

Escalabilidad y disponibilidad de almacenamiento onPremises

ADR - UTN - FRBA - 2020

Sistemas de almacenamiento de datos

Almacenamiento de datos

Evolución de los sistemas de almacenamiento corporativos:

- ▶ Sistemas y almacenamiento centralizado
- ▶ Cajas de discos conectados mediante controladoras de acceso directo (DAS)
- ▶ Cajas de discos accedidas mediante redes de almacenamiento (SAN)
- ▶ Cajas de discos accedidas mediante redes de datos IP (NAS)
- ▶ Sistemas de almacenamiento virtuales (vSAN)
- ▶ Servicios de almacenamiento en la nube.

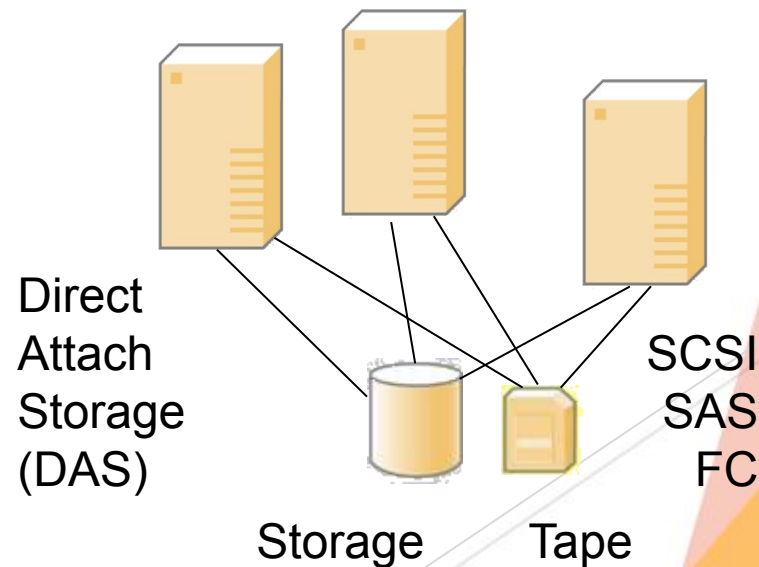
Almacenamiento de datos

Sistemas de almacenamiento accedidos de manera directa (DAS)

Accesible mediante controladoras e interfaces conectadas en forma directa a los servidores implementando protocolos de comunicación de **acceso por bloques**.

Protocolos:

- ▶ SCSI
- ▶ SAS
- ▶ FC



Almacenamiento de datos

Sistemas de almacenamiento accedidos de manera directa (DAS)

Ventajas:

- ▶ Ancho de banda garantizado
- ▶ Facilidad de configuración y mantenimiento

Desventajas:

- ▶ Escalabilidad limitada cantidad de puertos disponibles en la unidad de procesamiento que lo utiliza.
- ▶ Invisible para equipos no conectados físicamente.

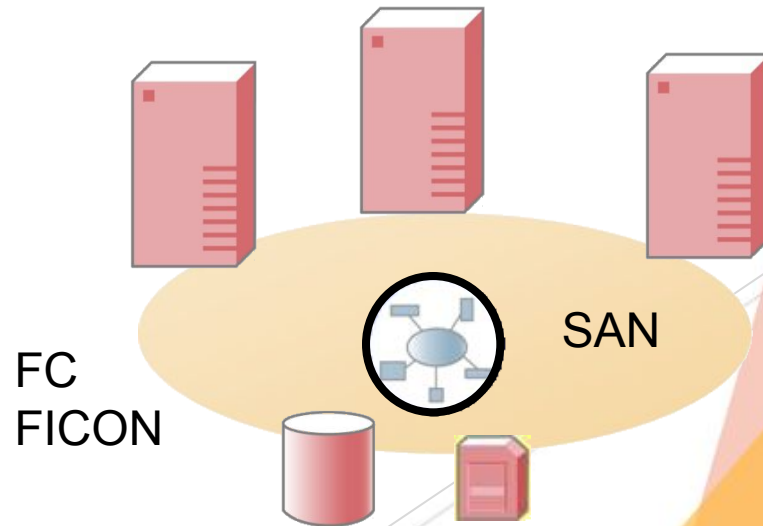
Almacenamiento de datos

Red de almacenamiento de datos (SAN)

Accesible mediante controladoras e interfaces conectadas a través de una red de almacenamiento con **protocolos de red para acceso por bloques**.

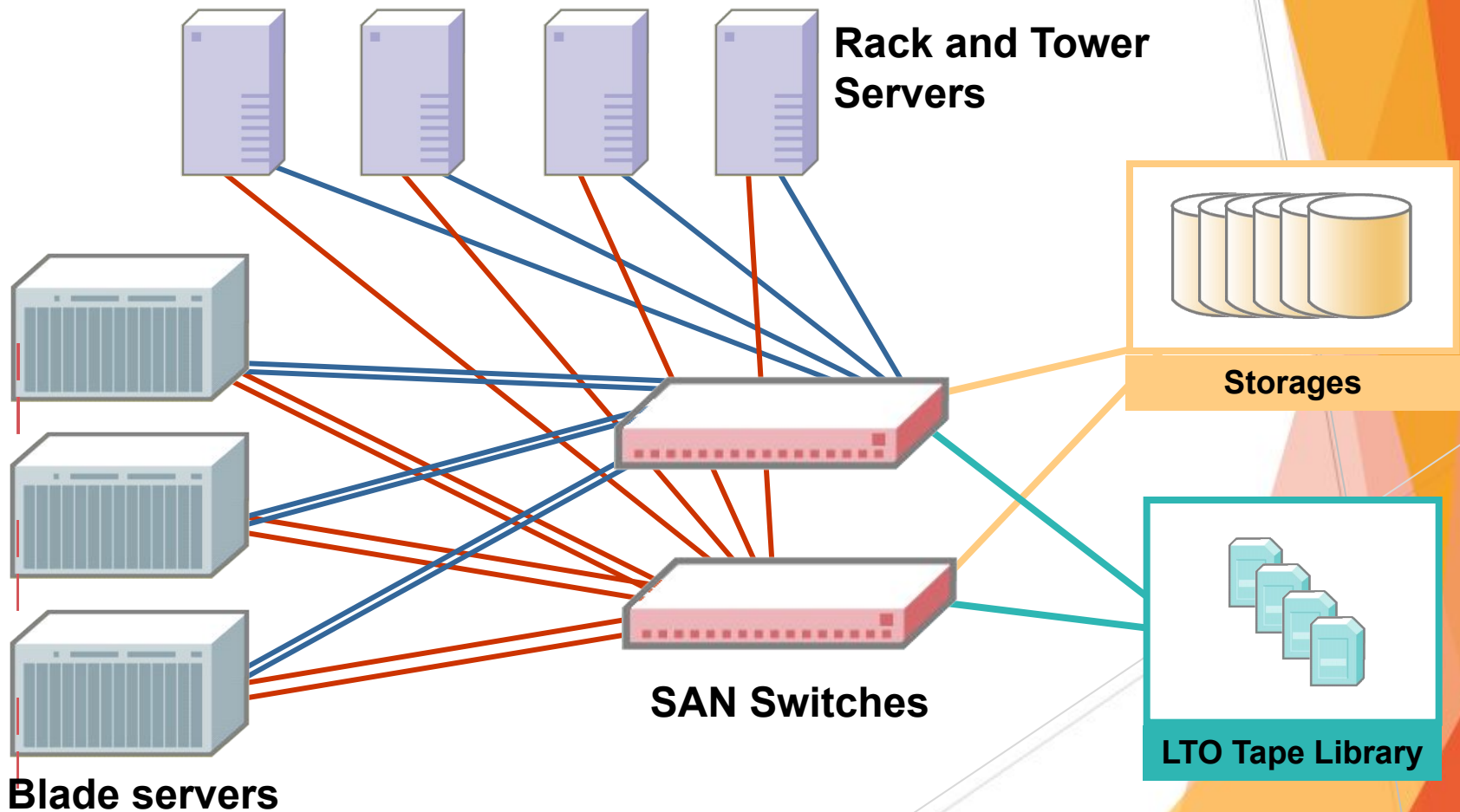
Protocolos:

- ▶ iSCSI
- ▶ FC
- ▶ FICON (IBM)



Almacenamiento de datos

Representación de una red SAN en configuración de alta disponibilidad



Almacenamiento de datos

Red de almacenamiento de datos (SAN)

Ventajas:

- ▶ Gran capacidad de ancho de banda
- ▶ Posibilidad de presentar los recursos a todos los miembros de la red SAN

Desventajas:

- ▶ Costos altos por la necesidad de componentes específicos y redundantes de la red.
- ▶ Mayor complejidad de administración, suma un componente(SAN) a la infraestructura.

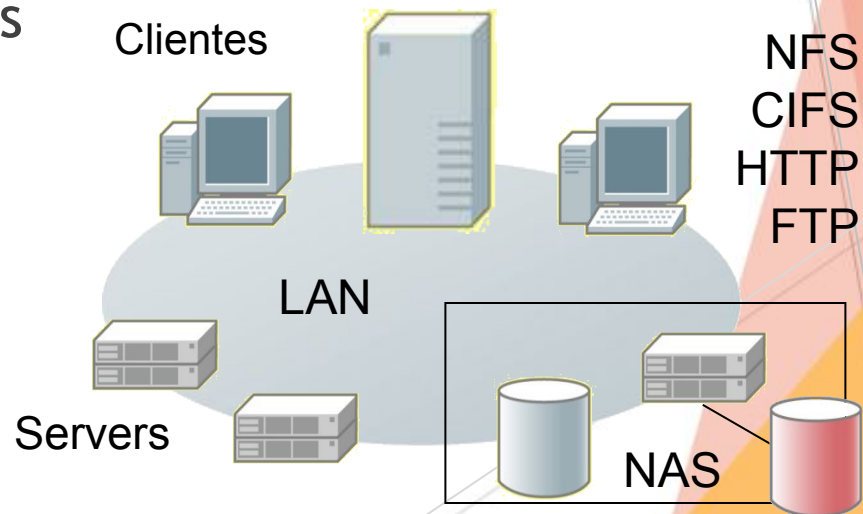
Almacenamiento de datos

Almacenamiento conectado en red (NAS)

Consiste en almacenamiento accesible por medio de redes IP de **transferencia de archivos**.

Los servicios son expuestos bajo protocolos:

- ▶ NFS
- ▶ CIFS
- ▶ HTTP
- ▶ FTP



Almacenamiento de datos

Almacenamiento conectado en red (NAS)

Ventajas:

- ▶ Administración del sistema de archivos delegada en el dispositivo NAS
- ▶ Posibilidad de acceso concurrente a los datos administrada por el dispositivo NAS
- ▶ Utilización/reutilización de arquitectura de red de datos existente

Desventajas:

- ▶ Menor rendimiento debido a las capas de abstracción
- ▶ Mayor riesgo por uso de ancho de banda en red de datos, si es compartida con la red de otros servicios.

The diagram consists of three overlapping ellipses. The top ellipse is pink and labeled 'File-level Storage' at the top and 'NAS' in the center. The bottom ellipse is green and labeled 'Block-level Storage' at the bottom and 'DAS' in the center. The intersection of these two ellipses is a brownish-tan color and labeled 'SAN' in the center. Various protocols are listed within the ellipses: 'SMB' and 'NFS' are in the pink ellipse; 'AOE', 'iSCSI', and 'FC' are in the brown intersection; 'Flash', 'PATA', 'SATA', 'SCSI', 'SAS', 'RAM', and 'FC' are in the green ellipse.

Almacenamiento de datos

Almacenamiento distribuido (horizontal)

Ventajas:

- ▶ Mayor disponibilidad, basada en la duplicación de datos distribuidos
- ▶ Mayor escalabilidad de capacidad y performance teórica (Sumando nodos al cluster de almacenamiento)
- ▶ Menor costo del equipamiento (hardware)

Desventajas:

- ▶ Mayor esfuerzo de ingeniería
- ▶ Menor tasa Volumen crudo/Volumen disponible

Disponibilidad de los datos



Disponibilidad de los datos

Disponibilidad (Availability):

Es la **posibilidad de poder continuar dando servicios** de storages ante un evento de falla en el hardware o software.

Ejemplo de elementos que aumentan la disponibilidad:

- ▶ Controladoras Redundantes
- ▶ Fuentes de energía y ventiladores Redundantes
- ▶ Switches redundantes (SAN/NAS)
- ▶ Tarjetas con dos puertos o dos tarjetas
- ▶ Protección RAID
- ▶ Discos de repuesto (spare)

Disponibilidad de los datos

Rendimiento (Performance):

Es la métrica usada para definir la velocidad de una sistema de almacenamiento.

Existen 3 métodos de medición:

- Input/Outputs per second (IOPS) - Bases de datos
- Throughput per second (MB/sec) - Streaming media
- Response Time - el tiempo que tarda en responder el almacenamiento un pedido de la aplicación (se mide en milisegundos [ms])

Disponibilidad de los datos

Soluciones para garantizar la disponibilidad de los datos

ON-LINE vs OFF-LINE

Factores para la toma de decisiones:

- Nivel de disponibilidad de los datos para su uso, en términos temporales.
 - Inmediato (on-line)
 - Diferido (off-line)
- Costos de inversión(CAPEX) y de explotación(OPEX).

Disponibilidad de los datos

Medidas de resguardo OFF-LINE

Disponibilidad de los datos

Dispositivos de almacenamientos en cinta

Tapes Drives manuales:

- ▶ 1 cabezal de lectura/escritura
- ▶ 1 cinta
- ▶ Operación manual



Tapes Drives semi-automáticos (autoloader):

- ▶ 1 cabezal de lectura/escritura
- ▶ 8 o 9 cintas
- ▶ Operación automática



Disponibilidad de los datos

Tipos de almacenamientos en cinta

Tapes Drives Tapes Drives automáticos:

- ▶ Varios cabezales de lectura/escritura
- ▶ Varios slots para cintas
- ▶ Operación automática



Tapes Drives Virtuales (librerías Virtuales VTL):

- ▶ Backup a disco que emula tecnología LTO
- ▶ Emula Varios cabezales de lectura/escritura
- ▶ Emula Varios slots para cintas
- ▶ Operación automática
- ▶ Mejora los tiempos para realizar el backup y el restore
- ▶ Utiliza técnicas de deduplicación / compresión para reducir espacio



Consultas...