Almacenamiento iSCSI

Bibliografía:

“Distributed Storage Networks. Architectures, Protocols and Management**”.**

Thomas C. Jepsen, John Wiley & Sons, Ltd.

Contenidos

1. Introducción.
2. Estándares sobre Almacenamiento IP.
3. iSCSI.

1. Introducción

- iSCSI es una tecnología de almacenamiento IP

- **¿En qué consiste el almacenamiento IP?** Transportar bloques de datos, provenientes o con destino a dispositivos de almacenamiento, utilizando Internet y el protocolo TCP/IP.

- Justificación del almacenamiento sobre IP:
  - Ubiquidad.
  - Disponibilidad.
  - Soluciones con una relación coste/rendimiento atractivas.

- Tecnologías de almacenamiento sobre IP:
  - Internet SCSI (iSCSI).
  - Fibre Channel sobre TCP/IP (FCIP).

2. Estándares

- SCSI sobre Internet (iSCSI):
  - Las primeras especificaciones sobre almacenamiento IP fueron desarrolladas por *IETF IP Sotarage Working Group*: RFC 3347 : *"Small Computer Systems Interface protocol over the Internet (iSCSI)"*.
  - *Standars for Internet SCSI (iSCSI)* (2003).
  - *iSCSI Naming Service* (2003).

3. iSCSI

- Características generales de iSCSI:
  - Mapea el modelo de procedimiento de invocación remota de SCSI sobre el protocolo TCP.
  - Las órdenes y respuestas del protocolo SCSI se transportan vía peticiones iSCSI y respuestas iSCSI.
  - Para alcanzar una mayor eficiencia iSCSI no utiliza de forma estricta la secuencia orden/data/estado/mensajes empleada por el protocolo SCSI, permitiéndolas combinaciones de estas fases:
    - Una orden y sus datos asociados pueden formar parte de un único mensaje.
    - Los datos y estado de la respuesta pueden combinarse en un único mensaje.

3.1. Arquitectura del protocolo iSCSI

- Arquitectura del protocolo iSCSI. Estructura por capas:
  - Capa SCSI: Bloque Descriptor de Orden (CDB).
  - Capa iSCSI: iSCSI PDU.
  - Capa TCP: una o más conexiones TCP (sesión).

3.1. Arquitectura del protocolo iSCSI

- Protocolo SCSI:
  - Protocolo que hace uso de un bus paralelo (SCSI bus).
  - Dependiendo del ancho del bus permite la conexión de 8, 16 o 32 dispositivos.
  - Cada dispositivo posee un ID (asociado a la interface controladora) y sobre este se definen "Unidades Lógicas" llamadas LUNs (asociadas a las unidades de almacenamiento).
  - Las LUNs pueden estar en una unidad *initiator* o *target*.
  - La unidad *initiator* es quién inicia una operación en el bus.
  - La unidad *target* es la que realiza la operación.
  - En el protocolo SCSI se permite las transferencias P2P entre un *initiator* y un *target*. Mientras se realiza la transferencia no se puede realizar otra
  - El protocolo posee un algoritmo de arbitraje que establece a quién se le cede el control del bus en cada momento.
  - Fases de uso de bus SCSI:
    - Fases de arbitraje y selección.
    - Fase de transferencia de información.
      - Fase de envío de mensaje.
      - Fase de orden.
      - Fase de lectura/escritura de datos.
      - Fase de recepción de mensaje.

3.2. Login iSCSI

- iSCSI utiliza un proceso de *login* para añadir una conexión TCP a una sesión, autenticar las partes que intervienen y negociar los parámetros de la sesión.

1. El nodo *initiator* se conecta a un puerto TCP para iniciar el proceso de *login*. Los nodos *targets* están a la escucha, en un puerto conocido, de solicitudes de conexión.
2. Se ejecuta el proceso de autenticación.
3. Se establecen los parámetros de la conexión.
4. Se entra en la fase de funcionalidades completas iSCSI: las órdenes y datos son enviados por el nodo *initiator* a las LUNs asociados a los nodos *targets.*

3.3. Principios de las sesiones i*SCSI*”

- Principio de *“Lealtad de orden iSCSI”*. Cuando una sesión se ha establecido, entonces se debe respetar este principio:

  “Para cualquier solicitud iSCSI enviada en una conexión, las correspondientes secuencias de datos y/o respuestas deben ser retornadas en la misma conexión”.

- La finalización de una conexión TCP se realiza mediante un mensaje de FIN.

- En un contexto de operación normal, la conexión TCP no se finaliza mientras se esté en la fase de funcionalidad completa.

3.4. Paquete iSCSI PDU

- Se trata de la estructura de datos utilizada por iSCSI para el envío de órdenes y datos, consta:

1. Segmento de encabezamiento básico.
   - Obligatorio.
   - Posee una longitud de 48 bytes.
     - Indica si el paquete debe ser entregado de forma inmediata (byte 0: bit 1)
     - Función del paquete PDU (byte 0: bits 2-7) (*initiator* y *targets* poseen conjuntos diferentes).
     - Final de secuencia (byte 1: bit 0).
     - Longitudes de las distintas partes del paquete (bytes 4-7).
     - Identificador de LUN (bytes 8-15).
2. Segmentos adicionales de encabezamiento (opcionales). Secuencia de números enteros de significado diverso.
3. Resúmenes de encabezamientos (*Header Digests*, opcionales). Se utilizan para controlar la integridad de los encabezamientos
4. Segmento de datos (opcional).
5. Resúmenes de datos (*Data Digests*, opcionales). Se utilizan para controlar la integridad de los segmentos de datos.

3.5. Códigos de operación del nodo *initiator* para paquetes de peticiones PDU

Código	Nombre	Significado
0x00	NOP - Out	Utilizado por el nodo *initiator* para verificar que una conexión/sesión está aún activa
0x01	Orden SCSI	El paquete PDU contiene un bloque descriptor de orden
0x02	Solicitud de función de tarea de gestión SCSI	Permite al *initiator* especificar el control de la ejecución de una o más tareas
0x03	Petición de *login*	Inicia el proceso de *login* con un nodo *target*
0x04	Petición de intercambio de información	Usado para intercambiar información
0x05	Salida de datos SCSI	Para operaciones de escritura
0x06	Petición de *logout*	Utilizado para realizar una finalización controlada de sesión
0x10	Petición SNACK	Solicitud de retrasmisión de respuestas numeradas, datos o paquetes PDUs R2T
0x1C-0x1E	Códigos específicos del fabricante	

3.6. Códigos de operación del nodos *targets* para paquetes de respuesta PDU

Código	Nombre	Significado
0x20	NOP - Out	Utilizado por el nodo *initiator* para verificar que una conexión/sesión está aún activa
0x21	Respuesta a orden SCSI	El paquete PDU contiene una respuesta de estado SCSI
0x22	Respuesta a solicitud de función de tarea de gestión SCSI	Respuesta enviada por el un nodo target después de realizar una o más tareas solicitadas por el nodo *initiator*
0x23	Respuesta a solicitud de *login*	Indica el progreso/finalización de un proceso de *login*
0x24	Respuesta de texto	Respuesta del nodo *target* a una petición del nodo *initiator* de intercambio de información
0x25	Entrada de datos SCSI	Para operaciones de lectura
0x26	Respuesta a petición de *logout*	Indica si la operación de *logout* se ha completado
0x31	Listo para transferir (R2T)	Indica que el nodo *target* está listo para transmitir
0x32	Mensaje asíncrono	Indica un estado especial en el nodo *target*
0x3C-0x3E	Códigos específicos del fabricante	
0x3F	Rechazo	Indica una condición de error iSCSI

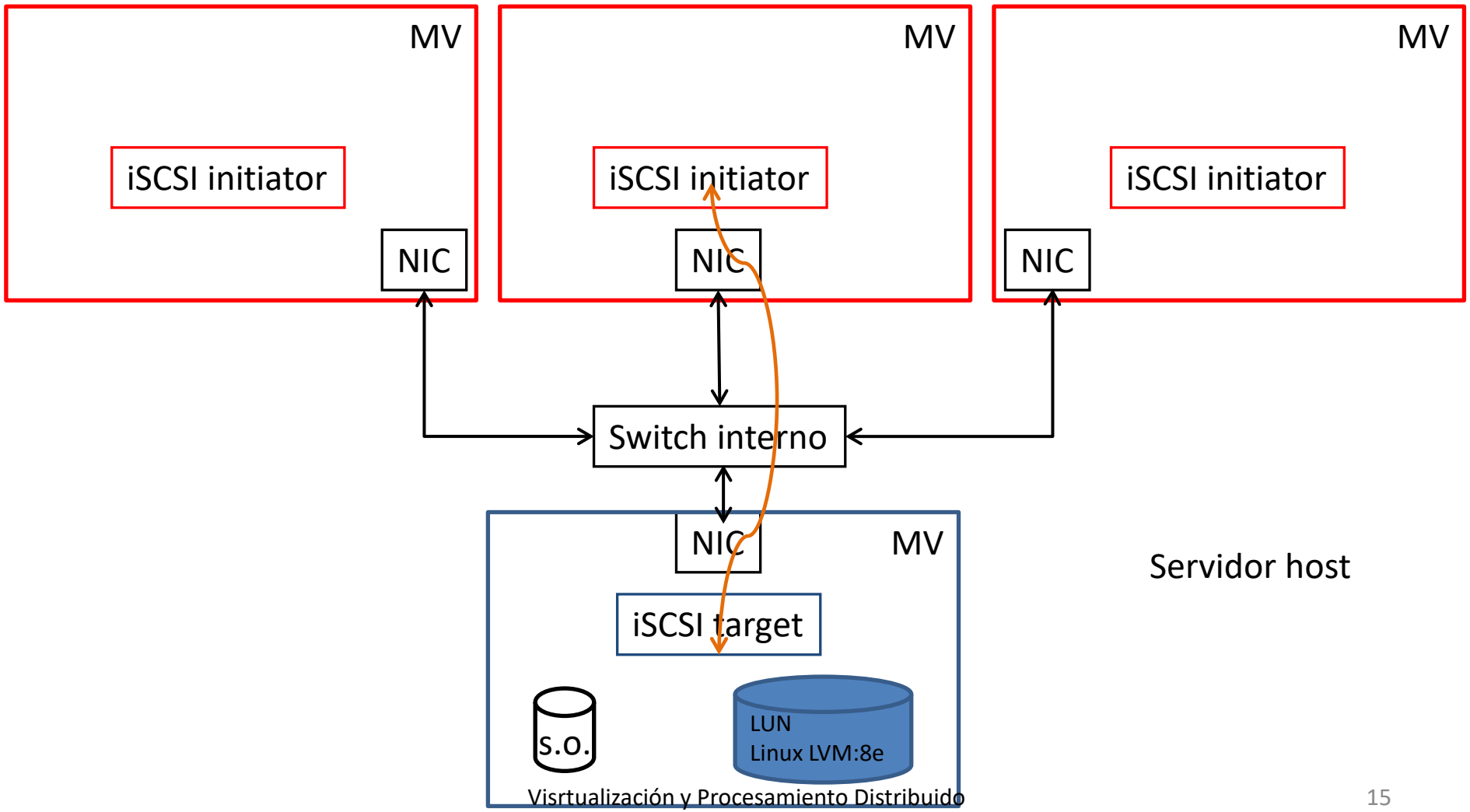
3.7. Nombres iSCSI

- Tanto los nodos *targets* como *initiator* utilizan nombres para identificar los nodos iSCSI. Un nombre de nodo iSCSI es el nombre del dispositivo SCSI del dispositivo iSCSI.
  - Los nombres están asociados a los nodos; no a las interfaces controladoras.
  - Los nombres deben identificar unívocamente a los nodos iSCSI a nivel global.
  - Un nodo *initiator* puede descubrir a nodos *targets* mediante una función de descubrimiento.
  - Los nodos iSCSI poseen sus direcciones iSCSI:
    - Nombre iSCSI.
    - + (opcionalmente) Dirección de transporte TCP (puerto).
  - Los nombres iSCSI puede tener alguno de los formatos:
    - Dirección IPv4.
    - Dirección IPv6.
    - Nombre del host (mivolumen.ejemplo.org).

3.8. Seguridad iSCSI

- El transporte de bloques de datos sobre una red IP trae como consecuencia la necesidad de utilizar mecanismos de seguridad para:
  - Asegurarse que una petición iSCSI proviene de un nodo legítimo.
  - Proteger los bloques de datos frente espionajes no autorizados.
- Se utilizan dos mecanismos en iSCSI para proporcionar seguridad:
  - A nivel de conexión : autenticación In-band
    - Nodos *targets* deben autenticar a los nodos *initiators*. Opcionalmente, los nodos *initiators* pueden autenticar a los nodos *targets*.
    - Un proceso de autenticación se realiza por cada conexión nueva iSCSI (*login* PDU).
  - IPSec (RFC2401). Para proteger los bloques de datos
    - Integridad criptográfica mediante el uso de códigos de autenticación de mensajes encriptados que se utilizan en cada paquete.
    - Autenticación de datos mediante el uso de códigos de autenticación de mensajes encriptados que se utilizan en cada paquete.
    - Confidencialidad mediante el envío de datos encriptados.

3.9. Almacenamiento compartido iSCSI en un sistema de virtualización



3.9. Almacenamiento compartido iSCSI en un sistema de virtualización

- **iSCSI initiator** – Establece la sesión y una vez cumplido el protocolo de comunicación, establece un identificador de comunicación envía por el canal comandos SCSI, encapsulados en paquetes IP. En la máquina virtual se debe ejecutar el software asociado a las funcionalidades de un nodo initiator.
- **iSCSI target** – Sirve unidades lógicas de almacenamiento LUN construidas sobre algún tipo de sistema de almacenamiento, como un disco, un array de discos, o sobre un volumen lógico LVM. Se define en el sistema servidor de disco. En la máquina virtual se debe ejecutar el software asociado a las funcionalidades de un nodo target.
- **Sesión de conexión de nodo initiator - nodo target –** Establece el canal por el que van a fluir los datos y los comandos entre el nodo initiator y el nodo target.