

DETERMINACIÓN DE LAS FUNCIONES Y SERVICIOS PRESTADOS POR LOS COMPONENTES DE LA SAN DEL CECAD DE LA UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS

Roberto Ferro E.

Ingeniero Electrónico, MSc en Teleinformática, docente de planta de la Universidad Distrital "Francisco José de Caldas", Coordinador del Programa de Ingeniería Electrónica de la Universidad Distrital "Francisco José de Caldas".
rferro@udistrital.edu.co

Carlos Mora R.

Ingeniero Electrónico de la Universidad Distrital, MS(c) Ciencias de la Información y las comunicaciones de la Universidad Distrital. Joven Investigador Colciencias. Bogotá, Colombia.
cemorar@correo.udistrital.edu.co

Bayron Villanueva O.

Ingeniero Electrónico Universidad Distrital "Francisco José de Caldas".
bfvillanuevao@correo.udistrital.edu.co

Tipo: Artículo de revisión

Fecha de Recepción: Abril 9 de 2010

Fecha de Aceptación: Mayo 14 de 2010

ESTABLISHMENT OF THE FUNCTIONALITY AND SERVICES PROVIDED BY SAN COMPONENTS AT CECAD (CENTER OF HIGH PERFORMANCE COMPUTING) – UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS

Abstract

In this article roles and services are assigned to all components of the Network Storage Center of High Performance Computing at Universidad Distrital Francisco José de Caldas. To achieve this goal, several tests were run involving the different elements of SAN, and also transfer rates were measured with the help of specialized tools.

Keywords: Storage, magnetic tape, hard disk, SAN, data transfer.

Resumen

En este artículo se asignan las funciones y servicios que deben prestar los componentes de la Red de Almacenamiento del Centro de Cómputo de Alto desempeño de la Universidad Distrital Francisco José de Caldas. Para alcanzareste objetivo se realizaron varias pruebas entre los diferentes elementos que conforman la SAN y se midieron las velocidades de transferencia con la ayuda de herramientas especializadas.

Palabras claves: Almacenamiento, cinta magnética, disco duro, SAN, transferencia de datos.

1. INTRODUCCIÓN

El uso óptimo de los elementos que conforman un centro de cómputo de alto nivel es el primer eslabón para lograr una

implementación exitosa y un desempeño sobresaliente; es importante conocer los elementos y las funciones que pueden realizar de forma más eficiente de acuerdo a sus capacidades y a las necesidades que se

busquen suplir. Aunque este artículo se encuentre enmarcado en la implementación de la SAN (Storage Área Network) para el CECAD (Centro de Computo de Alto Desempeño) de la Universidad Francisco José de Caldas, la definición de las funciones que debe realizar cada equipo que la conforma, es de gran importancia para optimizar el rendimiento y los servicios que presta dicha SAN.

El CECAD de la Universidad Distrital responde a la necesidad de realizar desarrollo e investigación, con el objetivo de mejorar la calidad de la educación de la comunidad estudiantil y habilitar una infraestructura que permita realizar proyectos de investigación a nivel de estudios de doctorado en Ingeniería.

El diseño y la conexión de la SAN fueron realizados y documentados en un trabajo de grado anterior a este artículo realizado por los mismos autores, por lo que estos aspectos no son competencia ni se tratan dentro de este artículo. [1].

La función principal de una SAN es la de proveer capacidad de almacenamiento y garantizar que la transferencia de datos entre los elementos de la misma se realice de una forma segura y robusta, además puede proveer de servicios de seguridad en los datos de alta disponibilidad en los mismos. [2]

Este documento muestra el análisis, la metodología y las pruebas que se hicieron para definir las funciones y servicios que se asignarían a los servidores y a los sistemas de almacenamiento que conforman la SAN, tomando en cuenta parámetros de velocidad de transferencia, capacidad de almacenamiento, interfaces de conexión y características de cómputo de dichos elementos.

2. METODOLOGÍA E IMPLEMENTACIÓN

Se realizaron pruebas de transferencia de archivos entre cada uno de los medios de almacenamiento con los que cuenta la SAN y en los distintos tipos de servidores que la conforman.

Se contó con dos tipos de servidores, el primero es un sistema IBM RISC 6000 S7A y el segundo es un Power6 520 de IBM, el cual fue virtualizado en dos máquinas distintas. Las pruebas fueron realizadas en uno de los servidores IBM RISC 6000 S7A y en una de las máquinas virtuales del servidor Power6 520, las características de estos equipos se pueden ver resumidas en la Tabla 1.[3], [4]

Tabla 1. Características Servidores SAN.

Servidor	Unidad procesamiento	Velocidad procesamiento	RAM	Sistema operativo
IBMS7A	12	262 Mhz	20GB	AIX 5.2
Power6	0.4	4.2 Ghz	3GB	AIX 6.1

Entre los medios de almacenamiento con los cuales se realizaron las pruebas se contó con discos locales de interfaz SCSI (Small Computer System Interface) y discos externos de conexión por FC** en cada servidor; discos de interfaz SSA (Serial Storage Architecture), y con cintas magnéticas de comunicación por FC (Fiber Channel) conectados ambos únicamente con el servidor S7A. Las características de estos medios de almacenamiento son descritas en la Tabla 2.

Los discos por conexión FC pertenecen al subsistema de almacenamiento DS5020, el cual físicamente posee 16 discos SATA (Serial Advanced Technology Attachment) los cuales fueron particionados en volúmenes lógicos de distintos tamaños y asignados a los distintos servidores que conforman la SAN[5].

Tabla 2. Características de los medios de almacenamiento pertenecientes a la SAN.

Servidor	Medio	Cantidad	Tamaño [GB]
IBM S7A	SCSI	6	4.5
	SSA	16	18.2
	SSA	16	36.4
	FC-SATA	16	200
	Cinta	180	100/200
Power6	SAS	1	146
	FC-SATA	10	500

Para realizar las conexiones entre los elementos de almacenamiento y los servidores se utilizaron dos switches de fibra cuyas características son mostrados en la Tabla 3 [6] [7].

Tabla 3. Características de los switches de fibra pertenecientes a la SAN.

Elemento	Puertos	Velocidad
Switch SilkWorm 2800	16	1 Gbps
Switch IBM 2498 B24	8	8 Gbps

La SAN está compuesta en su totalidad por distintos tipos de equipos listados en la tabla 4 y que se encuentran interconectados como se puede visualizar en la Fig. 1. Hay que tener en cuenta las velocidades de las conexiones entre cada uno de los enlaces para poder determinar las velocidades de transferencias máximas esperadas para cada prueba.

Tabla 4. Equipos que conforman la SAN.

Tipo	Nombre	Cantidad
Servidor	IMNS7A	2
Servidor	Powr6 520	1
Switch FC	SilkWorm 2800	1
Router SCSI-FC	StorageTek SN3250	2
Switch FC	IBM 2498 B24	1
Librería de Discos	IBM 7133 D40	2
Librería de Discos	IBM DS5020	1
Librería de Discos	StorageTek 1180	1

Para realizar las diferentes pruebas de transferencia entre cada uno de los medios de almacenamiento en cada uno de los dos tipos de servidores se hizo uso del comando cp del S.O AIX, el cual se encarga de realizar

una copia del archivo especificado hacia un destino como se muestra a continuación:

`cp [Archivo Origen] [Directorio Destino] [8]`

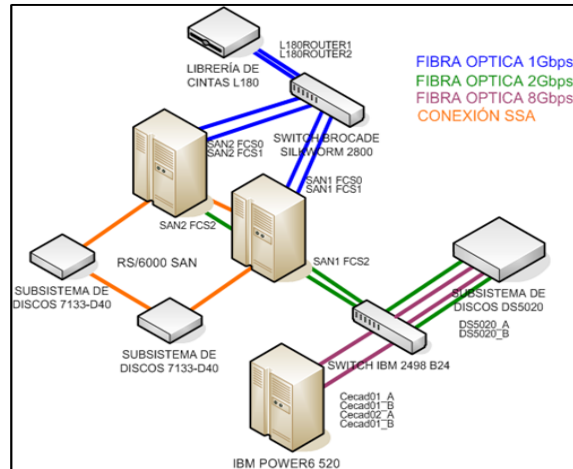


Fig. 1. Diagrama de la red de almacenamiento SAN.

En el Servidor IBM S7A, el cual tenía ligado cuatro medios de almacenamiento distintos (discos locales, discos SSA, discos FC-SATA y cintas magnéticas) se realizaron transferencias de archivos de distintos tamaños (200 MB, 500MB y 1 GB); desde cada uno de los tipos de almacenamiento ligados a dicho servidor entre sí y se realizó la comparación del rendimiento de cada uno de estos medios dentro del mismo servidor.

Adicionalmente se realizó la transferencia de un mismo archivo hacia dos cintas distintas conectadas a un mismo router para comprobar la tasa máxima de transferencia del Canal de Fibra ligado a éste último.

Además de las pruebas mostradas anteriormente se hizo hincapié en encontrar la velocidad máxima de lectura y escritura que el servidor podría obtener de los discos de interfaz de FC, para este fin se realizaron dos pruebas distintas para encontrar estos valores, la primera fue una prueba de múltiple acceso, en la cual se envió un archivo a un disco FC-SATA desde múltiples discos externos para hallar la velocidad máxima de escritura. La segunda prueba consistió

en realizar una transferencia de un mismo archivo desde un disco FC-SATA a múltiples discos externos para así comprobar la velocidad de lectura máxima posible en este medio. Estas mismas pruebas fueron realizadas para el Servidor IBM Power6 520 teniendo en cuenta que éste solo disponía de dos medios distintos de almacenamiento (discos locales SAS* y discos FC-SATA).

Para realizar todas las mediciones se utilizaron herramientas incluidas en el sistema operativo AIX que se pueden acceder a través de los comandos topas y nmon, los cuales entregan detalladamente un conjunto de estadísticas del sistema en general tomando muestras cada 2 segundos. [9] [10]

Dentro de estas estadísticas se pueden encontrar, la velocidad de escritura y lectura en los discos ligados a la máquina, el porcentaje de uso de dichos discos, la velocidad de transferencia de datos en las distintas interfaces del sistema, la memoria utilizada y el uso de CPU, entre muchas otras características que pueden ser consultadas detalladamente en los manuales de los comandos. En la Fig. 2 se puede ver un ejemplo de la ejecución del mandato topas.

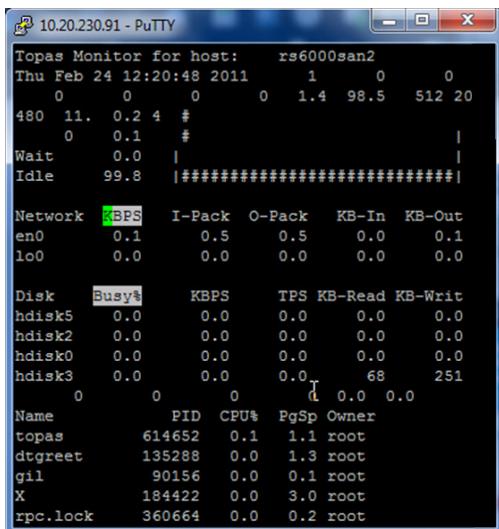


Fig. 2. Salida del comando topas.

Estas pruebas se realizaron con el fin de

poder visualizar cuales medios de almacenamiento resultaban ser los más eficientes, así como comparar la velocidad de las transferencias de datos en cada tipo de servidor, y de esta forma conocer la capacidad real de transferencia de cada uno de los elementos de la SAN y poder identificar la función principal que cada uno de estos elementos debería llegar a cumplir dentro de la misma.

3. RESULTADOS

En las tablas 5 y 6 se pueden observar las velocidades de transferencias máximas y medias que se lograron realizando una transferencia de un único archivo desde y hacia todas las interfaces de almacenamiento presentes en los servidores.

Tabla 5. Velocidades de transferencia media y máxima de los medios de almacenamiento del servidor S7A para pruebas con un archivo único

Medio	Velocidad de transferencia (KBps)	
	Media	Máxima
CINTAS	7710	7798
SCSI	9412	9892
SSA	28230	28398
FIBRA	32980	33200

Tabla 6. Velocidades de transferencia media y máxima de los medios de almacenamiento del servidor Power6 para pruebas con un archivo único.

Medio	Velocidad de transferencia (KBps)	
	Media	Máxima
SAS	96623	101192
Fibra	128830	134604

Para las pruebas de múltiple acceso y múltiple destino se tuvieron en cuenta las velocidades máximas y medias, sin embargo debido a los resultados obtenidos se dividieron para lograr un mejor entendimiento y poder realizar conclusiones más interesantes acerca del desempeño de los servidores. Las velocidades máximas obtenidas durante estas pruebas pueden encontrarse en las tablas 7 y 8.

Tabla7. Velocidades máximas de los medios conectados al servidor S7A

Medio	Velocidad de transferencia (KBps)	
	Multiacceso	Multidestino
SCSI	9892	9890
SSA	37243	37125
FC-SATA	91065	91065

Tabla 8. Velocidades máximas de los medios conectados al servidor Power6

Medio	Velocidad de transferencia (KBps)	
	Multiacceso	Multidestino
SAS	14076	96623
FC-SATA	175200	176656

En las Fig. 3, Fig. 4, Fig. 5 se puede detallar el comportamiento de la velocidad de transferencia media durante las pruebas de múltiple acceso y de múltiple destino para los dos tipos de servidores.

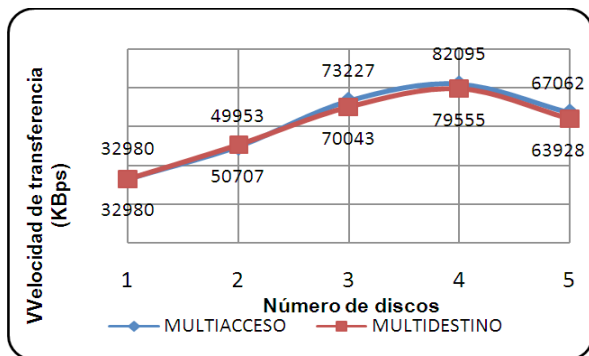


Fig. 3. Desempeño de los discos FC-SATA en las pruebas de múltiple acceso y múltiple destino en el servidor S7A

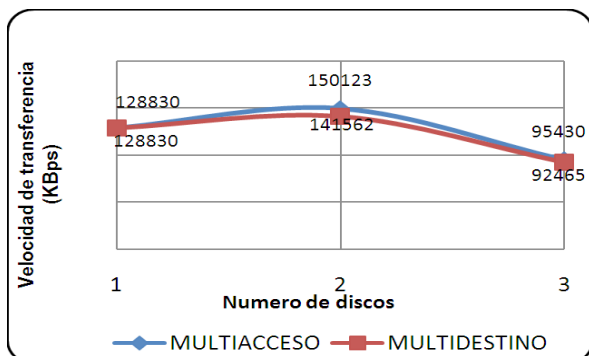


Fig. 4. Desempeño de los discos FC-SATA en las pruebas de múltiple acceso y múltiple destino en el servidor Power6.

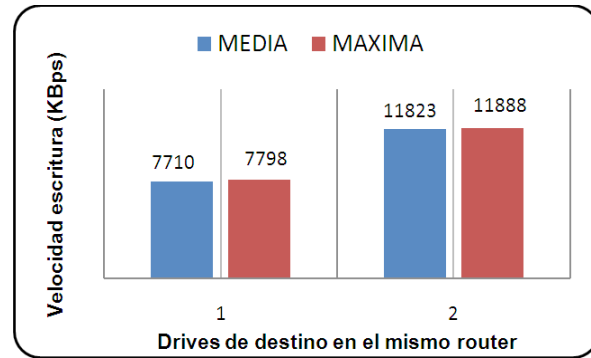


Fig. 5. Velocidad de escritura

4. DISCUSIÓN

El desarrollo de las pruebas arroja resultados interesantes que son analizados a continuación, tratando de comprender la razón de los comportamientos obtenidos y generando preguntas para nuevas investigaciones acerca de estos elementos.

Los valores máximos obtenidos para el envío de un único archivo en cada una de las distintas interfaces de discos y que pueden observarse en las Tablas 5 y 6, se obtuvieron en las pruebas realizada entre discos de la misma interfaz. Sin embargo para las pruebas hacia o desde interfaces que mostraban una capacidad de transmisión más alta no se encontró una gran disminución en la velocidad de transferencia.

Las pruebas realizadas demostraron que las velocidades de transferencia media y máxima en los distintos medios evaluados, resultaban ser independientes del tamaño total del archivo a ser transferido.

Las curvas de desempeño con respecto a las velocidades de transferencia media obtenidas para las pruebas de múltiple acceso y múltiple destino, mostraron un comportamiento similar en los dos servidores, se puede ver que la velocidad en la escritura y lectura de un disco va aumentando dependiendo de cuantos discos escriban o lean desde éste, aunque siempre existe un límite en el cual a partir de una cantidad de discos la eficiencia se ve disminuida consi-

derablemente. Sin embargo mientras para el servidor S7A es necesario leer o escribir en 4 discos para encontrar el punto de inflexión en su curva de desempeño, el servidor Power6 encuentra este punto únicamente al escribir o leer en 2 discos.

En las pruebas de múltiple acceso y múltiple destino realizadas sobre el servidor Power6, la eficiencia tanto en la lectura como en la escritura se ve reducida drásticamente cuando se comienza a leer o a escribir desde un tercer disco, siendo la velocidad de transferencia reducida en hasta un 36.5 %. Mientras que la eficiencia en la curva de desempeño de la misma prueba en el servidor S7A no se ve afectada tan pronunciadamente siendo la velocidad de transferencia reducida en un 18% a partir del quinto disco que procede a leer o escribir.

La velocidad de escritura máxima encontrada para los discos FC-SATA obtuvo variaciones significativas al realizar la transmisión desde los dos tipos de servidores, obteniendo resultados con una superioridad del 94% al realizar la transferencia desde el servidor Power6. Estos resultados no son los esperados ya que teóricamente con la velocidad de transmisión de 2Gbps que se tienen en los servidores S7A es posible alcanzar la velocidad de transferencia mostrada por las pruebas realizadas en el servidor Power6.

Según se pueden ver en las curvas de desempeño de múltiple acceso y múltiple destino (Fig. 3, 4 y 5) de los discos FC-SATA, la velocidad de escritura y de lectura en este tipo de discos se mantienen cercanas siempre la una a la otra siendo su diferencia nunca mayor al 6%, esto se cumplió para ambos tipos de servidores evaluados.

En los medios de almacenamiento en que se realizaron las pruebas, se comprobó que la velocidad de transferencia máxima se acerca a la velocidad máxima esperada, para dicho tipo de medio; esto se cumplió para

todos los medios excepto para los discos FC-SATA que teóricamente debería llegar una velocidad de hasta 300 MBps, pero que en la práctica no superaron los 180 MBps.

5. CONCLUSIONES

Teniendo como base los resultados obtenidos en las distintas pruebas realizadas a los medios de almacenamiento y los servidores de la SAN se puede ver claramente que el servidor IBM Power6 resulta idóneo para labores que requieran transferencias de archivos grandes a altas velocidades, como aplicaciones que usen el protocolo GridFTP. En cambio los servidores IBM S7A resultan más útiles encargándose de la administración de los backups en cinta de toda la SAN, ya que su velocidad de transferencia se encuentra muy limitada en comparación a la provista por el servidor Power6, pero resulta perfecta para trabajar con la velocidad impuesta por la librería de cintas.

De los resultados obtenidos se puede comprobar que el medio con la menor tasa de transferencia en todos los casos resulta ser siempre la cinta magnética, lo que implica esto es que ésta limitará la velocidad a la que trabajan los otros tipos de medios al tratar de acceder a ella, lo que hace ineficiente realizar la conexión del servidor Power6 a la librería de cintas.

A pesar de realizar tareas de múltiple acceso y múltiple destino nunca se logró alcanzar la velocidad de transmisión teórica que se puede logran en las conexiones de fibra, teniendo el mismo resultado en los dos tipos de servidores, demostrando que las tareas de transferencia se encuentran siempre limitadas por la velocidad de escritura que poseen los medios de almacenamiento.

Aunque la mayoría de resultados se encontraban dentro de los parámetros presupuestados, llamo la atención la baja capacidad de escritura que se encontró en la prueba de múltiple acceso a cinta debido a la conexión de más de un drive a los router

que realizan la traducción SCSI a FC. La velocidad de transferencia media de las cintas magnéticas suele ser muy estable aproximándose muy de cerca a la velocidad máxima en las que estas pueden trabajar lo que lo hace un medio muy confiable.

Para obtener mayores velocidades de escritura y lectura dentro de la SAN se recomienda realizar la actualización a discos de estado sólido (SSD), proponiéndose como un trabajo futuro el realizar una expansión de este artículo incluyendo este nuevo medio de almacenamiento.

Referencias Bibliográficas

- [1] C. Mora, B Villanueva, Adecuación, conexión y documentación de la red de área de almacenamiento para el CECAD de la Universidad Distrital Francisco José de Caldas. Universidad Distrital F.J.C, Abril de 2011.
- [2] J.Tate, F. Lucchese, R. Moore, Introduction to Storage Area Networks. Cuarta Edición, IBM Redbooks, p.p. 352 - 370, 1996
- [3] IBM, RS/6000 Enterprise Servers S70 and S7A: User's guide. Manuales IBM, p.p. 352, 1998.
- [4] B. Jacob, L. Dyjakon, K. Gun, IBM Power 520 and Power 550 (Power6) System builder. IBM Redbooks, p.p. 644 – 648, 2009
- [5] IBM, IBM System Storage DS5020 Storage Subsystem: Installation, user's and Maintenance Guide, 2 edition, p.p. 250, 2009.
- [6] Brocade Communication Systems. Brocade SilkWorm 2800: Hardware Reference manual. 1999.
- [7] IBM. IBM System Storage SAN24B-4 Express: Installation, Service and User Guide, p.p. 77, 2010.
- [8] IBM. AIX Commands Reference, Volume 1, a - c, p.p. 526, 2010.
- [9] IBM. AIX Commands Reference, Volume 4, n - r, p.p. 207, 2010
- [10] IBM. AIX Commands Reference, Volume 5, s - u, p.p. 386, 2010.

6. RECONOCIMIENTOS

Es indispensable expresar nuestro agradecimiento al Ing. Ricardo Andrés Torres por su colaboración y su enseñanza acerca del manejo y la administración de los equipos; al Ing. Pedro Fabián Pérez por mostrarnos la importancia del desarrollo de este proyecto y a todo el equipo del GICOG y del CECAD de la Universidad Distrital F.J.C por prestarnos las herramientas necesarias para desarrollar este documento.