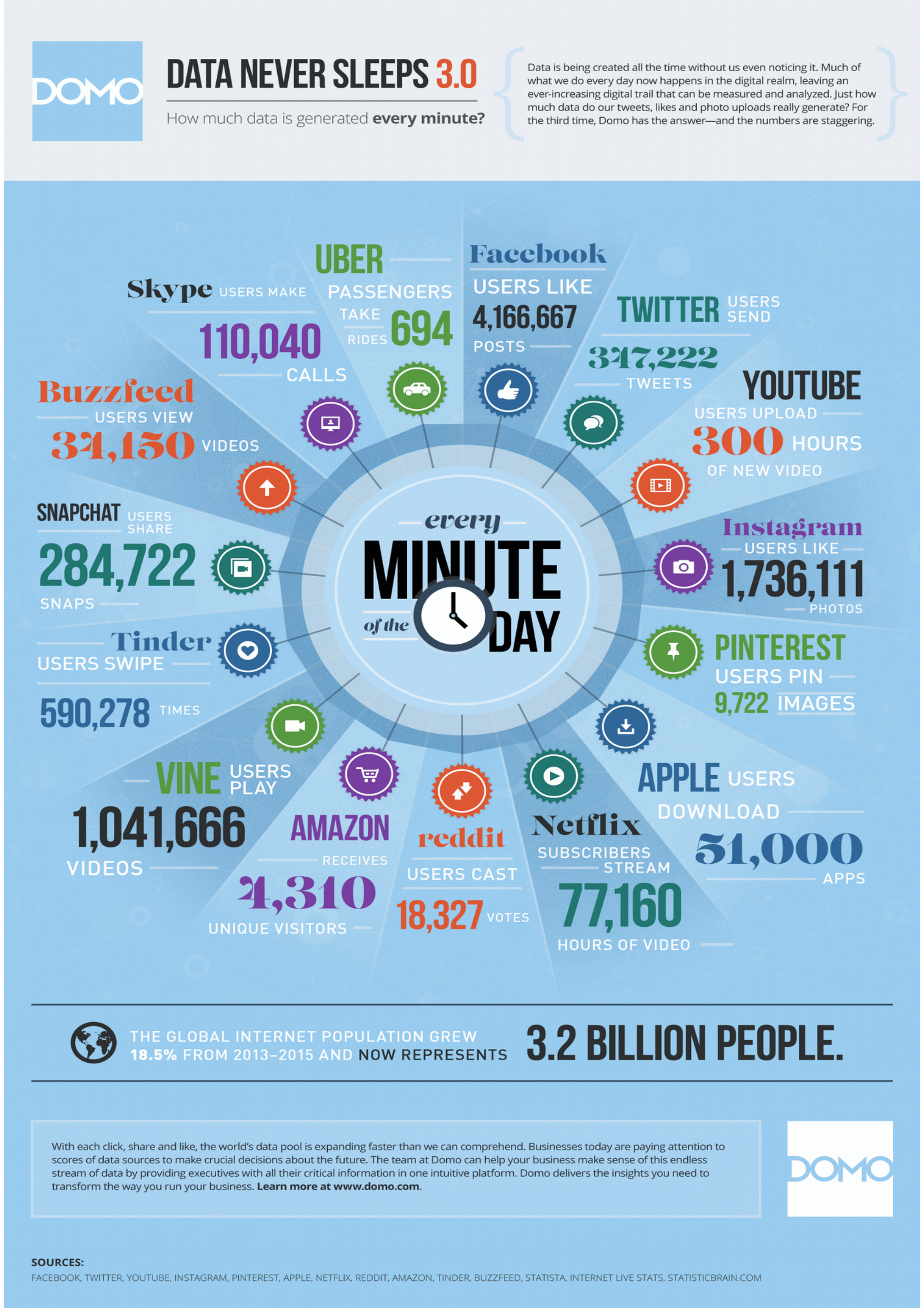
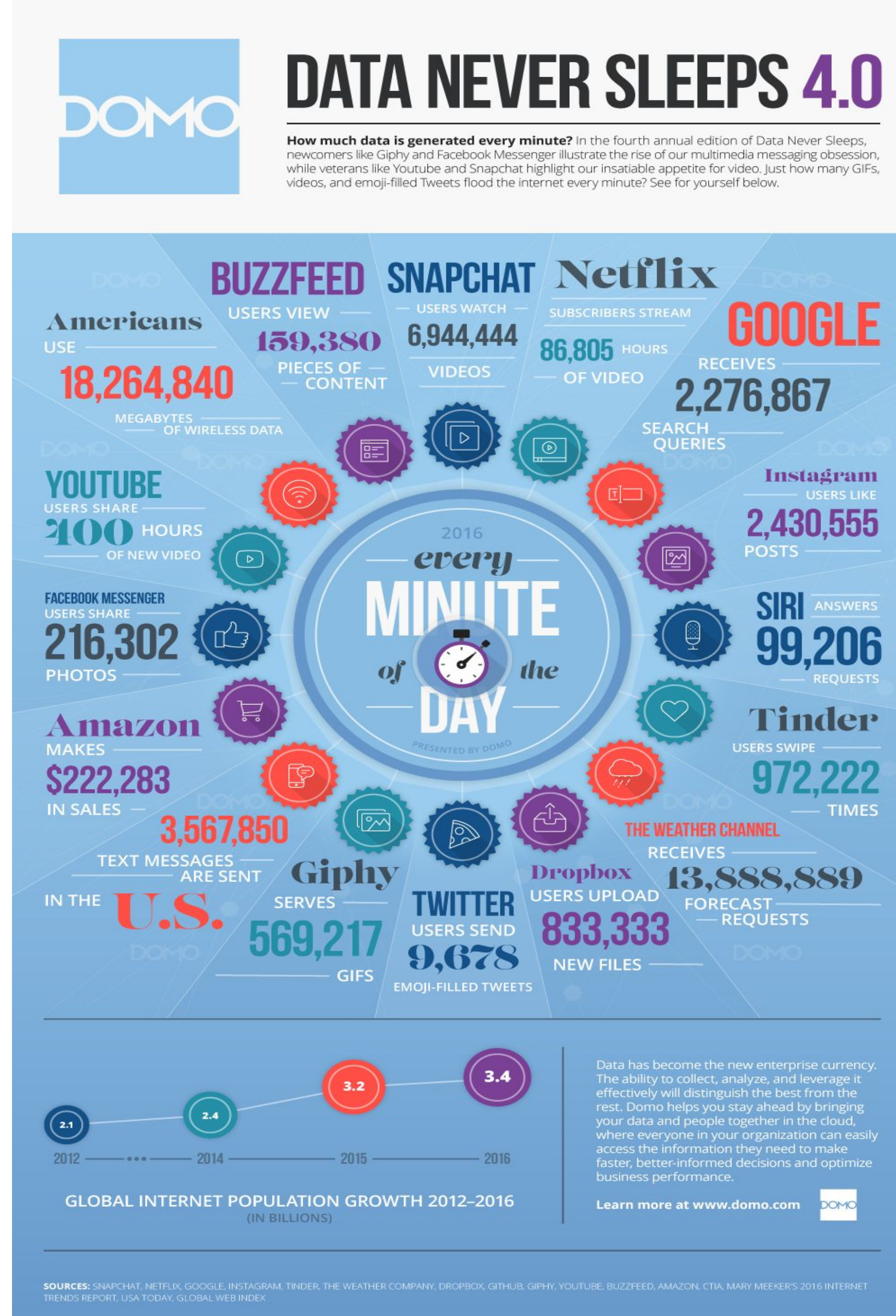
Storage Area Networks
Network Attached Storage

Gerardo

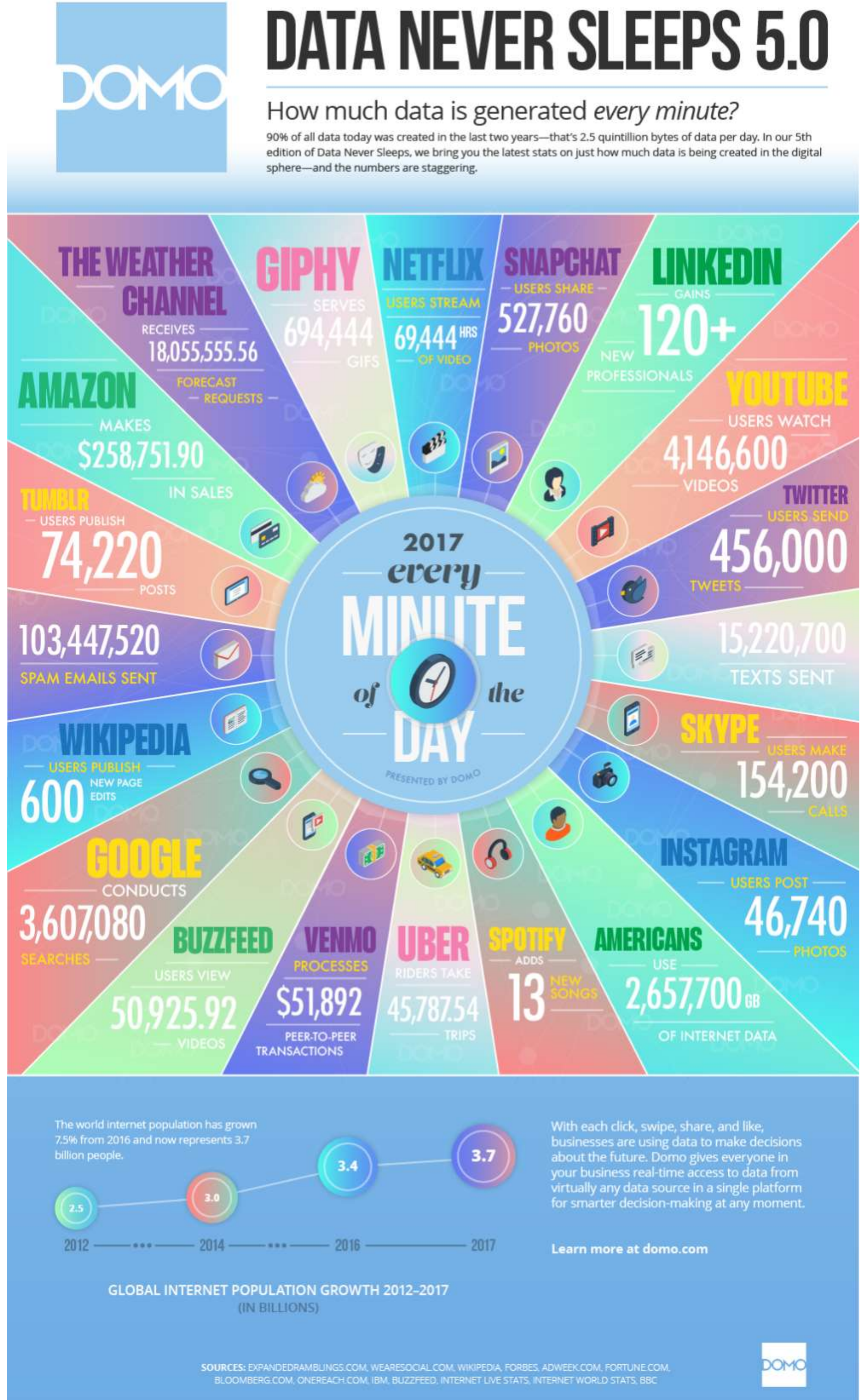
“Storage area networks are the future of enterprise storage, period. If your company is heading toward, or has already passed the terabyte mark in storage, it's a prime candidate for a SAN migration. If you are forecasting significant growth in storage requirements, you should develop your SAN strategy now”

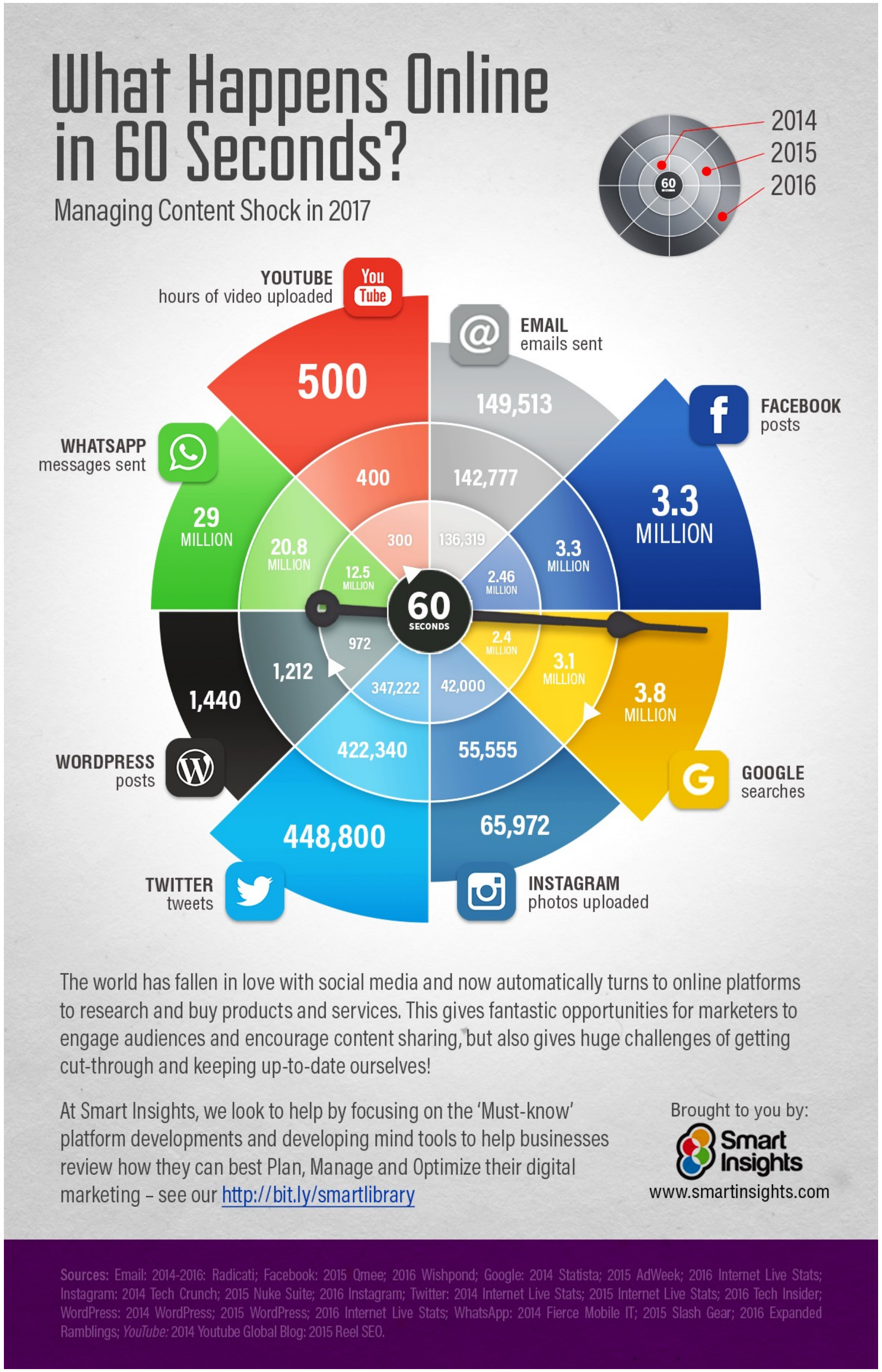
--Excerpted from “Building a Storage Area Network,” Dave Fetters





http://cdn.iflscience.com/images/4e100188-c00d-528c-9832-eb2052ad587b/content-1467213100-dataneversleeps.jpg
https://web-assets.domo.com/blog/wp-content/uploads/2015/08/15_domo_data-never-sleeps-3_final1.png
http://cdn.iflscience.com/images/28cfc245-c67b-50cd-9ce5-fb1500c8bfca/content-1501073543-17-domo-data-never-sleeps-5-01-1.jpg


DOMO
DATA NEVER SLEEPS 5.0
How much data is generated *every minute?*
90% of all data today was created in the last two years—that's 2.5 quintillion bytes of data per day. In our 5th edition of Data Never Sleeps, we bring you the latest stats on just how much data is being created in the digital sphere—and the numbers are staggering.
2017 every MINUTE of the DAY
PRESENTED BY DOMO
THE WEATHER CHANNEL RECEIVES 18,055,555.56 FORECAST REQUESTS
GIPHY SERVES 694,444 GIFS
NETFLIX USERS STREAM 69,444 HRS OF VIDEO
SNAPCHAT USERS SHARE 527,760 PHOTOS
LINKEDIN GAINS 120+ NEW PROFESSIONALS
YOUTUBE USERS WATCH 4,146,600 VIDEOS
TWITTER USERS SEND 456,000 TWEETS
15,220,700 TEXTS SENT
SKYPE USERS MAKE 154,200 CALLS
INSTAGRAM USERS POST 46,740 PHOTOS
AMERICANS USE 2,657,700 GB OF INTERNET DATA
SPOTIFY ADDS 13 NEW SONGS
UBER RIDERS TAKE 45,787.54 TRIPS
VENMO PROCESSES $51,892 PEER-TO-PEER TRANSACTIONS
BUZZFEED USERS VIEW 50,925.92 VIDEOS
GOOGLE CONDUCTS 3,607,080 SEARCHES
WIKIPEDIA USERS PUBLISH 600 NEW PAGE EDITS
103,447,520 SPAM EMAILS SENT
TUMBLR USERS PUBLISH 74,220 POSTS
AMAZON MAKES $258,751.90 IN SALES
The world internet population has grown 7.5% from 2016 and now represents 3.7 billion people.
2.5
3.0
3.4
3.7
2012
2014
2016
2017
GLOBAL INTERNET POPULATION GROWTH 2012–2017
(IN BILLIONS)
With each click, swipe, share, and like, businesses are using data to make decisions about the future. Domo gives everyone in your business real-time access to data from virtually any data source in a single platform for smarter decision-making at any moment.
Learn more at domo.com
SOURCES: EXPANDEDRAMBLINGS.COM, WEARESOCIAL.COM, WIKIPEDIA, FORBES, ADWEEK.COM, FORTUNE.COM, BLOOMBERG.COM, ONEREACH.COM, IBM, BUZZFEED, INTERNET LIVE STATS, INTERNET WORLD STATS, BBC
DOMO



https://www.smartinsights.com/internet-marketing-statistics/happens-online-60-seconds/

Evolution and Benefits of SANs

- In the pre-Internet era, storage management was not a big problem.
- The focus was on how to make data processing faster.
- Small Computer Systems Interface (SCSI) was born out of this need.
- As reliability and the popularity of computers and computer networks increased, the amount of data that had to be maintained on computers rose proportionately.
- Storage was embedded in servers.
- Server's storage capacity was severely limited because it could handle only a finite number of storage disks.
- Bottlenecks when several users tried to access information simultaneously.
- This gave birth to separate storage devices such as Just a Bunch Of Disks (JBODs), disk arrays, and tape libraries.
- These storage devices could effectively store huge amounts of data.
- However, with the growth of server farms, administrators realized that the interactions with these devices were extremely slow if too many users tried to access the same information from an individual server in the group.

NAS

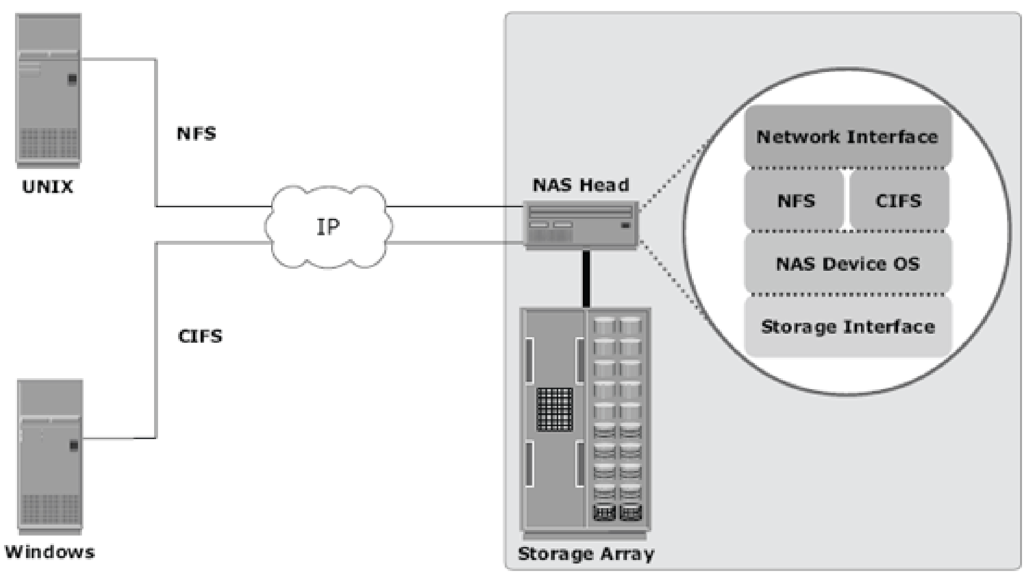
- Network Attached Storage.
- Instead of being connected to the servers, these NAS devices were directly connected to the network.
- Storage-related problems were lessened for network administrators, and proved to be an effective solution until the following drawbacks were realized:
  - Performance of NAS devices is severely limited by the available network bandwidth if the average data transfer rate is less than 100 Mbps.
  - Network traffic increases considerably when NAS devices are used in a network. This is because NAS devices can be involved in bandwidth-intensive data transmissions, such as backups. Because they share the same transmission medium as the clients and other servers on the network, excessive network traffic is generated, which can bring down overall network productivity and performance.
  - NAS devices are vulnerable to malicious attacks.
  - NAS devices can be difficult to manage because they are separate entities and are not logically tied together.

Benefits of NAS

- **Supports comprehensive access to information:** Enables efficient file sharing and supports many-to-one and one-to-many configurations. The many-to-one configuration enables a NAS device to serve many clients simultaneously. The one-to-many configuration enables one client to connect with many NAS devices simultaneously.
- **Improved efficiency:** Eliminates bottlenecks that occur during file access from a general-purpose file server because NAS uses an operating system specialized for file serving. It improves the utilization of general-purpose servers by relieving them of file-server operations.
- **Improved flexibility:** Compatible for clients on both UNIX and Windows platforms using industry-standard protocols. NAS is flexible and can serve requests from different types of clients from the same source.
- **Centralized storage:** Centralizes data storage to minimize data duplication on client workstations, simplify data management, and ensures greater data protection.
- **Simplified management:** Provides a centralized console that makes it possible to manage file systems efficiently.
- **Scalability:** Scales well in accordance with different utilization profiles and types of business applications because of the high performance and low-latency design.
- **High availability:** Offers efficient replication and recovery options, enabling high data availability. NAS uses redundant networking components that provide maximum connectivity options. A NAS device can use clustering technology for failover.
- **Security:** Ensures security, user authentication, and file locking in conjunction with industry-standard security schemas.

Componentts of NAS

- NAS head (CPU and Memory)
- One or more network interface cards (NICs), which provide connectivity to the network. Examples of NICs include Gigabit Ethernet, Fast Ethernet, ATM, and Fiber Distributed Data Interface (FDDI).
- An optimized operating system for managing NAS functionality
- NFS and CIFS protocols for file sharing
- Industry-standard storage protocols to connect and manage physical disk resources, such as ATA, SCSI, or FC

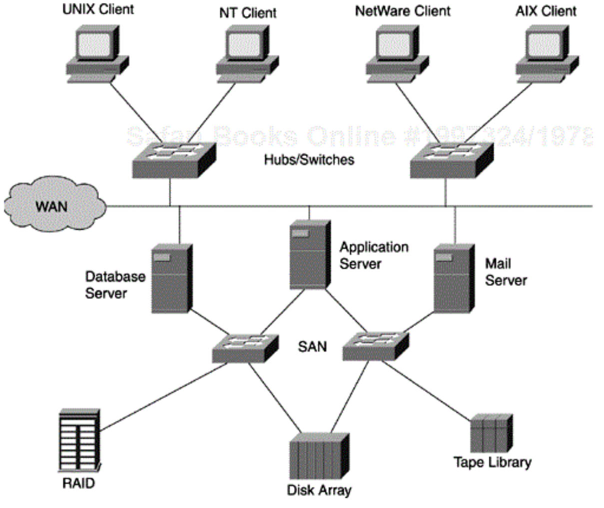


SAN

- Storage Area Networks.
- You can visualize a SAN as a separate high-speed storage network within a LAN.
- The entire concept of SANs was developed based on the following three factors:
  - The solution had to be a stable, one that would work well despite the predicted increase in Web-based transactions and businesses.
  - The solution had to help increase overall network productivity.
  - The solution had to not only work with small- and medium-sized networks, but also with large corporate networks that span the entire globe.

SAN

- All the storage resources are interconnected to each other by using general network elements (routers, hubs, switches, and gateways) which help in eliminating distance-based restrictions posed by the SCSI interface.
- The SAN is the storage behind the network servers.
- All storage devices are interconnected to form a separate network behind the network servers, which is well hidden from general users.
- All servers can access all storage resources, when required.
- Clients interact with servers over the LAN, whereas servers interact with storage devices over the SAN. Neither of these interactions (LAN and SAN) need to share network bandwidth, as was the case with earlier solutions such as NAS. This results in a marked increase in overall network productivity and satisfies one of the requirements for developing a workable solution.



SAN Advantages

- One of the advantages of the SAN is its scalability. You can add new storage resources to the network dynamically. Moreover, because a storage device is not associated with any server, adding a new storage device does not lead to adding a corresponding server. This is a boon in disguise for most network administrators because they have to manage fewer devices. This results not only in fewer management headaches, but also brings down the total cost of the network.

- Network performance also improves with SANs. Server/storage interactions do not use network bandwidth, which reduces network traffic considerably. This is because bandwidth-intensive bulk-data transfers, such as backups and database updates, occur within the SAN. As a result, traffic on the LAN, which is attached to the SAN, is not affected and users can continue with their normal work without having to feel the bandwidth pinch. Likewise, server performance is improved. Servers become free from data input/output (I/O) activities, which are highly CPU-intensive. As a result, servers can process client requests faster, which brings about a dramatic reduction in client response time.

SAN Advantages

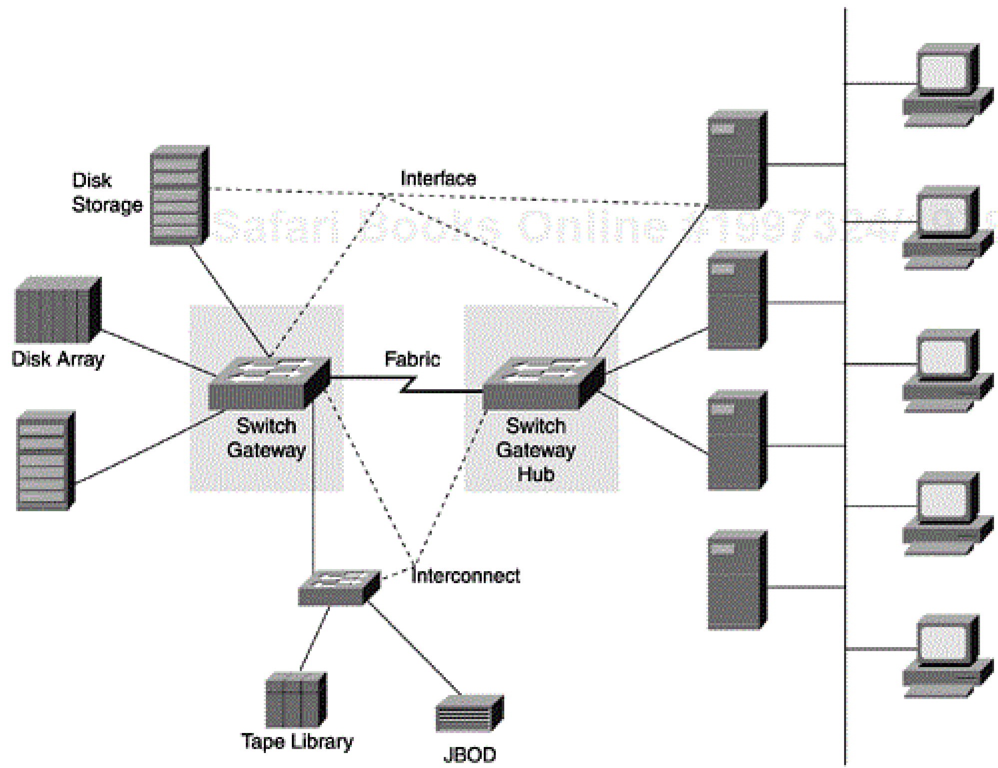
- Another advantage is improved data availability. In a SAN setup, multiple servers share access to the same data simultaneously. If a server becomes unavailable, other servers can take over, which results in improved data availability to the clients and high failovers. This helps in eliminating data-accessibility problems that arise due to a single point of failure (SPOF).

- The Fibre Channel infrastructure, which forms the base of the SAN, guarantees fast data transfer rates (from 100 Mbps to 1.0625 Gbps). This data rate is faster than data exchange rates in LANs to which the SAN is connected. This is especially beneficial in the case of e-commerce where fast data transactions are important to online companies that are aiming to gain favor from users.

SAN Components and Building Blocks

A wealth of devices and components make up a SAN. These include the following:

- Servers
- Storage
- Interfaces
- Interconnects
- Fabric
- Software and applications



SAN Servers

- Initially, SAN solutions supported a homogeneous server environment where the network consisted of servers with a common operating system. In a homogeneous environment, it is quite easy to connect these servers to the storage devices through the SAN and to share the data. However, a homogeneous environment is an unrealistic solution in the real world. This is because, over time, various corporate networks might have invested in a mix of servers according to their changing requirements.

- The current SAN infrastructure can support heterogeneous server environments. The mix of various server platforms supported includes versions of UNIX, Linux, Windows NT, Windows 2000, versions of NetWare, OS/390, AIX, Solaris, and HP-UX. The problem in using heterogeneous server platforms is that the file systems of the major server platforms, such as IBM mainframes, Windows NT, and UNIX, are incompatible with each other and are, therefore, incapable of sharing data. However, data can be converted with the help of specialized data conversion applications. For example, IBM and HP use software that converts data from mainframe to open system and vice versa. Also, SANs enable server clustering, which makes administration and troubleshooting a comparatively easier and less expensive scenario.

SAN Storage

- The storage infrastructure is where confidential and mission-critical data ultimately resides. Therefore, the storage infrastructure has to be fast, capable of storing large volumes (to prevent it from being scattered over various storage devices), easily accessible, manageable, and secure.
- SAN storage is consolidated, externalized from the primary network, and can be distributed across the corporate network based on requirements. The various storage devices supported by the SAN are high-capacity and high-performance devices. These include the following:
  - JBODs
  - Disk storage systems
  - Disk arrays
  - Tape libraries
  - Optical storage libraries

SAN Interfaces

- In the SAN environment, the shared storage repository is connected to SAN servers through a SAN interface, thus allowing the storage to be externalized from the servers. The most common SAN interfaces are SCSI, High-Performance Parallel Interface (HIPPI), bus and tag, Fibre Channel Arbitrated Loop (FC-AL), Enterprise System Connection (ESCON), and Fibre Connection (FICON). This section describes each of these interfaces.
- **SCSI** is a high-speed parallel interface that is used to connect high-speed SAN storage, such as JBODs, disk arrays, optical storage libraries, and tape libraries, to the SAN. Many vendors are currently using ANSI's SCSI-3 over the Fibre Channel infrastructure to achieve higher speeds, longer device distances, and a greater number of devices that can be connected through the interface.
- **SCSI-3** is a serial version of the earlier interface, which was a parallel interface.
- **HIPPI** is another ANSI standard interface that provides a point-to-point link for transferring data at speeds ranging between 100 Mbps and 200 Mbps over the Fibre Channel.
- **The FC-AL** is a high-speed (100 Mbps, 200 Mbps, 400 Mbps, and 800 Mbps) fault-tolerant Fibre Channel interface that can connect up to 126 SAN devices and is compatible with the SCSI interface.
- **The ESCON** interface is used to connect IBM's proprietary switches, known as ESCON directors, to other SAN components in half-duplex mode at a comparatively low speed of 17 Mbps.
- **FICON** is a high-speed (100 Mbps) interface that is used to connect FICON directors to other SAN components in full-duplex mode. FICON directors are IBM's proprietary next-generation Fibre Channel switches that provide additional management functions in addition to higher data transfer rates—100 Mbps—and full-duplex bidirectional communication.
- Compared to ESCON directors, FICON directors are about eight times faster. In addition, FICON directors can efficiently support physical connectivity up to 20 kilometers, com-pared to the mere three kilometers that is supported by ESCON directors.

SAN Interconnects

SAN interconnection devices, also known as SAN interconnects, are used to connect local and remote storage interfaces to various topologies. They also connect SAN interfaces to the SAN Fabric. The SAN Fabric is the interconnection of switches that is supported by hubs, gateways, and routers. Commonly used SAN interconnects are cables, connectors, adapters, extenders, hubs, bridges and multiplexors, routers, gateways, and switches.

Cables

As in a traditional networking environment, two types of cables are used in the SAN environment—copper and fiber-optic cables. Copper cabling is used for relatively short distances of 30 meters. Fiber-optic cables are used for longer distances that range from 2 to 10 kilometers. There are two types of fiber-optic cables:

Multi-mode fiber (MMF) cables that are used in connections spanning up to 2 kilometers

Single-mode fiber (SMF) cables that are used over longer distances spanning up to 70 kilometers

Connectors

Connectors are used in an environment where both copper and fiber-optic cabling is used. Connectors allow the interconnection of fiber-optic devices with copper devices. Gigabit Interface Converters (GBICs) are small interface modules that are used to connect copper or fiber-optic devices to hubs, switches, and adapters. Media Interface Adapters (MIAs) handle the conversion of copper-based connections into fiber-based connections and vice versa. MIAs can also convert a signal into the appropriate media type. MIAs are connected to most of the systems with the help of Fiber Channel transceiver units known as Gigabit Link Models (GLMs).

Adapters

Adapters are also commonly referred to as network interface cards (NICs) and Host Bus Adapters (HBAs). Adapters provide a physical interface between the host bus and SAN interfaces. To facilitate proper communication, the adapters support various upper-level protocols (ULPs), such as SCSI, TCP/IP, FICON, and ESCON. Generally, most adapters are shipped with device drivers that control their behavior.

Extenders

Similar to repeaters in a normal network, extenders are used to extend the length of physical media, both copper-based and fiber-optic.

Hubs

Similar to the hubs used in normal LANs, Fibre Channel hubs act as the central point of connection for various SAN devices. A typical Fibre Channel hub can support up to 126 nodes. Each port on a hub contains a Port Bypass Circuit (PBC), which opens or closes loops automatically to prevent a physical change or a device failure from affecting the functionality of other devices connected to the hub. A category of Fibre Channel hubs known as managed hubs also provides some of the management functionality offered by switches in a LAN environment. For detailed information on hubs, see .

Bridges and Multiplexors

Bridges facilitate communication between Fibre Channel interfaces (such as FICON and ESCON) and SCSI interfaces, which enables communication between the primary LAN and SAN. Bridges also provide connectivity between networks that operate on dissimilar protocols.

Multiplexors are specialized bridges that interleave signals from multiple devices and transmit them simultaneously through a single transmission medium. This helps in the effective utilization of the available network bandwidth.

Routers

Fibre Channel routers, which are also commonly known as storage routers, act much like routers that you use in normal networks. Similar to the normal routers, storage routers are responsible for transferring storage data between networks by using different transmission media and addressing schemes.

Gateways

As in normal networks, a Fibre Channel gateway is used to interconnect dissimilar LANs and distant networks over a wide-area network (WAN). However, a gateway may or may not perform protocol conversion.

Switches

Fibre Channel switches provide the same functionality as network switches. They are used to interconnect a large number of devices. However, the connected nodes do not share bandwidth, as is the case with hubs. Instead of broadcasting a signal to all its ports, a switch allows on-demand connections because it transmits the signal only to that port to which the destination device is connected. As a result, switches help considerably in reducing network traffic. They also help in the effective utilization of network bandwidth and increase the aggregate throughput.

SAN Software

The SAN software component includes the following types of management applications:

Applications to configure, maintain, and manage the SAN Fabric. For example, IBM's popular management software——Tivoli SANergy—allows simultaneous sharing of the same storage, file systems, and even the same files between multiple computers connected to a SAN. Another popular IBM management product is Tivoli Storage Network Manager, which discovers, displays, allocates, monitors, automates, and manages various components of the SAN Fabric and disk storage resources. BakBone's SmartClient allows network administrators to centrally control the attached media devices. Compaq also has a rich line of SAN management applications, including SANworks Enterprise Network Storage Manager and SANworks Storage Resource Manager.

Applications to help exploit SANs to the fullest to achieve maximum benefits. These applications include backup and recovery packages, volume managers that enable host-based and remote mirroring, disk striping, data replication, and other network management software. For example, IBM's Tivoli Storage Manager is a comprehensive SAN management software suite that offers SAN-enabled integrated enterprise-wide network backup, archive, and disaster recovery capabilities. Another software suite is BakBone's NetVault Dynamically Shared Drives (DSDs), which is a powerful LAN-free backup software tool.

File- or data-sharing software, extended file systems, and shared file systems. Generally, these applications use zoning and Logical Unit Number (LUN) masking techniques. Viacom's SV Zone Manager offers centralized and flexible access management to resources in the storage network. In addition, SV Zone Manager allows administrators and network managers to control user access to virtualized storage. Veritas SANPoint is another line of popular software products in this category.

Data Access over SANs

In a network, it is possible that more than one client (local or remote) will request access to the same data simultaneously. The setup of the SAN environment is such that multiple servers belonging to different platforms can access a storage resource at the same time. Data can be accessed from a storage device in the following ways:

- Physical partitioning of the storage disk volumes
- Logical partitioning of the storage disk volumes
- File pooling
- Sharing data

Physical partitioning of storage disk volumes is the simplest method of accessing data from the storage repository. In this method, disk volumes are assigned to each server for their exclusive use. After a disk volume is assigned to a server, it becomes inaccessible to other servers. This method is commonly used in a heterogeneous server environment.

With the logical partitioning of disk volumes, logical disk volumes are defined within the storage repository and assigned to their respective servers. A logical disk can either span across several storage resources or can be located on a single storage device. However, the storage controller must be capable of managing the logical grouping of the volumes and ensuring that one server cannot access another server's logical volume.

> The mountable namespace allows file systems to be added or removed from a partition set while the system is still running. Partitions can also be added to a partition set (or removed if the remaining partitions have enough space to contain all the data) while the system is running. The main advantage of mountable namespace is that it allows several mountable file systems to share the same pool of storage space, which enables the easy addition of more hard drives.

With the sharing data method, data can be accessed either in the form of data copy or the real-time sharing of data by multiple servers. Data sharing is accomplished in two ways, as noted in the following:

Sharing data copy—. According to this method, a copy of the requested data is assigned to the server. After the entire requested file is copied to the server, either the entire database (or file) is updated at the server-side at regular intervals or only the incremental updates are copied as and when any changes in the source copy occur. When the entire database or file is updated at regular intervals, the method is known as sharing the complete copy. The latter method is known as sharing the incremental copy. Complete copy sharing is frequently used by the backup and restoration applications, and the incremental copy is shared by data warehousing and data mining applications.

True data sharing—. According to this method, the same copy of the data is accessed by more than one server simultaneously. This method is considered to be the best because it generates less network traffic and allows the actual consolidation of storage. The servers can access the data in three ways. One-at-a-time access allows the requesting entities to read and update data in a sequential manner. The multiple read access method allows the requesting servers to read the data simultaneously. However, only one server at a time can update the data on the storage device. The multiple read/write access method allows multiple servers to read and update the data simultaneously. This method is practically obsolete because the simultaneous updating of data can cause severe conflicts related to data integrity.