

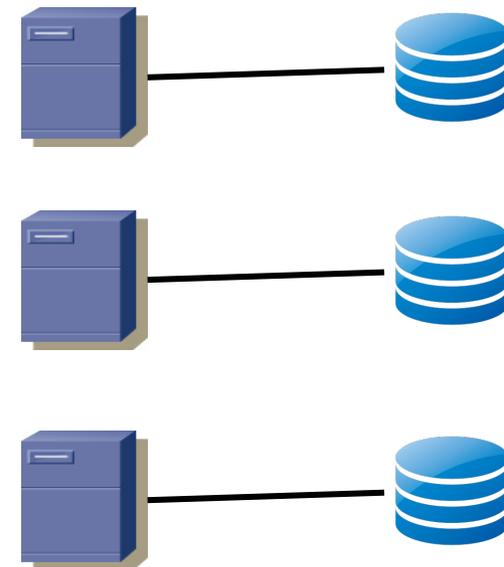
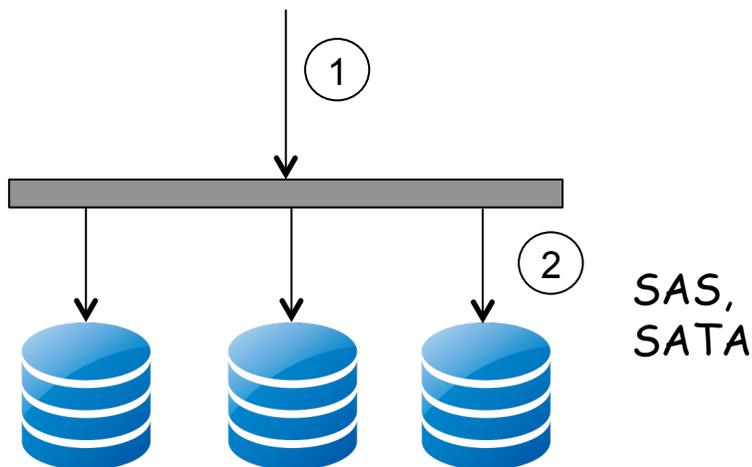


SAN \neq NAS



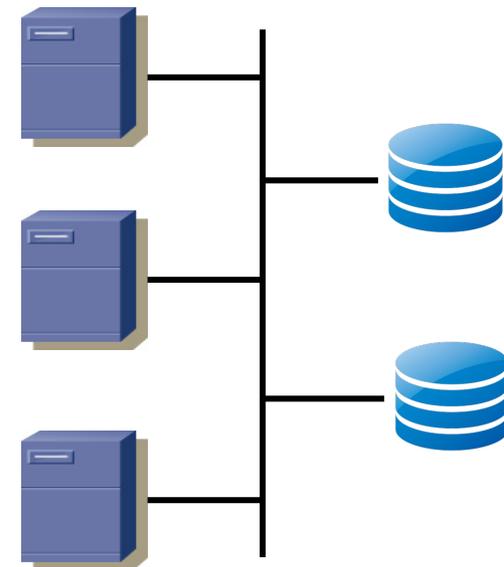
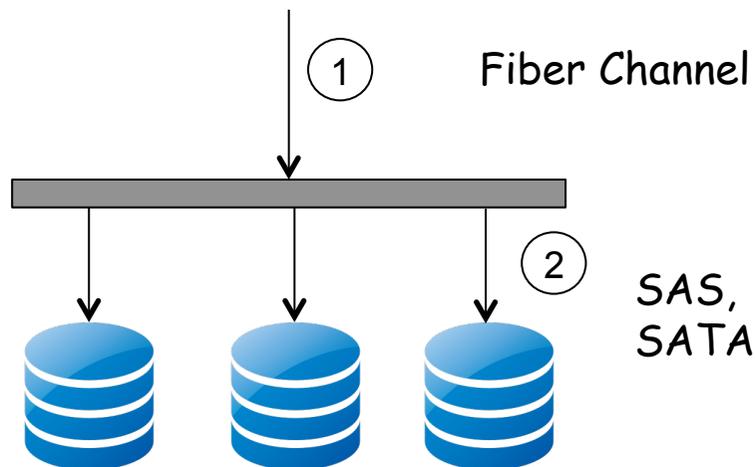
Acceso al almacenamiento

- El acceso al almacenamiento puede ser directo
 - *Direct Attached Storage (DAS)*
 - En ese caso cada servidor necesita su sistema de almacenamiento
 - Estos sistemas de almacenamiento externos están infrautilizados
 - Pueden ser por ejemplo para llevar a cabo *backups*
 - Un backup nocturno significa que el resto del día esos discos están inutilizados



Acceso al almacenamiento

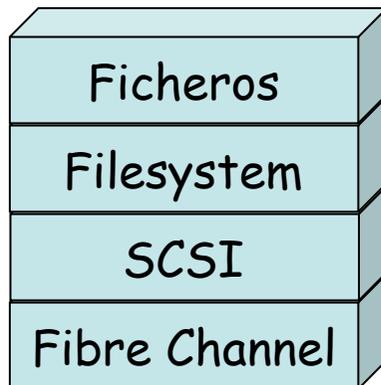
- El acceso al almacenamiento puede ser directo
- O puede ser a través de una red
 - Hablamos en ese caso de una *Storage Area Network (SAN)*
 - O de *Network Attached Storage (NAS)*
 - Estamos hablando del *front-end* de acceso a los discos
 - El *back-end* es común que sean discos SAS o SATA



SAN ≠ NAS

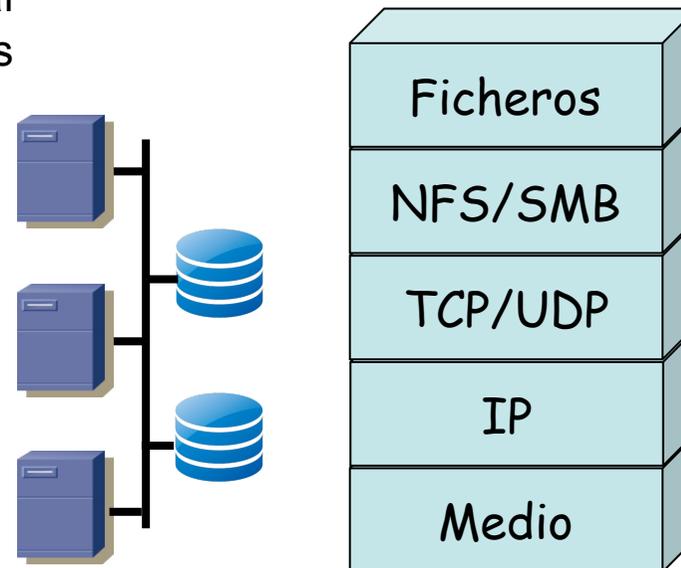
En una SAN

- Se accede de forma serie a bloques de disco
- Normalmente mediante comandos SCSI-3
- Los protocolos están optimizados para baja latencia y nulas pérdidas
- La solución de transporte más habitual es Fibre Channel
- Acceso de varios servidores al mismo volumen requiere sistemas de ficheros especiales



En una solución NAS

- Se accede a ficheros
- Se suele transportar sobre una tecnología LAN (o LAN + IP)
- Los protocolos no garantizan baja latencia ni nulas pérdidas (su recuperación aumenta la latencia)
- NFS, SMB/CIFS, AFP, etc



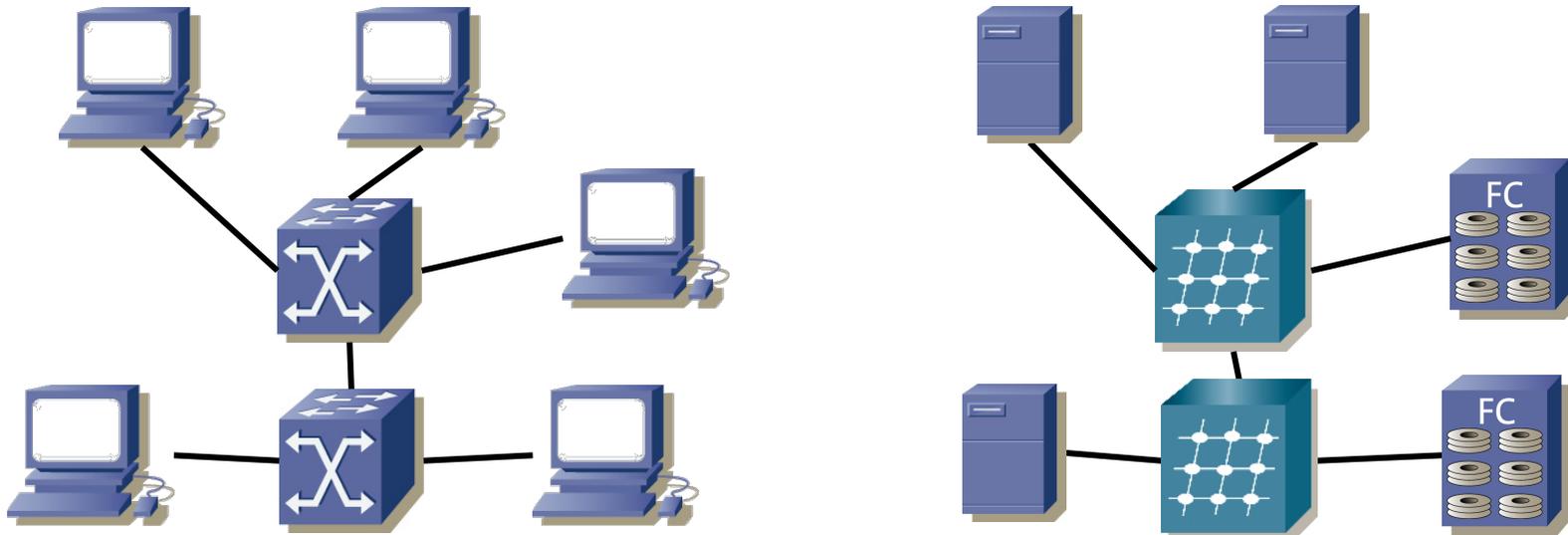
SAN vs NAS

- Rendimiento haciendo backups
 - Mayor en SANs, generalmente más rápida
 - En algunos casos se pueden mover datos del disco de un servidor a una cabina sin intervención de la CPU
 - En sistemas de ficheros con gran cantidad de ellos es más eficiente una copia del dispositivo *raw*
- Bases de datos
 - Algunas requieren un acceso al disco a nivel de bloques por rendimiento e integridad



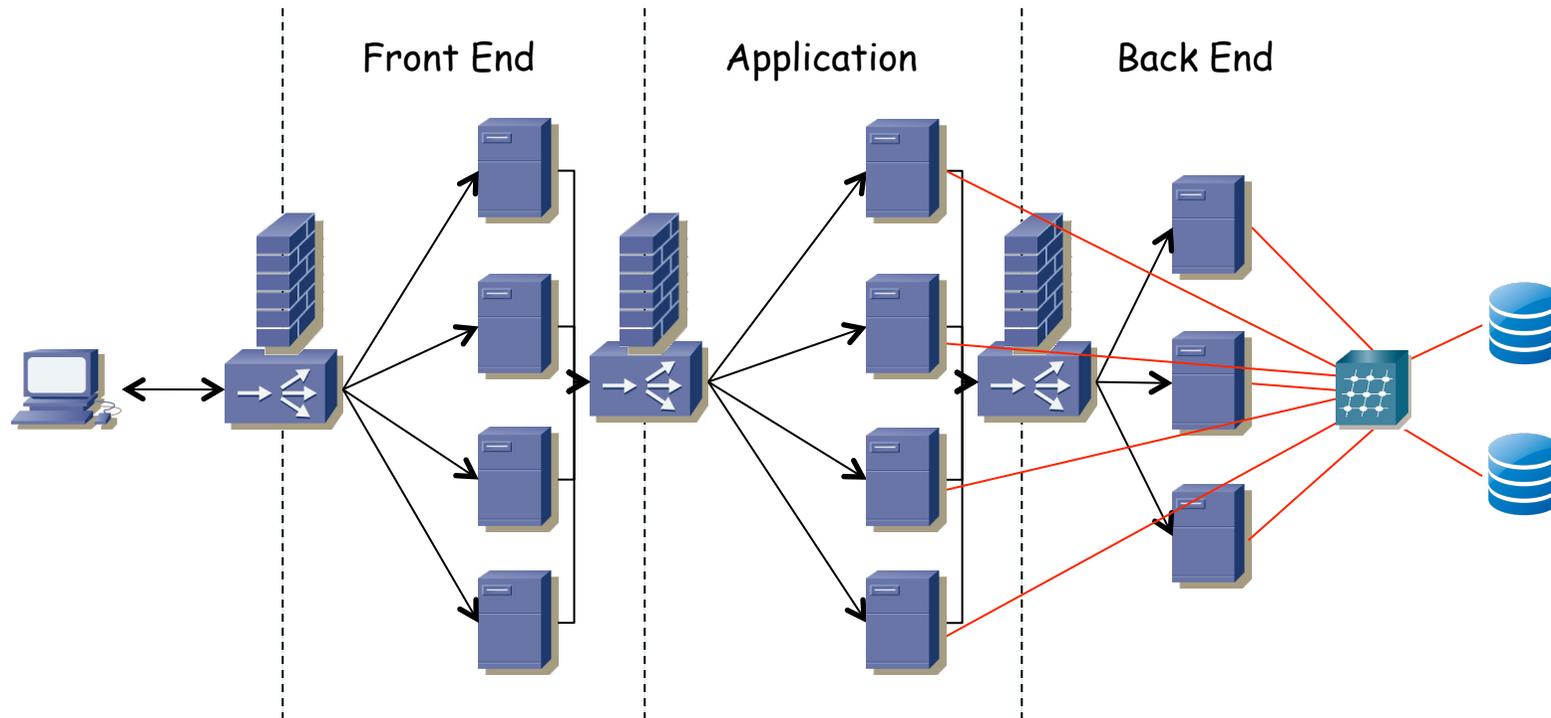
¿ Dónde encaja la SAN ?

- Estamos hablando de una tecnología de red
- Cuenta con sus propios conmutadores
- Se dice que forman un *fabric*
- Cuenta con su propia pila de protocolos
- El interfaz en el host es el *Host Bus Adapter (HBA)* que sería el equivalente a la NIC en una LAN



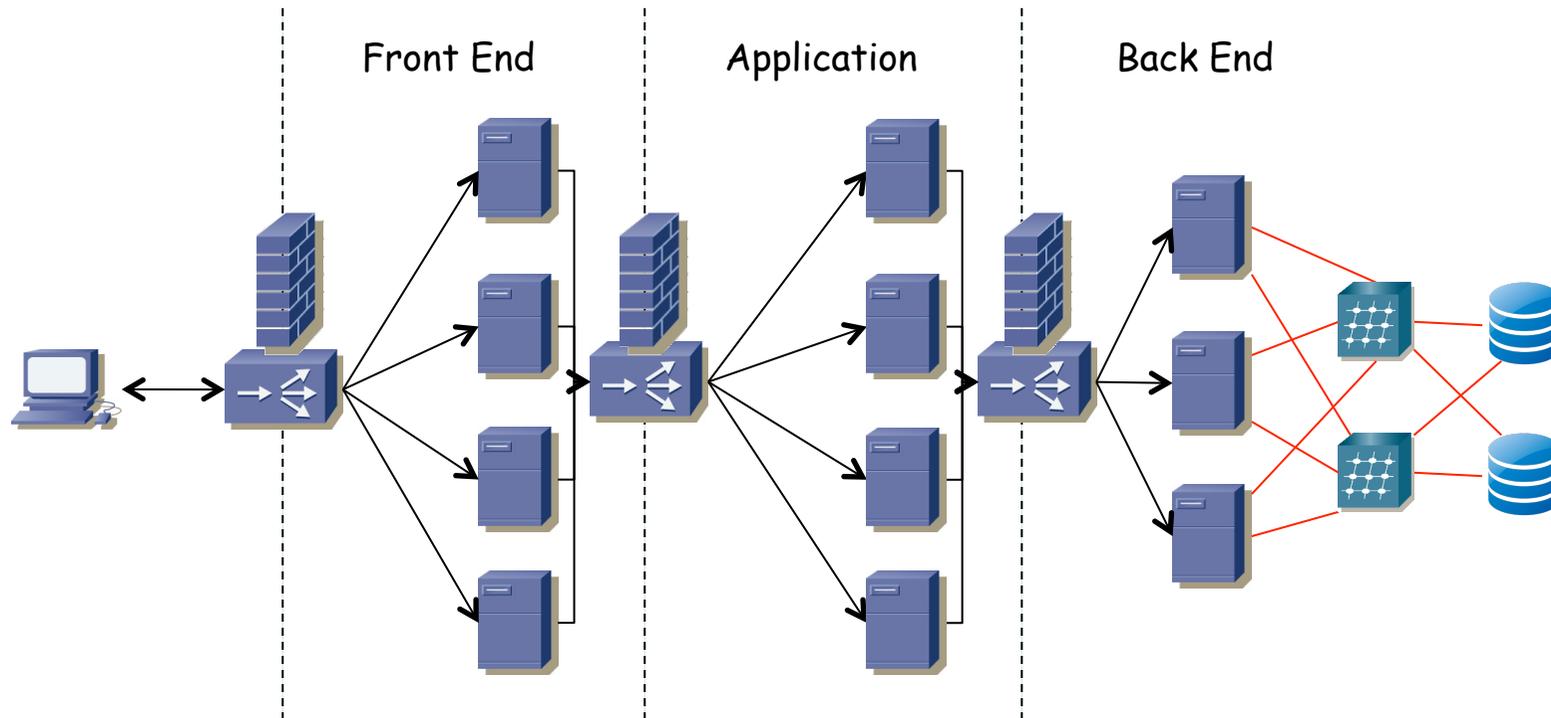
Es decir...

- Y recordemos que lo estamos poniendo en el backend del servicio porque es donde suelen estar los datos
- Pero nada impide que los servidores de cualquier otra capa empleen almacenamiento SAN
- O incluso en la misma SAN
- ¡ Incluso en los mismos discos !



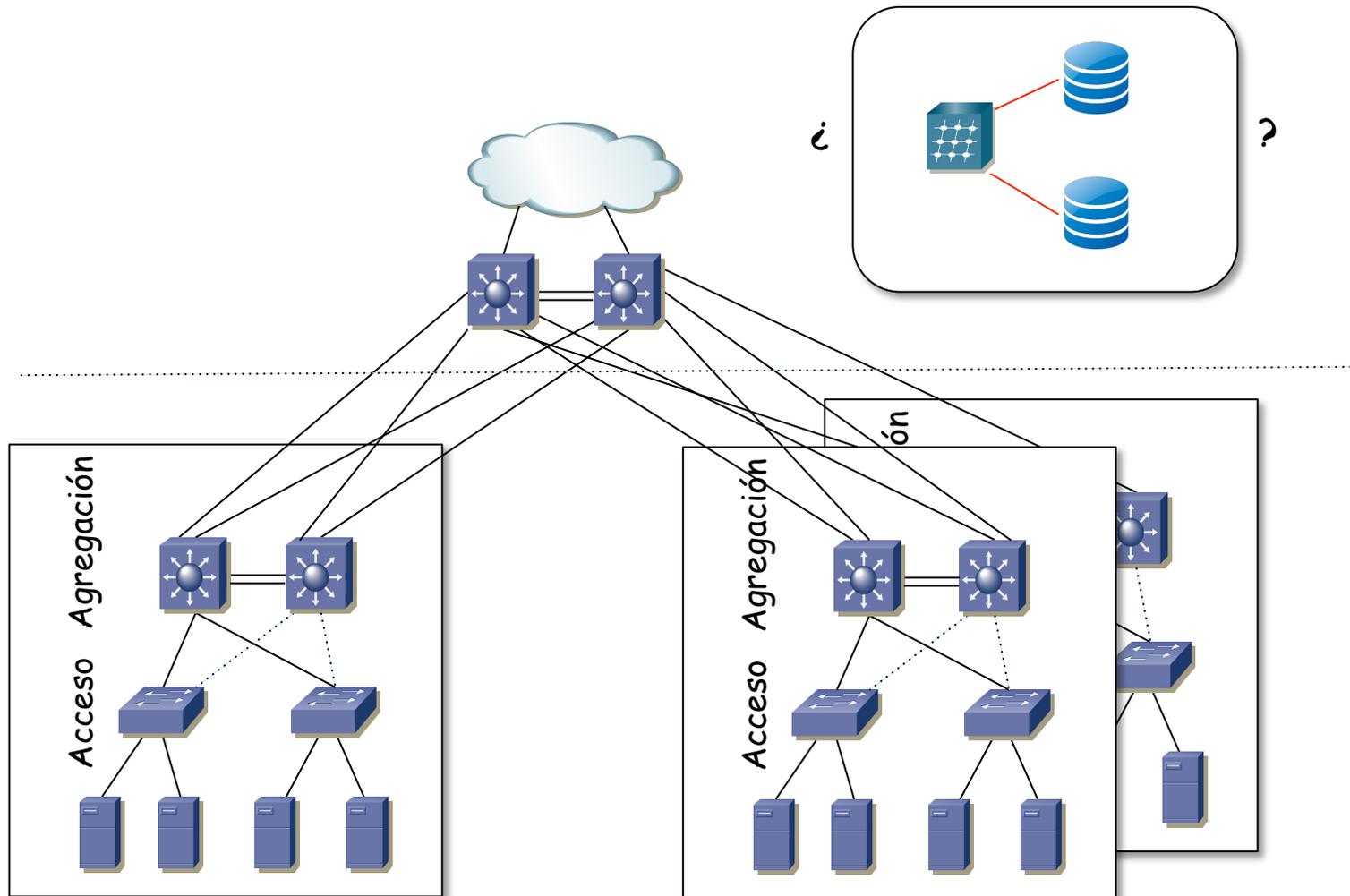
High Availability

- Y no hay que olvidarse de la redundancia
- (Retiro los interfaces de los servidores de la capa de aplicación por claridad)



Diseño

- ¿Dónde encaja esa(s) SAN(s) en la red del data center?
- Es independiente, pero hablaremos más sobre esto

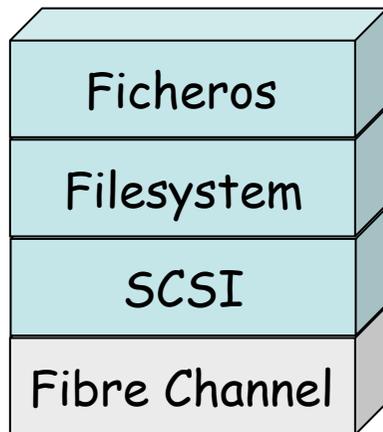


Fibre Channel

FC: Arquitectura

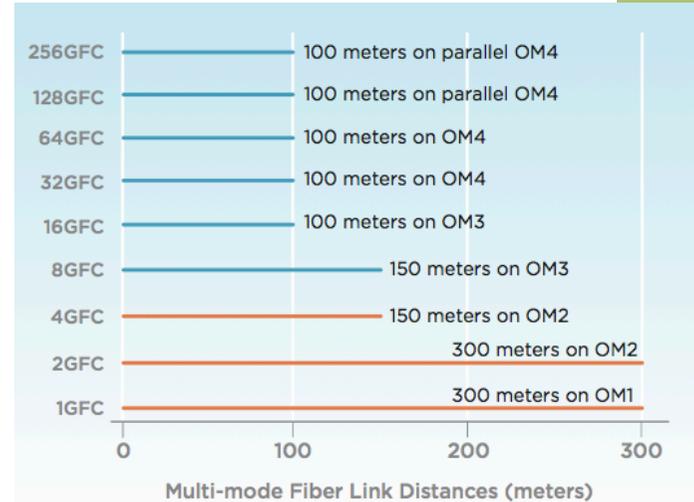
Fibre Channel

- Desarrollo comenzado a finales de los 80s
- Lo normal es que sea sobre fibra pero puede ser sobre cobre
- “*Fibre*” es el *spelling* británico en lugar del americano para “fibra”
- Soluciona el transporte pero no fija lo que transporta
- Así, puede transportar comandos SCSI pero también IP o ATM
- Inicialmente una de sus ventajas era su alta velocidad en comparación con las tecnologías LAN de la época

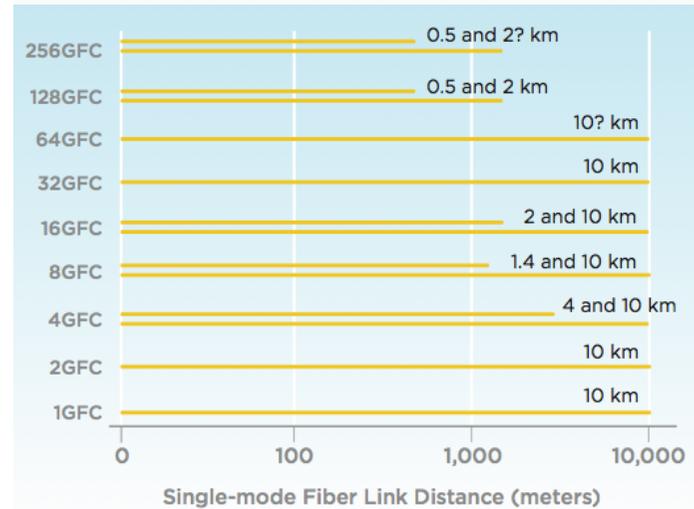


Fibre Channel

- Para los nodos el comité T11 del INCITS tiene estandarizado: 1GFC, 2GFC, 4GFC, 8GFC, 16 GFC, 32GFC y 128GFCp (que son 128Gbps mediante 4 canales)
- Para la conexión entre conmutadores tiene 10GFC, 20GFC, 40GFCoE (FC over Ethernet), 100GFCoE y 128GFCp
- En el *roadmap* se habla de hasta 1TFC en el mercado en 2020

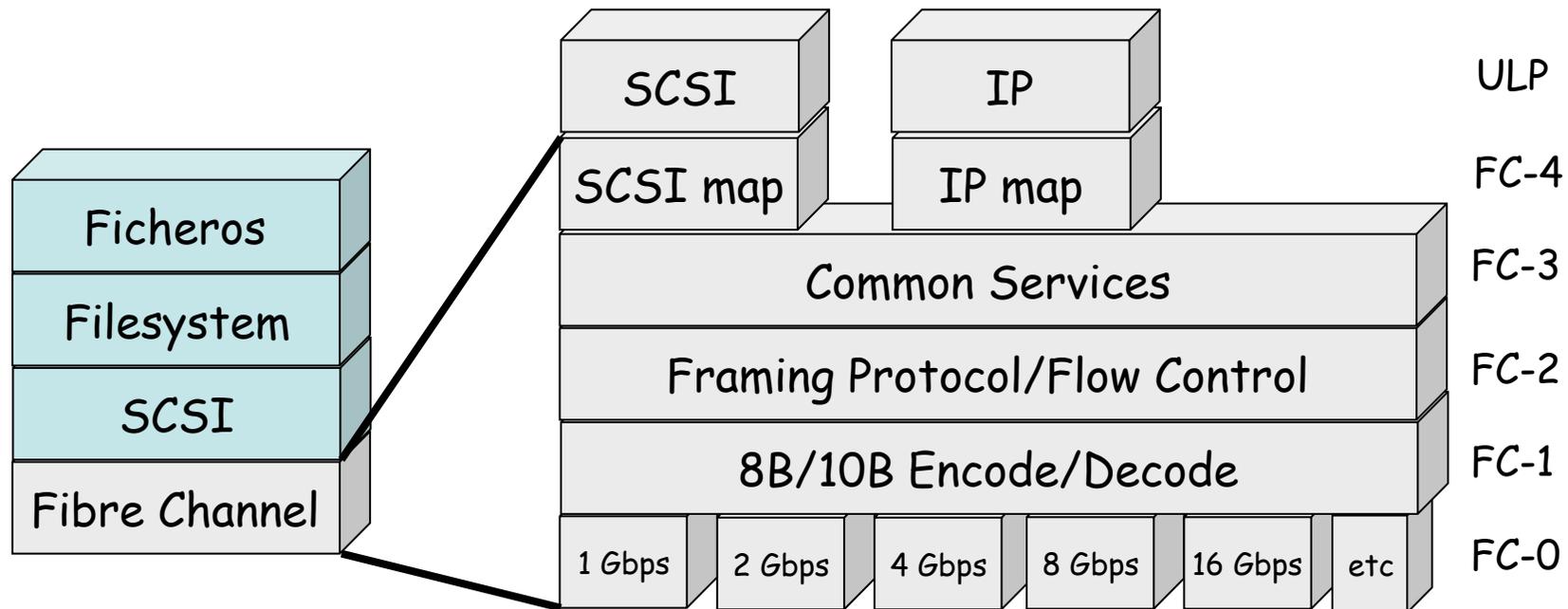


<http://fibrechannel.org/roadmap/>



Fibre Channel

- FC-0: Capa física
- FC-1: Codificación, sincronización, control de errores
- FC-2: Formato de trama, señalización para gestión
- FC-3: Ofrece un conjunto único de servicios aunque por debajo haya varios puertos físicos (*name, login, address manager, alias server, fabric controller, management, key distribution, time*)
- FC-4: Capa de adaptación para protocolos superiores como puede ser SCSI o IP (ULP = Upper Level Protocol)

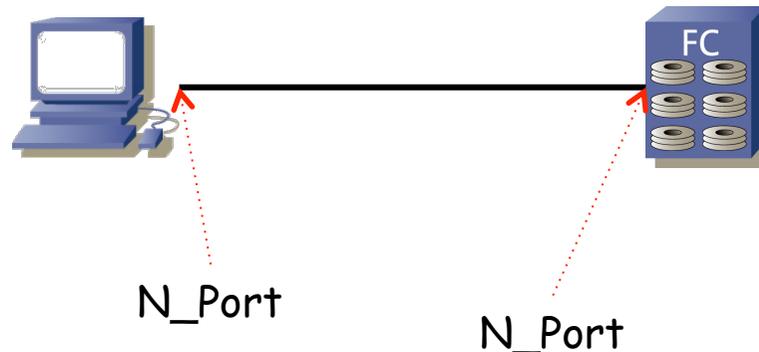


FC: Topologías

Topologías

Point-to-Point

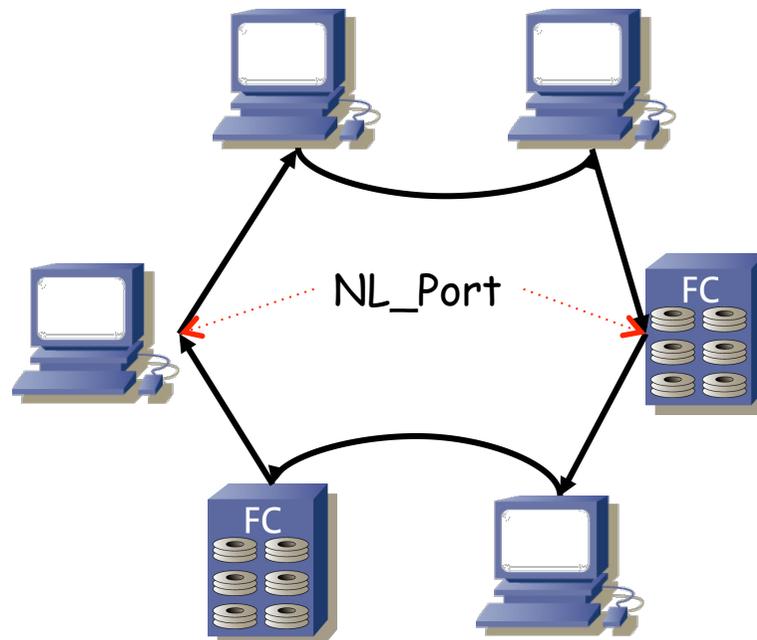
- No hace falta un conmutador o crear “una red”
- Igual que en una Ethernet, podríamos tener simplemente un enlace punto-a-punto
- Ganamos algunas de sus características técnicas pero desde luego no la de compartir el uso de los discos
- Es una conexión directa entre los puertos de 2 nodos, que se vienen a llamar “N_Ports”



Topologías

Arbitrated loop (FC-AL)

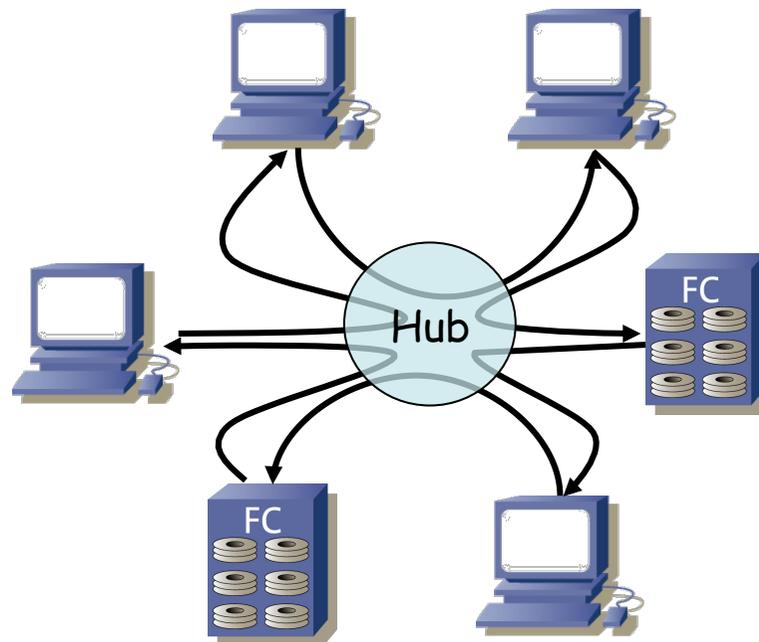
- Anillo compartido
- 2-126 dispositivos pero muchos menos por rendimiento
- Los puertos se llaman “L_Ports” (NL_Port o FL_Port)
- Se negocia cuál de los nodos actúa como *master*
- Un nodo establece un circuito entre dos puertos que monopoliza el anillo para poder comunicarse
- No puede hacerlo de nuevo hasta haber dado turno a todos los demás
- (...)



Topologías

Arbitrated loop (FC-AL)

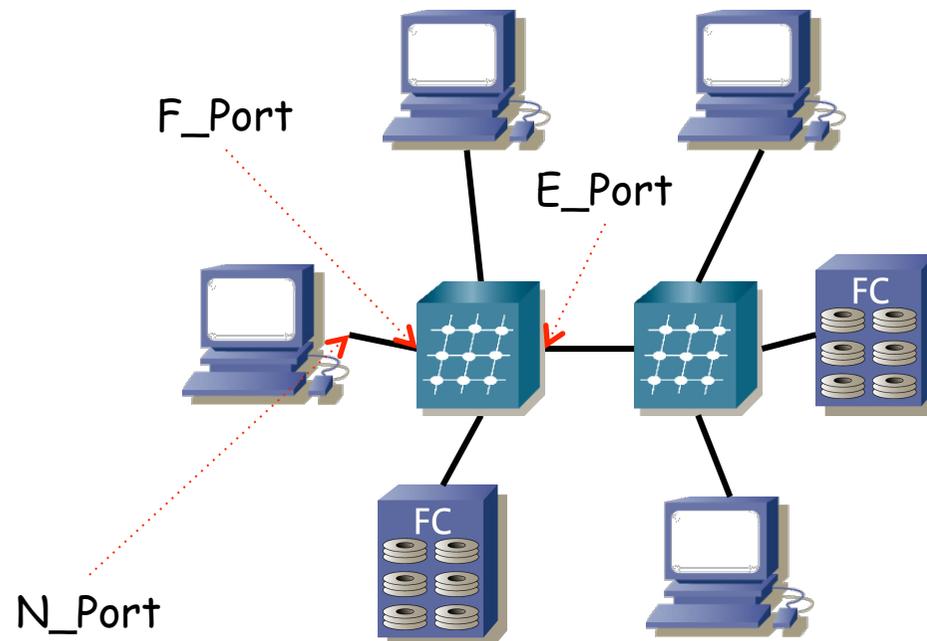
- Anillo compartido
- 2-126 dispositivos pero muchos menos por rendimiento
- Los puertos se llaman “L_Ports” (NL_Port o FL_Port)
- Se negocia cuál de los nodos actúa como *master*
- Un nodo establece un circuito entre dos puertos que monopoliza el anillo para poder comunicarse
- No puede hacerlo de nuevo hasta haber dado turno a todos los demás
- También hub



Topologías

Crosspoint switched (fabric)

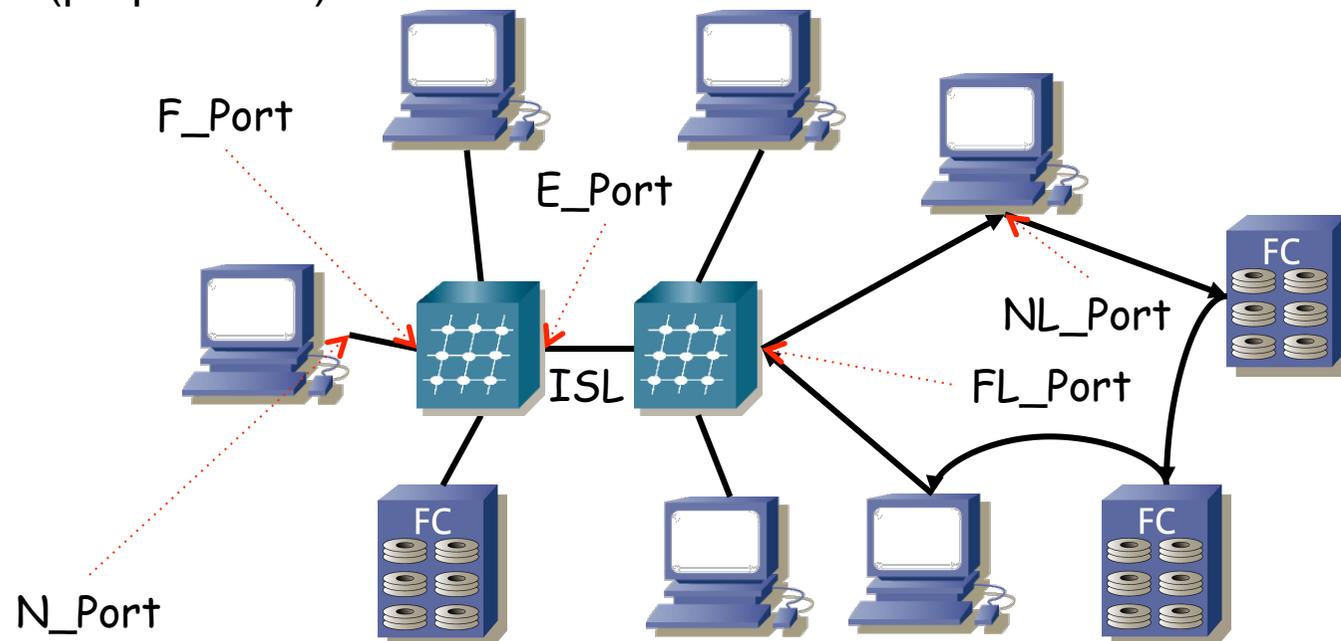
- Uno o más conmutadores, interconectando múltiples nodos
- El direccionamiento permite en teoría hasta 2^{24} nodos
- “F_Port” (Fabric Port) a los nodos, “E_Port” (Expansion) entre switches
- ISL = *Inter-Switch Link*
- (...)



Topologías

Crosspoint switched (fabric)

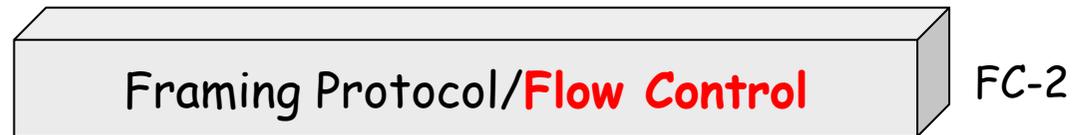
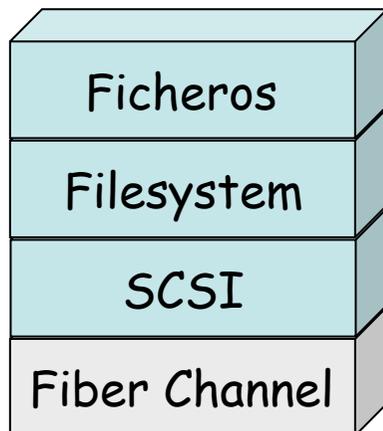
- Uno o más conmutadores, interconectando múltiples nodos
- El direccionamiento permite en teoría hasta 2^{24} nodos
- “F_Port” (Fabric Port) a los nodos, “E_Port” (Expansion) entre switches
- ISL = *Inter-Switch Link*
- “FL_Port” para conectar a un FC-AL
- “G_Port” puerto genérico que se comporta según lo que se le conecte
- Y otros (propietarios)



FC: Service Classes

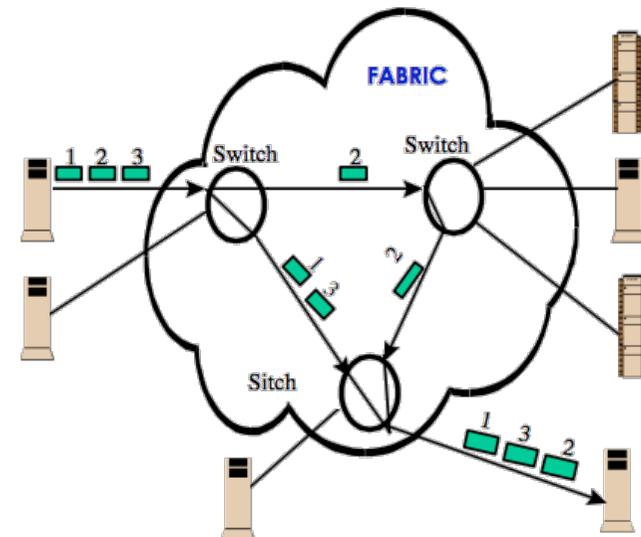
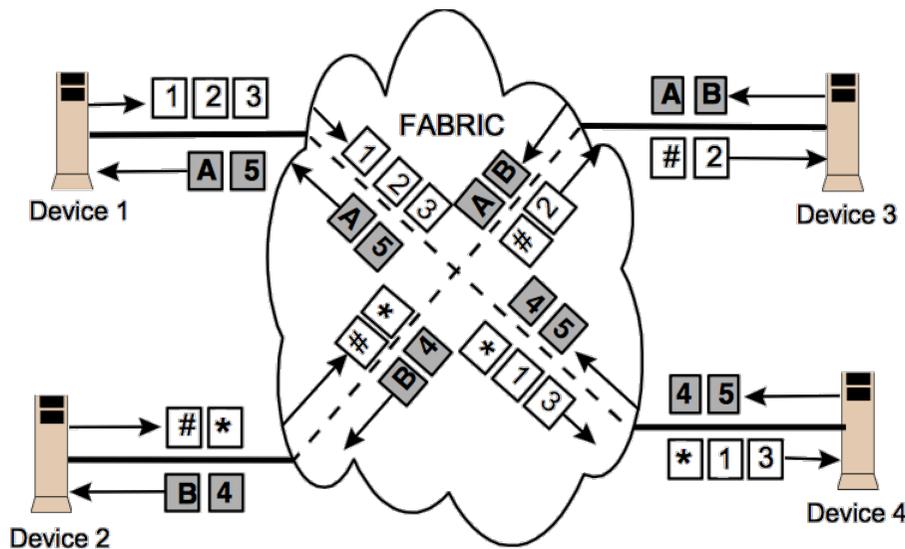
Service Classes

- Diferentes “clases de servicio” (Class 1, Class 2, etc)
- Con circuitos virtuales (dedicados o compartidos) o datagramas
- Con ACKs y NACKs o sin ellos
- Con garantía de orden o no
- Con control de flujo en cada salto y/o extremo a extremo
 - *Credit Based*
 - Básicamente una ventana deslizante
 - Mejora el rendimiento en situaciones de alta carga



Service Classes

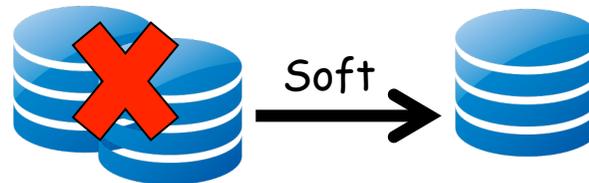
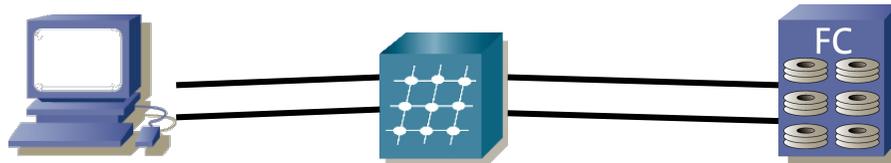
- Class 2 (*Multiplex*)
 - Sin conexión, no garantiza entrega en orden, emplea ACKs y NACKs
 - Se pueden estar recibiendo en un puerto tramas de varios nodos, así que comparten la capacidad (multiplexación)
 - Control de flujo en cada salto y extremo a extremo
- Class 3
 - Sin conexión, no garantiza entrega en orden, sin ACKs ni NACKs, lo debe hacer el ULP
 - Hace control de flujo solo salto a salto
 - Podría haber pérdidas con elevada congestión



FC: Otros servicios

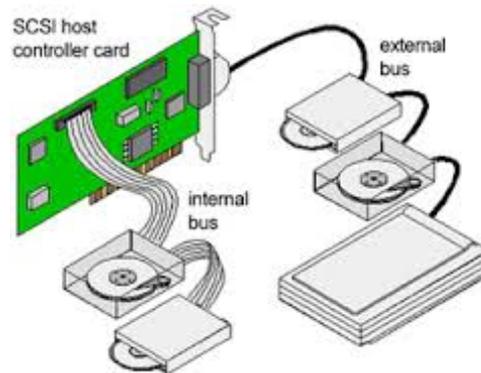
Multipath

- Permite caminos redundantes
- Varios HBAs en host y varios interfaces en sistema de almacenamiento
- FC dispone del protocolo FSPF (*Fibre Channel Shortest Path First*)
- No bloquea caminos alternativos sino que puede repartir carga
- O pueden servir como caminos redundantes ante fallos
- El sistema operativo del host requiere software que le permita ver los dos accesos al volumen como un solo volumen



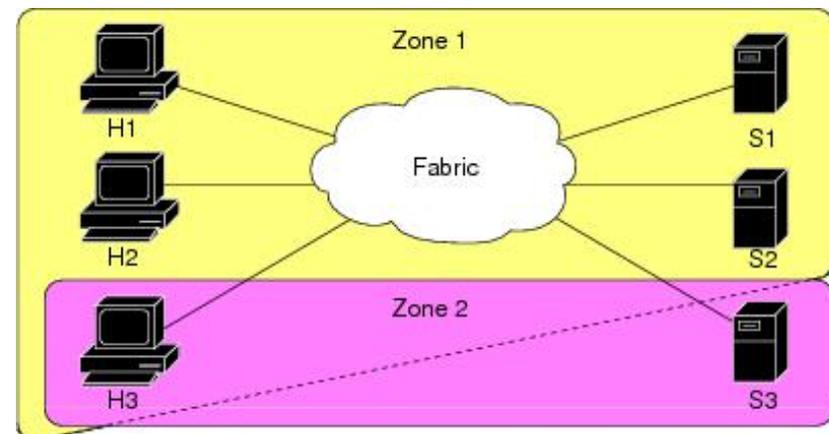
Masking/Zoning

- Tradicionalmente en un bus SCSI solo había un elemento que enviaba comandos al resto
- En realidad el protocolo permite a un dispositivo comunicarse con varios hosts
- Pero muchos sistemas operativos esperan tener el control exclusivo de cualquier dispositivo que ven en el bus



Masking/Zoning

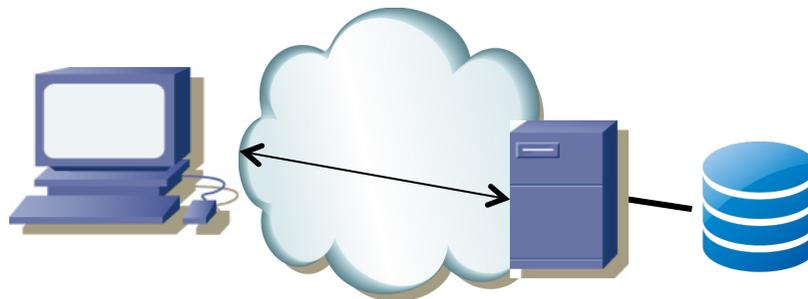
- En una SAN múltiples hosts pueden acceder al mismo dispositivo de almacenamiento
- Se puede limitar mediante *zoning*
- Una zona es como una SAN privada virtual de forma que se limita la comunicación a los dispositivos en la misma
- En el caso *Hard zoning* se crea la zona con una lista de puertos físicos de conmutador
- En el caso *Soft zoning* se emplean las direcciones de los dispositivos (WWNs, *World Wide Names*) para limitar el acceso
- *LUN masking* funciona por encima y permite que aunque el host tenga acceso a la cabina no lo tenga a todas las LUNs
- Muchas soluciones propietarias



NAS

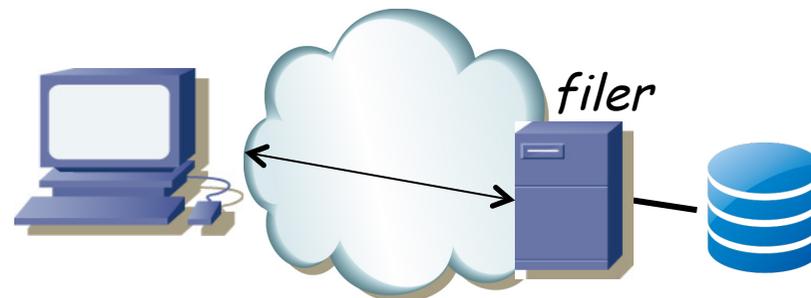
Acceso a ficheros

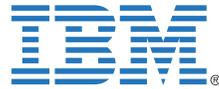
- Solución cliente-servidor, hoy en día sobre TCP/IP
- Integrado con el sistema operativo en el cliente
- Eso permite que las aplicaciones vean el sistema de ficheros remoto como si fuera local
- Ejemplos de protocolos: SMB, NFS, AFP
- Otras alternativas menos frecuentemente integradas en el cliente: FTP, HTTP (WebDav), SSH/SFTP
- El protocolo da acceso a ficheros y directorios, no a bloques



Acceso a ficheros

- Inicialmente el servidor es un hardware de propósito general que comparte parte de su disco
- Protocolos más habituales
 - SMB
 - NFS

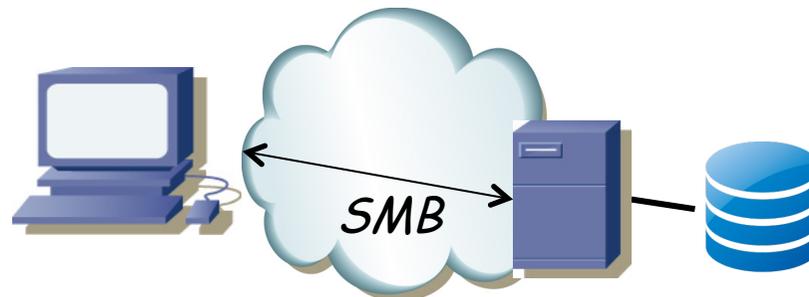




SMB



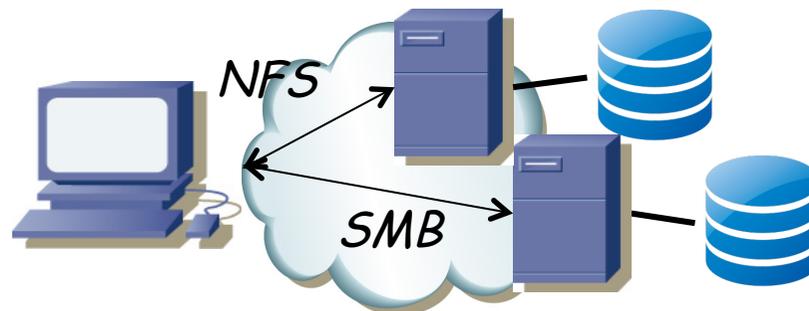
- *Server Message Block*
- Desarrollado por IBM y Microsoft
- Principalmente en sistemas Windows
- Empleado para resolución de nombre, navegar recursos compartidos, compartición de ficheros, acceso a impresoras, autenticación, etc
- Sobre TCP, UDP y otros
- SMB, CIFS (*Common Internet File System*), SMB2, SMB3, ...



NFS

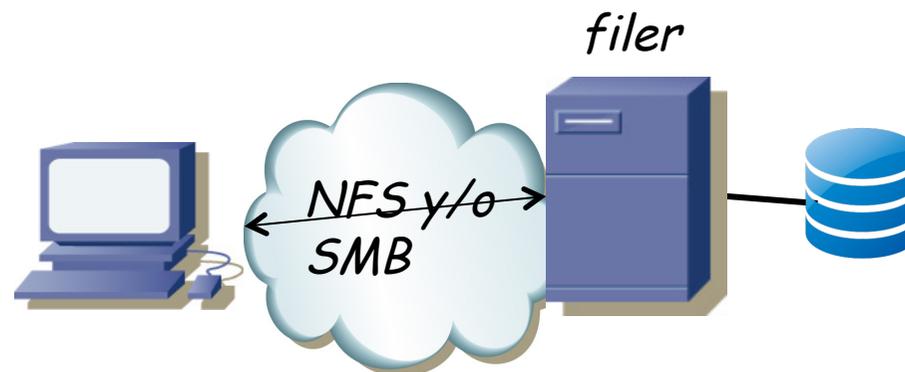


- *Network File System*
- Principalmente en sistemas UNIX (desarrollado por Sun Microsystems)
- Emplea los protocolos XDR (External Data Representation, RFC 1832) y RPC (Remote Procedure Call, RFC 1831)
- Muy común para centralizar los *home directories*
- mountd, nfsd
- NFSv2 (RFC 1094) sin estado (UDP), tamaños de 32 bits
- NFSv3 (RFC 1813) sin estado (UDP aunque también TCP), tamaños de 64 bits
- NFSv4 (RFC 3530) con estado (TCP), seguridad
- NFSv4.1 (RFC 5661) introduce *parallel NFS* (pNFS)



NAS

- NAS = *Network Attached Storage*
- Inicialmente el servidor es un hardware de propósito general
- Hoy en día hardware dedicado (*appliances*) para esta tarea
- Optimizado para ello
 - Simplifica la gestión
 - RAIDs, discos *hot-swappable*
 - Capaces de compartir los mismos recursos mediante varios protocolos simultáneamente (SMB + NFS)
 - Mejoras de rendimiento en la implementación del soft servidor
- (...)



NAS

- NAS = *Network Attached Storage*
- Inicialmente el servidor es un hardware de propósito general
- Hoy en día hardware dedicado (*appliances*) para esta tarea
- Optimizado para ello
 - Simplifica la gestión
 - RAIDs, discos *hot-swappable*
 - Capaces de compartir los mismos recursos mediante varios protocolos simultáneamente (SMB + NFS)
 - Mejoras de rendimiento en la implementación del soft servidor
- El disco en vez de ser local puede estar en una SAN

