

Introducción a los Sistemas de Almacenamiento Distribuido

Bibliografía:

Thomas C. Jepsen, “Distributed Storage Networks. Architectures, protocols and Management”, cap. 1, pp. 1-22, Edit. John Wiley & Sons, 2003.

Objetivos

- Proporcionar una visión general de los sistemas de almacenamiento distribuidos
- Definir un glosario de términos relacionados con las redes de almacenamiento distribuido.

Contenidos

1. Definición de red de almacenamiento distribuido.
2. Objetivos de las redes de almacenamiento distribuido.
3. Definición de red de almacenamiento.
4. Evolución histórica.
5. Terminología relacionada.

1. Red de almacenamiento distribuido.

Definición

- Una red de almacenamiento distribuido (*Distributed Storage Network - DSN*): es una infraestructura que permite el acceso compartido de múltiples sistemas (ordenadores de sobremesa, servidores, dispositivos móviles de usuario, etc.) al recurso de almacenamiento de un sistema información (SI).

2. Red de almacenamiento distribuido.

Objetivos

- Mejorar los siguientes aspectos del recurso de almacenamiento de un SI:
 - Disponibilidad de los datos.
 - Posibilitar el manejo compartido de los datos.
 - Eficiencia en el manejo de los datos.
 - Independencia de tecnologías y fabricantes.

3. Red de almacenamiento.

Definición

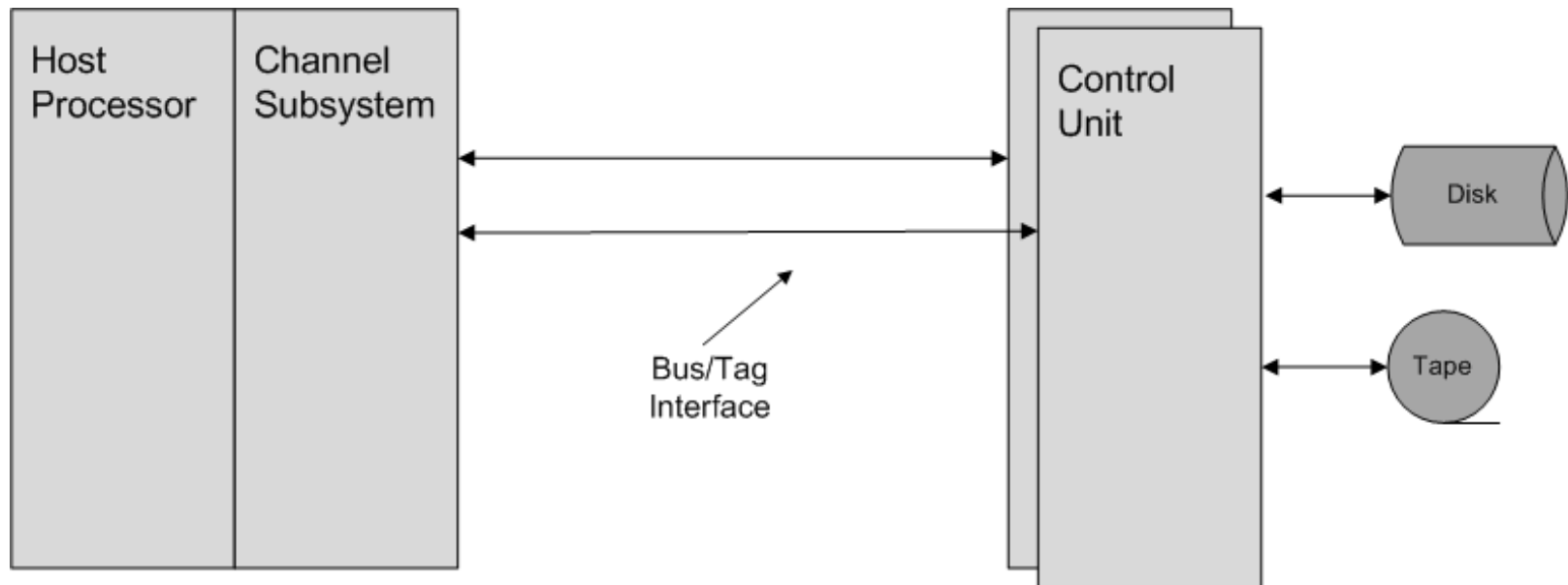
- **Red de almacenamiento** (*Storage Network*).
Red especializada en el transporte de datos entre distintos tipos de dispositivos.
 - Transporte bloques de datos reconocibles por distintos sistemas operativos y dispositivos de almacenamiento
 - Utiliza protocolos especializados para el transporte de datos (SCSI, ESCON, FC, etc.)

4. Red de almacenamiento.

Evolución histórica

- Redes de almacenamiento de *Mainframe* (1960s).
 - Un procesador utiliza un subsistema específico de E/S para comunicarse con los dispositivos de almacenamiento
 - Uso de buses paralelos para la conexión de los dispositivos
 - Limitaciones
 - Ancho de banda
 - Distancia

4. Red de almacenamiento. Evolución histórica



4. Red de almacenamiento.

Evolución histórica

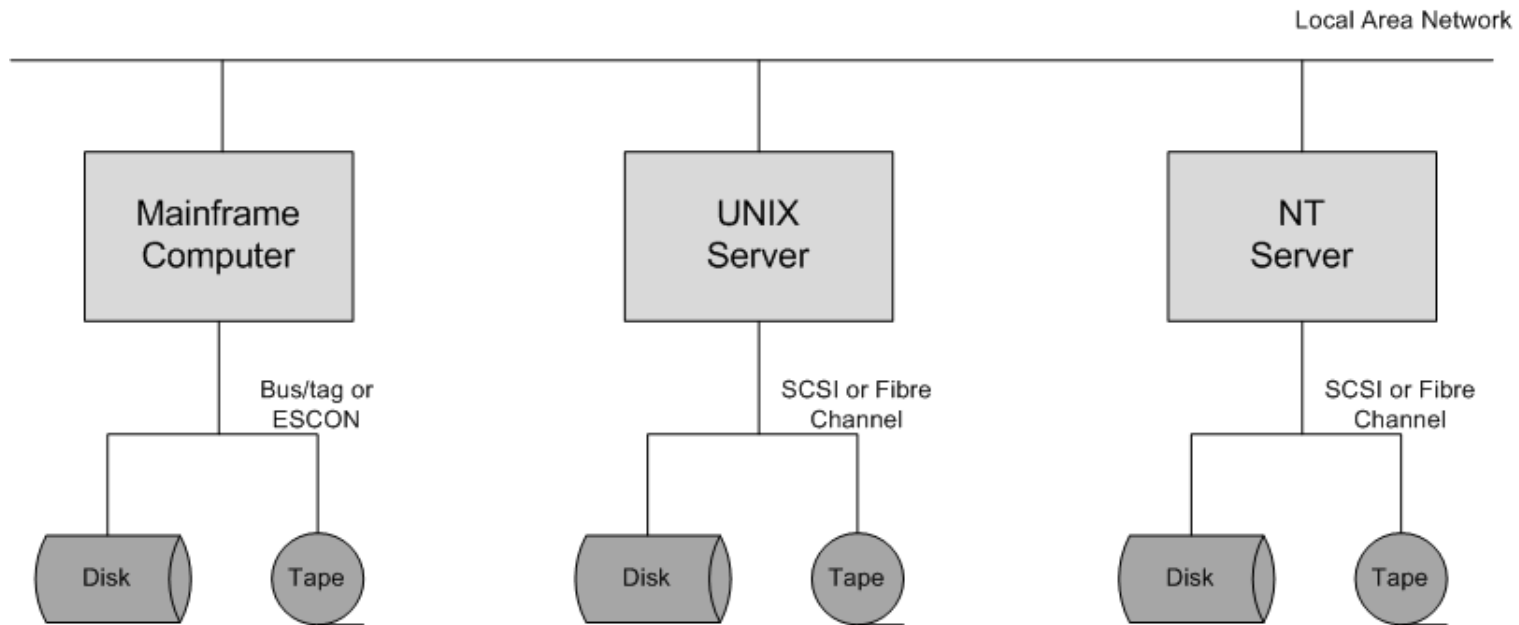
- Almacenamiento para miniordenadores (1970s) y ordenadores personales (1980s)
 - Necesidad de aumentar el grado de integración de los dispositivos de almacenamiento:
 - Capacidad
 - Densidad de almacenamiento
 - Tecnologías:
 - *Integrated Drive Electronics/AT Attachment (IDE/ATA)*
 - *Small Computer Systems Interface (SCSI)*
 - Limitaciones
 - Dificultad para el manejo de grandes cantidades de datos
 - Distancia

4. Red de almacenamiento.

Evolución histórica

- El problema de la “islas de almacenamiento” (*storage islands*), limitaciones:
 - Accesibilidad a los datos
 - Eficiencia del almacenamiento
 - Dificulta el manejo de los datos
 - Dificulta la compartición de datos
 - Productividad

4. Red de almacenamiento. Evolución histórica



Storage
Islands

Fuente [1]

5. Terminología relacionada

- **Red de área local, *Local Area Network (LAN)*.**
Es una red para la comunicación entre computadoras.
 - Orientada a soportar una gran variedad de servicios:
 - Transferencia de archivos
 - Llamada a procedimientos remotos
 - Gestión y administración de redes
 - Etc.

5. Terminología relacionada

- **Red de banda ancha, *Wide Area Network (WAN)*.**
Es una red concebida para dar cobertura de servicios a larga distancia.
 - *Uso de protocolos específicos:*
 - *Asynchronous Transfer Mode (ATM) (1980).*
 - *Dense Wavelength Division Multiplexing (DWDM) (1985).*
 - *Synchronous Optical Network (SONET)(1990s).*
 - *Utiliza arquitecturas como:*
 - *Backbone: topología basada en un bus al que se conectan todos los nodos.*
 - *Ring: los nodos se conectan formando un anillo.*

5. Terminología relacionada

- ***Storage Area Network (SAN)***. Red de almacenamiento que proporciona:
 - *Acceso compartido a los datos*
 - *Integra múltiples tipos de dispositivos de distintos fabricantes*
 - *Utiliza diferentes protocolos (pila de protocolos) que permiten un espacio de almacenamiento:*
 - *Eficiente*
 - *Fiable*
 - *Estandarización de funcionalidades (backup, recovery, mirroring, migration)*

5. Terminología relacionada

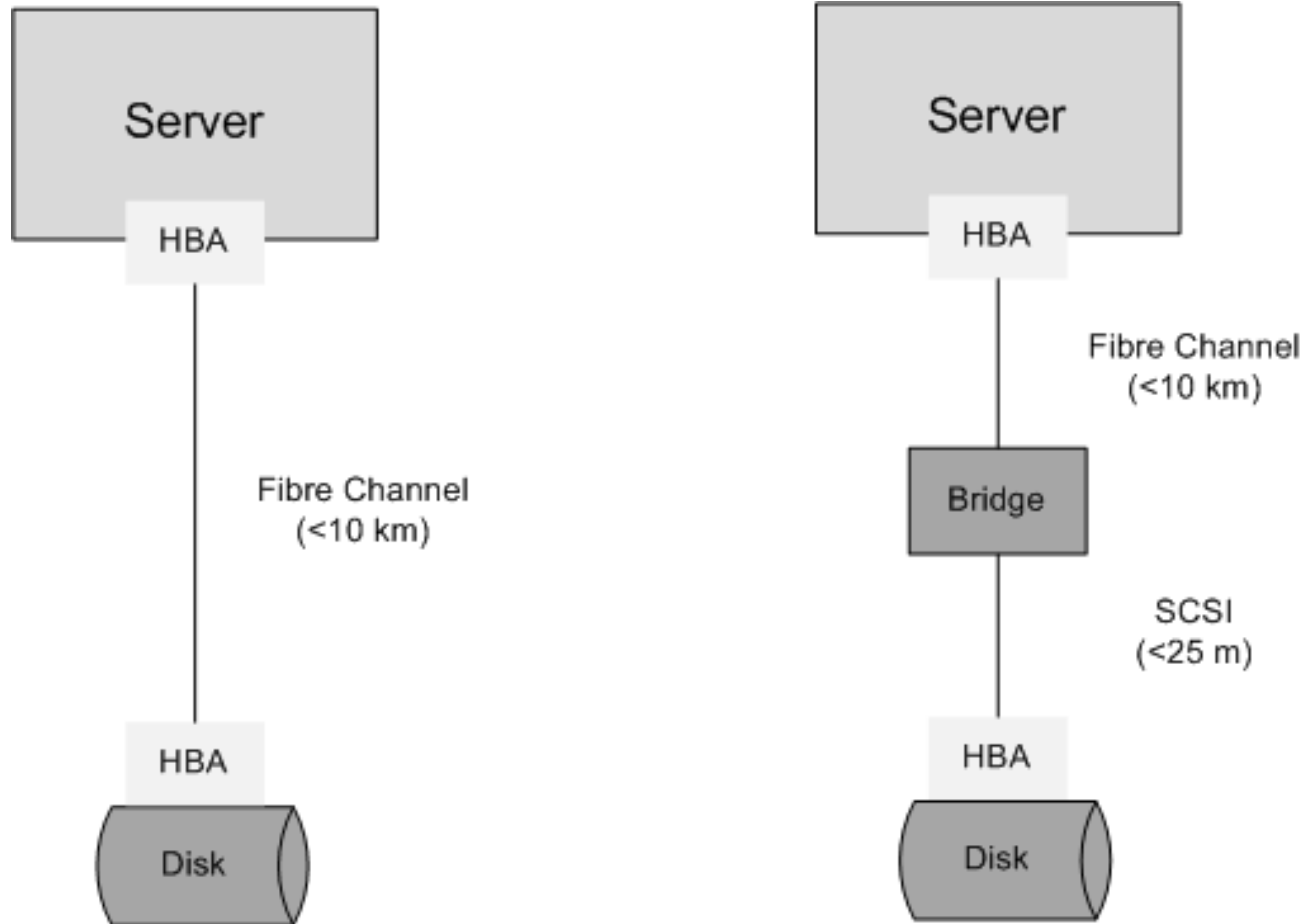
- **Un repaso de conceptos: Niveles OSI**

Nivel	Denominación	Descripción	Dispositivo
1	Físico	Medio físico Cómo se transmite la información	
2	Enlace de datos	direccionamiento físico Acceso al medio Detección de errores Distribución ordenada de trama Control del flujo	Bridge Switch
3	Red	Enrutamiento de los paquetes	Router Firewall
4	Transporte	Transporte de los paquetes de datos	
5	Sesión	Controlar y mantener el enlace establecido entre dos computadores	
6	Presentación	Representación de la información	
7	Aplicación	Acceso de las aplicaciones a la red	

5. Terminología relacionada

- **Tipología punto-a-punto en SAN.** Cada dispositivo de almacenamiento se conecta directamente a un host.
 - Protocolos: SCSI y FC.
 - Limitación: distancia.
 - SCSI < 10 mts.
 - SCSI/FC < 10 kmts.
 - Dispositivo característico: Bridge

5. Terminología relacionada

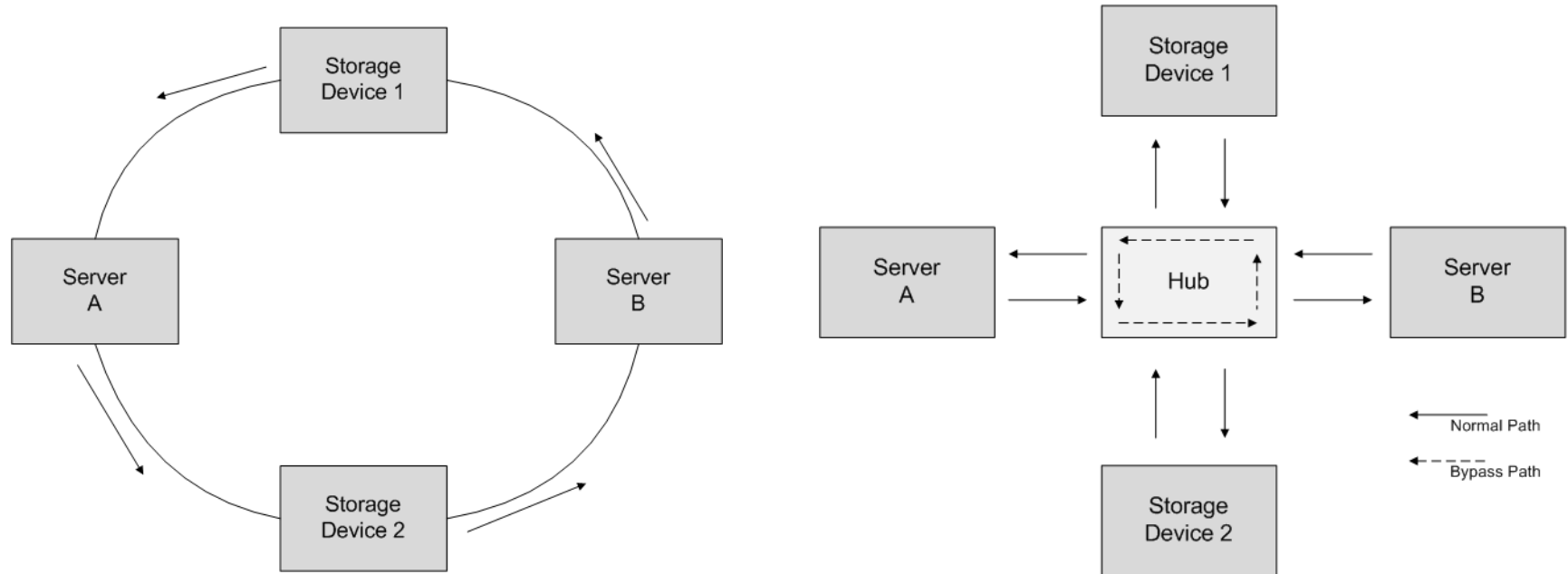


Fuente [1]

5. Terminología relacionada

- **Tipología Bucle arbitrado, *Arbitrated Loop Topology (FC-AL)***. Los dispositivos formando un bucle en el que el puerto de salida de un dispositivo se conecta al del entrada del siguiente.
 - Protocolos: FC
 - Hasta 127 nodos pueden conectarse
 - La lógica de arbitraje establece qué par de puertos toman el control del bucle y evita que un par de nodos acaparen el bucle
 - Cuando se establece una conexión, todo el ancho de banda del bucle se dedica a ella
 - Variante: FC-AL con Bridge

5. Terminología relacionada

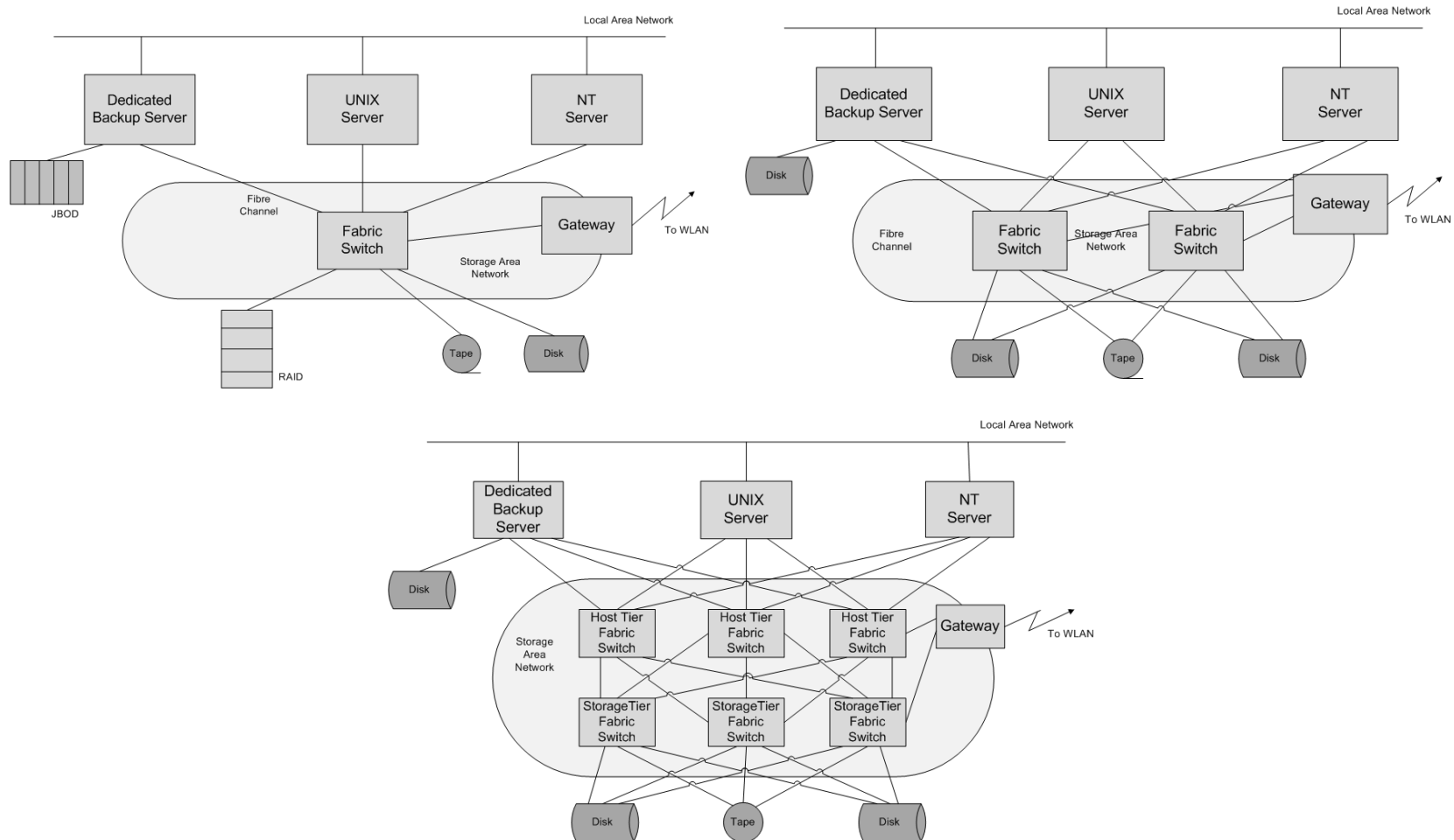


Fuente [1]

5. Terminología relacionada

- **Tipología basada switch, *Switch Fabric Topology*.** Los dispositivos se conectan a un *switch* y las conexiones entre dispositivos se llevan a cabo a través de él.
 - Permite múltiples conexiones concurrentemente usando todo el ancho de banda
 - Tipología típica de los sistemas SAN
 - Se pueden utilizar varios *switches* con el objeto de:
 - Mejorar la confiabilidad de las conexiones (varias rutas de conexión entre dispositivos): conectando los dispositivos a 2 o más *switches*
 - Mejorar la escalabilidad (número de nodos de SAN): utilizando conjuntos de *switches* organizados en niveles

5. Terminología relacionada

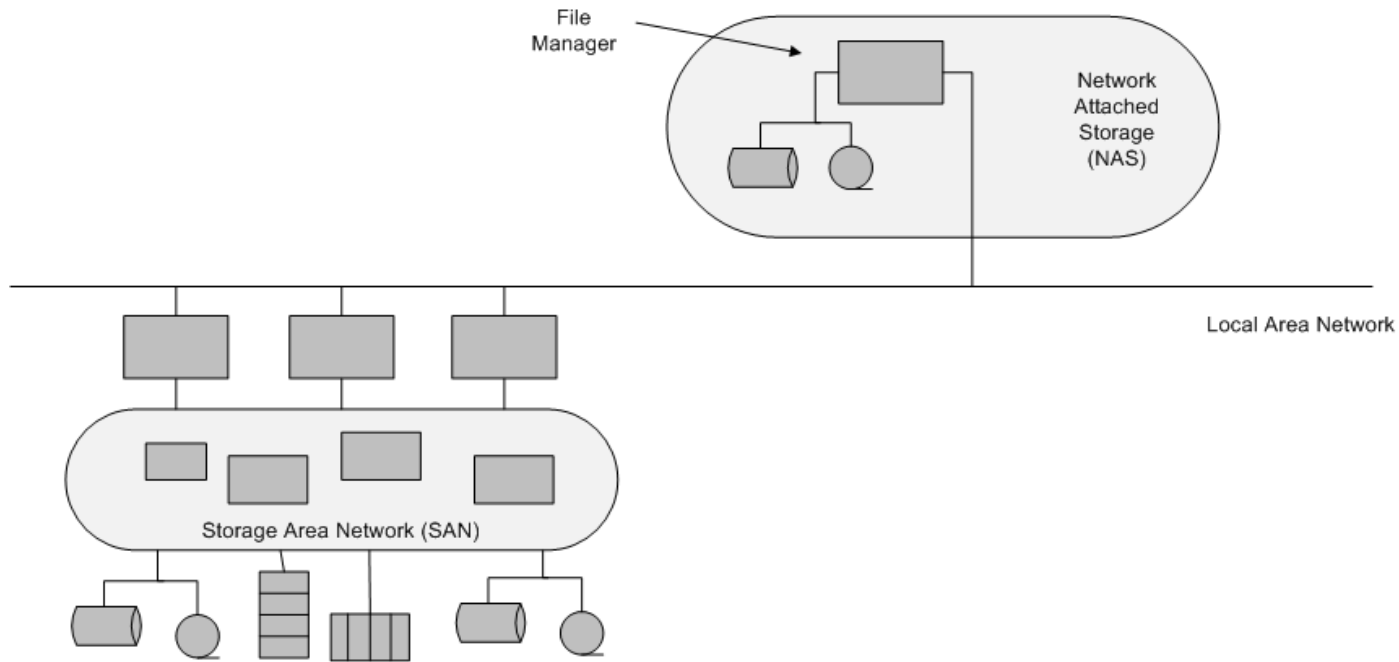


Fuente [1]

5. Terminología relacionada

- ***Network attached Storage*** (NAS). Tecnología que consiste en conectar un gestor de archivos a una LAN.
 - Está orientado a transferencia de archivos
 - Utiliza protocolos LAN: Ethernet, Gigabit Ethernet

5. Terminología relacionada



NAS

Fuente [1]

5. Terminología relacionada

- **Almacenamiento compartido.** *Espacio de almacenamiento* accesible a procesadores que pueden ser de distinto tipo.
- Tres formas de conseguirlo:
 - Almacenamiento particionado. Utilizando particiones separadas para cada tipo de sistemas operativo soportado.
 - Compartición basada en copias de datos. Los datos compartidos se copian en los distintos sistemas que acceden a ellos.
 - Utilizando almacenamiento virtual. Los datos compartidos se almacenan en un sistema de almacenamiento y una capa de virtualización posibilita el acceso de sistemas operativos de distinto tipo.

5. Terminología relacionada

- **Virtualización de almacenamiento.**
Tecnología que proporciona un espacio virtual de almacenamiento (disco virtual, volumen virtual).
 - Las aplicaciones acceden al espacio virtual y no a dispositivos físicos.
 - Técnica utilizada para proporcionar un espacio compartido.

5. Terminología relacionada

- **Array redundante de discos independientes, *Redundant Array of independent Disks (RAID)*.** Tecnología que utiliza conjuntos de discos y técnicas de:
 - *Striping* para conseguir rendimiento.
 - *Mirroring* para conseguir confiabilidad.
 - Técnicas de chequeo y corrección de errores (ECC).
- El uso combinado de estas técnicas da lugar a los distintos tipos de RAID

5. Terminología relacionada

Tipo de RAID	Striping	Mirroring	ECC	Funcionalidad
RAID 0	SI	NO	NO	Aumenta el rendimiento No proporciona confiabilidad
RAID 1	NO	SI	NO	No mejora el rendimiento de un disco Mayor grado de confiabilidad
RAID 2	SI	NO	SI	Mejora el rendimiento de un disco Mejora la confiabilidad de un disco. ECC utiliza varios discos.
RAID 3	SI	NO	SI	Mejora rendimiento de un disco. Proporciona confiabilidad a menor coste que RAID 1. Se dedica un solo disco para ECC. No permite solapar operaciones de E/S.
RAID 4	SI	NO	SI	<i>Striping</i> : utiliza cadenas largas de bits con paridad fija. Un registro de E/S se escribe en un disco. Se pueden solapar operaciones de lectura. No se pueden solapar operaciones de escritura.
RAID 5	SI	NO	SI	<i>Striping</i> : utiliza cadenas largas de bits con paridad flotante. Se pueden solapar operaciones de lectura. Se pueden solapar operaciones de escritura. El mejor para sistemas mutiusuarios.
RAID 10	SI	SI	NO	Arrays de <i>strips</i> en el que cada strip es un RAID 1. Mejor rendimiento que RAID 1.

5. Terminología relacionada

- ***Just a Bunch of Disks (JBOD)*** o también denominado ***spanning***. Tecnología que utiliza conjuntos de discos para proporcionar un sólo volumen de datos. Su objetivo es combinar varias unidades de disco para proporcionar un mayor espacio de almacenamiento.

5. Terminología relacionada

- **Replicación de disco, *disk mirroring*.** Técnica que tiene por objetivo duplicar los datos en diferentes lugares de almacenamiento.
 - Replicación síncrona, ***Synchronous mirroring***. La replicación se realiza cuando el dato cambia
 - Replicación asíncrona, ***Asynchronous mirroring***, la replicación se realiza a posteriori haciendo uso de un historial de cambios en el dato.
 - Replicación centrada en el procesador, ***processor centric mirroring***. El computador es el responsable de realizar las escrituras en el disco primario y en el disco espejo.
 - Replicación centrada en el almacenamiento, ***storage centric mirroring***. El computador escribe en el disco primario y el controlador es el responsable de realizar la replicación en el disco espejo.