

13
2ej



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
ARAGON

ADMINISTRACION Y CONTROL DEL AMBIENTE DE
DISCOS MAGNETICOS PARA MAINFRAMES IBM

TESIS CON
FALLA DE ORIGEN

TESIS PROFESIONAL

QUE PARA OBTENER EL TITULO DE:
INGENIERO EN COMPUTACION
P R E S E N T A ;
MARIO RAMIREZ FUENTES

ASESOR: ING. JUAN GASTALDI PEREZ



MEXICO, D. F.

1992



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO.

INTRODUCCION	2
CAPITULO I. ANTECEDENTES	8
Términos y Conceptos Basicos	9
Evolución histórica de la tecnología de Hardware	11
Los procesadores de transistores	12
El transistor	12
El desarrollo de unidades periféricas	14
El acceso al azar	16
Una idea que se convirtió en punto de referencia para la ingeniería	17
Penetración de IBM en el mercado	19
Cintas Magnéticas	19
Discos Magnéticos	20
Novedades IBM para la industria	21
CAPITULO II CONFIGURACIONES DE DISPOSITIVOS IBM DE ALMA- CENAMIENTO MAGNETICO	25
Tecnología de almacenamiento magnético	26
Tecnología de discos fijos 3380	27
Interfases, unidades de control y canales	30
Requerimientos físicos de la configuración de discos .	32
Requerimientos de espacio físico	32
Requerimientos de energía eléctrica	33
Requerimientos de aire acondicionado	34
CAPITULO III CENTRALIZACION DE LAS FUNCIONES DE ADMINIS- TRACION DE ESPACIOS MAGNETICOS	36
Entendiendo la administración magnética centralizada .	37
Las metas de la administración de espacios centralizada	39
Estructura organizacional	41
Administración de espacios como parte de soporte técnico	42
Administración de espacios como parte de Operación	43
Administración de espacios como parte de Planeación y Control	43
Administración de espacios como parte Independiente de la estructura	44

Efectividad de grupo	45
Proyecto de planeación de actividades	45
Los grupos de usuarios	47
Proporcionando servicio al usuario	49
Acuerdo de niveles de servicio	50
Políticas y procedimientos	58
Configuraciones óptimas de hardware	60
Asignación de trabajos a los miembros del grupo de espacios magnéticos	62
Estructura de responsabilidades de grupo	62
Responsabilidades por equipos de computo	63
Responsabilidades por concepto de nivel de servicio ..	63
Responsabilidades por área de especialización	63
El futuro de la administración de espacios	66
El papel del administrador de espacios magnéticos ...	66
El papel del usuario	67

CAPITULO IV ALMACENAMIENTO MAGNETICO EN DISCO 68

Almacenamiento físico de los datos	69
Metodos de acceso	70
Localización de archivos	71
Organización Secuencial	72
Organización Particionada	72
Organización Secuencial-Indexada	73
Organización de Acceso Directo	73
Organización VSAM (Virtual Storage Access Method) ...	74
Características de los archivos VSAM	74
Tipos de archivos VSAM	76
Catalogos VSAM/ICF	77
Definición de archivos VSAM	79
Organizaciones de archivos	83
Definición de los nombres de archivos	83
Categorías de los archivos	86
Diferencias en la administración VSAM y NO-VSAM	87
Ubicación de archivos NO-VSAM bajo MVS	88
Ubicación de archivos VSAM	89
Respaldo de Información	90
Frecuencia y número de versiones de respaldo	90
Proceso de recuperación en caso de desastre	92
Desarrollo de un plan de respaldo	93
Definiendo un plan de recuperación	95
Control y uso del espacio en disco	97
Administración del espacio magnético	97
Problemas potenciales de la <u>administración</u> de espacio	99

Los usuarios pueden ayudar en la administración del espacio en disco	100
Ubicación de espacio secundario	100
Tamaño de bloque eficiente	101
Liberación de espacio desperdiciado	102
Borrado de archivos al momento de su cierre	103
Ambientes de discos	103
Determinación de los ambientes de discos	106
Identificación de los grupos de archivos y de sus requerimientos	108
Elección de los nombres esotéricos de los ambientes ..	109
Atributos del montaje y uso de discos	111

CAPITULO V MEMORIA CACHE, DASD FAST WRITE Y DUAL COPY.. 113

La importancia de un tiempo de respuesta más rápido	114
Conceptos de memoria CACHE	117
¿ Como se realiza una función de lectura en memoria CACHE ?	117
¿ Como se realiza una función de escritura en memoria CACHE ?	118
Comparación entre sistemas con la facilidad "CACHE" y sin la facilidad "CACHE"	119
Tamaño y utilización de la memoria CACHE	120
Ventajas de la memoria CACHE	122
La facilidad de "DASD FAST WRITE"	124
Conceptos y funcionamiento de DASD FAST WRITE	124
La facilidad de "DUAL COPY"	126
Conceptos y funcionamiento de DUAL COPY	126
Ventajas de DUAL COPY	126

**ADMINISTRACION Y CONTROL DEL AMBIENTE DE DISCOS MAGNETICOS
PARA MAINFRAMES IBM**

INTRODUCCION

La era actual se caracteriza por generar, recopilar y procesar información. La cantidad de estas actividades que hoy en día se realizan no tiene precedente en la historia. Más aún, el volumen aumenta a un ritmo vertiginoso y continuará así en el futuro.

Las organizaciones públicas, privadas y los ciudadanos de muchas naciones se ven afectados y beneficiados cada vez más por la tecnología computarizada que almacena, maneja y procesa datos. Paulatinamente se aprecian más en forma directa, los beneficios de esa tecnología. Una gran parte de la información que se recopila no se computariza aún. El costo del equipo y de los programas para su manejo, decrece a un ritmo muy rápido, al tiempo que la cantidad de información que almacenan y administran las computadoras crece desmesuradamente. Puede preverse ya, que una vasta porción de los requerimientos actuales de datos se almacenarán más económicamente en archivos y bancos de datos de computadoras que en papel. No solamente se almacena información impresa o tabular; esta tecnología en desarrollo maneja cada vez más: dibujos con líneas fotografías, grabaciones de voz y otros tipos de información.

Dos son las áreas principales que constituyen la piedra angular en esta era de la explosión de la información: la tecnología de bancos de datos y la de telecomunicaciones.

A partir de los últimos años de la década de los sesenta muchas organizaciones gubernamentales, industriales y educativas desarrollaron grandes sistemas de información (manejo de cuentas, de partes, de mercados, etc). Más recientemente, se ha iniciado la integración de diversos sistemas de aplicaciones computarizadas y sus archivos asociados, dentro de sistemas de información más capaces y provechosos centrados en bancos de datos, con vías a satisfacer a las comunidades de usuarios que crecen cada vez más y son aún más sofisticadas. Han surgido muchos retos debido a la necesidad de integrar dentro de bancos de datos, archivos y datos diversos con mayor efectividad de costos, para usuarios con diferentes intereses, prioridades y requerimientos. Entre estos retos se tienen necesidades de estructuras de datos y posibilidades de organización de archivos, más capaces que las que proporcionan los sistemas de cómputo estándares a través de lenguajes como COBOL, FORTRAN, PL/1, RPG, etc.; independencia de los datos entre programas de aplicación y estructuras particulares dadas al banco; habilidad para que los

usuarios compartan datos; irredundancia de la información; alto nivel de rendimiento y eficiencia; y además seguridad integral.

En los últimos años ha surgido una tecnología esparcida y diversa para estructurar bancos de datos integrados y sistemas con esa orientación, que comprende conocimientos, herramientas, prácticas y sistemas comerciales de programación (software). En los años ochentas se presenciaron una gran cantidad de mitos en la tecnología del manejo de datos; se desarrollaron numerosos sistemas generalizados de manejo de archivos y bancos para satisfacer los requerimientos y retos de las décadas de los próximos años. En la actualidad, cada vez es mayor el número de usuarios que disfrutan los beneficios de la tecnología y conocimientos disponibles sobre bancos de datos, mismos que siguen incrementándose.

Tecnología del equipo (hardware)

La tecnología de la computación electrónica ha proporcionado dos tipos básicos de almacenes para datos; la memoria o almacén principal, en la cual puede mantenerse con efectividad de costos y en forma temporal cantidades relativamente pequeñas de datos (a lo sumo unos cuantos millones de bits) para procesarse; y los dispositivos de almacén externos tales como cintas, discos, tambores, sistemas de almacén masivos, etcétera, en los cuales pueden guardarse con un costo razonable, mayores cantidades de datos en forma permanente.

La memoria principal proporciona acceso a los datos con velocidades que se miden en microsegundos, o nanosegundos, mientras que el almacén externo, principalmente electromecánico, proporciona acceso más lento por varios órdenes de magnitud, en el rango de los milisegundos. El almacén externo es varias veces menos costoso que el almacén principal.

La pareja ideal de una Unidad Central de Procesamiento (UCP), (en inglés CPU) sería un sistema que proporcionara un almacén ilimitado, en el cual cualquier parte pudiera ser accesible sin demora al programa o usuario que la solicitara. Con la tecnología actual, a este ideal se le aproxima únicamente, la memoria principal pero su costo la limita a ser de tamaño pequeño con ve-

locidades de acceso de microsegundos y nanosegundos. Los crecientes archivos de datos muchos de los cuales contienen ya un billón de bits de información (o sea 1,000.000.000.000) pueden almacenarse con efectividad de costo en dispositivos de almacén externo exclusivamente.

Esos dispositivos son significativamente más lentos que la memoria principal (en el orden de 1/1000) pero también mucho más baratos.

Los dispositivos externos de almacén de acceso directo (en inglés DASD) y los bancos de datos en línea son el centro de los sistemas de información con esa orientación. Es asombroso el crecimiento de las aplicaciones en línea para los datos que se tienen almacenados.

Las capacidades y limitaciones de la tecnología disponible de almacenes masivos, conforman la naturaleza de la tecnología de manejo de datos y de las facilidades que se proporcionan a los usuarios.

Se espera que al menos al inicio de esta década de los noventas, la tecnología actual mejore significativamente en términos de mayores densidades de grabación, menores tiempos de acceso y de costo por bit. Sin embargo es poco probable que ocurran mejoras o cambios radicales. Permanecerá la brecha entre las velocidades de operación de memoria-CPU y los dispositivos de almacén externos. La forma y naturaleza de la tecnología actual de software para el manejo de datos no se verán afectadas radicalmente por los avances en equipo que se vislumbran. Así, los fundamentos y la tecnología de manejo de datos que se presentan y se prevén, existirán aún por muchos años; es por ello que la administración de esta tecnología, principalmente de los recursos de almacenamiento magnético continuará siendo una de las principales actividades rectoras de cualquier empresa de cómputo.

De esta manera resulta por demás esencial comprender la tecnología de almacenamiento para poder apreciar y diseñar sistemas de administración de datos óptimos. Para ello debemos primero situarnos en la época que nos está tocando vivir, es decir la época de los grandes computadores requeridos para la --

sobrevivencia de cualquier empresa que maneje volúmenes considerables de información, y evidentemente, sobrevivir no consiste sólo en adquirir e instalar un sofisticado equipo de cómputo, sino primordialmente en la manera en que se operara y administrará este equipo y todos sus recursos, siendo uno de sus principales recursos aquel destinado para el almacenamiento de todos los datos vitales de la empresa.

Día con día los requerimientos de almacenamiento de datos en línea van extendiéndose en todas las aplicaciones. Dado que los costos del almacenamiento magnético ha ido decreciendo a medida que la tecnología ha ido desarrollando dispositivos de cada vez mayor capacidad, podría parecer muy sencillo solucionar todas las implicaciones del crecimiento del volumen de información manejada con el simple hecho de comprar nuevo equipo, es decir, hardware que vaya al día con la tecnología.

Sin embargo es muy válido preguntarse ¿Cuánto nos está costando este hardware realmente?. A pesar de que el costo por megabyte de almacenamiento en línea en disco continúa decreciendo, los costos para la administración de los mismos se han disparado con el tiempo. Simplemente pensemos en el hecho de que comprar equipo adicional implica mayores requerimientos en el procesamiento de los datos, tales como mayor consumo de energía, mayor espacio físico, y sobre todo mayor cantidad de gente para instalar, mantener y administrar este nuevo equipo.

Podríamos entonces pensar que ante estos altos costos para el almacenamiento, lo mejor sería detener el crecimiento del equipo de cómputo y arreglárselas con el ya existente. Definitivamente esta **NO es una buena idea**; el crecimiento es excelente!!, es una muestra de "salud" en el negocio, es la evidencia de que los recursos de cómputo están adaptándose a las necesidades reales del cliente (el usuario) que es lo más importante para la empresa. La meta de la administración de los recursos de cómputo nunca debe ser la de limitar su crecimiento, sino conseguir el control más eficiente en su administración. No perdamos de vista que el almacenamiento de datos al igual que el personal y el dinero, es un recurso que se debe administrar con suma eficiencia; el uso ineficiente de este recurso resulta al corto plazo el más caro de todos.

La clave del éxito consiste en balancear los requerimientos del usuario (tiempos de respuesta, disponibilidad y espacio) con los dispositivos instalados para satisfacer estos requerimientos, meta que se pretende alcanzar con el presente trabajo de una manera teórica pero aplicable en cualquier empresa que cuente con el equipo citado, aunque con pequeñas variantes entre empresa y empresa acordes con su muy individual ambiente de cómputo.

CAPITULO 1
ANTECEDENTES

1.1.- Términos y Conceptos Basicos

Qué significa manejo o administración de datos?. Según Williams [1], es un término muy general y amplio que se refiere al manejo de la tecnología de computadoras requerida para organizar, almacenar y manipular datos. La unidad más pequeña de datos que se considera por lo general es el dato elemental, llamado también campo o 'atributo' por ejemplo el número de un empleado. Un conjunto de datos elementales constituye un registro lógico o entidad. Un tipo de registro lógico es un registro con una constitución particular de datos elementales; por ejemplo un registro de empleado constituido por un número de empleado, un nombre y una dirección. Un archivo es una colección de ocurrencias de un mismo tipo de registros; por ejemplo un archivo de registro de empleados.

Una base de datos es un conjunto de ocurrencias de varios tipos de registros en el cual tanto los tipos de registros como sus ocurrencias están interrelacionados mediante relaciones específicas. El término base de datos es sinónimo de banco de datos. El archivo tradicional de COBOL es un banco de datos con un solo tipo de registro. Banco de datos es un término de moda, razón por la cual muchas personas se refieren a un solo archivo grande como un banco de datos. Un banco reside típicamente en dispositivos externos de almacén de acceso directo.

La manera en que se lleva a cabo la representación, almacenamiento y manejo de la información mediante la computadora, está relacionada completamente, con los dispositivos de cómputo periférico involucrados. Por lo cual, y por todo lo anteriormente mencionado este trabajo de investigación se enfoca a los dispositivos de almacén externos, sobre los cuales se organizan las masas de información y otras aplicaciones.

La contraparte ideal de una Unidad Central Procesadora (UCP) sería un sistema que proporcionara un almacén ilimitado, o que pudiera ponerse cualquier sección de éste a disposición inmediata del programa o usuario que lo socilitara. Observando la tecnología actual, solamente la memoria o almacén principal se aproxima a este ideal; su costo la limita a un tamaño finito y requiere de cierto intervalo de tiempo de acceso.

Las facultades y limitaciones de los dispositivos masivos electromagnéticos de almacén que se han creado, conforman la naturaleza de la tecnología de manejo de datos y las facilidades que se dan a los usuarios. Es esencial comprender los dispositivos de almacenamiento para poder entender, apreciar, diseñar y usar sistemas de manejo de datos.

Diversos dispositivos llamados periféricos de entrada/salida (E/S) proporcionan varios tipos de almacenen para datos a parte del almacén principal. Los dispositivos de E/S incluyen lectoras/perforadoras de tarjetas, impresoras, máquinas de escribir y teclados, unidades de cinta magnética, unidades de acceso directo, equipo de teleproceso, unidades de entrada de datos (por ejemplo, unidades de teclado a disco), videos, sistemas para almacén en microfilm, equipo para control de procesos, y otros. No obstante, el grueso de los datos que manipulan los programas de aplicación, se guarda en dos tipos de dispositivos a las cuales se hacen referencia como dispositivos de almacén externo: unidades de cinta magnética y dispositivos de almacén de acceso directo (discos magnéticos). Notese que los dispositivos de almacén externos, en su sentido más amplio, incluyen definitivamente a sistemas que almacenan cualquier tipo de información de manera legible para la computadora: por ejemplo, los sistemas de almacén en microfilm.

Varios atributos son típicos de los dispositivos para almacén. Estos incluyen capacidad y posibilidad de dar direcciones a los datos, o cómo se localizan los datos, el tiempo de acceso de ellos, la tasa de transferencia de datos, las ventajas y limitaciones físicas y el costo. Los dispositivos se seleccionan mediante la evaluación de estos atributos contra los requerimientos de la aplicación o sistema.

1.2.- Evolución histórica de la tecnología de hardware.

Al parecer con el primer procesador electrónico, los expertos valoraron que inicialmente sólo cuatro o cinco grandes industrias en los Estados Unidos hubieran podido realmente estar interesadas en el uso de los mismos.

El primer procesador construido en serie fue instalado en 1951 y, contra los pronósticos originales, se fueron volviendo muy populares, gracias a la introducción de nuevas técnicas, nuevas unidades y nuevos métodos de programación.

De acuerdo con el manual de Xephon [2], en 1953, el número de procesadores instalados en todo el mundo era de aproximadamente 100 unidades. En 1958, en los Estados Unidos solamente, se podían contar ya 2,500 procesadores en uso.

Al final de la considerada Primera Generación de procesadores en 1958, estas máquinas habían conquistado ya la confianza de las compañías que los utilizaban.

En esos años, México tenía instalados 2 procesadores: un Sistema IBM-650 en la Universidad Nacional Autónoma de México y un sistema IBM-632 en una institución comercial.

Consideradas al principio más como instrumentos de cálculo para investigación científica y universitaria que como máquinas útiles a nivel operativo de empresas y compañías, los procesadores fueron superando su capacidad para elaborar información general, pero con las limitaciones debidas a su técnica de construcción y de programación, todavía no muy refinadas.

El uso de estos procesadores hoy en día ya no representa una "aventura" para las empresas y compañías que los instalan, sino que se convierte en una necesidad para resolver los problemas operativos.

1.2.1.- Los procesadores de transistores.

Al final de los años 50, los tubos al vacío son sustituidos por transistores en los circuitos aritméticos y lógicos de las unidades centrales. Nace de esta manera lo que es reconocido como la "Segunda Generación" de los procesadores electrónicos. Con el uso de los transistores y el perfeccionamiento de las máquinas y de los programas, el procesador electrónico se vuelve más rápido, más económico y se instalan millares de unidades en todo el mundo.

Debido al desarrollo económico mundial y al crecimiento continuo de las empresas, el empleo de nuevas técnicas de organización y dirección empresarial, los procesadores pasan de realizar tareas fundamentalmente contables y estadísticas a emprender las aplicaciones más complejas, en los campos más variados.

1.2.2.- El transistor.

Inventado en 1948 por los norteamericanos J. Bardeen, W. M. Brattain y W. Shockley, por el cual obtuvieron el premio Nobel por el mismo, el transistor es un dispositivo electrónico formado por un cristal de silicio o de germanio, al cual se han introducido átomos de diferentes materiales. De acuerdo con ciertos valores de tensión eléctrica que se le apliquen, el transistor tiene la capacidad de conducir o no conducir corriente eléctrica y representar así el "1" o el "0" que son necesarios para el funcionamiento interno de las máquinas.

Con respecto a los tubos de vacío, los transistores representan una serie de ventajas notables: tienen un costo real de fabricación mucho menor y una velocidad mucho mayor, aproximadamente 10 veces; son capaces de pasar del estado "1" al estado "0" en pocas décimas de millonésimas de segundo.

Las dimensiones del transistor son de algunos milímetros, en comparación con los centímetros de los tubos de vacío, permitiendo así construir máquinas con decenas de millares de circuitos complejos contenidos en un espacio muy reducido.

También aumenta la seguridad de funcionamiento porque los transistores operan "en frío", evitando así las fallas causadas por sobrecalentamiento, que se registraban con bastante frecuencia en los tubos de vacío. En la práctica, la duración promedio de un transistor es de 90,000 horas aproximadamente, equivalentes a más de 10 años de trabajo continuo.

Posteriormente los transistores son remplazados por los circuitos integrados cuya filosofía de funcionamiento continúa siendo la misma: 1=encendido, 0=apagado.

1.3.- El desarrollo de las unidades periféricas.

El desarrollo notable de los procesadores y las aplicaciones en este período no se debe únicamente a las características de las mejoras de las memorias auxiliares y de las unidades de entrada y salida de datos.

Las memorias de discos magnéticos, por medio del empleo de cabezas magnéticas múltiples, montadas en soportes en forma de peines, letras o números. Varias unidades de disco pueden ser conectadas al mismo procesador, dando así una capacidad de "memorización" de varios miles de millones de caracteres; es posible registrar, por ejemplo, todas las características anagráficas de los habitantes de una ciudad con un millón de habitantes y tenerlo constantemente actualizado de acuerdo a nacimientos, matrimonios, defunciones, etc.

En adición a los discos magnéticos fijos se introducen unidades en las que los grupos de discos son removibles y pueden ser sustituidos fácilmente con otro grupo de discos, en pocos segundos.

Cada paquete de discos puede guardarse en un archivero como si se tratara de un volumen normal, y su intercambiabilidad asegura una reserva prácticamente ilimitada de datos listos para ser procesados.

Cuando se planean los requerimientos de memoria para un sistema, ayuda el considerar el almacenamiento de una computadora en la forma de una pirámide. En la cima está la memoria principal del procesador, en ocasiones llamada memoria primaria o almacén y en la base las unidades de discos y de cintas.

Todos los datos y programas deben pasar a través de la memoria principal para ser procesados.

Sin embargo, como sería prohibitivamente caro usar la memoria principal para almacenar todos los datos que necesita un sistema de almacenamiento secundario, se utilizan las cintas y los discos magnéticos.

Durante muchos años, la cinta magnética aún ampliamente usada hoy en día, era la forma más común de memoria auxiliar. Este medio de almacenamiento fue tan obvio que, durante años, cuando los productores de cine querían decir "computadora" a menudo le dieron vuelta a sus cámaras para obtener fotos en close-up de rollos de cinta girando rápidamente.

IBM inventa la columna de vacío para las cintas magnéticas en 1950 que proporcionaba un colchón mecánico que impedía que la cinta se rompiera al arrancar o parar. Con este aumento en la confiabilidad, la cinta se convirtió en un medio de almacenamiento popular.

En 1952, la primera unidad de cinta magnética de IBM, la 726 podía almacenar 100 caracteres de información por pulgada de cinta.

Para 1968, la densidad de la cinta había aumentado a 1,600 caracteres por pulgada. Esta casi se cuadruplicó en 1973 que permitió al modelo 3420-8 almacenar 6,250 caracteres por pulgada.

Las mejoras en la velocidad también aumentaron. En 1952 la unidad de cinta 726 podía leer y escribir información al ritmo de 75 pulgadas por segundo. En comparación, la 3420 modelo 8 que se puso de moda en la década de los 80 s, lee o escribe a 200 pulgadas por segundo.

1.4.-El acceso al azar.

Pero a pesar de estos notables avances, la cinta magnética tiene una desventaja que es tan evidente hoy como lo era hace 30 años: la información puede ser grabada o leída sólo serialmente; por ejemplo, si una pieza de información deseada está en el extremo más lejano de la cinta, toda la cinta tiene que ser recorrida para rescatarla. A lo anterior hay que añadir el alto tiempo operativo requerido para el montaje de la cinta en la unidad.

IBM reconoció este problema en 1955, cuando introdujo al mercado una unidad de almacenamiento auxiliar alternativo: el disco magnético RAMAC (método de acceso al azar). Tenía un brazo que leía o escribía y que podía tener acceso a la información de cada uno de sus 50 discos giratorios en una fracción de segundo.

A fines de la década de los 60 s, los ingenieros de IBM buscaron un nuevo diseño para una cabeza que pudiera leer y escribir a una distancia muy cercana y que pudiera fabricarse en grandes cantidades bajo un proceso relativamente simple.

El resultado de sus esfuerzos fue una nueva tecnología conocida como "winchester" que estableció un patrón dentro de la industria.

Introducida en 1973, con la unidad de disco 3340, el diseño de la cabeza lectora/escritora de esta nueva tecnología, permitió el ensamble a sólo 18 millonésimas de una pulgada sobre la superficie del disco. La unidad de disco 3340 podía almacenar 1.7 millones de bits por pulgada cuadrada.

Si una marea creciente está rodeando la industria de las computadoras, en ninguna parte es más evidente que en el área de las unidades de entrada/salida. Porque entrada/salida (E/S) cubre esa galaxia en explosión continua de impresoras, pantallas, terminales, terminales de efectivo en los bancos, terminales para los supermercados y rastreadores heliográficos.

1.5.- Una idea que se convirtió en punto de referencia para la ingeniería.

Mucho antes de que el área que existía al sur de San Francisco fuera llamada "El Valle de Silicio", IBM estableció su primer laboratorio en la costa del Oeste, en un pequeño edificio en el centro de San José, California, para trabajar en algunas de las tecnologías que no se investigaban en el Este.

Fué en 1952 que un tema requería desesperadamente de atención: un método para que se dispusiera de datos para el procesamiento por computadora cuando estos se solicitaran.

En esa época existían tres formas de datos para dárselos a las computadoras y los tres tenían sus desventajas. La cinta magnética tenía que ser procesada en serie, manteniéndose así el caro sistema de procesamiento central en espera, mientras que en cientos de metros de cinta magnética se buscaban los datos y en el caso de las tarjetas perforadas, estas tenían que ser ordenadas antes de ser procesadas, causando retrasos similares.

Las memorias de tambor que almacenaban datos al azar, aunque demostraron ser más versátiles eran también muy costosas.

Las tarjetas magnéticas, las placas, las cintas, las bandas y los alambres, todos fueron examinados por los ingenieros del laboratorio de IBM de San José California, bajo la dirección de Reynold B. Johnson, un ex-profesor de secundaria de las materias de matemáticas y ciencias, el cual dirigía el laboratorio. En enero de 1953 Johnson tomó una decisión crítica. El laboratorio estudiaría los discos como una alternativa de almacenamiento magnético.

Al inicio reinó un gran escepticismo y el equipo de Johnson tuvo que enfrentarse a algunos complejos problemas: mantener un espaciamiento de una milésima parte de una pulgada entre la cabeza grabadora y el disco; barnizar la superficie del disco para que tuviera la suficiente suavidad para hacer lo anterior posible y diseñar un servo mecanismo eléctrico para acceso preciso de disco a disco y de pista a pista.

No fue un problema fácil, pero el 6 de mayo de 1955 IBM anunció la tecnología del RAMAC 350, un nuevo tipo de dispositivo de almacenamiento cuyo mecanismo electromagnético de lectura/escritura podía recuperar información de cualquiera de los 50 discos (uno sobre el otro) de 24 pulgadas que daban vueltas en menos de un segundo. Los ingenieros se dieron cuenta de que habían elaborado una tecnología radicalmente nueva que llevaría al procesamiento de datos en línea. Pero no podían haberse dado cuenta de que también habían creado una nueva industria. Una que para 1983 era de casi 125 compañías que generaban 12.5 mil millones de dólares derivado de la tecnología básica 350. Recientemente la Sociedad Americana de Ingenieros Mexicanos designó al RAMAC 350 (que se convirtió en una parte integral de los sistemas 305 en 1956) como "un acontecimiento histórico internacional de ingeniería mecánica". El primer producto de cómputo en recibir dicho premio.

1.6.- Penetración de IBM en el mercado.

Existen en la actualidad una serie de empresas internacionales dedicadas al desarrollo y comercialización de grandes cantidades de software y hardware tales como HITACHI, IBM, MEMOREX, UNYSIS, etc, tan solo por citar algunas de ellas; sin embargo no todas han logrado conquistar y penetrar el mercado a tan grande escala como la hecho IBM en los últimos años.

IBM es una de las empresas dedicadas al desarrollo de hardware y software a gran escala más importantes del mundo. Su crecimiento y penetración el mercado de la computación ha rebasado cualquier expectativa incluso de los propios directores de hace algunas décadas.

Su mercado abarca desde kits decodificadores simples hasta grandes macroprocesadores como el caso de los 3090; desde un juego o programa educativo hasta complejos sistemas operativos como el MVS.

Practicamente no existe campo de la cibernética que IBM no haya explotado o se encuentre en proceso de investigación. En los subsecuente se tratarán algunas de las inovaciones tecnológicas IBM que han tenido una aceptación considerable dentro del mercado.

1.6.1 Cintas magnéticas.

Uno de los productos más recientes en el mercado de las cintas magnéticas es el nuevo subsistema IBM 3480, anunciado en marzo de 1984. Esta nueva generación de unidades de cinta reemplaza el familiar carrete de cinta por un cartucho de fácil manejo que es aproximadamente la cuarta parte del tamaño de un carrete estándar de 10.5 pulgadas. Tiene una densidad de grabación de 38,000 caracteres por pulgada, seis veces la densidad de las cintas tradicionales. La información del procesador central puede ser transferida a la cinta 3480 o recuperada de ellas a una velocidad de hasta 3 millones de caracteres por segundo, más del doble de la velocidad a la que las cintas lo hacían. Esta velocidad ha sido lograda con ayuda de una nueva cabeza que graba en película delgada, siendo esta unidad una de las primeras en utilizar la tecnología de cabeza de película delgada.

1.6.2 Discos magnéticos.

El disco que estableció el estándar para los sistemas pequeños fue también en 1970 para cargar microprogramas en un controlador de disco, el disco flexible o "floppy" que se parece a un disco de fonógrafo. La Computadora Personal puede almacenar hasta 1.4 millones de caracteres de información en un nuevo disquette de alta capacidad.

En la medida que la capacidad de almacenamiento se ha incrementado, su costo ha descendido. En 1956, un millón de bytes almacenado en un RAMAC/350 costaba 130 dólares hoy el usuario paga aproximadamente 1.25 dólares por la misma cantidad de almacenamiento en discos modelo 3380.

No sólo ha crecido la capacidad de almacenamiento de los archivos de disco durante los años, bajando los costos de almacenamiento, sino que se han dado también ganancias impresionantes en la velocidad de acceso a los datos almacenados en los discos. La información en un disco es almacenada en pequeñas regiones magnetizadas en vías concéntricas. La información se escribe en el disco y se recupera de él por una cabeza grabadora que se mueve radialmente a través del disco que da vueltas.

Al principio las cabezas grabadoras eran armadas con material de ferrita. Eran convertidas en electromagneto enviando corriente a través de un alambre que las envolvía. Pero para fines de la década de los 70's usando procesos tales como la fotolitografía y el blindaje, los ingenieros de IBM habían desarrollado cabezas grabadoras en "películas delgadas" Esta tecnología permite que el archivo del disco IBM 3380 lea y escriba datos a la velocidad de 3 millones de caracteres por segundo.

El almacenamiento de discos se ha vuelto infinitamente más útil en las últimas décadas por los avances en la arquitectura de los sistemas. Por ejemplo, muchos de los sistemas de las computadoras de hoy utilizan un concepto que se desarrolló a través de muchos años y fue introducido por primera vez en gran escala en el S/360 de IBM en 1972. Conocido como almacenamiento virtual, permite que la información sea intercambiada automáticamente entre la memoria principal y el almacenamiento del disco. En efecto, el intercambio ocurre tan rápidamente que el almacenamiento de disco parece ser una extensión de la memoria principal. Estos conceptos de diseño de sistemas avanzados han

beneficiado al usuario permitiéndole obtener una velocidad de computación más rápida y sin embargo el costo es comparable al más barato de los almacenamientos en disco.

Las alternativas para la memoria no terminan con la cinta auxiliar y con las unidades de disco. Muchos usuarios necesitan almacenamiento a largo plazo y a costo bajo de una gran calidad de datos como archivos de seguros o documentos del gobierno. Los diseñadores de los sistemas de computación de hoy siguen buscando mejores formas de almacenar la información para los usuarios cuyos requerimientos van de lo minúsculo a lo masivo.

1.6.3. Novedades de IBM para la industria.

El manual "The MVS/XA Storage Management Library" de IBM [3] proporciona la siguiente información sobre desarrollo de productos IBM:

SUBSISTEMAS Y LENGUAJES		RPG FORTRAN	APL HASP ASP	IMS PL/1 CICS	JES2 JES3 DB2	SQL RACF	
SISTEMAS OPERATIVOS	OS/360	CP/CMS	VM/370 MVS/370	VM/XA MVS/XA		MVS/ESA	
I/O	DASD	REMOVABLE DISK	LASSER PRINTER WINCHESTER DISK		CARTRIDGE TAPE		
MEMORIA REAL		DYNAMIC MEMORY CELL		CAPACIDADES DE 64KB A 256 MB		SRAM SINCH WAFER	
DISEÑO	S/360	MULTI- PROCESSOR	VIRTUAL STORAGE	DYADIC RISC			
ARQUITEC- TURA	S/360	S/370	SNA	XA	SAA	ESA/ 370	
	1960	65	70	75	80	85	88

Tal vez la principal razón por la que IBM tiene una gran presencia tanto en el mercado de mainframes como de periféricos (dentro de estos últimos principalmente dispositivos de almacenamiento magnético) (ver FIGS. 1.1, 1.2, 1.3, 1.4) es que responde de una manera eficiente a los requerimientos principales de una institución entre los que se pueden mencionar:

- Alto potencial de crecimiento
- Necesidad de Integrar la Información Institucional
- Distribución y Acceso a toda la información institucional
- Mezcla compleja de procesos y usuarios
- Altos volúmenes de datos
- Integridad y seguridad de todos los recursos de Informática
- Necesidades complejas de conectividad.

A continuación se enlistan las áreas en las que IBM ha tenido una particular influencia:

Administración Pública
Instituciones Bancarias
Petroquímica
Manufactura
Distribución
Proceso
Seguros
Transporte
Hospitales

Supermercados

Comunicación Masiva

Educación

Servicios

Finalmente en la figura 1.5 podemos apreciar la distribución de los ingresos anuales de IBM por continente y por conceptos de ventas y servicios.

INDUSTRIA DEL PROCESAMIENTO DE INFORMACION
COSTOS DE LOS CLIENTES

43%	SERVICIOS PROFESIONALES	AMERICA	56%		
	SERVICIOS DE TELECOMUNICACIONES		ASIA - PACIFICO	20%	
	SOFTWARE				
	MANTENIMIENTO		EUROPA MEDIO ORIENTE AFRICA	35%	
57%	EQUIPO DE TELECOMUNICACIONES	U.S.A			40%
	IMPRESORAS				
	MEMORIA				
	ESTACIONES DE TRABAJO				
	PROCESADORES				

Fig. 1.5

CAPITULO 2

CONFIGURACIONES DE DISPOSITIVOS IBM DE ALMACENAMIENTO MAGNETICO

2.1 Tecnología de almacenamiento magnético.

No cabe duda que han sido los discos magnéticos los dispositivos que han sufrido las mayores transformaciones en los últimos años comparándolos con otros dispositivos tales como las cintas y los cartuchos magnéticos. En ellos se han observado mejoras muy significativas en aspectos tales como densidades de grabación, tamaño físico, y tiempos de respuesta sumamente pequeños debido al incremento de velocidad de giro de los discos y del movimiento de las cabezas lectoras.

La tecnología de cintas en cambio, aunque ha sufrido transformaciones y mejoras, ha permanecido mas estable durante los últimos años, conservándose el tamaño del grosor de la cinta en media pulgada a pesar de la gran variedad de densidades de grabación, formas y velocidades de enrollado.

Sin embargo, cualesquiera que sea el tipo de dispositivo utilizado, cualquier subsistema de almacenamiento debe de cubrir integralmente los siguientes requisitos:

- . Almacenar datos de manera "permanente" o por el tiempo que sea necesario.
- . Mover datos a y desde el CPU por medio de comandos o instrucciones.

Si no cumple con estos 2 criterios, entonces el producto no es un subsistema de almacenamiento magnético.

Una vez cumplidas las condiciones básicas, el dispositivo de almacenamiento debe de considerar algunos aspectos relevantes:

- . El acceso de los datos debe de cumplir las necesidades de las aplicaciones que los utilizarán.
- . Los tiempos de acceso de los datos debe ser aceptable para la aplicación.
- . El precio debe ser aceptable.

La tecnología se ocupa de mejorar estos 3 últimos aspectos, ya que podemos observar que una perforadora de tarjetas y un operador trabajando en conjunto cumplen con las 2 premisas básicas, pero difícilmente cubrirán el segundo conjunto de requisitos que hoy día son ya considerados como indispensables para la mayoría de las empresas.

2.1.1 Tecnología de discos fijos 3380.

Un dispositivo de almacenamiento directo (DASD, Direct Access Service Device) consiste básicamente de un disco giratorio con una superficie magnética y una cabeza de lectura/escritura que se puede mover radialmente sobre la superficie del disco. Los datos son grabados en pistas (tracks) circulares que se encuentran distribuidos a distancias fijas unas de otras alrededor del eje. En un mismo eje varios discos son colocados con cabezas independientes para la lectura de cada uno de ellos. Normalmente los datos son grabados en ambas caras del disco aunque las superficies superior e inferior de todo el conjunto no son usadas para almacenar datos debido a que existe peligro de polvo y suciedad que pudieran causar errores de lectura y escritura. Todas las pistas de todas las superficies a una misma distancia radial son conocidas como un cilindro. (Figura 2.1)

La lectura o escritura de los datos en el disco es una operación que consiste de 3 pasos: primero, la cabeza de lectura/escritura se posiciona sobre la pista apropiada; al tiempo utilizado para esta primer operación se le conoce como tiempo de búsqueda y puede involucrar un movimiento que va desde la pista cero a unos pocos cientos mas de pistas; una vez que la cabeza está posicionada adecuadamente, puede pasar algún tiempo adicional mientras el registro requerido rota delante de la cabeza de lectura/escritura; esta demora se llama demora rotativa o tiempo de latencia y puede implicar tanto como una rotación completa; el componente final es el tiempo de transferencia de datos y representa el tiempo necesario para llevar a cabo la transferencia real de los datos entre el disco y la computadora y es un tiempo que comparado con el de búsqueda y el de latencia es sumamente menor.

EL DISCO MAGNETICO

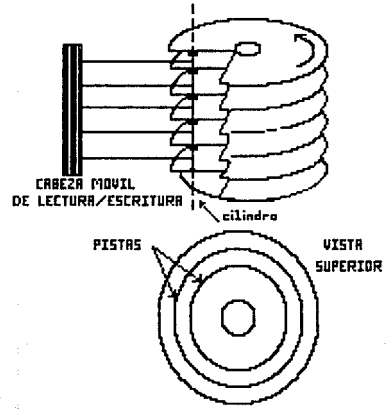


fig 2.1

Los discos de acceso directo modelo 3380 cuentan con un tiempo promedio de respuesta de 32 milisegundos, los cuales podemos descomponer como se muestra en la figura 2.2 en una configuración de discos y un ambiente de cómputo bien afinado (esta separación de tiempos puede variar drásticamente en configuraciones de discos desbalanceadas):

TIEMPO DE RESPUESTA DE UN DISPOSITIVO 3380			
ESPERA EN EL PROCESADOR (5 ms)	TIEMPO DE SERVICIO DEL DISCO (27 ms)		
(5 ms)	TIEMPO DE BÚSQUEDA (18 MS)	TIEMPO DE LATENCIA (8.8 ms)	TRANSF. DE DATOS (0.2ms)

fig. 2.2

El tiempo de búsqueda a su vez se puede dividir en tiempo de posicionamiento de la cabeza (seek time) y lo que es propiamente el desplazamiento de la cabeza sobre la superficie (search time).

Estos datos fueron tomados del manual de Proaks [5] para una empresa de aceros del estado de California en los Estados Unidos de Norteamérica en un ambiente de cómputo bien afinado que manejaba dispositivos 3380.

Con este tiempo de repuesta, un disco 3380 tiene una capacidad de transferencia de datos de 3 millones de bytes por segundo y su tiempo de rotación es 16.7 milisegundos por vuelta (60 giros por segundo !).

En cuanto a capacidades, estas varían de modelo a modelo en la familia 3380; la diferencia entre un modelo y otro consiste básicamente en un incremento de la cantidad de cilindros por dispositivo; hasta esta fecha IBM ha fabricado 3 modelos 3380: los de densidad sencilla, los de densidad doble y los de triple densidad. En estos tres modelos se mantiene la relación de 15 pistas (tracks) por cada cilindro, variando el número de cilindros de modelo a modelo como se muestra en la figura 2.3:

CAPACIDADES DE LOS DISCOS 3380

	DENSIDAD SENCILLA	DENSIDAD DOBLE	DENSIDAD TRIPLE
CILINDROS/DISCO	885	1770	2655
PISTAS/CILINDRO	15	15	15
BYTES/PISTA	47,476	47,476	47,476
MEGABYTES/DISCO	630	1260	1890

fig. 2.3

Generalmente dos volúmenes de discos conforman lo que es conocido como un HDA (Hard Disk Unit) y un mueble o unidad completa de discos contiene dos HDA's, es decir, cuatro volúmenes 3380; un "string" o "cadena" de discos consta de 4 unidades completas lo que nos da un total de 16 volúmenes 3380 por cada "string" de discos lo que en capacidad se ve reflejado desde poco más de 10 hasta los 30.24 Gigabytes por string dependiendo del modelo .

2.2 Interfases, unidades de control y canales.

Es interesante observar la cantidad de dispositivos externos que se pueden conectar a una computadora, entre ellos, los de almacenamiento secundario. En una configuración compleja de dispositivos de almacenamiento, sería sumamente difícil para el ordenador central controlar todo el código de instrucciones requeridos por los dispositivos de almacenamiento con los que realizan el posicionamiento de sus mecanismos de lectura/escritura para el acceso de datos y mas aún si se tienen diferentes modelos ya que cada dispositivo tiene patrones diferentes (por ejemplo un disco IBM modelo 3380 y otro MEMOREX modelo 3350).

La solución a este difícil problema consiste en proporcionar un mecanismo entre la computadora y sus dispositivos periféricos para que "traduzca" el código y el procesador central se pueda dedicar a otras actividades mas prioritarias. Esta unidad de interfase es conocida como "unidad de control" o simplemente "controlador". Cada controlador es diferente y depende del modelo y tipo de dispositivo; su función consiste en traducir las características únicas de un dispositivo específico a una forma común.

Conforme la computadora es mas grande y veloz (como es el caso de una 3090) y conforme se añaden más y más dispositivos de E/S a la estructura principal, el controlar cada una de las operaciones de E/S llega a convertirse en un verdadero problema. La solución a esto es el canal, que en lo esencial es una minicomputadora de propósito especial que se conecta entre uno o mas controladores y la computadora (fig. 2.4).

Las funciones del canal consisten en contar los caracteres y actualizar las direcciones entre la memoria principal, tomando sobre si la responsabilidad de controlar una operación de E/S y liberando a la unidad central de proceso para que ésta realice otras tareas.

La mayoría de las computadoras tienen dos tipos diferentes de canales; el primero llamado un multiplexor, esta proyectado para conectarse a dispositivos de E/S de baja velocidad tales como lectoras de tarjetas e impresoras; el canal multiplexor permite

CONEXION DE LOS DISPOSITIVOS DE E/S A LA COMPUTADORA CENTRAL

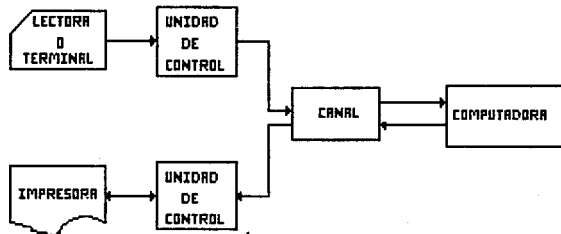


fig. 2.4

que docenas de dispositivos de E/S de baja velocidad se manejen concurrentemente, traslapando o multiplexando sus operaciones y realizando varias operaciones de E/S simultáneamente. El segundo tipo de canales es conocido como canal selector, al cual se le conectan dispositivos de E/S de "alta" velocidad (comparados con las impresoras o lectoras de tarjetas, ya que si los comparamos con los ciclos internos de la computadora son en realidad bastante lentos) tales como las unidades de cintas magnéticas y las de discos magnéticos (fig. 2.5); al contrario de lo que sucede con el canal multiplexor, un selector está proyectado para manejar la transmisión de datos entre la memoria principal y un solo dispositivo de E/S a la vez. Debido a que estos dispositivos son relativamente rápidos, las ventajas del traslaje o el multiplexado de las operaciones de E/S no son las mismas. En efecto, un canal selector sirve como una trayectoria de datos de alta velocidad que se conecta a un solo dispositivo de E/S con la computadora y en la cual los datos se transmiten en una forma conocida como "modo ráfaga".

Ahora que se han analizado los diferentes componentes de una configuración de un sistema de cómputo, ya pueden ser conectados y administrados (fig. 2.6), particularmente los próximos capítulos serán enfocados al análisis de la administración de los dispositivos de E/S de acceso directo, es decir, de los discos magnéticos.

En la figura 2.6 podemos observar que el sistema comienza con la computadora en si misma, la cual está dividida en tres partes: la unidad central de proceso (CPU), los registros y la memoria; dentro de la memoria se encuentran tanto instrucciones de programa como datos; la CPU a su vez se divide en dos partes: la unidad de control y la unidad aritmética y lógica. La unidad de control es la responsable de sacar una instrucción de la memoria y codificarla (en otras palabras, de determinar lo que va a hacer el programa). El control se transfiere luego a la unidad aritmética y lógica, la cual ejecuta la instrucción especificada; a continuación se devuelve el control a la unidad de control en donde se repite el proceso.

Externamente a la computadora, se encuentran conectados los dispositivos de E/S de la manera en que ya se ha discutido.

TIPOS DE CANALES PARA I/O

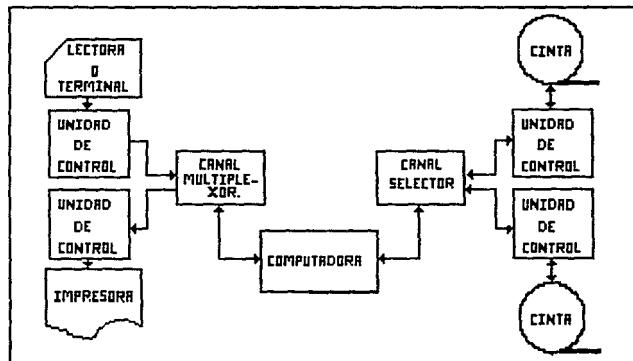


Fig. 2.5

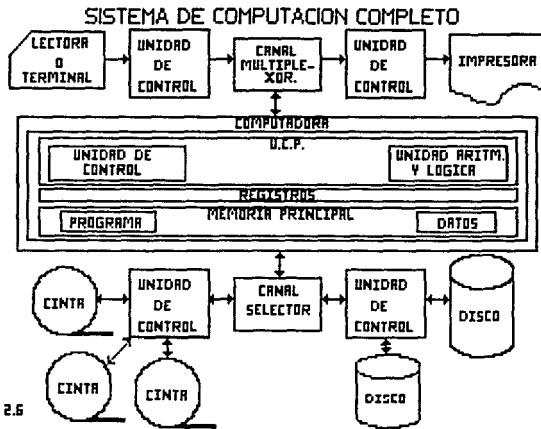


Fig. 2.5

2.3. Requerimientos físicos de la configuración de discos.

Así como son de suma importancia aspectos tales como el performance, la capacidad y la disponibilidad de los discos una vez que son instalados en una configuración, la propia instalación física juega un papel predominante para la consecución de todos los anteriores; el ambiente físico en el cual es instalado el equipo puede influir en su comportamiento drásticamente.

A continuación se mencionarán brevemente los principales aspectos que deben ser considerados antes de proceder a la instalación de la configuración; no se realizará sin embargo un análisis detallado de los mismos ya que no está contemplado dentro del alcance del presente trabajo de investigación y el sólo estudio de aspectos tales como los requerimientos de seguridad física de la instalación podrían constituir un nuevo trabajo completo de investigación, quizás aún mas grande que el que se está presentando.

2.3.1. Requerimientos de espacio físico.

En una configuración de un sistema de cómputo, los requerimientos de espacio físico incluyen:

- Accesibilidad, resguardo y seguridad del equipo.

El equipo requiere ser accesible al personal que se encarga de intervenirlo, ya sea por mantenimientos preventivos o correctivos o para su propia operación; debe estar al mismo tiempo protegido contra las impurezas del medio ambiente o contra daños operativos accidentales; debe finalmente cubrir las normas mínimas de seguridad que garanticen la salud e integridad física del personal que tiene contacto con el equipo.

- Espacio físico ocupado por el equipo.

El sitio que se elija para instalar el equipo deberá ser lo suficientemente grande para toda la configuración que sea colocada.

- Espacio auxiliar de almacenamiento.

El local seleccionado para instalar el equipo no deberá limitarse a ser suficiente para que quepa el sistema sino que debe considerar requerimientos futuros de crecimiento, por lo menos para los próximos tres años.

- Conveniencia operativa.

El equipo deberá estar lo suficientemente cerca para que pueda ser atendido por el operador o el técnico en caso de requerirlo y a la vez lo suficientemente lejos (o aislado) para que no dañe al personal en caso de desastre.

2.3.2. Requerimientos de energía eléctrica.

El equipo de cómputo requiere una fuente de poder lo suficientemente grande como para que soporte a toda la configuración, libre de interferencias o perturbaciones del medio ambiente. Generalmente las compañías de electricidad son las que se encargan de proporcionar la totalidad de la potencia requerida por el equipo de cómputo, sin embargo es aconsejable contar con una fuente alterna conocida como "no-break" que sustituya en forma instantánea al suministro original en caso de presentarse alguna interrupción; recordemos que un segundo de interrupción es suficiente para tener grandes pérdidas o errores con la información que se estaba procesando en ese momento. También es recomendable que se cuente con un sistema de energía adicional e independiente al del equipo de cómputo para ser utilizado por el sistema de alumbrado y de aire acondicionado, esto último para evitar una posible sobredemanda de potencia que pudiera poner en peligro a toda la instalación misma.

Deberá tenerse presente además el riesgo de posibles problemas por malas conexiones de tierras y proteger al equipo contra fuentes de electricidad estática.

2.3.3 Requerimientos de aire acondicionado.

Un sistema de control de la temperatura y humedad del medio ambiente resulta indispensable para el adecuado funcionamiento del equipo de cómputo; recordemos que cualquier dispositivo del sistema genera calor y el aire acondicionado debe mantener el medio ambiente dentro de los límites tolerables por el equipo de cómputo.

Algunos de los factores que deben considerarse para determinar los requerimientos de aire acondicionado son los siguientes:

- Disipación total de calor por el equipo de cómputo.
- Número de personas en el área.
- Requerimientos de alumbrado del área.
- Cantidad de aire fresco introducido.
- Recalentamiento del aire en circulación.
- Conducción del calor a través de muros y ventanas.
- Altura entre el piso y el techo.
- Área total del terreno ocupado por el equipo de cómputo.
- Número y posición de las puertas.

Todos los requerimientos físicos para la instalación de un equipo de cómputo deben ser considerados muy seriamente antes que cualquier otra cosa; una vez que se cuenta con una área controlada y bien instalada se puede proceder a comenzar a distribuir las cargas de trabajo en los dispositivos de almacenamiento de una manera óptima; en los capítulos siguientes se presentarán diversas propuestas para una adecuada distribución y administración de la información.

Una vez instalada la configuración de discos físicamente, se procede a su administración desde un punto de vista lógico, es

decir, balancear cargas de información y administrar los datos contenidos en los dispositivos de almacenamiento para garantizar así los niveles de servicio que reclaman todos los usuarios de nuestros días.

En los capítulos subsecuentes se plantean diversas alternativas para conseguir lo anterior, partiendo de la base de una administración centralizada de los recursos de almacenamiento magnético, tema de análisis del siguiente capítulo.

CAPITULO 3

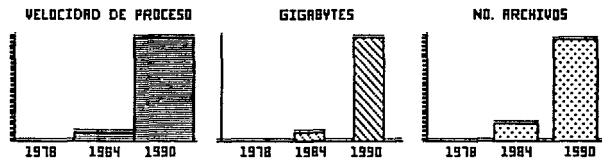
**CENTRALIZACION DE LAS FUNCIONES DE ADMINISTRACION DE ESPACIOS
MAGNETICOS**

3.1 Entendiendo la administración magnética centralizada.

El acelerado crecimiento de los recursos de cómputo incluyendo aquellos de almacenamiento magnético presenta nuevos retos para la mayoría de los centros de cómputo de nuestros días. El uso de dispositivos de almacenamiento de acceso directo (DASD's) se está incrementando a una tasa mayor al 50% por año en muchas organizaciones mientras que los requerimientos de velocidad y memoria de los procesadores están creciendo a una tasa del 60% por año; la figura 3.1 muestra la tendencia de crecimiento de los recursos de almacenamiento y de la capacidad de los procesadores durante los últimos años. Aunado a esto, los usuarios demandan cada día una mayor eficiencia para la utilización de sus datos; esto significa que el número de personas requeridas para administrar los espacios magnéticos se incrementa de manera paralela con el crecimiento de los dispositivos de almacenamiento. Bajo esta perspectiva el costo de los recursos humanos requeridos para administrar los recursos magnéticos puede comprometer los beneficios del crecimiento de los centros de cómputo, a menos que haya cambios en las técnicas de administrar estos recursos ya que como se comentó en la introducción de este trabajo, la gente es el recurso más caro de cualquier empresa de cómputo.

Algunos problemas cotidianos requieren de urgente atención, entre algunos de los más comunes que se pudieron detectar se encuentran los siguientes:

- Los recursos de almacenamiento no son siempre utilizados eficientemente, con desperdicios que frecuentemente llegan hasta el 50%.
- La productividad se ve impactada por frecuentes faltas de espacio en disco que originan reprocesos de las aplicaciones.
- Los usuarios de aplicaciones interactivas se quejan de los tiempos de respuesta del sistema.
- Los archivos no siempre están disponibles cuando se necesitan.
- Los problemas de seguridad de datos están creciendo con el incremento del número de usuarios.



RECURSO	1978	1984	1990
VELOCIDAD PROCESO (MIPS)	4	29	289
GIGABYTES (DE DISCOS)	9	102	1157
NUMERO DE ARCHIVOS	4.5K	44K	250K

**TENDENCIAS DE LOS RECURSOS DE
LOS CENTROS DE COMPUTO.**

Fig. 3.1

La implementación de la administración de espacios magnéticos centralizada ayuda considerablemente a la erradicación de estos problemas de hoy en día. Esta implementación no es fácil ni rápida y menos aún si son los usuarios quienes en la actualidad realizan estas funciones. Según Cockrell [7], el primer objetivo principal del administrador de espacios magnéticos debe ser ayudar a los usuarios a entender cómo la centralización de las funciones de administración de espacios puede resultarles sumamente benéfica.

En algunas instalaciones la implementación de un grupo de administración de espacios magnéticos constituye una de las principales prioridades para poder liberar a los usuarios de las responsabilidades que implica controlar sus archivos y su propio espacio en disco y aumentar su productividad dentro de las funciones claves del puesto que desempeñan. Sin embargo podría presentarse el caso de que los usuarios se resistieran a transferir sus responsabilidades y a ser sometidos a un proceso de control. En este caso se deberá entender primero las causas de su resistencia y después ayudarles a darse cuenta que la centralización de las funciones de administración de espacios representa un enorme beneficio para ellos.

"El cambio por sí mismo puede ser sumamente difícil y si la gente no entiende las razones y las necesidades del cambio se puede volver más complicado aún" (Cockrell [4]). Una comunicación abierta y honesta así como material e información suficientes constituyen la única alternativa para iniciar el proceso de centralización de funciones de administración de espacios de la mejor manera y sobre todo con el convencimiento de las áreas usuarias. Además de entender estos problemas, el usuario necesita conocer los costos y los beneficios que un control centralizado significa en áreas tales como el uso de espacio, la disponibilidad de espacio, el performance y la instalación de dispositivos.

3.1.1. Las metas de la administración de espacios centralizada.

Centralizar la administración de espacios magnéticos es el primer paso para asegurar que las necesidades de crecimiento magnético están respaldadas por una metodología ordenada y con

efectividad de costos; además es el punto inicial para un sistema con recursos magnéticos administrados. Estos son beneficios que quizás en un inicio no sean muy claros pero que se harán evidentes después de un tiempo relativamente corto. Las técnicas de control que se analizarán en los siguientes capítulos, están encaminadas a conseguir una óptima utilización del espacio en disco, de la disponibilidad de los archivos, y del performance. Sin embargo para poderlas llevar a cabo se requiere un grupo especializado en estas técnicas y que tenga el control y la responsabilidad de las mismas.

Para conseguir estas 3 metas básicas el administrador de espacios magnéticos debe encaminar sus esfuerzos en actividades de 3 áreas generales:

- Administrar los archivos.
- Administrar los ambientes de discos.
- Configurar un subsistema de almacenamiento.

La figura 3.2 muestra la interacción de estas actividades con las metas principales del administrador de espacios.

Así pues, uno de los propósitos fundamentales del grupo de administración de espacios magnéticos es proporcionar servicios de almacenamiento a los usuarios de la computadora con el objeto de incrementar su productividad. A medida que la administración de espacios en disco recibe mayor atención en el procesamiento de datos así como de los usuarios ejecutivos, la administración de espacios magnéticos se convierte en una interfase básica entre los usuarios y el sistema.

La función del administrador de espacios no es simplemente administrar datos en un centro de cómputo sino que debe crear y mantener una comunicación abierta y positiva con la comunidad de usuarios y con otras áreas del centro de cómputo de la organización.

**TECNICAS Y
METAS DEL
ADMINISTRADOR DE ESPACIOS.**

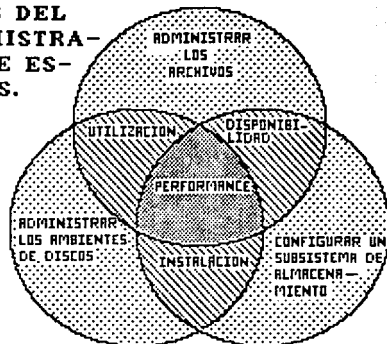


Fig. 3.2

De acuerdo con Hoffman [8], una manera en la que el administrador de espacios magnéticos puede concebir su trabajo es en términos de 3 misiones básicas: estar a la vanguardia del departamento, participar como miembro integral del equipo de procesamiento de datos de la organización y proporcionar servicio a los usuarios. Hacer de estas concepciones una meta de alta prioridad es un reto que requiere esfuerzo y dedicación. (Figura 3.3).

PRIORIDADES DEL ADMOR. DE ESPACIOS

<u>ADMINISTRADOR</u>	<u>NOTAS</u>
- <i>LES ASIGNAR</i>	<i>PRIORIDAD 1</i>
- <i>EL GRUPO DE REQUISICION DE ESPACIOS ASIGNADOS</i>	<i>PRIORIDAD 2</i>
- <i>EL CENTRO DE PROCEDIMIENTO DE LEYES</i>	<i>PRIORIDAD 3</i>

Fig. 3.3

3.2 Estructura organizacional.

Establecer un grupo de administración de espacios magnéticos dentro del organigrama institucional no es una tarea fácil. Sin embargo, el hecho de consolidar a la gente involucrada en la administración de espacios dentro de un departamento asignado para tales funciones tiene grandes ventajas tanto para el centro de cómputo como para la comunidad de usuarios.

Un grupo de administración de espacios deberá insertarse dentro de la estructura del centro de cómputo de manera tal que sea consistente con el estilo de la organización y resuelva las necesidades del "negocio". A continuación se presentarán diferentes opciones para ubicar al grupo de administración de espacios que en la actualidad se aplican dentro de la estructura de diferentes centros de cómputo, pero antes consideremos que un centro de cómputo se subdivide en 3 niveles básicos: la **área operativa**, la de **soporte técnico** y la de **planeación y control** (Figura 3.4); es posible que los términos utilizados para describir funciones similares varíen entre empresa y empresa o que las interpretaciones de estas 3 áreas difieran respecto a las que aquí se plantean por lo cual se detallan estas enseguida:

Operación.-

Este grupo realiza actividades encaminadas a arrancar, parar, checar y controlar el procesamiento de los programas dentro del procesador; es responsable también del mantenimiento del hardware instalado, de atender los requerimientos de impresión de reportes y de montajes de cinta. Típicamente un grupo operativo incluye operadores de consolas, de impresoras, de cintas, técnicos en redes de teleproceso y comunicaciones y otros.

Soporte técnico.-

Este grupo puede abarcar una gran variedad de actividades técnicas tales como programadores del sistema, diseño de redes, administración del performance, afinación del sistema, instalación de productos, seguridad de datos, administración de bases de datos y administración de cambios. Frecuentemente varios departamentos están incrustados dentro del grupo de soporte técnico para funciones específicas.

AREAS BASICAS DE UN CENTRO DE COMPUTO

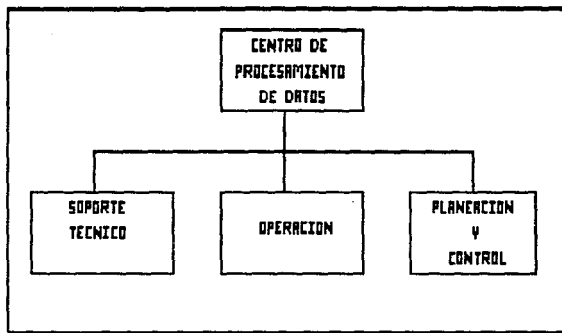


Fig. 3.4

Planeación y control.-

La función de este grupo está enfocada a los aspectos del "negocio" del procesamiento de datos tales como consideraciones financieras, planeación y utilización del equipo instalado, planeación de recursos y procedimientos y prácticas de control. Esta misión es frecuentemente realizada por un sólo departamento a menos que el centro de cómputo sea muy grande.

3.2.1. Administración de espacios como parte de Soporte técnico.

Insertar al grupo de administración de espacios magnéticos dentro del área de soporte técnico puede reforzar las relaciones de cooperación con los grupos de programación del sistema y administración de bases de datos. (Fig 3.5).

Este modelo de organización tiene ventajas estratégicas para las instalaciones que se están encaminando hacia el prototipo del ambiente de administración de espacios ideal (control del ambiente por el propio sistema) porque éste acentúa la interrelación entre el administrador de espacios y el programador del sistema. Su cercanía también con el administrador de las bases de datos puede promover servicios consistentes y adecuados de almacenamiento a los grupos usuarios.

Algunas tareas que se podrían ver especialmente beneficiadas por este tipo de modelo organizacional son:

- Afinación y performance del sistema
- Administración de catálogos
- Cambios e instalación de software y hardware

3.2.2. Administración de espacios como parte de Operación.

Cuando el administrador de espacios está unido organizacional-

ADMINISTRADOR DE ESPACIOS DENTRO DE SOPORTE TECNICO

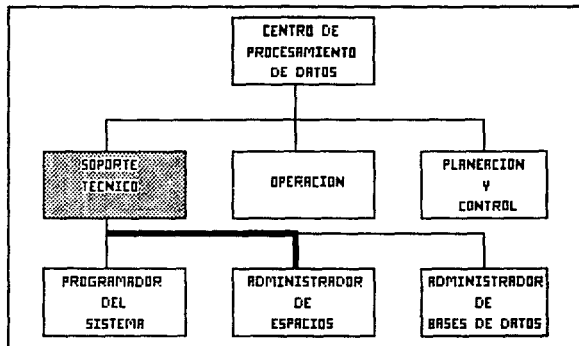


Fig. 3.5

mente al área de operación, el reforzamiento se puede centrar alrededor de los requerimientos diarios de los procesos de producción. Una ventaja importante que se da fuera de la propia estructura es una sensibilización creciente de las áreas usuarias resultante del trato diario con las áreas operativas.

Puede existir un menor énfasis en los aspectos técnicos y estratégicos de la administración de espacios magnéticos pero otros aspectos cobran una especial importancia tales como:

- Procedimientos de respaldo y recuperación de información
- Frecuente comunicación con los usuarios.

La figura 3.6 muestra el modelo de este tipo de estructura.

3.2.3. Administración de espacios como parte de Planeación y Control.

Ubicar al administrador de espacios con el grupo de planeación y control (Figura 3.7) podría tener el efecto de que se reforzaran los aspectos de planeación y de "negocio" en la función. Una fuerte conciencia de los costos y una especial atención a la planeación son aspectos críticos para el éxito de la administración de espacios magnéticos, por lo que este arreglo tiene su mérito y habría que evaluar muy bien antes de pensar en descartarlo.

Con esta opción al igual que con el modelo del administrador de espacios dentro del área operativa se requerirá un especial esfuerzo para mantener la importante relación con el programador del sistema y con otros grupos de soporte técnico.

Las siguientes son ejemplos de tareas que cobran particular importancia para el administrador de espacios cuando se encuentra dentro de este grupo:

ADMINISTRADOR DE ESPACIOS DENTRO DE OPERACION

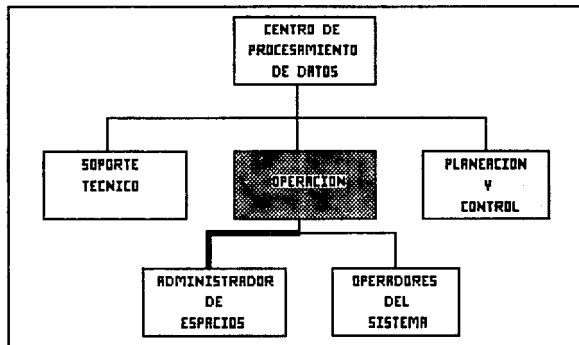


Fig. 3.6

ADMINISTRADOR DE ESPACIOS DENTRO DE PLANEACION Y CONTROL

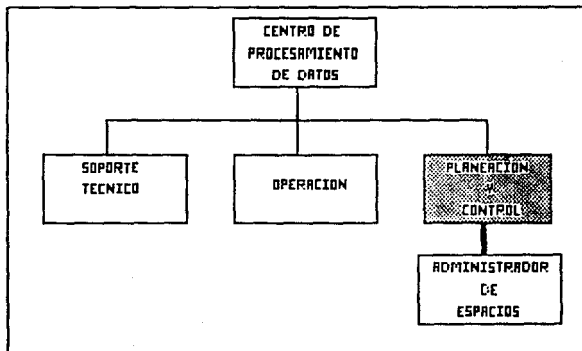


Fig. 3.1

- Planeación de la capacidad
- Definición de políticas
- Adecuado orden y configuración de hardware y software.

3.2.4 Administración de espacios como parte independiente de la estructura.

No es necesario alinear la administración de espacios magnéticos con una área específica del centro de proceso de datos. No es absurdo el pensar al grupo de administración de espacios como una área más de la estructura original planteada. (Fig 3.8). Este modelo enfatiza la creciente importancia de las funciones de administración de espacios magnéticos e incrementa el criterio y visión del grupo.

Sin las influencias organizacionales que vienen de las dependencias con otros grupos, los aspectos de la administración de espacios que se refuerzan son afectados por los intereses de los miembros del equipo y por las necesidades específicas del "negocio".

Estas 4 alternativas de estructuras son válidas y viables de instalarse en un centro de cómputo y deberá ser la propia empresa la que decida cual se ajusta mejor a sus necesidades para tomar la decisión de colocar al grupo de administración de espacios magnéticos en un lugar o en otro.

Claro está que independientemente de en donde sea ubicada el área de administración de espacios dentro de la organización, todas las actividades que se mencionaron como particularmente reforzadas por estar en uno u otro lado deberán de ser contempladas como objetivos del área.

ADMINISTRADOR DE ESPACIOS INDEPENDIENTE

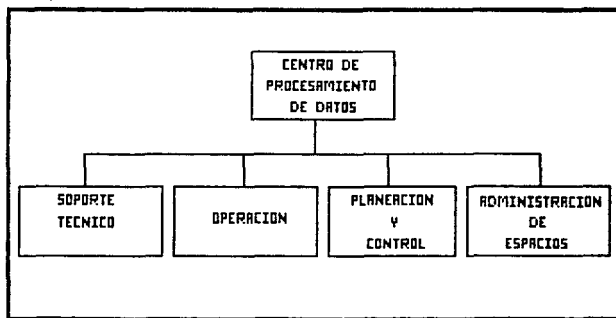


Fig. 3.8

3.3 Efectividad de grupo.

Se ha estado ya comentando de manera general sobre las metas y responsabilidades del grupo de administración de espacios magnéticos. Para poder asignar adecuadamente el trabajo a los miembros del equipo y conseguir así la efectividad de equipo se requiere un entendimiento más detallado de las tareas y responsabilidades que cada miembro del grupo espera cumplir. A lo largo de las siguientes páginas se hablará ampliamente al respecto.

Si el grupo de administración de espacios va a ser creado por primera vez, todos los puntos que se tratarán deberán ser tomados en cuenta; si ya se ha comenzado con la centralización de estas funciones, algunos aspectos pueden ser omitidos.

3.3.1 Proyecto de planeación de actividades.

Planear y poner fechas para la implementación de las nuevas técnicas.

La planeación y coordinación de actividades y cambios son esenciales y requieren de experiencia en administración de espacios y familiaridad con otros aspectos de soporte técnico. Estas actividades requieren un fuerte entendimiento técnico del resto de las tareas, capacidad para identificar dependencias, habilidades de comunicación y negociación así como de organización. Se aconseja poner fecha a las actividades requeridas para la implementación de las nuevas técnicas. Dependiendo del tamaño de la instalación, y de la agresividad de las funciones, esta actividad de planeación puede tomar del 10 al 100% del tiempo de una persona.

En esta etapa se deberá definir la misión del grupo ya que es primordial entender hacia donde se va con las nuevas funciones de administración de espacios antes de comenzar la asignación de tareas y comenzar a trabajar. Un ejemplo de modelo de misión del grupo se presenta en la figura 3.9.

EJEMPLO DE MISION DEL GRUPO

DEPARTAMENTO DE ADMINISTRACION DE ESPACIOS MAGNETICOS (DEFINICION DE SU MISION)

RESPONSABILIDADES

El grupo de Administración de Espacios Magnéticos es responsable de:

- Proporcionar servicios de almacenamientos magnéticos que incrementen la productividad de los usuarios del centro de cómputo.
- Establecer técnicas que proporcionen los niveles de servicio requeridos con la mayor relación costo-efectividad.
- Hacer uso eficiente de los recursos... etc

METAS INMEDIATAS

- Control total de todos los recursos de almacenamiento magnético.
- Generación de "storage-pools" con una configuración eficiente de E/S.
- Políticas completas y adecuadas de administración de archivos.
- Entendimiento y aceptación de las políticas por el usuario.

METAS A MEDIANO Y LARGO PLAZO

- Automatización de las actividades manuales de admon. de espacios.
- Independencia del usuario de los dispositivos de almacenamiento.

Fig. 3.9

Asegurar la instalación de todos los productos de software requeridos.

La implementación del plan para centralizar la administración de espacios puede implicar la instalación de nuevos productos diseñados para tal fin como por ejemplo el DSFSM (Data Set Free Space Manager), el RACF (Resource Access Control Facility), ADAM (Automatic Disk Analysis Management), etc.

La instalación de este software debe ser coordinada con los programadores del sistema e instalado por personal calificado. Es importante trabajar en conjunto con los programadores del sistema para planear la instalación de los productos y para que emitan sus puntos de vista en el plan general de instalación de las nuevas funciones del centro de cómputo.

Planear una configuración óptima de dispositivos de almacenamiento.

Un plan de hardware para conseguir un uso efectivo de los dispositivos de almacenamiento magnético, un adecuado performance requiere habilidad para proyectar y documentar planes de capacidad para la instalación en el proyecto de planeación de actividades.

El plan de hardware es parte del plan de instalación de subsistemas de almacenamiento por lo que es importante trabajar muy de cerca con el grupo de planeación de hardware para la instalación del centro de procesamiento de datos.

Negociar las dependencias del plan con otros departamentos dentro y fuera del centro de cómputo.

El acuerdo y la cooperación de toda la organización del centro de procesamiento de datos (soporte técnico, operación, etc) y de las áreas usuarias son factores vitales para el adecuado funcionamiento de la propia organización. La combinación de los aspectos técnicos con los administrativos es requerida dentro del grupo de administración de espacios. Es posible asignar responsabilidad total para negociaciones a una sola persona, o quizás

sea mas adecuado asignar áreas de negociación a diferentes personas del grupo, por ejemplo, alguien se puede encargar de la operación y otras con usuarios específicos.

3.3.2 Los grupos de usuarios.

Una buena relación de trabajo entre las áreas usuarias y el administrador de espacios es extremadamente importante. Para conseguir esta meta se aconseja seguir los siguientes pasos:

Identificar a todas las áreas usuarias y a sus representantes.

El primer paso para trabajar en armonía con los departamentos usuarios es identificarlos a todos así como a sus representantes para poder establecer el flujo de sus requerimientos, comentar sus necesidades y negociar niveles de servicio en beneficio del propio usuario.

Resulta también mucho más simple para el administrador de espacios realizar las negociaciones con una sola persona que representa a toda un área y que será el frente para entrega y recepción de solicitudes de servicio que estar atendiendo de manera personalizada a cada uno de los miembros de los departamentos usuarios.

Conocer los requerimientos del usuario.

Una entendimiento de cómo se había estado realizando la administración de espacios en el pasado es crucial para satisfacer todos los requerimientos de espacio, disponibilidad y performance de las áreas usuarias. Esta actividad requiere experiencia y comprensión de los conceptos de administración de espacios así como habilidad de relación interpersonal.

Establecer un programa de educación al usuario.

A medida que las responsabilidades del usuario cambian, es

necesario proporcionar una educación y entrenamiento continuos a los departamentos usuarios. Los representantes de cada área pueden ayudar a identificar aspectos específicos que requieran enfatizarse. El diseño de cursos de capacitación incluyendo el material requiere experiencia, un claro entendimiento de las funciones de administración de espacios y habilidades de interrelación personal. Cuando la educación del usuario es una carga pesada, varios miembros del departamento pueden compartir esta responsabilidad existiendo un coordinador para lo mismo.

Establecer mecanismos que sirvan de interfase con las áreas usuarias.

Es muy importante contar con un mecanismo perfectamente definido para las interacciones diarias con los usuarios. Esto incluye mecanismos tales como establecer un escritorio y asignar una persona en el mismo para recepción de solicitudes de servicio y reportes de problemas, o bien algún tipo de formato escrito en el que se especifique el requerimiento. Independientemente de la técnica usada, el usuario debe entender como procesar sus requerimientos y debe recibir un servicio cortés y eficiente por parte del administrador de espacios.

Dependiendo del tamaño y estructura del departamento de administración de espacios, se podría pensar en una rotación periódica de la función de recepción de solicitudes y atención a problemas entre todos los miembros del grupo, o bien asignar administradores de espacios para atender áreas específicas dentro de la población de usuarios, por ejemplo, un administrador de espacios dedicado exclusivamente a atender un sistema específico. Esta responsabilidad requiere de habilidades administrativas y de organización así como un entendimiento básico de las prácticas diarias de administración de espacios.

Realizar una guía de usuario que contenga todas las políticas y procedimientos.

Los procedimientos para obtener servicios del administrador de espacios deberán estar perfectamente documentados y ser los suficientemente claros para los usuarios. El usuario debe entender sus responsabilidades para nombrar y ubicar sus archivos,

los criterios establecidos para migración, respaldo y borrado de archivos, los criterios de seguridad de datos; la metodología para recuperación en casos de desastre; cómo especificar sus necesidades y cómo interactuar con el administrador de espacios. Esta tarea de documentación requiere un entendimiento básico de las prácticas de administración de espacios y habilidades de comunicación escrita.

Acordar niveles de servicio con los grupos usuarios.

La documentación y la negociación de un convenio que establezca niveles de servicio del administrador de espacios con los usuarios es una tarea que requiere experiencia y un fuerte entendimiento de los conceptos de administración de espacios así como habilidades en técnicas de negociación.

3.3.2.1 Proporcionando servicio al usuario.

Para conseguir la verdadera efectividad de grupo se debe de partir de la premisa de que "la meta básica de la administración de espacios magnéticos deberá ser la de proporcionar servicio a los grupos de usuarios" (Hoffman [8]). La administración de espacios debe ser para contribuir al logro de objetivos de la empresa más que ser por sí mismo un objetivo de la empresa. La finalidad que se persigue con la centralización de la administración de espacios es la de mejorar la eficiencia de la empresa por medio de un incremento en la productividad de los usuarios.

Para que lo anterior se pueda llevar a cabo, el administrador de espacios debe conocer el grado de satisfacción del usuario con respecto al servicio que se le está proporcionando. No es únicamente por medio de la medición de los niveles de servicio (que se verá con mayor detalle mas adelante) como encontraremos el sentir del usuario con respecto al servicio proporcionado; otros aspectos de las relaciones cotidianas de trabajo son también muy importantes; por ejemplo, vale la pena investigar las siguientes cuestiones que en condiciones normales no son contempladas en los convenios de niveles de servicio:

¿Es el grupo de administración de espacios reconocido como un grupo de amigable y que presta ayuda cuando se le requiere?

¿Responde el departamento rápidamente a los problemas emergentes?

¿Entienden los usuarios las razones de las políticas para obtener un servicio del grupo o éstas son simplemente vistas como obstáculos burocráticos?

Este tipo de cuestiones son los aspectos finos del servicio al usuario y es de suma importancia estar enterados de este sentir para si fuera necesario, se ejecuten las acciones que contribuyan a mejorar la imagen del grupo con las áreas usuarias. Una opción para conocer este sentir es la realización periódica de encuestas de imagen.

Mantener canales de comunicación abierta con el usuario reforzará una relación de trabajo positiva y fructífera. Las negociaciones con el usuario no deberán circunscribirse a las políticas preestablecidas de administración de espacios., no obstante que nos deberemos de asegurar que contamos con una metodología documentada de servicio así como un proceso de aprobación de excepciones, de lo contrario nos veremos inundados con "requerimientos especiales" de las áreas usuarias.

3.3.2.2 Acuerdo de niveles de servicio.

Basicamente un acuerdo de niveles de servicio es un contrato que especifica los niveles de servicio que el proveedor (en este caso el administrador de espacios) ha acordado con el usuario. El acuerdo de niveles de servicio consiste de una descripción cuantitativa de los servicios requeridos por el usuario de acuerdo a la disponibilidad existente de recursos (Fig. 3.10). En el caso de la administración de espacios magnéticos, los servicios que requieren especificarse son requerimientos de espacio en disco y en cinta, periodos de disponibilidad de los archivos, performance y seguridad. Es vital que los niveles de servicio se establezcan en terminos de objetivos que puedan ser cuantificados, es decir, que exista una fuente para obtener la información del servicio y poder así medirlo periódicamente.

El administrador de espacios deberá trabajar muy de cerca junto con otras áreas del centro de cómputo antes de establecer un nivel de servicio a conceptos difíciles de cuantificar tales como el performance, pues debe recordarse que los subsistemas de

**ACUERDO DE
NIVELES DE
SERVICIO**

NECESIDADES DEL USUARIO
MAS
POLITICAS DE ADMINISTRACION
DE ESPACIOS MAGNETICOS

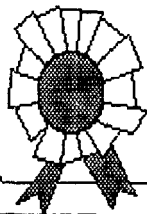


Fig. 3.10

almacenamiento magnético son sólo uno de los muchos componentes que pueden afectar el nivel de performance de las aplicaciones. El administrador de espacios deberá comprometerse a brindar el nivel de servicio pactado considerando que este nivel está basado en la cantidad de recursos de almacenamiento con que se cuenta y con la configuración que se estableció para satisfacer las necesidades de los grupos usuarios. El administrador de espacios **NO podrá** comprometerse a un servicio que técnicamente le sea imposible proporcionar.

El proceso de negociación del acuerdo de niveles de servicio es una función en la que el administrador de espacios se deberá especializar. La aceptación formal de un convenio y su aprobación por los altos niveles de la dirección es una función administrativa que deberá realizar el administrador de espacios.

Si el cumplimiento de los niveles de servicio involucran a otros puestos técnicos u operativos tales como el programador del sistema o el operador del área de cintas, estos deberán incluirse en la firma del convenio.

En las siguientes hojas se presenta un ejemplo escrito de un convenio de niveles de servicio celebrado en una empresa de la vida real.

\ (Inicio)

CONVENIO DE NIVELES DE SERVICIO
DEL ADMINISTRADOR DE ESPACIOS MAGNETICOS
PARA EL PROCESAMIENTO DE DATOS

1 DE ENERO DE 1989 AL 30 DE JUNIO DE 1989

Aprobado por:

Firma

Fecha

(Representante del departamento usuario 'X')

Firma

Fecha

(Representante del departamento usuario 'Y')

Firma

Fecha

(Representante del departamento de Admon. de espacios magnéticos)

Firma

Fecha

(Representante del departamento de Admon. de bases de datos)

PROLOGO

Objetivo y alcance:

El propósito de este documento es definir los niveles de servicio de los subsistemas de almacenamiento magnético para el procesamiento de datos. Con base en los requerimientos del usuario, los niveles de servicio abarcan los conceptos de espacio en disco, permanencia de archivos, performance y seguridad.

Este documento describe también el procedimiento de cambios, responsabilidades de las áreas involucradas, la manera como se reportarán los resultados de los niveles de servicio, las fuentes para medir los niveles de servicio y las políticas de servicio del grupo de administración de espacios magnéticos

Procedimiento de cambios:

Tanto el grupo de administración de espacios magnéticos como las áreas firmantes del presente documento pueden solicitar una revisión de los umbrales de los niveles de servicio para cualquier concepto en cualquier momento que lo deseen informando el hecho a todas las áreas involucradas y convocando a una junta en donde se expondrán las justificaciones y las propuestas de modificación con todos los detalles. Cualquier renegociación de los actuales niveles de servicio implicará una actualización del presente documento.

Se realizará una validación semestral del presente documento por todas las áreas firmantes a fin de mantenerlo vigente y realizar las actualizaciones que se consideren pertinentes, después de lo cual serán distribuidas copias del documento actualizado a todas las áreas participantes.

Responsabilidades:

El grupo de espacios magnéticos será responsable de la administración de los recursos de almacenamiento para garantizar los niveles de servicio que se establecen en este convenio y de resolver cualquier problema que en un momento dado comience a afectar la calidad de los mismos. Es responsable así mismo de promover la actualización semestral del presente documento.

Los usuarios deberán sujetarse a las políticas de ubicación de archivos descritas en el manual de guía del usuario del cual se anexa un ejemplar (un documento ficticio que no será descrito en este ejemplo de convenio de servicio). De esta manera recibirán un servicio mas eficiente de los recursos de almacenamiento magnético y contribuirán al logro de los umbrales propuestos para tales servicios.

DETALLES DEL SERVICIO

Misión del grupo de administración de espacios:

La misión del grupo de administración de espacios magnéticos consiste en proporcionar un servicio de alta calidad en la administración de todos los archivos de las áreas usuarias. Esto incluye el espacio en disco y en cinta para proporcionar servicios tales como respaldo y recuperación de archivos a fin de asegurar el mas alto performance así como la integridad y seguridad de la información.

NIVELES DE SERVICIO POR AMBIENTE

Los archivos de todas las diferentes áreas usuarias son organizados dentro de grupos de archivos que comparten características similares llamados "ambientes" o "storage-pools". Considerando los requerimientos de todos los usuarios, el administrador de espacios ha definido 3 ambientes de discos con niveles de servicio específicos para cada uno que cubren las exigencias mínimas planteadas por las diversas áreas. Para una información mas detallada del propósito de cada ambiente, el tipo de archivos que en ellos residirán y cómo ubicar archivos en un ambiente en particular, favor de consultar la guía del usuario anexa.

Los ambientes de discos definidos fueron:

- . Pool primario
para archivos batch, TSO y bases de datos de pruebas
- . Pool de base de datos
para archivos de CICS e IMS

- . Pool temporal
para archivos temporales, es decir, archivos cuyo nombre comience con & o con &&.

Esta sección describe los niveles de servicio para los conceptos de espacio, disponibilidad, performance y seguridad para cada uno de los ambientes.

ESPACIO

Pool primario

Este ambiente contendrá discos que contarán con un 20% de espacio libre todas las mañanas para ubicación de nuevos archivos.

Diariamente DFDSS reclamará automáticamente el espacio no utilizado de los archivos ubicados y borrará los archivos que hayan expirado o que ya no sean necesarios. Semanalmente DFDSS borrará todos los archivos no catalogados.

Por las noches, DFHSM migrará a cinta todos aquellos archivos que no hayan sido referenciados durante los últimos 15 días en un volumen de nivel 1. Los archivos de nivel 1 en cinta no referenciados durante 50 días serán reubicados a una cinta de nivel 2. Los archivos de nivel 2 en cinta que no sean referenciados durante 30 días serán dados de baja.

Diariamente se medirá el índice de fragmentación de los archivos. Si este rebasa el 40% serán deragmentados por medio ... etcétera.

Pool de bases de datos.

La disponibilidad de espacio para este ambiente será responsabilidad del departamento de administración de bases de datos.

Pool temporal.

Este ambiente contará diariamente con un espacio de por lo menos un 30% por disco y ...etc.

DISPONIBILIDAD

Pool primario

Todas las noches DFHSM respaldará en cinta todos los archivos que han cambiado y no se les ha respaldado durante los últimos 2 días. Dos versiones en cintas existirán por cada archivo las cuales serán borradas después de un año de no ser referenciadas.

Los archivos que no hyan podido ser espaldados en cinta por DFHSM a causa de estar abiertos durante el proceso de respaldo, serán respaldados en cinta por el administrador de espacios a la mañana siguiente. Además todos los discos pertenecientes a este ambiente serán respaldados en cinta una vez a la semana de acuerdo a ...
etc...

Pool temporal.

No se realizarán respaldos en cinta de los archivos ni de los discos de este ambiente. Todos estos archivos serán borrados cada 2 días o bien ... etc...

.
.
.

etc...

PERFORMANCE

Dado que el tiempo de respuesta involucra una gran cantidad de factores independientes a los dispositivos de almacenamiento, tales como el propio procesador, los dispositivos de comunicación , etc, este concepto será tratado en otro acuerdo independiente que contemple niveles de servicio para tiempos de respuesta.

El tiempo de restauración de los archivos del pool primario respaldados en cinta no deberá ser mayor a los 90 segundos incluyendo el tiempo de montaje. (Un convenio adicional sobre tiempos de montaje se negociará con el área de operación).

SEGURIDAD

La facilidad de RACF estará disponible para todos los usuarios. Será responsabilidad del administrador de RACF crear y mantener todos los perfiles de seguridad requeridos.

El administrador de espacios vigilará y reportará todos los intentos de acceso no autorizados de manera que ... etc...

CONSIDERACIONES ESPECIALES

. Todos los archivos que inicien con el nombre "MALTUD02.ASCUB" en el pool primario serán respaldados en cinta todas las noches siempre y cuando hayan sufrido cambios durante el día. Estas versiones en cinta tendrán una vigencia de 400 días.

. Los grupos de generación del pool primario serán respaldados en cinta hasta que ... etc...

REPORTE DE NIVELES DE SERVICIO

El grupo de administración de espacios obtendrá diariamente los niveles de servicio ofrecidos para cada concepto y lo viernes de cada semana se enviará un reporte sumariado a los representantes de las áreas usuarias.

Con base en los reportes obtenidos se podrán levantar reportes de alertamiento para el archivo que no cumplió con alguno de los niveles pactados de acuerdo a los siguientes 4 niveles:

Alerta Nivel 1- El nivel de servicio estuvo fuera del rango aceptable en alguno de los días de la semana que acaba de concluir. El administrador de espacios deberá desarrollar un plan de acción que corrija la situación y ... etc...

Alerta Nivel 2- El nivel de servicio estuvo fuera durante las últimas 2 semanas previas ...etc...

Alerta Nivel 3- etc...

Alerta Nivel 4- etc...

\ (Fin)

3.3.3 Políticas y procedimientos.-

Una de las tareas mas penosas para el administrador de espacios es convencer a los usuarios de la conveniencia de crear políticas para la ubicación de archivos y utilización de los recursos magnéticos. Es también motivo de frecuentes fricciones el implantar un control de solicitudes y servicios por el método de documentación escrita. El usuario se pregunta ¿para que diablos llenar un formato y justificar el motivo de mi solicitud por escrito?... una llamada telefónica sería mas que suficiente... Qué grupo tan burocrático!!... etc.

El administrador de espacios deberá realizar labor de convencimiento desde la negociación de los niveles de servicio, explicando al usuario lo indispensable que resulta para el grupo de espacios magnéticos el contar con un control y con una metodología de servicio y deberá pedir la opinión de todos los usuarios a fin de encontrar el procedimiento que mas convenga a ambas partes; no necesariamente se requerirá del uso de papel para el trámite de solicitudes de servicio ya que se podría pensar en instalar un procedimiento automatico tal como una especie de correo electrónico o algún otro; lo importante es que la solicitud quede documentada en algún sitio.

A continuación se presentará un resumen de las políticas mínimas indispensables que deberán instalarse en un ambiente de proceso de datos aunque muchas de ellas se discutirán mas a detalle posteriormente.

Crear estándares para los nombres de los archivos.

El uso de convenciones para asignar nombres a los archivos es esencial para poder establecer los niveles de servicio requeridos por el usuario. Contar con estándares para los nombres de archivos permitirá hacer un uso productivo de los recursos de almacenamiento magnético disponibles.

Esta actividad requiere un amplio entendimiento del funcionamiento de productos tales como RACF, TSO, DFDSS, ADAM, y en general de los productos que permiten hacer valer los estándares establecidos para nombrar los archivos.

Estándares de ubicación de espacio.

Estandarizar la ubicación de archivos en disco y cinta y controlar la codificación de ciertos parámetros y subparámetros en las declaraciones DD de los JCL's es una técnica sumamente efectiva en la administración de espacios magnéticos.

Familiaridad con el sistema y los productos (software) para administrar espacios magnéticos es requerida para asegurar los estándares que se establezcan.

Definir políticas de administración de los archivos.

Las políticas para administrar todas las fases del ciclo de vida de los archivos son importantes para proporcionar los niveles de servicio adecuados. Es necesario determinar como se controlará las actividades de migración, respaldos en cinta, borrado o migración de los archivos no utilizados, autoridad y seguridad para el acceso de datos, identificación y planes de recuperación de los archivos vitales para casos de desastre, etc.

Para cada una de estas actividades, las políticas especificarán como se realizará la función (por ejemplo, la seguridad de los archivos por medio de RACF). Todas estas políticas deberán estar documentadas y ser aprobadas y distribuidas a todas las áreas usuarias. Se deberá establecer también un proceso aprobado para el manejo de excepciones.

La definición de estas políticas requiere de un claro entendimiento de los conceptos de administración de espacios, de las metas y misión del grupo y, sobre todo, de las necesidades del usuario.

Negociar niveles de servicio con las áreas operativas.

Un acuerdo formal escrito con las áreas de operación asegura que todos los requerimientos de recursos magnéticos por los diferentes usuarios sean proporcionados exitosamente así como los niveles de servicio comprometidos.

Esta actividad requiere de una clara comprensión de la administración de espacios magnéticos, así como habilidades de interrelación personal y de técnicas de negociación.

3.3.4 Configuraciones óptimas de hardware

Evaluar la capacidad del equipo de almacenamiento instalado.

Parte del proceso para integrar un ambiente administrado es examinar el equipo actual de almacenamiento magnético y entonces determinar si es suficiente (discos, cintas, controladores, cache, etc).

Realizar esta actividad requiere de un conocimiento de las capacidades del equipo instalado y de las metas específicas del grupo de administración de espacios.

Configurar "strings" de discos.

El subsistema de almacenamiento magnético debe ser configurado de manera tal que exista un balance del performance, la disponibilidad de dispositivos y la utilización del espacio para que cubra con las expectativas de servicio negociadas.

Realizar esta actividad requiere de una clara comprensión de las capacidades de los dispositivos, conceptos de programador del sistema y de las metas del grupo de administración de espacios.

Asegurarse de considerar las limitaciones físicas.

Si se piensa expandir el hardware de almacenamiento (que es lo más seguro en el corto y mediano plazo), existen requerimientos de espacio, energía y aire acondicionado que son necesarios considerar.

Este aspecto requiere experiencia con los aspectos físicos de instalación y movimiento de dispositivos. El centro de cómputo debe contar por lo menos con una persona (dentro o fuera del departamento de espacios magnéticos) asignada para estas funciones.

3.4 Asignación de trabajo a los miembros del grupo de espacios magnéticos.

Después de tener perfectamente identificadas las actividades necesarias para realizar la centralización de las funciones de administración de espacios magnéticos de manera efectiva, el siguiente paso es explorar algunos de las posibles opciones para asignar trabajo a todos los miembros del departamento de espacios magnéticos.

Antes de comenzar es muy importante tener en mente que no existe un método bueno o malo para describir y asignar trabajo de administración de espacios. Simple y sencillamente se plantearán ideas, conceptos y ejemplos de la vida real que pueden ayudar a crear descripciones de trabajo para cubrir necesidades específicas. Además, las funciones estarán directamente influenciadas por la posición organizacional en que se encuentre ubicado el departamento de espacios magnéticos. Otra influencia de gran importancia para la asignación de tareas al grupo es el grupo en sí mismo, es decir, sus características personales tales como dominio técnico, necesidades de desarrollo profesional y por supuesto el número de personas asignadas al grupo.

3.4.1 Estructura de responsabilidades del grupo.

En las hojas anteriores se ha hablado mucho sobre las metas y responsabilidades de un grupo de administración de espacios magnéticos. Las funciones del grupo no incluyen sólo la administración del espacio, del performance, de la disponibilidad, etc, sino otros aspectos de igual importancia tales como el trabajo y la interacción con grupos usuarios, el establecimiento de políticas y el control e instalación de nuevos dispositivos de hardware.

Existen pues diversos criterios para distribuir la carga de trabajo entre los integrantes del equipo de administración de espacios. A continuación se presentan algunas ideas de las muchas que se pueden utilizar para este fin.

3.4.1.1. Responsabilidades por equipos de cómputo.

Este criterio consiste en distribuir las actividades por subsistemas del equipo de almacenamiento. Por ejemplo, una persona se puede encargar de la administración de todos los conceptos de servicio de las unidades de cintas, otra se encargará de todos los discos dedicados a las aplicaciones línea, otra a la memoria CACHE (memoria electrónica de alta velocidad; se dedica un capítulo completo a su estudio), etc. Si se decide utilizar este criterio, el líder del grupo deberá asegurarse de que el grupo trabaje apegado a políticas y procedimientos que los usuarios puedan fácilmente reconocer y entender.

3.4.1.2. Responsabilidades por concepto de nivel de servicio.

Esta técnica se refiere a asignar a cada miembro del grupo la responsabilidad de hacer cumplir uno o varios de los niveles de servicio acordados con las áreas usuarias; por ejemplo, una persona tiene la responsabilidad total de garantizar el espacio suficiente en disco, mientras que otra se encargará de los procedimientos de respaldo y recuperación de archivos, etc. Ciertas tareas administrativas también pueden distribuirse por persona, como por ejemplo el establecimiento de estándares para los nombres de los archivos o la realización del convenio de niveles de servicio; claro está que para la realización de un documento de este tipo la persona asignada tendrá que relacionarse con todos los miembros del equipo (además de los usuarios) para recoger sus sugerencias y puntos de vista, sin embargo sólo esa persona será responsable de la conclusión y autorización de dicho convenio.

3.4.1.3 Responsabilidades por área de especialización.

Algunas de las funciones de administración de espacios requieren habilidad y experiencia en ciertas disciplinas, por ejemplo la configuración de hardware o la seguridad de datos. En tales áreas se pueden asignar representantes que serán especialistas únicamente de esas funciones. A continuación veremos algunas de las posibles áreas de especialización de un grupo de administración de espacios, sin que esto signifique que sean las únicas que existen ni las mejores:

Planeación de la capacidad.

Aunque muchas empresas han creado departamentos específicos para la realización de estas funciones, el administrador de espacios bien puede tener bajo su cargo esta responsabilidad por ser el más involucrado con la utilización diaria de espacios magnéticos y conocer las tendencias de crecimiento de los mismos.

Performance de los dispositivos de almacenamiento.

Las actividades relacionadas con la configuración de hardware así como de su afinación requieren de una gran capacidad técnica así como de una estrecha cooperación con el grupo de programadores del sistema. Si no se cuenta con una persona con la experiencia suficiente, se puede de cualquier modo asignar alguien que de inicio sirva como interfase al técnico en configuraciones de equipos de cómputo.

Seguridad de datos .

Un administrador de seguridad de datos dentro del grupo de espacios magnéticos ayudará a asegurar prácticas consistentes de protección e integridad de la información vital del sistema. Muchas empresas también cuentan con un departamento específico para realizar las funciones de seguridad de datos, sin embargo resulta más económico y más conveniente dejar esta responsabilidad al grupo de administración de espacios.

Disponibilidad y recuperación de archivos.

Si la naturaleza de la empresa obliga a contar con una metodología bien definida de recuperación en caso de desastre, se puede definir un puesto especial en el departamento de espacios magnéticos que considere todas las actividades requeridas. Esta función debe incluir una estrategia de educación al usuario para sensibilizarlo sobre la importancia de la protección y recuperación de archivos. Si no se define una persona para la realización de estas funciones, ellas se pueden distribuir a todo el grupo de espacios magnéticos como una más de sus responsabilidades.

Hemos pues visto varias estrategias de distribución de funciones al grupo de espacios magnéticos. Sin embargo la decisión final dependerá de factores tales como la posición organizacional dentro de la empresa, de las habilidades y experiencia de las personas así como de la cantidad de recursos de almacenamiento con que se cuente. Sin embargo, sea cual sea el criterio utilizado, el líder del grupo deberá asegurarse de fomentar una comunicación franca y abierta con y entre todos los miembros del grupo, así como de realizar reuniones periódicas (por lo menos cada 15 días) con el grupo a fin de analizar y discutir problemas encontrados para las actividades de cada uno de los miembros así como el avance de las mismas.

3.5 El futuro de la administración de espacios.

Se ha discutido ampliamente como el grupo de administración de espacios centralizada mejora de manera inmediata el uso adecuado de los recursos de almacenamiento así como su performance y consigne que la instalación de nuevos dispositivos sea una tarea mas simple para el personal de soporte técnico y una actividad transparente para el usuario.

El administrador de espacios es esencial hoy en día no sólo para resolver los problemas de una administración descentralizada sino también para comenzar a desarrollar la estrategia de un ambiente de datos auto-administrado en un futuro cercano. El administrador de espacios se convertirá en la conexión entre el procesamiento físico y lógico de los datos. Será la interfase entre el usuario y el sistema. En la figura 3.11 se observa como los beneficios básicos de separar los conceptos físicos y lógicos de los datos son que las consideraciones lógicas (responsabilidades del usuario) serán independientes de las características técnicas de los dispositivos de almacenamiento; los subsistemas de almacenamiento podrán ser administrados eficientemente por el sistema vistos como un "pool único" con los niveles de disponibilidad y performance optimos.

Bajo este esquema, el usuario describirá sus necesidades de almacenamiento de datos en términos de atributos lógicos; el administrador de espacios establecerá las políticas de instalación y vigilará y controlará la administración de espacios realizada por el sistema; el sistema se encargará de administrar los recursos de almacenamiento magnético.

3.5.1 El papel del administrador de espacios magnéticos.

A pesar de que las actividades del administrador de espacios irán cambiando a medida que el sistema vaya asumiendo sus funciones, la presencia de este puesto seguirá siendo indispensable para garantizar la disponibilidad de los espacios magnéticos realizando funciones mas enfocadas al análisis que a la operación, por ejemplo:

SEPARACION FISICA-LOGICA DEL FUTURO

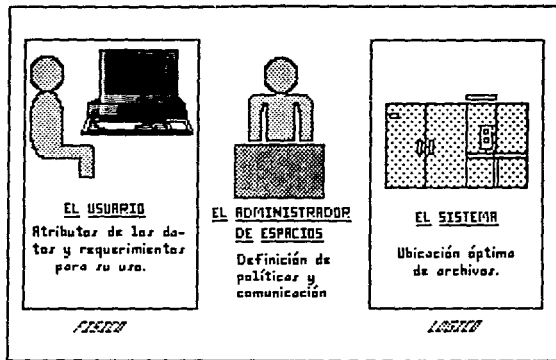


Fig. 3.11

. Servir como una interfase entre el sistema y el usuario para la definición y actualización de políticas en las que se basará el sistema para tomar sus decisiones de ubicación y depuración de archivos.

. Controlar la administración realizada por el sistema analizando los reportes generados y ajustando los criterios cuando sea necesario.

. Proporcionar soporte al usuario para que trabaje de manera efectiva con el ambiente auto-administrado por el sistema.

. Proporcionar el soporte en problemas tales como fallas de los dispositivos de almacenamiento.

Dada la reducción de actividades manuales de administración de espacios que se prevee en el futuro, el perfil del puesto requerirá de mayores habilidades técnicas que en la actualidad ya que se convertirá en un puesto completamente técnico y analítico. Entendiendo cómo administra el sistema los recursos de almacenamiento (así como el programador del sistema entiende el funcionamiento del sistema operativo), el administrador de espacios podrá sacar la mayor ventaja de las capacidades de los recursos.

3.5.2 El papel del usuario.

Con los sistemas del futuro, el usuario ya no tendrá que preocuparse en entender a los dispositivos de almacenamiento magnético ni de sus modelos o características técnicas. Desde el punto de vista del usuario, el almacenamiento físico será percibido como un recurso completamente adecuado, capaz de proporcionar los requerimientos especificados. Con una administración magnética realizada por el sistema y bajo el control del administrador de espacios, el usuario podrá estar seguro de que sus datos están seguros y bien administrados, permitiéndole dedicarse completamente al cumplimiento de sus responsabilidades básicas.

CAPITULO 4

ALMACENAMIENTO MAGNETICO EN DISCO

4.1 Almacenamiento físico de los datos.

Es muy fácil imaginar cómo se almacenan los datos físicamente en tarjetas perforadas. Toda columna de tarjeta contiene un carácter. Cada tarjeta contiene un registro único completo; un paquete de tarjetas forma un archivo. Puesto que cada tarjeta perforada tiene exactamente 80 columnas de largo, cada registro contendrá exactamente 80 caracteres. Se dice que tales registros son de longitud fija.

Algunos dispositivos en disco magnético están divididos en "sectores"; en otras palabras, la superficie está dividida en un cierto número de sectores de longitud fija (Fig. 4.1). Normalmente cada sector contiene un solo registro. Un tamaño común de sector en muchos disquetes de microcomputadoras es de 256 bytes. Esto significa que entre el almacenamiento principal y el secundario se transfieren registros de longitud fija de 256 bytes cada vez que el programa emite una instrucción read(lectura) o write (escritura).

Para conservar el espacio en disco o cinta magnéticos los datos suelen agruparse en bloques. El agrupamiento en bloques también mejora la eficiencia de un programa reduciendo la cantidad de lecturas y escrituras físicas (búsqueda en un disco o arranque y parada en una cinta) reduciendo así el tiempo requerido para leer y escribir los datos. El agrupamiento en bloques implica el poner varios registros lógicos en un solo registro físico grande (Fig. 4.2). Debemos observar cuidadosamente la diferencia entre registros lógicos y registros físicos; el registro físico es el bloque entero, mientras que el registro lógico está formado por los datos necesarios para completar una sola iteración en un programa.

No todos los registros tienen una longitud fija, lo cual es útil para muchas aplicaciones. En este caso la longitud del registro (esto se especifica con el atributo LRECL en un parámetro del lenguaje de control) normalmente es parte del propio registro.

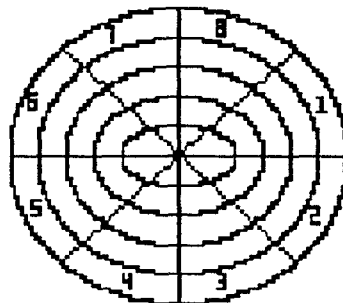


Fig. 4.1 PAQUETE DE DISCOS DIVIDIDO EN SECTORES



Fig. 4.2 DATOS DE LONGITUD FIJA AGRUPADOS EN BLOQUES EN CINTA MAGNETICA.

4.2 Métodos de acceso.

El almacenamiento y la actualización de datos son prácticas muy comunes en las empresas, grandes y pequeñas, en escuelas, en hospitales; para realizar tales tareas, las organizaciones necesitan acceder y manipular los datos de diversas maneras.

No es suficiente solamente con almacenar los datos en el dispositivo secundario; debe ser posible recuperarlos. Una vez creado el archivo, ¿cómo se le encuentra de nuevo cuando se le requiera? Y una vez localizado el archivo, ¿cómo se pueden encontrar los registros individuales que se necesitan?. Estas son las preguntas con las que trata el Método de Acceso.

Cuando los bancos de datos crecen y crecen como ha sucedido en la actualidad, se requiere tener algún método para poder acceder eficientemente la información. Las rutinas que controlan la transferencia de datos entre la memoria interna del computador y los dispositivos de memoria secundaria constituyen precisamente el método de acceso.

Obviamente un archivo puede contener una cantidad tan grande de registros que no sería práctico almacenarlos a todos en la memoria real de la computadora a un mismo tiempo. A medida que los registros son procesados, estos son escritos en archivos ubicados en dispositivos de memoria auxiliar (discos o cintas básicamente). Estos archivos según IBM [4] pueden ser en términos generales de dos tipos:

SECUENCIALES

DIRECTOS

Estos dos tipos de archivos corresponden directamente a dos métodos de acceso: procesamiento secuencial (también llamado procesamiento batch) y procesamiento directo. No debemos confundir el tipo de acceso del archivo con las organizaciones de los archivos que pueden ser de diversos tipos.

En un sistema operativo mainframe, por ejemplo IBM se tienen disponibles muchas clases diferentes de métodos de acceso. Para el caso de acceso secuencial se cuenta con el Método de Acceso Secuencial con Lista de Espera (Queued Sequential Access Method (QSAM)) que realiza automáticamente la formación de bloques y la eliminación de los mismos y un Método de Acceso Secuencial Básico (Basic Sequential Access Method (BSAM)) que deja estas responsabilidades al programador. El Método de Acceso Secuencial Indexado (Indexed Sequential Access Method (ISAM)) también tiene las versiones con líneas de espera (QISAM) y el básico (BISAM). Para acceso directo sólo se proporciona una versión básica (BDAM). Se entra también a los archivos virtuales por medio de VSAM que es el Método de Acceso de Almacenamiento Virtual (Virtual Storage Access Method). El BTAM, QTAM y VTAM permiten que el programador entre a los archivos utilizando líneas de telecomunicaciones.

4.2.1. Localización de archivos.

A un carrete o cartucho de cinta magnética se le llama un volumen. Típicamente cada archivo de cinta magnética tiene un sólo volumen aunque hay archivos de volumen múltiple y volúmenes con múltiples archivos

Debido a que los operadores no pueden leer el título del contenido de la cinta, se utilizan etiquetas externas para su identificación.

Un solo paquete de discos o un solo tambor también es conocido como volumen; la expresión volumen se utiliza para hacer la referencia a una sola unidad física de algún medio de almacenamiento de datos.

En cinta magnética, los archivos individuales van precedidos por una etiqueta, También los archivos de acceso directo necesitan etiquetas. En lo típico en un volumen de acceso directo (DASD), las etiquetas de todos los archivos de este volumen se agrupan al inicio del mismo, comunmente en el primero o segundo cilindro. Esta tabla de contenido de volumen conocida como VTOC (Volume Table Of Contents) identifica cada archivo e inidica en

donde comienza, en términos del cilindro real y la dirección de la pista entre otras cosas.

Sin embargo en la mayoría de los sistemas hay muchos volúmenes de acceso directo diferentes, tal como se mencionó con anterioridad, varios DASD se instalan típicamente en conjunto y en ellos varios manipuladores comparten una unidad de control común. Antes de encontrar la tabla del contenido de un volumen, primero se debe localizar el volumen correcto. Un método simple sería buscar en todas las VTOC's de los volúmenes hasta localizar el archivo deseado. Esto funcionaría en un sistema pequeño, pero en sistemas grandes donde hablamos de más de 50 volúmenes de acceso directo en línea, buscar en uno por uno llevaría demasiado tiempo. La solución consiste entonces en generar un catálogo o índice en el que estén listados los nombres y la ubicación de cada uno de los archivos del sistema. Este catálogo se almacena en un volumen con una ubicación conocida por el sistema y cuando un programa desea un archivo particular, se lleva el catálogo a la memoria principal, el computador lo lee y busca en él los datos requeridos para la localización ágil del volumen que contiene el archivo deseado.

4.2.2. Organización Secuencial.

En un archivo secuencial, los registros son organizados únicamente considerando sus localizaciones físicas sucesivas dentro del archivo. Los registros están generalmente, mas no necesariamente, en secuencia de acuerdo a sus llaves (campos de control). Estos registros son usualmente leídos o actualizados en el orden en el que aparecen en el archivo. Los registros individuales no pueden ser localizados rápidamente y generalmente un registro intermedio no puede ser dado de baja sin haber recorrido primero los registros que le anteceden. Esta organización es utilizada básicamente cuando la mayoría de los registros son procesados cada vez que se accesa el archivo.

4.2.3. Organización Particionada.

Un archivo particionado es aquel que se divide en varias unidades llamadas "miembros". Cada miembro cuenta con un nombre

Único y puede ser llamado por su nombre para procesarse. Nuevos miembros pueden añadirse o darse de baja del archivo cuando se requiera. Los registros dentro de cada miembro son almacenados secuencialmente.

La organización particionada es usada principalmente para el almacenamiento de datos secuenciales, tales como programas, subrutinas y tablas. Por ejemplo, una biblioteca de subrutinas puede ser un archivo particionado cuyos miembros sean tales subrutinas.

4.2.4. Organización Secuencial-Indexada.

El archivo secuencial-indexado es similar a un archivo secuencial con la particularidad de que el procesamiento más rápido es posible, así como es más rápida la localización de registros individuales para procesos no secuenciales. Los registros son almacenados por medio de índices asociados con el archivo. De esta manera, un pedazo de área es reservada para la adición de nuevos registros eliminando la necesidad de tener que rescribir todo el archivo.

Cuando este método es utilizado, el software dedicado a las operaciones de E/S se encarga del manejo de estas operaciones sin que el programa que lo va a acceder requiera una sección especial para controlar esta situación.

4.2.5. Organización de Acceso Directo (DA).

Un archivo organizado de manera directa se caracteriza por la relación total entre la llave de un registro y la dirección de ese registro en el disco o cinta. Esta relación es establecida por el usuario. Este método para organizar los registros es generalmente utilizado para archivos cuyas características no permiten el uso de organizaciones secuenciales o secuenciales-indexadas, o bien para archivos en los que el tiempo requerido para localizar un registro en particular debe ser mínimo.

4.2.6. Organización VSAM (Virtual Storage Access Method)

Cuando IBM anunció su serie de computadoras sistema/370, que fue una serie en la cual se implementó una técnica de administración de memoria conocida como memoria virtual, también se introdujo una nueva estructura de archivo: Método de Acceso de Almacenamiento Virtual (VSAM) permitiendo un novedoso criterio de organización de datos.

El programa que funciona y realiza las funciones del método de acceso se llama AMS (Access Method Service) y este a su vez puede ser manipulado por el usuario a través de una interface conocida como IDCAMS con la que puede realizar de manera relativamente fácil las siguientes funciones:

- Definir y borrar catalogos VSAM ó ICF
- Definir y borrar archivos VSAM
- Cargar datos en los archivos
- Listar información de los catálogos o de los archivos
- Convertir archivos ISAM a VSAM
- Mover archivos de un Sistema Operativo a otro.

El método de acceso para almacenamiento virtual supera ampliamente todas las organizaciones vistas con anterioridad, tanto en sus tiempos de respuesta como en la manera en que optimiza el uso de espacio.

Es por ello que se ha decidido dedicar una sección completa al estudio detallado del funcionamiento y de la estructura de este tipo de archivos.

4.2.6.1 Características de los archivos VSAM.

VSAM no es simplemente un conjunto de rutinas de un método de

acceso. Es una poderosa facilidad para manipular datos dentro de una instalación. Entre las principales características podemos encontrar:

- . Viabilidad de almacenar la información independientemente del modelo de disco o cinta que se utilice.
- . Rutinas para acceso secuencial y directo, así como por llave, dirección relativa o número de registro relativo.
- . Opciones para optimizar el rendimiento
- . Catálogos para definir fácilmente los archivos
- . Un programa multifunción de servicio (AMS, Access Method Services) para definir y dar mantenimiento a los catálogos y archivos.

Bajo el método VSAM, los datos se almacenan en bloques de longitud fija. Dentro de un bloque, los registros individuales se almacenan en secuencia utilizando alguna clave (número de cuenta, número de seguro social, etc), o bien, en la secuencia de entrada (que es el orden en el que se reciben los registros). Si suponemos una cierta secuencia de clave, que es el criterio que probablemente sea el más común, se llevará un índice que indique el último registro de cada bloque. De esta manera VSAM se parece mucho al Método de Acceso Secuencial Indexado (ISAM), y sin embargo según IBM [4] lo supera por algunas mejoras significativas:

- El método ISAM con sus índices de cilindro y de pista, está diseñado en forma específica para los archivos de disco magnético; VSAM en cambio utiliza una dirección de "byte relativo", análoga a la dirección de memoria principal para acceder los datos y con ello este método de acceso se vuelve como ya se había comentado previamente, independiente del dispositivo.

- Otra de las debilidades principales de ISAM es la forma como se manejan las altas y las bajas al archivo. VSAM dispersa el "espacio libre" en todo el archivo físico y su soporte lógico (software) incluye el código necesario para reordenar los registros y los indicadores, lo cual hace mucho más eficiente la actualización de archivos y la recuperación de datos.

- Sin duda que el mayor beneficio potencial de VSAM se deriva de la relación entre éste y la memoria virtual en si misma ya que permite que el programador direcciona las instrucciones del programa y las áreas de datos que en ese momento residen en algún dispositivo de almacenamiento secundario como si estuvieran en realidad en la memoria real. El programa controlador de la memoria virtual se encarga de convertir estas direcciones en direcciones reales; es como si en el direccionamiento base/desplazamiento se eliminara un paso mas.

4.2.6.2.- Tipos de archivos VSAM.

Son tres las diferentes organizaciones de archivos VSAM que podemos utilizar: KSDS (Key-Sequenced Data Set), ESDS (Entry-Sequenced Data Set) y RRDS (Relative Record Data Set). A continuación mencionaremos las principales características de cada tipo.

KSDS (Key-Sequenced Data Set)

- Los registros son ubicados utilizando un campo llave (key-field).
- El acceso a los registros se realiza internamente por llave a través de un índice o bien utilizando directamente la "dirección relativa de byte" (RBA, Relative Byte Address).
- Sus espacios libres son reutilizados para la inserción de nuevos registros o para cambiar sus longitudes.
- El espacio liberado por el acto de borrar registros es de inmediato etiquetado como un "espacio disponible" (free space)

ESDS (Entry-Sequenced Data Set)

- Los registros se graban de acuerdo al orden en que llegan
- El acceso a los registros se realiza unicamente por medio de la RBA.

- Los nuevos registros se van insertando en un espacio reservado para tal al final del archivo.

- Un registro no puede ser borrado pero si remplazado por otro que tenga exactamente la misma longitud.

RRDS (Relative Record Data Set)

- Sus registros son de longitud fija basados en una posición relativa.

- El acceso es a través de un número relativo de registro

- Los espacios vacíos entre los registros pueden contener nuevos registros de longitud fija.

- Cuando un registro es borrado queda un bloque vacío pero el espacio no es reclamado como disponible.

De nueva cuenta no se puede afirmar tan fácilmente que una organización sea mejor que otra, ya que ello dependerá de las características de los datos que contendrá, es decir, si nos encontramos con un archivo de gran volatilidad y continuas búsquedas de registros, quizás ni la organización ESDS ni la RRDS sean las más adecuadas sino la KSDS en este caso.

4.2.6.3. Catálogos VSAM/ICF.

En las últimas versiones de los sistemas operativos de IBM-mainframe (1985) todas las definiciones de archivos VSAM están almacenadas en catálogos que pueden ser de tipo VSAM o ICF (Integrated Control Facility). Estas 2 estructuras de catálogos son muy semejantes aunque ICF elimina varias ineficiencias que se presentaban en la estructura VSAM pero no serán mencionadas aquí.

En primer lugar debemos de contar con un catálogo maestro (MASTERCAT), el cual usualmente (mas no necesariamente) contiene únicamente apuntadores ("entries") para un segundo nivel de catálogo llamado catálogo de usuario (USERCAT). Este catálogo de usuario contendrá los apuntadores a todos los archivos VSAM y no

VSAM. Estos dos tipos de catálogos son básicos para la implementación de un sistema que manejará archivos VSAM y se definirán vía IDCAMS. La figura 4.3 ejemplifica con claridad las conexiones entre archivos y catálogos, todo dentro de un catálogo maestro.

Los catálogos de usuario están plenamente justificados tanto desde un punto de vista de facilidad en la administración de los archivos, como para agilizar los tiempos de respuesta de los mismos. Podemos imaginarnos qué sucedería si sólo contáramos con un catálogo maestro al que estuvieran conectados unos 500 archivos (las instalaciones grandes de cómputo actuales manejan más de 10,000 archivos); ¿cuánto tiempo le tomaría al método de acceso la simple localización de tal cantidad de archivos siendo estos solicitados simultáneamente para operaciones de E/S?. Evidentemente el tiempo de respuesta para los archivos se vería muy empobrecido por la sólo acción de encontrar los archivos dentro del catálogo maestro.

Es por lo anterior que resulta por demás recomendable la definición de catálogos usuarios que contengan grupos de archivos que compartan alguna característica en común; por ejemplo se puede definir un catálogo que contenga archivos de todos las aplicaciones línea tales como los CICS, (Customer Information Control System, es un administrador de terminales, programas transacciones y archivos utilizado para aplicaciones en línea) y otro que contenga los archivos de los procesos batch de actualización. De esta manera el método de acceso solo tendrá que buscar al archivo solicitado en el USERCAT que le corresponda.

Como siempre el equilibrio es importante en todo y tendremos que cuidar el no caer en un número tan exagerado de catálogos de usuarios que resulte contraproducente, pues ahora AMS además de localizar al archivo en el propio catálogo usuario tendría que recorrer una larga lista de catálogos usuarios para identificar aquel que contuviera el archivo solicitado.

Proaks [5] recomienda que un número de 4 catálogos por cada 1000 archivos es una cifra razonable, pero ello dependerá nuevamente de las características del ambiente de cómputo.

EJEMPLO DE UNA CONFIGURACION DE ARCHIVOS

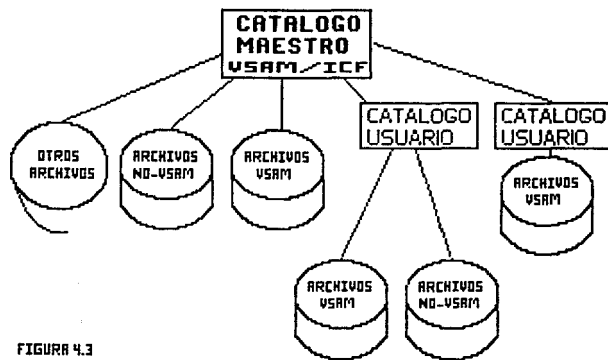


FIGURA 4.3

4.2.6.4. Definición de archivos VSAM.

El siguiente paso es ya la definición propiamente del archivo VSAM nuevamente utilizando IDCAMS. Esta definición especificará el catálogo al que pertenecerá el archivo más no la posición exacta en el disco, así como la cantidad de espacio que ocupará (tamaño del archivo) independientemente de que sea o no indexado, el tamaño de sus registros y otro tipo de parámetros orientados hacia el performance del archivo y que comentaremos mas adelante.

CONCEPTO DE "CLUSTER"

Para un archivo KSDS, un cluster es la combinación de un componente llamado DATA y otro llamado INDEX. El cluster provee forma de tratar ambos componentes como si fuesen uno sólo con el nombre del cluster precisamente, pero cada parte puede imprimirse de manera separada, ya que se puede dar a cada componente un nombre diferente y procesarlos por separado.

En el caso de archivos ESDS y RRDS también aplica el concepto de cluster con la diferencia de que no cuentan con un componente del tipo INDEX, sólo cuentan con su cluster y su componente DATA. Una idea que facilita enormemente el control para la administración de este tipo de archivos VSAM es nombrar los componentes DATA e INDEX exactamente igual que su cluster, agregando unicamente al final de cada nombre los prefijos DATA e INDEX respectivamente; así por ejemplo, un archivo KSDS cuyo nombre fuera USERDEPT.NOMINA.CHEQUES tendria como componentes:

USERDEPT.NOMINA.CHEQUES.DATA
y USERDEPT.NOMINA.CHEQUES.INDEX respectivamente.

El nombre del cluster se especifica al definir el archivo. Un nombre de cluster que contenga mas de 8 caracteres (como en el ejemplo anterior) deberá segmentarse en grupos de caracteres separados por un punto; cada grupo puede contener máximo 8 caracteres. Al primer bloque de 8 caracteres del nombre se le denomina Alias; en el ejemplo el alias seria USERDEPT.

Si en la definición del cluster no se especifica un nombre para los componentes DATA e INDEX, VSAM asigna automáticamente nombres propios a cada componente de acuerdo a la siguiente regla:

CLUSTERNAME.Tbbbbbbb.DFDyyddd.TAAAAAAA.Tbbbbbbb

en donde:

CLUSTERNAME es el primer calificador del nombre del cluster

yyddd es la fecha (año,día)

aaaaaaa y bbbbbbb son valores random generados por el sistema

Los archivos VSAM son definidos con el comando DEFINE CLUSTER. Cuando un cluster es definido, VSAM incluye en el catálogo las siguientes entradas para describir al archivo:

- Una entrada CLUSTER que define al cluster como un componente independiente.
- Una entrada DATA que describe al componente DATA del cluster
- En caso de ser un archivo KSDS, una entrada INDEX que describe al componente INDEX del cluster.

Todos los atributos del cluster son almacenados en el catálogo. La información almacenada en el catálogo proporciona los detalles necesarios para administrar el archivo o para acceder el cluster o los componentes individuales.

Toda la información descriptiva necesaria, así como opciones de seguridad, integridad y "performance" es especificada cuando se define el cluster. Esta información puede aplicar al componente DATA, al INDEX o a ambos

Información descriptiva:

- Tipo de organización (KSDS, ESDS, RRDS) según los parámetros INDEXED/NONINDEXED/LINEAR/NUMBERED.
- Longitud promedio y máxima de los registros que contendrá con el parámetro RECORDSIZE. Este parámetro no se especifica en los archivos lineales (RRDS).
- Longitud y posición de la llave en los registros de un archivo tipo KSDS con el parámetro KEYS.
- Nombre y "password" del catálogo en el que será definido el archivo con el parámetro CATALOG.
- Etiqueta del(os) disco(s) en que será ubicado el espacio para el cluster con el parámetro VOLUMES.
- Cantidad de espacio a reservar para el archivo según los parámetros CYLINDERS/TRACKS/RECORDS.
- Especificación si el archivo va a ser recatalogado o creado por primera vez con el parámetro RECATALOG.
- Cantidad mínima de espacio de buffer para operaciones de E/S en memoria real con el parámetro BUFFERSPACE

Información de "performance"

- Especificación para decidir si los registros pueden o no sobrepasar lo establecido por los intervalos de control con el parámetro SPANNED.
- Indicar explícitamente el tamaño de los intervalos de control, en vez de dejar a VSAM asignarlo con el parámetro CONTROL-INTERVALSIZE.
- Definir si se preformatean o no las áreas de control durante la carga inicial del archivo con el parámetro SPEED/RECOVERY.
- Para un cluster indexado (KSDS) se puede incluir además:
 - . Que se dupliquen los registros de índices a fin de reducir

los retardos por rotación del disco con el parámetro REPLICATE.

- . Colocar el grupo secuencial del componente index adyacente a la control-area de la componente data.
- . Ubicar el componente INDEX en un volumen(es) diferente(s) al del componente DATA con el parámetro VOLUMES.
- . La cantidad de espacio libre en los intervalos de control y áreas de control del componente DATA con el parámetro de FREESPACE.

Información de integridad y seguridad

- Identificación de rutinas propias de autorización para verificar que un requerimiento tiene derecho a acceder el archivo con el parámetro AUTHORIZATION.
- Verificar que las operaciones de escritura se han completado y que los datos ya pueden ser leídos con el parámetro WRITECHECK.
- Definir de que manera va a ser compartido el archivo entre programas y sistemas simultáneamente con el parámetro SHAREOPTIONS.

Una excelente idea para incrementar el "performance de los archivos VSAM" es distribuirlos en la mayor cantidad posible de discos, ya que ello repartirá los requerimientos de E/S en varios dispositivos y reducirá sensiblemente la utilización de los mismos. Claro está que esto implica un riesgo mayor en caso de falla física de alguno de los discos en los que esté distribuido el archivo.

4.3 Organizaciones de archivos.

En la actualidad, IBM [3] basado en recientes estudios ha concluido que un promedio de una persona es requerida para administrar 10 gigabytes de almacenamiento en disco, lo cual resulta dramático en las empresas modernas que cuentan con un sistema de computo complejo, ya que se ha podido también observar que sus capacidades de almacenamiento varían desde los 40 hasta los 300 gigabytes o más, lo cual implica una fuerte erogación económica en recursos humanos, ya que de acuerdo a este estudio, en teoría se requiere de entre 4 y 30 personas para administrar exclusivamente el espacio magnético del equipo de computo!!.

¿Cómo enfrentar este problema?. Es claro que no existe una solución general en cuanto a que la organización de datos en archivos pueda diferir considerablemente de una manera u otra.

En esta parte del trabajo se establecerán los pasos y las estrategias que se pueden seguir y el tipo de decisiones que se pueden tomar para que una instalación de este tipo pueda implementar nuevas facilidades que contribuyan a incrementar sensiblemente este rendimiento de 10 gigabytes/persona propuesto por IBM.

4.3.1 Definición de los nombres de los archivos.

El primer requisito para optimizar la administración de los archivos es contar con estándares en cuanto a los nombres que se les asignen a los mismos.

Existen diversas opiniones en cuanto a los criterios para definir adecuadamente estos estándares; de acuerdo con IBM [4] sólo 2 bastan para conseguir un control adecuado:

ORIENTACION.

Diversas orientaciones pueden considerarse para establecer una convención en cuanto a los nombres de los archivos. Existen sin embargo 2 clases principales en cuanto a la orientación de los -

archivos, estas son las de negocios o aplicaciones y las técnicas, y aunque hay muchas razones para seleccionar una u otra, se prefiere una orientación hacia las aplicaciones ya que es mas simple de continuar a través del tiempo, a diferencia de la orientación técnica que implica considerar una serie de características de los archivos (usualmente el rendimiento, el tiempo de respuesta, etc) que seguramente tendrán cambios radicales a través de los años.

Sin embargo ciertos nombres de archivos serán inevitablemente técnicamente orientados, como es el caso de los archivos SYS1's requeridos en la instalación de los sistemas operativos de IBM (MV, MVS, MVS-XA, ESA, etc), o aquellos archivos en los que el sistema ejerce algún control técnico, tales como las áreas de paginación y de "swap". (áreas en disco utilizadas por el sistema para el manejo de la memoria virtual)

Con esto se trata de explicar que no se debe adoptar una actitud inflexible en cuanto a elegir una u otra orientación y utilizar ésta para absolutamente todos los archivos del sistema; una combinación de ambas resulta óptima de acuerdo a las necesidades de la empresa.

NIVELES.

Como ya se ha mencionado, la flexibilidad es importante y las convenciones para establecer los nombres de los archivos deberán ser mínimas tratando unicamente de diferenciar perfectamente los datos de las aplicaciones de mayor peso de la empresa. Por razones técnicas, tales como la administración de catalogos y la seguridad en esta administración, es muy conveniente asegurarse de que el primer nodo o prefijo del nombre del archivo defina por sí solo el proceso o aplicación que lo utiliza; ello permitirá dividir a los procesos en bloques de datos no compartidos (a menos que diferentes procesos con distintos objetivos requieran acceder un mismo archivo); es decir, que los archivos que pertenezcan a los usuarios deberán tener un nombre con un cierto primer prefijo, mientras que los que pertenezcan a un proceso de producción deberán tener otros nombres con otros primeros prefijos, y de igual forma los archivos que sean accedados por procesos línea (CICS).

Un ejemplo de convención podría ser:

- . USRnnnn si el archivo pertenece a un usuario de una area nnnn para su propio uso
 - . DPTnnnn si el archivo es de uso a nivel departamental
 - . BATCHnn si es un archivo a utilizarse por procesos BATCH
 - . ONLnnnn si es un archivo para aplicaciones "ON-LINE"
- etc...

Nótese sin embargo que existen implicaciones en cuanto el esfuerzo requerido para el mantenimiento de un sistema de seguridad en el catálogo (Aquí con seguridad nos referimos a los accesos permitidos a los archivos por los usuarios y los procesos).

No sería muy aconsejable incluir mayores detalles de información en el nombre de los archivos para no hacer las convenciones demasiado complejas, excepto en los siguientes casos:

- 1- Acumulaciones lógicas de archivos que necesaria e indivisible mente tienen que ir juntas, por ejemplo, el caso de los archivos VSAM en los que se tiene un nombre para el CLUSTER, - otro para la DATA y otro para el INDEX (si es KSDS), siendo aconsejable que el DATA y el INDEX tengan el mismo nombre - que su CLUSTER añadiendo unicamente al final del nombre 'DATA' ó 'INDEX' según sea el caso.
- 2- Si una aplicación es implementada para múltiples usuarios, o si la aplicación misma cuenta on un número razonable de subdivisiones que realizan diferentes funciones, entonces se aconseja estandarizar también el segundo prefijo del nombre para diferenciar funcionalmente a los archivos en el caso - específico de esa aplicación.

En el caso del sistema operativo MVS una limitación que sería cuestionable aceptar es aquella impuesta por ISPF, el cual prefiere nombres de 3 nodos, y aunque no es obligatorio, si podría aplicarse en el caso de bibliotecas de usuarios con la -

forma USERID.X.Y. ISPF (Interactive System Program Facility) es un sistema que permite al usuario interactuar con el sistema operativo en tiempo real para funciones tales como edición, creación y ejecución de programas.

4.3.2 Categorías de los archivos.

El propósito de categorizar los archivos es el de simplificar y crear una metodología de procedimientos para la administración de los datos. Para poder pensar en una posible automatización de la administración de espacios magnéticos se requiere forzosamente que el autómata pueda tomar decisiones respecto a algún criterio preestablecido para los archivos, de preferencia criterios generales y no particulares para el archivo.

Es posible categorizar los archivos de acuerdo a los datos que contienen de muchas maneras. La mayoría de las clasificaciones que se presentan a continuación son válidas y son importantes.

La primera y principal clasificación de archivos está en razón del tiempo que se requiere que permanezcan los datos de dichos archivos; de esta manera podemos clasificar a los archivos de acuerdo a su contenido de la siguiente manera:

- Datos transitorios
Accesados por períodos relativamente cortos y sin ninguna utilidad posterior.
- Datos permanentes
Usados y actualizados continuamente y siempre válidos y útiles.
- Datos cíclicos
Usados y utilizados periódicamente, válidos y útiles sólo de una actualización a otra.
- Datos "generacionales"
Creados pero ya no actualizados posteriormente, siempre válidos pero posiblemente decreciendo en valor a través del tiempo.

Una segunda clasificación de archivos esta en función de la aplicación o proceso por los que son accedados:

- .- Datos "Online".
Usados por sistemas de acceso y actualización en línea.
- .- Datos de "tiempo compartido"
Usados por sistemas de tiempo compartido tales como TSO (Time Sharing Option).
- .- Datos "Batch"
Usados por programas batch tanto en modo lectura como escritura así como en actualización.
- .- Datos del sistema
Archivos de servicio utilizados por el sistema operativo, -tales como los de paginación o los de spool.

Una tercera clasificación es considerando su frecuencia de acceso:

- .- Muy frecuentemente accedados
- .- Frecuentemente accedados
- .- Poco accedados
- .- Casi nunca accedados

Finalmente existe la categorización de acuerdo a la importancia o lo crítico que resulta para la empresa la aplicación que accesa los archivos. Esta es una categorización que tiene que poner en una mano el efecto de no disponibilidad al servicio, es decir, si el servicio puede continuar en escencia a pesar de que los datos no estén disponibles, y en la otra mano el impacto al negocio de continuar la no disponibilidad de la información

4.3.3 Diferencias en la administración de VSAM y NO-VSAM.

Como ya se mencionò anteriormente. los archivos VSAM son requeridos en cualquier ambiente de datos IBM, sin embargo las técnicas para su administración varían considerablemente respecto a los archivos no VSAM. La razón principal de esto radica en su

interface de servicios de acceso (IDCAMS) la cual se utiliza para toda la administración y manipulación de estos archivos.

La mayor restricción que IDCAMS genera a estos archivos VSAM, es la necesidad de ubicarse o definirse única y exclusivamente a través de IDCAMS precisamente, en vez de hacerlo vía JCL (Job Control Language es el lenguaje de control para la ejecución de programas y procesos dentro del sistema operativo) o por un comando de "ALLOCATE" propio de TSO. Una segunda condición para la ubicación de VSAM es que cualquiera de ellos debe estar forzosa-mente direccionado a un catálogo, aunque ello no significa una tarea muy penosa si se siguen ciertas reglas muy simples que se mencionaran posteriormente, y en cambio ayuda bastante a su administración.

Si se desea implantar un control para la ubicación de archivos, un proceso similar deberá seguirse con IDCAMS para la ubicación de archivos VSAM, el cual deberá contemplar criterios tales como la selección del volumen en el que residirá y la cantidad de espacio que requerirá. Ciertamente que esta no es una tarea fácil y se antoja pensar que ISPF debería de contar con una interface más directa para la ubicación de este tipo de archivos. Para ambientes no-MVS resultaría prácticamente imposible la generación de un ambiente de archivos VSAM ya que se requeriría de un enorme esfuerzo por parte de los programadores y de los analistas del sistema para simular una interfase IDCAMS modificada.

Es por ello que que Cockrell [7] considera que IBM debería proporcionar una mejor interfase para la generación y administración de este tipo de archivos tan importantes.

4.3.4. Ubicación de archivos NO-VSAM bajo MVS.

Cuando se solicita un espacio magnético en disco, tradicionalmente se hace en unidades de cilindros, pistas o bloques, especificando el volumen explícitamente, ya sea por su etiqueta o bien por referencia de algún otro archivo. Lo anterior tiene ciertas implicaciones que es importante considerar:

1.- Las especificaciones y las características lógicas del dispositivo son generalmente bien conocidas.

2.- Las decisiones explícitas son tomadas con base en un conocimiento a detalle de donde están ubicados todos y cada uno de los archivos que ya contiene el disco. Ello generalmente repercute en la asignación de uno o varios discos a un sistema o a un grupo de sistemas particulares.

3.- Cuando el tipo de dispositivo es cambiado o los sistemas son reagrupados se requiere también una conversión completa de los JCL's que accedían a este dispositivo explícitamente, con todo el error potencial que esto implica ya que la posibilidad de una omisión está latente.

4.- El administrador de discos magnéticos invierte una gran cantidad de su tiempo "vigilando" las cargas de trabajo en los discos así como otros indicadores de "performance" así como haciendo las mejoras cuando hay cambios importantes en la reagrupación de aplicaciones.

5.- El administrador de las bases de datos invierte una gran cantidad de su tiempo "vigilando" el estado de la misma para liberar espacio y aumentar el "performance".

4.3.5. Ubicación de archivos VSAM.

La ubicación de archivos VSAM se enfrenta con todos los problemas anteriores y varios más, aunque las soluciones son diferentes. Para entender todas las permutaciones de ubicación de espacio en archivos VSAM es necesario conocer muy bien AMS (Access Method Services). La ubicación forzosamente debe direccionar uno o varios volúmenes por medio de IDCAMS, a diferencia de la ubicación de archivos NO-VSAM que si pueden ser ubicados sin especificar explícitamente la etiqueta de un volumen.

4.4. Respaldo de información.

El respaldo de un archivo es una copia del mismo hecha con el propósito de recuperarlo en caso de daño o pérdida. El respaldo y recuperación de archivos es sin embargo utilizado también para recuperación de discos completos.

Utilizar cintas o discos para respaldar información depende de qué tan rápido se requiera la recuperación de la misma además de otros factores tales como costos de los datos, costos de operación, espacio físico disponible, requerimientos de energía eléctrica y de aire acondicionado, tamaño de los archivos y finalmente si se desea que los archivos respaldados sean portables o no.

Tradicionalmente, el proceso de respaldo se ha realizado por lo menos de 2 maneras:

. Respaldo por aplicación .

Archivos específicos son respaldados en ciertos momentos o cuando se cumple alguna condición específica.

. Respaldo por sistema.

Respaldos completos del volumen o disco entero en intervalos regulares (diario, semanal o mensual)

4.4.1. Frecuencia y número de versiones de los respaldos.

La frecuencia y el número de versiones de la información respaldada son factores determinados primordialmente por los requerimientos de recuperación de los usuarios. En consecuencia esto deberá ser negociado precisamente con el usuario antes de tomar cualquier determinación al respecto.

Lo anterior no significa que no contemos con una serie de factores técnicos que nos pueden ayudar a definir la frecuencia de los respaldos y los cuales se mencionan a continuación:

- Los recursos y el tiempo requerido para resincronizar los datos después de una recuperación vía respaldo.
- La frecuencia a la cual cambia la información. Por ejemplo, si un archivo cambia su contenido diariamente, se aconseja un respaldo diario.
- La facilidad relativa para reconstruir el archivo en caso de pérdida. Por ejemplo, es más fácil reconstruir una biblioteca objeto que una fuente, por lo que es más importante respaldar la fuente que el objeto.
- La pérdida de tiempo por respaldo relativa a la terminación de algún trabajo.
- Tamaño de la ventana de respaldo. Ventana de respaldo se refiere al tiempo disponible para ejecutar los respaldos sin afectar otros procesos batch o línea.

Los factores que determinan el número de versiones de respaldo son usualmente:

- La probabilidad de ocurrencia de errores lógicos en los datos fuente. Usualmente los errores lógicos no son detectados durante el respaldo, lo que conlleva a una propagación de errores a los propios respaldos. Por ejemplo, si un usuario accidentalmente borra un cierto miembro de un archivo particionado y no se dá cuenta antes de que el respaldo se realice, este usuario requerirá una versión más antigua del respaldo para recuperar su miembro. Otro ejemplo podría ser cuando una aplicación cambia diariamente y el periodo máximo para detectar algún error es de una semana.
- La importancia de los datos. Sabemos que en cualquier procedimiento de respaldo existe el riesgo de que el propio respaldo se dañe o se pierda al mismo tiempo que los datos originales. Algunos datos son tan vitales o tan caros para regenerarse que varios respaldos duplicados deben realizarse para reducir el riesgo de pérdida total.

- Costo de almacenamiento de respaldos. El costo de mantener copias múltiples de respaldo es determinado por el costo de almacenar los "datos extras", de manera que el costo que la empresa pueda pagar para reservar estos datos extras influirá decisivamente en el número de copias que se puedan generar.

Se deben así mismo establecer ciertas convenciones para nombrar los calificadores de mas bajo nivel que identifiquen la frecuencia y la versión del respaldo, de acuerdo a los criterios que se mencionaron en el tópico de "nombres de los archivos", por ejemplo DSIFRE.DIARIO.discoxx.v00001, en donde el alias DSIFRE indica que es un archivo respaldado, DIARIO indica que se respalda diariamente, discoxx será el nombre del disco respaldado y V00001 indica que es la versión 1.

4.4.2 Proceso de recuperación en caso de desastre

A medida que las organizaciones almacenan mayor cantidad de su información en computadoras, la necesidad de proteger esta información de los desastres se incrementa. Aunque los desastres son raros, cuando ocurren, frecuentemente la propia existencia de la empresa depende de la habilidad de sus centros de proceso para recuperarse de los daños. Es por ello de suma importancia que un centro de proceso cuente con un plan completo de recuperación en caso de desastre.

Las fases para la planeación de una recuperación completa en caso de desastre son amplias y sumamente complejas por lo que salen del alcance de este trabajo. Un plan completo de recuperación en caso de desastre incluye entre otras cosas un amplio número de tareas, tales como procedimientos de evacuación del personal, medidas para proteger al equipo del daño, métodos para mantener al personal informado de las medidas de seguridad y varios más. En esta sección sólo se tratará el aspecto de recuperación de archivos en caso de desastre.

Los respaldos para recuperación en caso de desastre requieren de consideraciones muy especiales que normalmente no aplican a los otros tipos de respaldos.

La primer diferencia básica entre los respaldos regulares y aquellos utilizados por desastre es que estos últimos deben ser transportados a un lugar diferente al centro de proceso, lo que comunmente se denomina bóveda fuera de sitio. Qué tan remoto será este lugar dependerá del tipo de desastre para el que se está preparando; por ejemplo, en caso de un incendio limitado al centro de cómputo, guardar las cintas en el edificio de al lado bastará; en cambio en el caso de un gran terremoto, la bóveda fuera de sitio debería situarse a muchos kilómetros de distancia. En consecuencia, la distancia del sitio de recuperación puede variar notablemente entre empresa y empresa. Pero el simple hecho de que debe ser un lugar diferente dá especial significancia al procedimiento de respaldo y recuperación.

Dado que se requerirá transportar los archivos respaldados a un sitio distinto, ellos se deberán almacenar en un medio portable, es decir, en cinta magnética, aunque otra alternativa podría ser el transportar los datos directamente al centro alterno mediante una red de comunicaciones eficiente.

Se deben también considerar las diferencias entre el centro de proceso origen con el centro destino para evitar problemas de incompatibilidad. Más aún, se debe considerar cómo soportará el centro alterno la carga de trabajo adicional en caso de desastre del centro origen. Para manejar estas y otras consideraciones, se deberá desarrollar tanto el plan de respaldo como el de migración.

4.4.3 Desarrollo de un plan de respaldo.

Los puntos principales que deberemos considerar en un plan de respaldo son:

¿Cuáles son las aplicaciones críticas que requieren respaldarse?

La identificación de las tareas críticas de una organización resulta con frecuencia ser un trabajo muy difícil. El simple hecho de definir el término "crítico" es un problema delicado, ya que todos los usuarios consideran sus aplicaciones como críticas. Un problema adicional es que ciertas aplicaciones se convierten en críticas solo en ciertos periodos del año.

Como siempre, el factor principal para determinar que tan crítica es una aplicación es sin duda su valor dentro del negocio, es decir, qué tanto impactaría las finanzas de la empresa la pérdida de la aplicación en cuestión. Tal decisión deberá involucrar a los altos niveles gerenciales de la organización.

¿Cuáles son los archivos asociados con las aplicaciones vitales?

Una vez que hemos identificado las aplicaciones críticas, requeriremos reconocer los archivos asociados con ellas. Ello incluye a todos los archivos directamente relacionados con las aplicaciones (tales como archivos maestros y de actualizaciones), así como los archivos requeridos para soportar la aplicación. Podemos incluir además archivos que expliquen cómo recuperar la aplicación.

¿Con qué frecuencia se deben obtener los respaldos y cuántas versiones es conveniente mantener?

Como ya se comentó previamente, estas respuestas se deberán determinar involucrando a los representantes de los diversos grupos de usuarios. Tendremos además que contemplar que ciertos respaldos de archivos deberán ser sincronizados, por ejemplo, pensemos que un archivo de salida de la aplicación "A" es posteriormente utilizado como entrada de la aplicación "B", se tendrá entonces que considerar los respaldos de ambas aplicaciones para prevenir errores lógicos en los datos cuando estos sean recuperados.

¿Cómo minimizar las dependencias "ambientales" de los archivos?

Nuestro objetivo deberá ser siempre el de minimizar las dependencias ambientales de los archivos. Por ejemplo, asegurarnos de que los nombres esotéricos de dispositivos (y no la etiqueta del dispositivo), sean usados en la mayor medida de lo posible para evitar dependencias con los dispositivos de almacenamiento auxiliar. Claro está que no siempre se podrán evitar todas las dependencias ambientales, ya que por ejemplo, si nuestro ambiente utiliza JES3 y el ambiente donde se recuperará usa JES2, (Job Entry Subsystem es un subsistema del sistema operativo que se encarga de la administrar la entrada y salida de procesos para ejecución en el sistema) todos los JCL's tendrán que ser modificados durante el proceso de recuperación.

¿Qué documentación debe proporcionarse para la recuperación?

Considerando que posiblemente la gente especialista no podrá estar presente para ayudar durante el proceso de recuperación, se debe proporcionar la documentación suficiente que permita enterarse a cualquier persona de lo que se ha hecho.

Como ya se había comentado, se pueden generar archivos con instrucciones para la recuperación. Para los nombres de estos archivos instructivos se recomienda nuevamente seguir alguna convención para el nombre, por ejemplo, poner como último calificador del nombre la palabra INSTRUCT; ello los hará más simples de identificar.

¿Cómo mantener un inventario de los archivos respaldados?

Es indispensable establecer algún método para mantener la historia de cuáles archivos han sido respaldados, cuándo ha ocurrido ello, cuántas versiones existen, cuáles respaldos se encuentran en el sitio de recuperación, cuáles se encuentran aún en la instalación, etc.

4.4.4. Definiendo un plan de recuperación

Las principales preguntas que deberán responderse para planear un buen plan de recuperación son:

¿Qué modificaciones requerirán realizarse a los JCL's?

Identifiquemos todas las modificaciones necesarias a los JCL's de las aplicaciones a recuperar antes de procesarse en el nuevo ambiente, por ejemplo, cambios debidos a diferencias entre JES2 y JES3, renombre de ciertos archivos para evitar nombres duplicados con el centro de recuperación, cambios de los nombres de los ambientes de discos si los del centro de recuperación son diferentes, etc.

¿Cómo se garantizará la seguridad de los archivos?

De ser posible, el sitio de recuperación deberá definir nombres genéricos en RACF (Resources Access Control Facility es

un subsistema opcional diseñado por IBM para administrar la seguridad de la información) para proporcionar seguridad en los accesos de los archivos.

¿Cómo serán catalogados los archivos?

El sitio de recuperación deberá definir los alias necesarios para catalogar los nuevos archivos, además de ciertos grupos de generación (GDG's) en caso de requerirse. Los GDG's son archivos que sirven como base para generar varias versiones del mismo archivo pero realizadas en diferentes fechas, tales versiones tienen el mismo nombre y sólo varían en el último calificador que es el que indica el número de versión del archivo.

¿Cuenta el sitio de recuperación con los discos suficientes y el hardware adecuado?

Naturalmente, el sitio de recuperación debe de contar con CPU's y hardware en general compatible con el ambiente a recuperar, así como espacio en disco suficiente para soportar la transferencia.

¿Cuáles son las aplicaciones no críticas en el sitio de recuperación?

En caso de que alguna de las aplicaciones del lugar de recuperación requiera reprogramarse o cancelarse a fin de poder soportar la carga de trabajo adicional, las tareas no críticas deberán ser identificadas.

¿Cómo sabremos que el plan de recuperación realmente funcionará?

La única forma de saber si el plan de recuperación funciona o no es probándolo. No tenemos que esperar a que un desastre se presente para aplicarlo. Podemos probar nuestro plan de recuperación programando pruebas de transferencia periódicas (por lo menos 2 al año) y si existe una alta rotación de personal, más frecuentemente aún.

4.5. Control del uso de espacio en disco

La distinción entre archivos activos e inactivos es muy importante para la adecuada administración de los mismos.

Los **archivos activos** son aquellos que los usuarios y aplicaciones accesan por lo menos una vez desde su creación y tienen un uso planeado. Estos archivos tienen un uso regular. Los **archivos inactivos** son imágenes de los archivos activos y tienen un uso no planeado en condiciones normales; en estos se incluyen los archivos de respaldo, los archivos "generacionales", y archivos de bóveda fuera de sitio para situaciones de desastre. Los archivos inactivos pueden no estar catalogados y son accedados por diversos métodos de inventario o por ciertos comandos del sistema.

Es importante distinguir los archivos activos de los inactivos ya que la manera en que se administran es muy diferente. En primer lugar, estos 2 tipos de archivos tienen diferentes requerimientos de niveles de servicio. Por ejemplo, los requerimientos de "performance" de los archivos activos son mucho más elevados que los de los archivos inactivos; los usuarios pueden usualmente tolerar algún retraso en el acceso de archivos inactivos. El espacio y la disponibilidad son también muy diferentes en cada caso.

Las diferentes necesidades de los archivos activos y los inactivos determinan su posición en los subsistemas de almacenamiento magnético. Debido precisamente al alto nivel de performance requerido por los archivos activos, ellos residen normalmente en disco mientras que los archivos inactivos usualmente residen en cinta. Se tendrá que trabajar en conjunto con el usuario para encontrar los requerimientos para ambos tipos de archivos.

4.5.1. Administración del espacio magnético

Cuando hablamos de administrar el espacio en disco nos resultará más fácil entender los problemas y sus soluciones si consideramos que existen 2 perspectivas diferentes en este tópico: el usuario y el administrador del espacio.

El principal objetivo del usuario es minimizar la ocurrencia de fallas sin importar la cantidad de recursos que tenga que utilizar. Las fallas consumen el tiempo del usuario mientras puede determinar la causa de las mismas y resolverlas. Ante esta perspectiva el usuario se realiza preguntas del siguiente tipo:

- ¿Dónde puedo guardar mis datos?
- ¿Puedo conseguir el espacio que necesito?
- ¿Cuánto tiempo este espacio me será disponible?
- ¿Qué opciones debería especificar?
- ¿Qué puedo hacer si se me acaba el espacio?

Se puede observar que estas preguntas están encaminadas básicamente a la eliminación de posibles fallas por espacio, y no a la optimización del performance de sus archivos

Naturalmente que el administrador de espacios deberá estar también interesado en reducir al máximo el número de fallas por espacio en disco. Sin embargo otro de sus objetivos principales es optimizar al máximo el uso de este espacio. Se debe pues de asegurar de que algún usuario o grupo de usuarios no abusen del recurso en detrimento del resto de usuarios. En consecuencia un administrador de espacios magnéticos se hace el siguiente tipo de preguntas:

- ¿Cómo puedo controlar la ubicación de archivos?
- ¿Cómo puedo minimizar la sobrecapacidad del espacio en disco?
- ¿Cómo puedo eliminar los archivos ubicados pero no utilizados?
- ¿Cómo puedo saber cuando un archivo ya no se requiere?

A continuación se proponen ciertos criterios y métodos que conllevarán a un equilibrio de los objetivos del usuario en relación a minimizar fallas por espacio en disco y los del administrador de espacios para optimizar el uso de los mismos.

4.5.1.1. Problemas potenciales de la administración de espacio.

Todos los archivos tienen un ciclo de vida, y unos mas rápido que otros, pero todos pasan por las siguientes etapas:

- 1.- Planeación o establecimiento de las características del archivo.
- 2.- Ubicación física del archivo.
- 3.- Acceso y utilización del archivo.
- 4.- Borrado del archivo por obsolescencia.

Durante la etapa de planeación, el usuario establece la cantidad de espacio que requerirá el archivo y otras opciones tales como sus características de la DCB (Data Control Block), parámetro con el que se especifica la longitud de los registros (LRECL), el tamaño del bloqueaje (BLKSIZE), la organización de los registros (DSORG) y el formato de los registros (RECFM). Después el usuario (generalmente), ubica su archivo, y a medida que el tiempo pasa y el archivo va llenándose de información y el espacio asignado originalmente puede ir extendiéndose si es que fue especificada una ubicación secundaria en el parámetro de SPACE y finalmente, el archivo es dejado de utilizar o es dado de baja.

Muchos problemas pueden presentarse durante este proceso. Por ejemplo, durante el proceso de planeación, el usuario en la mayoría de los casos desconoce la cantidad exacta de espacio que se requerirá y seguramente ubicará un espacio mayor al que realmente requería. Si la ubicación es dada por medio de JCL (Job Control Language), otros usuarios pueden copiarlo y de esta manera multiplicar el problema. Si la ubicación se hace vía TSO, el usuario puede adquirir vicios de ubicación de archivos sin estar conciente del desperdicio de este recurso tan importante.

Otros usuarios conocen las ventajas potenciales de un mejor performance al ubicar espacio en archivos secuenciales muy grandes con unidades de cilindros en vez de tracks y pueden intencionalmente sobre-solicitar espacio a pesar de que solo utilicen un 1% del mismo.

El proceso de borrado de los archivos puede nunca ocurrir. Los usuarios ubican sus archivos simple y sencillamente porque en ese momento los necesitan, pero generalmente nunca tienen el tiempo para darlos de baja cuando ya no los requieren o bien porque se les olvida o porque ni siquiera se enteran de cuándo el archivo dejó de ser necesario.

De no existir un proceso o una área encargada de analizar y corregir este tipo de situaciones, los problemas de falta de espacio en disco se repetirán con continua frecuencia.

Si esta área de administración de espacios magnéticos educa a los usuarios y establece y difunde políticas para la ubicación y el borrado de archivos, ellos se convertirán en un eje de apoyo hacia los objetivos del administrador de espacios de proporcionar siempre el espacio suficiente a sus usuarios.

4.5.2. Los usuarios pueden ayudar en la administración del espacio en disco.

A continuación se presentan algunas recomendaciones que se han obtenido mas por experiencia que por algún proceso teórico, que definitivamente contribuyen a la optimización del uso de espacio en disco en un ambiente MVS complejo. VSAM queda excluido aquí ya que este tipo de archivos fue comentado a detalle en hojas anteriores.

4.5.2.1. Ubicación de espacio secundario.

El valor primario de espacio en disco, debería de especificar la cantidad total de espacio requerido inicialmente. El espacio secundario permite la extensión automática del espacio a medida que el archivo crece y no se desperdicia espacio con una ubicación primaria excesiva.

Para liberar al usuario de la tarea de calcular la cantidad de espacio secundario a especificar, y para asegurar que ésta es siempre especificada, es aconsejable codificar una "EXIT" para -

sustituir esta especificación, tal como la ya existente en MVS/XA y conocida como IFGOEXOB para asignar un espacio secundario si no se especificó ninguno y actualizar el espacio si esto fuera necesario.

4.5.2.2. Tamaños de bloque eficientes.

La elección de un tamaño adecuado de bloque en el parámetro BLKSIZE de la DCB puede mejorar considerablemente la utilización de espacio así como el performance del archivo.

Los tamaños de bloque que son muy pequeños gastan mucho espacio, por ejemplo, en los "huecos" conocidos como "gaps entre los bloques", es decir, los espacios no utilizados que quedan entre bloque y bloque de datos. Consideremos que en discos 3380, un bloque de tamaño de 80 bytes utiliza unicamente el 14% de la capacidad del track. La cantidad de espacio residual se decrementa a medida que el tamaño de bloque se aproxima a un track.

Diversos muestreos realizados durante la presente investigación, demostraron que en un ambiente mixto de discos en el que se cuenta con modelos 3330, 3350 y 3380, un tamaño de bloque de 6233 bytes resulta ideal. El acto de bloquear los archivos proporciona la mayor independencia de los dispositivos de acceso directo ya que se logra un excelente uso de espacio sin importar el modelo de disco que se esté utilizando. Por ejemplo, asumiendo un archivo de registros de longitud fija de 80 caracteres, un tamaño de bloque de 6160 utiliza optimamente el 94% del espacio en discos 3330, el 96% de los 3350 y el 90% de los 3380.

Encontramos también que en un ambiente con discos 3350 y 3380 exclusivamente, el tamaño de bloque ideal es de 9076 bytes ya que asumiendo nuevamente el ejemplo de un archivo con registros de longitud fija de 80 caracteres, un bloqueaje de 9040 bytes utiliza el 94% de los discos 3350 y el 95% de los 3380.

En el caso de contar unicamente con discos modelo 3380, un bloqueaje de 23476 bytes utiliza el 99% del espacio de los dispositivos, aunque es importante puntualizar que si los tamaños de bloque son tan significativamente incrementados, se puede

requerir incrementar también los tamaños de la región de los programas que utilizan estos archivos.

Finalmente, debemos considerar que los archivos de tipo particionado (bibliotecas) con miembros pequeños, no se verán significativamente beneficiados por el hecho de aumentar su bloqueaje.

Además de especificar tamaños de bloque eficientes para los nuevos archivos, sería conveniente evaluar los de los archivos ya existentes para lo cual existen diversos productos en el mercado tales como el ISMF (Interactive Storage Management Facility) con el cual el usuario puede obtener fácilmente los tamaños óptimos de bloque, o mejor aún, existe otro producto llamado BLOCK-MASTER de manera automática cambia el bloqueaje de los archivos con tamaños de bloque no óptimo al momento en que el archivo inicia una operación de E/S.

4.5.2.3. Liberación del espacio desperdiciado.

La mayoría de los usuarios alojan mas cantidad de espacio del que realmente necesitan. En algunas ocasiones esto es debido a que utilizan JCL's viejos y no ajustan las cantidades de espacios reclamadas dentro de los JCL's a sus necesidades actuales. Otras veces se debe simplemente a que evitan la posibilidad de tener fallas por falta de espacio suficiente en sus archivos. Es por esto que existe un parámetro que pocos utilizan en sus parámetros de espacio en disco; esta opción se incluye dentro del parámetro de SPACE en el JCL y permite eliminar el espacio no utilizado por el archivo después de que éste es cerrado; es decir, si se hace un reclamo de 100 cilindros y sin embargo sólo se utilizaron 25, con la opción de RLSE el resto de los 75 cilindros serán declarados como disponibles dentro del disco.

Existen sin embargo algunos archivos en los que no debe utilizarse la opción de RLSE:

- Archivos que son cerrados en un paso y reabiertos nuevamente para salida en un paso posterior.

- Archivos secuenciales en los que deberá reservarse espacio para futuras actualizaciones.

- Archivos VSAM (en esta organización no existe la facilidad para liberar el espacio no utilizado).

4.5.2.4. Borrado de los archivos al momento de su cierre.

El parámetro de JCL FREE=CLOSE origina que el archivo en el cual se especifique sea borrado inmediatamente después de que es cerrado en vez de tener que esperar a que termine el paso completo, siendo esto último la mayor ventaja con respecto a la opción de DELETE en terminación normal dentro del parámetro de DISP (disposición del archivo).

Por supuesto que una excepción a este parámetro son los archivos que son cerrados en un paso y reabiertos para actualización en un paso posterior.

Para los archivos VSAM no existe un equivalente de FREE=CLOSE.

4.5.3 Ambientes de discos.

El administrador de espacios tal y como se ha mencionado previamente tiene como meta fundamental garantizar espacio magnético principalmente en disco, a fin de que los procesos puedan ejecutarse sin fallas por faltas de espacio en disco.

No debemos olvidar que nos encontramos en una época de grandes avances en las áreas de cómputo y que existe una fuerte tendencia hacia la automatización. El administrador de espacios no puede quedar al margen por lo que ya existen en el mercado una serie de paquetes que le hacen la vida menos amarga al eliminarle funciones manuales operativas y permitirle dedicar mas tiempo al diseño y al análisis. Existen paquetes tales como el DFHSM y el DFDSS ambos de IBM, y otros tales como el ADAM (Automatic Disk Analisis Management) de otra empresa que con base en una serie de reglas y parámetros especificados por el administrador de es-

pacios, realizan las funciones de borrado, migración, compresión, etc de archivos de manera automática cuando han cumplido alguna cierta condición.

Se puede observar entonces que para poder pensar en una automatización de este tipo de funciones operativas del administrador de espacios se requiere de haber definido estándares para la ubicación de archivos. Los estándares deben comenzar desde el nombre hasta el tamaño y las frecuencias de acceso. Sin embargo, en un ambiente de cómputo de grandes proporciones (es decir, que cuenta con 1000 o más archivos) los archivos pueden variar tanto entre sí en sus condiciones (tamaños, permanencia en disco, tipo de aplicación, etc) que resulta indispensable el hecho de crear ambientes de discos con características bien definidas para el tipo de archivos que van a contener.

Cuántos discos deba contener cada ambiente dependerá completamente de las necesidades de la empresa y de la cantidad de espacio que ocuparán los archivos que vivirán en cada ambiente. Tendremos entonces que definir políticas de ubicación de archivos para cada ambiente. (En el capítulo anterior se analizaron con detalle los aspectos administrativos que implica el establecimiento de este tipo de políticas).

Un "storage pool" es un conjunto predefinido de discos usados para almacenar datos con características comunes. Tales grupos se especifican en una biblioteca de carga del sistema operativo que es leída al momento de su arranque (IPL) para reconocer los diferentes ambientes de discos o "storage pools" del sistema. Esto permite combinar las necesidades de almacenamiento en disco con las características físicas de los dispositivos sin que el usuario requiera conocer o entender la configuración de hardware del centro de cómputo, pues basta con que especifique dentro del parámetro de UNIT de sus JCL's el nombre esotérico del grupo de discos que contienen las características deseadas; el resto del trabajo lo realiza el sistema operativo asignando alguno de los discos del grupo y accedando o grabando la información requerida.

La definición de "storage pools" o ambientes de discos presenta una serie de significativas ventajas:

Aumento de la productividad del usuario.

Con ambientes de discos bien definidos y administrados por un grupo profesional de administradores de espacio, los usuarios incrementan su productividad en actividades encaminadas a lograr un mejor uso de sus propios recursos y herramientas, además de que el espacio es mejor utilizado y en consecuencia habrá menos probabilidades de fallas por falta de espacio en disco.

El espacio en disco es usado mas eficientemente.

Los usuarios que requieren grandes cantidades de espacio pueden compartirlo mas efectivamente con los usuarios que requieren menos. Además, con ambientes de discos, el administrador de espacios puede hacer uso de los paquetes existentes en el mercado para administrar el espacio de forma mas eficiente.

Mayor efectividad en la administración del "performance".

Ubicar archivos dentro de diferentes "storage pools" utilizando nombres esotéricos permite a MVS (el sistema operativo) balancear la actividad de E/S dentro de un grupo grande de dispositivos. El "performance" o "actuación" de los discos se vuelve mas consistente reduciendo así sus tiempos de respuesta.

Mayor efectividad en la administración de disponibilidad.

La agrupación de datos dentro de ambientes permite desarrollar procedimientos de respaldo y recuperación mas eficientes así como utilizar las herramientas existentes en el mercado para facilitar estas actividades al administrador de espacios.

La instalación de nuevos discos es mas fácil.

Nombres esotéricos en los campos de UNIT permite cambiar los dispositivos de la instalación sin que se requieran cambios en los JCL's de los usuarios. Dado que los usuarios no están concientes

de los discos existentes en cada "storage pool", ellos no se ven afectados por la eliminación o inclusión de nuevos discos en la configuración.

La ubicación de archivos es mas sencilla.

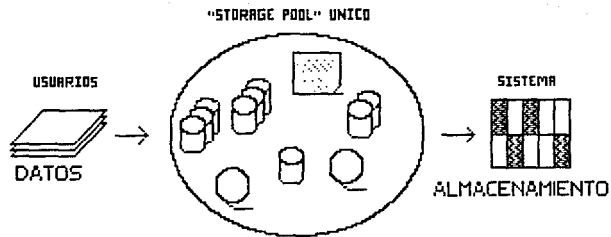
Cuando se establecen nombres esotéricos en los JCL's, los nuevos archivos pueden ser automáticamente direccionados al "storage pool" apropiado. Si el sistema operativo controla la ubicación de los archivos, el usuario es liberado de la penosa tarea de especificar las características del dispositivo, tales como el número de serie del volumen, tamaño del bloque, y direcciones del cilindro o del track en que será ubicado el archivo.

La generación de ambientes de discos es el primer paso para separar las necesidades lógicas de los datos de las características físicas de los dispositivos.

4.5.3.1 Determinación de los ambientes de discos.

En un ambiente ideal de administración de espacios, todos los datos residirían en un solo ambiente de discos. El sistema determinaría la ubicación de los datos y automáticamente evaluaría el "performance", la disponibilidad de los datos, el espacio disponible y la seguridad. Basado en requerimientos predefinidos del usuario el sistema compaginaria las necesidades lógicas de los datos con las características físicas de los dispositivos instalados y aseguraría que cada disco fuera usado de la manera mas eficiente. (figura 4.4).

Desafortunadamente, los ambientes actuales de administración de espacios magnéticos no han alcanzado aún este ideal. A pesar de que los sistemas operativos se vuelven mas inteligentes día con día, no es tan sencillo manejar todos los aspectos de la automatización de la administración de espacios. Ningún producto en el mercado es capaz de administrar todos los diferentes tipos de archivos de un ambiente de cómputo.



**AMBIENTE IDEAL DE
ADMINISTRACION DE
ESPACIOS MAGNETICOS**

Fig. 4.4

De esta manera, no podemos hablar de un ambiente único, sin embargo es viable hablar de varios ambientes, cada uno con sus propios criterios y procedimientos de administración.

Existen diversas formas de determinar el mejor conjunto posible de ambientes para los datos de un centro de cómputo:

1) Una manera es listando los archivos por aplicación con las diversas áreas usuarias e identificar los requerimientos de espacio, permanencia, performance y seguridad. Manejar los datos por aplicación puede ser útil para evaluar futuros crecimientos pero puede ser también muy engañoso. Recordemos que la administración de espacio en disco es un asunto que comprende todos los datos del centro de cómputo y hacerlo sobre bases aplicativas puede resultar en una solución parcial.

2) Otra manera para determinar ambientes es listar los archivos y clasificarlos por su estructura (por ejemplo archivos de bases de datos o de sistema, VSAM o particionados, etc) y dividirlos en ambientes de acuerdo a las herramientas con que se cuente para administrarlos. Esto es una buena idea para cuantificar las limitaciones en el ambiente de datos actual y para identificar los archivos que requieren un tratamiento especial. Debido a que las capacidades de las actuales herramientas de administración de espacio en disco son una de las razones por las que se requiere más de un ambiente de discos, podemos estimar el número de ambientes de acuerdo al número de herramientas o técnicas disponibles o requeridas.

3) Un tercer método para subdividir en ambientes los archivos, es tomando en cuenta sus consideraciones técnicas de performance, permanencia y espacio ocupado y entonces eliminar aquellos archivos que no cumplan con estos requerimientos.

Cualquiera que sea el criterio para establecer ambientes de discos es importante no olvidar el concepto de "ambiente único" para que cuando las herramientas de administración de espacios continúen evolucionando estemos en condiciones de adaptar todo nuestro ambiente global de datos sin demasiados esfuerzos a los avances que la tecnología ofrezca.

4.5.3.2 Identificación de los grupos de archivos y de sus requerimientos.

Identificar todos los tipos de datos en la instalación y cuantificar sus requerimientos de performance, seguridad, permanencia y espacio es una tarea bastante difícil.

Un ejemplo de grupos de archivos podría ser:

- Archivos interactivos
- Archivos VSAM
- Archivos permanentes
- Archivos ISAM
- Archivos multivolumen
- Archivos muy grandes
- Archivos temporales
- Archivos del sistema
- Archivos de bases de datos

Después de haber identificado los tipos de archivos de la instalación y de considerar sus restricciones y requerimientos se está en condiciones de desarrollar un conjunto apropiado de ambientes de discos en el sistema. Ya se mencionó con anterioridad que el número ideal de ambientes es uno solo puesto que mientras menor número de ambientes se tenga, mas eficiente se volverá la administración de estos ambientes; esto no es una contradicción; los ambientes son lamentablemente necesarios hoy en día para la administración de espacios, sin embargo un número grande de ambientes dificulta su control y las mejoras en performance y espacio se ven limitadas por la complejidad del ambiente.

Los centros de cómputo que manejan cantidades muy grandes de archivos (mil o mas archivos), irremediamente requerirán de la definición de varios ambientes. Un ejemplo que hoy en día funciona en mas de una empresa es el que se muestra en la figura 4.5.

En este ejemplo cada ambiente contiene los siguientes tipos de archivos:

EJEMPLO DE AMBIENTES DE DISCOS

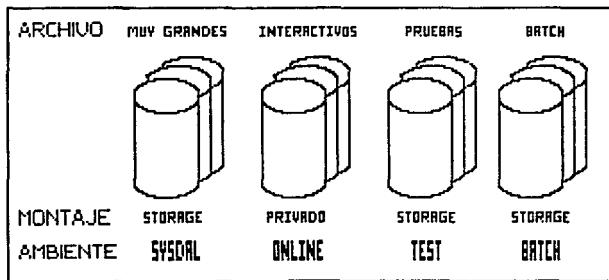


Fig. 4.5

Muy grandes

Archivos secuenciales de mas de 100 cilindros.

Interactivos

Archivos (generalmente VSAM) utilizados por aplicaciones con respuesta en línea tales como los CICS

Pruebas

Archivos utilizados por aplicaciones dedicadas a probar y desarrollar nuevas facilidades.

Batch

Archivos de procesos batch que requieren un buen performance.

Cualesquiera que sean los ambientes generados, es muy conveniente realizar revisiones periódicas para su revaluación. Si se realizó una buena planeación para esta estructura, es poco probable que sufra modificaciones o se requiera agregar nuevos ambientes. Lo que definitivamente si esta sufriendo cambios fuertes es la cantidad de archivos que caen dentro de cada ambiente como lo muestra la figura 4.6 donde observamos una clara tendencia hacia el incremento de las aplicaciones en línea (interactivas) así como hacia la desaparición de las aplicaciones batch.

4.5.3.3 Elección de los nombres esotéricos de los ambientes.

El nombre esotérico identifica al grupo de discos que comparten características en su contenido y uso. La figura 4.7 ilustra la diferencia entre nombres esotéricos e identificación de volúmenes (VSN, Volume Serial Number). Previamente se comentó que los nombres esotéricos se definen en el sistema operativo durante su inicialización mediante la macro instrucción UNITNAME y son referenciados por los usuarios en el parámetro UNIT de JCL durante la ubicación de algún volumen.

DISTRIBUCION DE ARCHIVOS EN UNA INSTALACION TIPICA

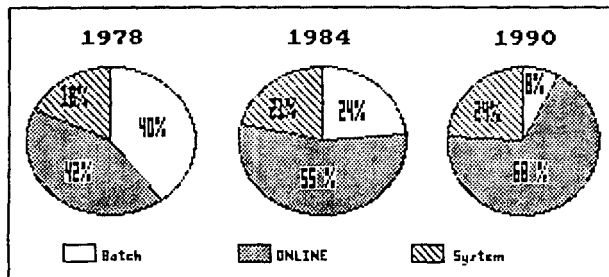
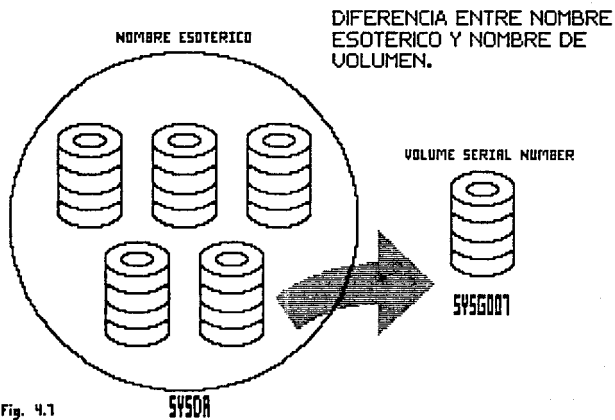


Fig. 4.6



La teoría dice que cuando se especifica un nombre esotérico en el parámetro de UNIT al ubicar un nuevo archivo, se está dejando al sistema decidir el mejor disco dentro de ese ambiente para ubicar dicho archivo. MVS verifica la actividad actual de operaciones de E/S de todos los discos del ambiente indicado por el nombre esotérico y selecciona el volumen montado como STORAGE con el menor tiempo de retraso (tiempo de encolamiento mas tiempo de canal). Sin embargo la distribución nunca es perfecta, por lo que resulta conveniente efectuar un constante "monitoreo" de actividad y utilización de discos y canales y realizar las afinaciones que resulten convenientes.

Los nombres esotéricos de los ambientes deben ser de 1 a 8 caracteres y el nombre por si mismo debe indicar el tipo de archivos que va a contener. Es una buena idea crear al igual que en los archivos, estándares para los nombres de los ambientes y que los nombres de los discos ahí contenidos sigan el mismo patrón que el ambiente. Un ejemplo de convención en cuanto a los nombres de los ambientes puede ser el siguiente:

PRMDA : Discos que contendrán archivos permanentes.
VSMDA : Para volúmenes que contienen solo archivos VSAM.
CICBDA : Discos con archivos y bases de datos de aplicaciones línea (CICS).
DB2DA : Discos con bases de datos de DB2.
SCRDA : Discos con archivos temporales (scratch).
PAGE : Discos con archivos de paginación
SPLDA : Discos de spool.
SYSDA : Discos con archivos del sistema operativo.
etc ...

y un ejemplo para los VSN's de los discos podría ser tomar las primeras 3 letras del ambiente seguidas por la dirección física del disco. Por ejemplo un disco cuya dirección fuera 6F1 y perteneciera al ambiente PRMDA tendría como etiqueta PRM6F1.

Lo anterior es solo un ejemplo de los patrones que se pueden seguir para estandarizar los nombres de los ambientes y de los discos, sin embargo cada empresa puede elegir los propios que mas se adecuen a sus necesidades.

4.5.4 Atributos del montaje y uso de discos.

Los atributos de montaje de un disco se utilizan para indicar al sistema operativo que ese dispositivo en particular deberá ser montado en línea y estar disponible para ubicación de nuevos archivos. Los atributos de uso indicarán el tipo de ubicación de archivos que será permitido.

Cuando el sistema es arrancado, los atributos de montaje y uso de cada disco son especificados en el miembro VATLST00 de la biblioteca de sistema SYS1.PARMLIB. Una vez que el volumen es montado, este permanece montado en línea hasta que el operador de algún comando de UNLOAD o de VARY OFFLINE en la consola del operador. El volumen puede ser montado nuevamente usando el comando de MOUNT. Si el disco no es especificado en VATLST00, éste se montará público al arranque del sistema.

Los discos pueden ser montados con cualquiera de los siguientes 3 atributos:

- PUBLIC:** Para volúmenes que contendrán sólo archivos temporales, o volúmenes para ser usados como discos de trabajo del sistema.
- PRIVATE:** Para volúmenes que contienen archivos permanentes y aislados. Los discos privados pueden ser ubicados solamente si se especifica explícitamente el VSN en el requerimiento de ubicación de algún archivo dentro del JCL.
- STORAGE:** Volúmenes que pueden contener archivos permanentes o -- temporales y que son ubicados aún sin especificar el -- VSN, bastando con especificar el nombre esotérico en el requerimiento de ubicación dentro del JCL.

Los discos pueden ser montados como public, private o storage, mientras que las unidades de cinta sólo pueden ser montadas como públicas o privadas.

Debemos estar conscientes de que por más que se optimice la información en los dispositivos de almacenamiento, los tiempos de respuesta de estos dispositivos continuarán limitando las expectativas de tiempos de respuesra semejantes a los proporcionados por la unidad central de proceso. Claro está que de no existir un correcto balance en las cargas de trabajo por dispositivo y una adecuada administración, el sistema de cómputo caería en un completo caos.

Sin embargo, ante la urgente necesidad de tiempos de respuesta menores a los que actualmente ofrecen los dispositivos de almacenamiento, ha comenzado a surgir en los últimos años un nuevo concepto de almacenamiento de datos: la memoria CACHE; este es el tema de análisis del siguiente capítulo.

CAPITULO 5

MEMORIA CACHE, DASD FAST WRITE Y DUAL COPY

5.1. La importancia de un tiempo de respuesta mas rápido.

En el mundo competitivo de los negocios, el incremento de productividad es un camino clave para lograr las metas. El reto es, encontrar herramientas de trabajo ágiles que permitan a los usuarios incrementar su productividad, lo cual redundará en la elaboración de sus trabajos en una forma más rápida y eficiente.

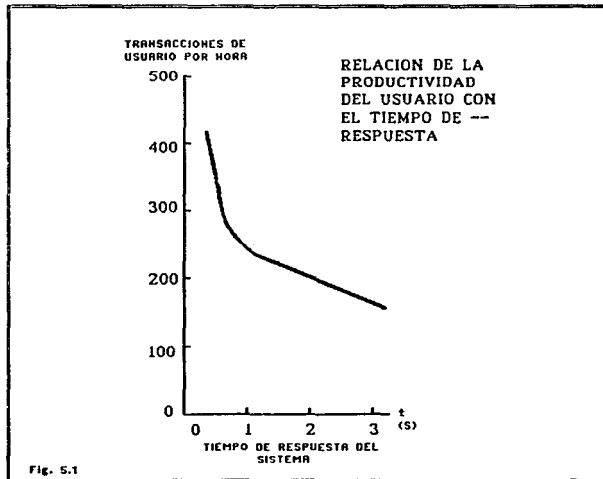
Los usuarios de las computadoras pueden ser mas eficientes si el sistema de computo es más eficiente por si mismo. Una medición de la eficiencia de un sistema, es la medición de su tiempo de respuesta (el tiempo de respuesta es el intervalo en el que el usuario da entrada a un comando y hasta que recibe la respuesta en su terminal), mayor eficiencia en el sistema se ve reflejado en un tiempo de respuesta mas rapido.

Un estudio realizado por Arvind J. Thadhani [9] encontró que el número de transacciones que un usuario profesional completa, se incrementa notablemente con las mejoras del tiempo de respuesta, principalmente cuando los tiempos de respuesta son menores a un segundo. La figura 5.1 ilustra lo anterior con claridad.

Este dramático incremento en la productividad del usuario, cuando los tiempos de respuesta estan por debajo de un segundo, se explica facilmente, ya que cuando el usuario digita un comando esta pensando en lo que tiene que realizar posteriormente, si el tiempo de respuesta es muy lento se rompe esta secuencia de pasos mentales, y le toma tiempo al usuario retomar la idea que tenia al momento de digitar el comando, reflejándose esto en una disminución de la productividad.

Como ya lo observamos, la creciente necesidad del procesamiento de datos, así como de la automatización de gran parte de las actividades que nuestra sociedad tiene que realizar, obliga a que la evolución de los sistemas de cómputo continúe a un ritmo acelerado.

Dentro del ambiente IBM, y debido a los grandes volúmenes de información que tienen que manejar las principales empresas del



mundo, así como la necesidad de ponerla al alcance de todos los usuarios, existe una gran preocupación por mejorar los tiempos de respuesta, lo cual es requerido para mantenerse a la vanguardia en el mundo de los negocios.

Se observa con todo esto la gran importancia que tienen los dispositivos de almacenamiento auxiliar, principalmente las unidades de disco, donde existen las principales pérdidas de tiempo, en cuanto al procesamiento de datos se refiere, ya que debido a la gran capacidad de información que puede ser almacenada en cada unidad de disco se generan grandes cuellos de botella, requiriendo así mayor velocidad en estos dispositivos periféricos.

Hasta hace apenas 2 años, se hablaba de capacidades de memoria principal, en "MAIN FRAMES", de 16 MB. a lo sumo, velocidades de procesamiento máximas de 10.5 MIPS (Millones de instrucciones por segundo), capacidades de almacenamiento en discos magnéticos de 317.5 MB. En la actualidad todas estas cifras han sido considerablemente incrementadas, como ejemplo tenemos en algunos modelos de computadoras IBM, capacidades de 126 MB. de memoria real, con posibilidades de crecer aun más, velocidades de procesamiento de más de 60 MIPS, así como la posibilidad de almacenar en discos magnéticos hasta 2650 MB. por cada unidad de disco tal como se comentó ampliamente en el capítulo 2.

Como se observa, con el incremento en la capacidad de almacenamiento de información en discos magnéticos, se requiere equipos de entrada/salida de mayor velocidad de acceso, así como algoritmos más eficientes de acceso y administración de dichos dispositivos (lo cual ya ha sido discutido ampliamente), principalmente discos magnéticos, ya que sin esto, los tiempos de respuesta de cualquier sistema de cómputo se verían seriamente afectados muy a pesar de tener los procesadores de mayor velocidad existentes en el mercado.

A la par del desarrollo obtenido en el Hardware, las compañías diseñadoras de sistemas y equipos de cómputo se han preocupado por mejorar sus sistemas operativos a fin de disminuir al máximo los tiempos de respuesta, un ejemplo claro es IBM, compañía que desarrolló e instaló los subsistemas de canales con reconfiguración dinámica, lo cual permite a cualquier requeri--

miento de entrada/salida de un sistema de computo acceder a sus dispositivos perifericos hasta por 16 trayectorias diferentes, desconectándose la trayectoria donde se realizó la operación de entrada mientras es localizado el dato requerido, permitiendo así que la trayectoria pueda ser utilizada por otros requerimientos durante el tiempo de busqueda del dato. Una vez encontrado este, el disco se reconecta al sistema por cualquier trayectoria disponible para terminar el requerimiento, con lo cual existe la posibilidad de que las trayectorias entre discos y procesadores puedan atender más requerimientos, sin depender del tiempo o utilización que en ese momento tengan los dispositivos de almacenamiento auxiliar, y no solo por una trayectoria como ocurría hasta hace unos años.

Es cierto que los dispositivos periféricos, han tenido un desarrollo acorde al ambiente informático, sin embargo para disminuir más aun la posibilidad de altos tiempos de respuesta en los sistemas de computo, provocados por los accesos a discos, surgió el nuevo concepto de memoria cache. Por medio de los controladores 3990, IBM proporciona nuevas y novedosas facilidades de cache tales como el "Cache Read", "DASD FAST WRITE" y "Dual Copy".

5.2 Conceptos de memoria CACHE.

La memoria cache es una memoria electrónica de alta densidad que está ubicada en la unidad de control y por lo tanto, es compartida por varias unidades de disco.

La principal función de la memoria cache es almacenar datos de gran utilización y que normalmente residen en las unidades de disco, dentro de la memoria electrónica del controlador (Memoria Cache), con lo cual tenemos la posibilidad de transferir datos entre esta memoria y el canal a la velocidad del canal, sin tener que acceder hasta la unidad de disco magnético.

Si consideramos que la velocidad de transferencia del canal es de 4.5 MegaBytes por segundo y la de los discos es de 3.0 MegaBytes por segundo, se logra inicialmente un ahorro del 50% en el tiempo de transferencia de información así pues podemos asegurar que es más rápida la transferencia de datos entre la memoria cache y el canal, que entre el disco y el canal, ya que además de mejorar en un 50% el tiempo de respuesta por las velocidades de transferencia, ya no existe la demora o atraso en los tiempos de respuesta que originan los Seek's de los discos (tiempo que tarda el disco en posicionar sus cabezas lectoras en los datos requeridos), así como los originados por el rotacional delay del propio disco, (esto es el tiempo que tarda en dar una revolución el disco para leer o escribir los datos), así pues es notorio este beneficio, ya que con la memoria cache se disminuye la utilización de la unidad de control al igual que la utilización de las unidades de disco.

5.2.1 ¿Cómo se realiza una función de lectura en memoria CACHE?

Si una copia de los datos está en cache cuando el procesador inicia un requerimiento de lectura (en inglés "read hit"), la unidad de control transfiere el dato deseado de la memoria cache al canal; si la copia del dato no está en cache (en inglés "read miss"), la unidad de control envía un requerimiento de datos directamente del canal a la unidad de disco y, al mismo tiempo, escribe el dato (más la pista completa del disco en donde estaba contenido el dato) dentro de la memoria cache, previendo su posi-

ble utilización en los requerimientos posteriores hacia los datos de la pista ubicada en cache para que mas adelante sean leídos en cache y se conviertan en un READ HIT.

5.2.2. ¿ Cómo se realiza una función de escritura en memoria CACHE ?

Si una copia de los datos esta en la memoria cache cuando el canal inicia un requerimiento de escritura ("write hit"), la unidad de control escribe el dato directamente al disco y, al mismo tiempo, escribe el dato en cache. El registro en cache es actualizado por si este vuelve a ser referenciado, sin embargo, antes de que la unidad de control pueda señalar la operación como completa, éste debera verificar que el dato fué satisfactoriamente escrito en la unidad de disco; así pues el registro es escrito simultaneamente en cache y en la unidad de disco y con la señal de " fin de dispositivo", enviada por la unidad de disco, se da por terminada la operación.

La escritura en disco esta provista de una integridad de datos ya que la copia del registro en cache y disco es identica.

Cuando el registro a ser actualizado no es encontrado en cache, la condición es llamada "ESCRITURA NO ENCONTRADA" (En ingles WRITE MISS), por lo que el registro es actualizado directamente en disco y ya no es escrita en cache.

5.3. Comparación entre sistemas con la facilidad "CACHE" y sin la facilidad "CACHE".

Es posible realizar una comparación del performance (rendimiento), entre sistemas que cuentan con memoria cache y sistemas sin memoria cache, estos valores del performance son obtenidos con modelos matemáticos y no deberán considerarse bajo algún ambiente específico. En un intento por asegurar la confiabilidad de los resultados y las comparaciones, se asumen valores fijos para ciertas variables del sistema, los cuales son suministrados a los modelos matemáticos; por este motivo algunos valores asumidos pueden no aplicar para un ambiente dado de un sistema de producción.

Quando se estan evaluando mejoras al subsistema de discos, es importante remarcar que estas mejoras no deberan forzosamente verse reflejadas en el rendimiento del sistema de computo, ya que existen otras variables tales como la velocidad del procesador, el tamaño de la memoria principal, sistema operativo y la red de comunicaciones las cuales pueden limitar el rendimiento del sistema, a pesar de que el subsistema de discos haya sido optimizado.

Tomando en consideración lo anterior, se obtuvieron los siguientes resultados que se muestran en la siguiente tabla:

OPERACION	TIEMPO DE SERVICIO EN UN CANAL DE 3 MB/SEC
Read Hit	2.8 ms.
Read Miss	20.4 ms.
Write Hit	20.4 ms.
Write Miss	20.4 ms.

El termino "Miss", significa que no fué encontrada copia dentro de la memoria cache del registro requerido, Asi pues el termino "Read Miss" ocurre cuando el sistema intenta leer un re-

gistro en cache y este no es encontrado y de igual forma el termino "Write Miss" indica que la actualización de un registro no pudo ser realizada en memoria cache.

Como se puede observar en la tabla, el "Read Hit" es la operación más rápida de las cuatro que se ejemplifican, la operación de "Read Miss", es un ejemplo del tiempo que se toma una lectura en sistemas sin memoria cache.

5.3.1. Tamaño y utilización de la memoria CACHE.

El tamaño de la memoria cache es de 32, 64 128 Y 256 MegaBytes los cuales estan divididos en segmentos de 16 KB, Que son asignados de acuerdo al tamaño de la información que es requerida por el sistema; si el tamaño de la información requiere de más memoria cache, esta le sera asignada en segmentos de 16 KB hasta satisfacer el requerimiento. Varios algoritmos de "Cacheado" (Asignación de cache a cierta información), determinan como se manejará el cache para cada tipo de operación de I/O. Dentro de los algoritmos de "Cacheado" se incluyen el normal, secuencial, ByPass-Cache y la inhibición de carga a cache, los cuales podran ser manejados óptimamente de acuerdo a los metodos de acceso que sean utilizados en la instalación.

El algoritmo de uso normal de cache es utilizado hasta que no le sea definido otro algoritmo via software al sistema. El "cacheado normal" significa que la información permanecera en cache después de ser referenciada en una operación de lectura y permanecerá en cache hasta que los requerimientos de utilización de memoria cache, y el algoritmo de "datos no referenciados ultimamente" permitan que nuevos datos sean "encimados".

El "Cacheado Normal" aplica para ambas operaciones de lectura y escritura, y es el algoritmo más apropiado para metodos de acceso aleatorio, en donde son incluidos los metodos Basic Direct Access Method (BDAM), el metodo de acceso por llave del Virtual Storage Access Method (VSAM), el metodo particionado o de directorio Partitioned Access Method (PAM), entre otros.

Para metodos puramente secuenciales como lo son el Basic Sequential Access Method (BSAM), el Queued Sequential Access Method (QSAM) asi como el proceso secuencial de VSAM es invocado el algoritmo de "Cacheo Secuencial".

El acceso de archivos de una manera secuencial, es detectada por el programa de canal del sistema operativo con lo cual el sistema asigna el algoritmo de "cacheo secuencial", y asi el procesador no solo trae a memoria cache una pista del disco, sino que trae cinco pistas incrementando los "Read Hits" en cache y con ello el rendimiento de nuestroa sistemas.

El método de "Bypass-cache" no utiliza la memoria cache, esto es, los requerimientos de I/O son directamente en el disco.

El método de "Inhibición de carga de cache", mantiene una copia de lo que actualmente esta en cache y evita que otras copias posteriores sean cargadas en cache.

5.4 Ventajas de la memoria CACHE.

Existen grandes ventajas que son detectadas en forma inmediata, después de activar la memoria cache, algunas de estas ventajas se mencionan a continuación:

A) Disminución de la utilización de los discos.

Debido a que los accesos a unidades de discos magnéticos son menores con la instalación de memoria cache, la utilización o tiempo de ocupación de las unidades de disco se ve disminuida, lo cual se logra transfiriendo bloques de información requerida por el sistema, de la unidad de discos a la memoria cache ubicada en la unidad de control; a partir de este momento y para los subsecuentes requerimientos de esta información, el viaje no se tendrá que realizar hasta la unidad de disco, sino que el requerimiento viaja solo hasta la unidad de control, por lo que el dispositivo queda listo por si el sistema requiere información contenida en la misma unidad de disco magnético.

B) Disminución en la utilización de las unidades de control.

Lo anterior se logra gracias a las velocidades en que se transfiere la información entre las unidades de control y los canales, olvidándose, la unidad de control, de tener que administrar los requerimientos de salida de información desde las unidades de discos, ya que la información fué encontrada en la memoria cache y regresada de inmediato.

C) Velocidades de transferencia iguales a las del canal.

Debido a las grandes velocidades de procesamiento que tienen actualmente los "Main Frames", se hace necesaria una transferencia de datos igualmente rápida, la tendencia mostrada es innegable, y prueba de esto es la posibilidad de transferir información solo entre la unidad de control y el canal, olvidándose de tener que hacer un viaje hasta la unidad de disco.

D) Disminución en tiempos de trabajos con acceso al azar a discos.

La disminución del tiempo de proceso, en procesos con archivos de acceso aleatorios (Random), Se logra gracias a los algoritmos de administración de la memoria cache, con los cuales se puede detectar cuál es la información que con mayor frecuencia se utiliza y mantenerla dentro de memoria cache, ya que existe una gran probabilidad de que ésta se vuelva a utilizar.

E) Increíbles mejoras en los tiempos de procesos secuenciales.

Quando el sistema de memoria cache detecta que el tipo de proceso que se esta llevando a cabo es secuencial, en lugar de transferir pista por pista del disco hacia memoria cache, lo hace de 5 en 5 pistas, consiguiendo así que casi en el 100% de los requerimientos de información ésta sea encontrada en memoria cache.

F) Disminución en los tiempos de proceso de las áreas de "SORT".

Las áreas de "SORT", son áreas en disco que son utilizadas para realizar ordenamientos de manera ascendente o descendente de una gran cantidad de registros, por lo que los procesos que utilizan estas áreas realizan un excesivo I/O; con base en esto, si las áreas de sort son asignadas en memoria cache, el 98% del proceso se realiza dentro de la memoria cache, a velocidades de procesamineto muy superiores a las que se podrian obtener en unidades de disco.

G) Posibilidad de asignar la memoria CACHE de acuerdo a las necesidades.

Debido a que el tamaño de la memoria cache es finita, tenemos la posibilidad de asignar la memoria cache solo a los discos que consideremos vitales en nuestra instalación o bien a los archivos que de acuerdo a su metodo de acceso o importancia de estos mismos puedan ser incluidos en la memoria cache.

5.5 La facilidad de "DASD FAST WRITE".

Ya observamos en las paginas anteriores, parte de los trabajos realizados por la compañía IBM para mejorar el rendimiento de los sistemas de cómputo; en primera instancia se mencionó lo relacionado a sistemas con memoria cache, medio por el cual, la gran mayoría de los requerimientos de lectura en discos, en lo que respecta al tiempo de respuesta, se ven notablemente reducidos, sin embargo se observa que las operaciones de escritura pueden ser aun mejoradas. Por este motivo IBM lanzó al mercado durante 1990, el concepto de "DASD FAST WRITE" (escritura rápida en disco).

5.5.1 Conceptos y funcionamiento de DASD FAST WRITE.

El concepto DASD FAST WRITE es una facilidad que nos permite disminuir notablemente los accesos a las unidades de disco magnético, ya que los requerimientos de escritura son enviados hacia la memoria del controlador, el cual se encarga de terminar el requerimiento de escritura, permitiendole a los programas continuar trabajando sin tener que esperar hasta que la información quede grabada en disco.

DASD FAST WRITE mejora el rendimiento del subsistema de discos, ya que no es requerido grabar la información en disco de una forma inmediata, DASD FAST WRITE almacena simultaneamente la información en memoria cache y en la memoria no volátil (NVS, Non-Volatile Storage) incluida en el controlador.

Debido a que una copia de la información es puesta en la memoria no volátil, la unidad de control al momento de que la información es registrada en memoria cache y en la memoria no volátil, envía simultaneamente instrucciones de fin de canal y fin de dispositivo, permitiendo a los programas continuar trabajando sin esperar a que el dato sea escrito en la unidad de disco. La información permanece en memoria cache y en la no volátil hasta que el dato es escrito en disco para liberar memoria cache y así puedan continuar trabajando las demás aplicaciones con esta facilidad.

Una gran parte de estos requerimientos de escritura operan directamente con memoria cache sin tener que viajar hasta el disco, resultando en el mismo rendimiento que el observado en las operaciones de lectura. La figura 5.2 ilustra lo anterior.

Resumiendo las grandes ventajas que nos brinda la facilidad de DASD FAST WRITE, podemos listar las siguiente:

- Permite que las operaciones de escritura sean ejecutadas a velocidad de cache.
- Provee una integridad de datos similar a cualquier escritura en discos.
- Elimina los requerimientos inmediatos de escritura de información al disco.
- Libera al canal para otras actividades y operaciones adicionales.
- Mantiene una copia de la información en la memoria no volátil del controlador, hasta que ésta sea vaciada al disco.

DASD FAST WRITE

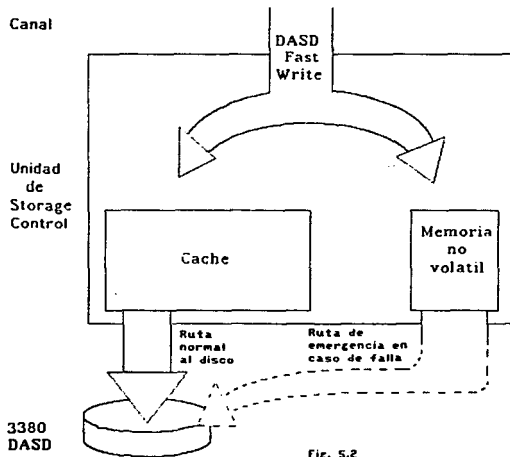


Fig. 5.2

5.6 La facilidad de "DUAL COPY".

Hemos observado los grandes avances obtenidos en lo que respecta a mejorar los tiempos actuales de acceso a disco, ya sea por requerimientos de lectura o de escritura; se ha comentado también sobre los incrementos en la capacidad de almacenamiento que las unidades de disco magnetico tienen dia a dia, haciendo a cualquier sistema de computo muy dependiente de estos dispositivos, ya que conforme pasa el tiempo, es mucho mayor la cantidad de información que puede ser almacenada dentro de un dispositivo magnetico, quedando en gran riesgo esta información al momento de ocurrir una falla en estos dispositivos; si pusieramos como ejemplo a una unidad de disco magnetico que contenga al sistema operativo o bien bibliotecas vitales del sistema, el riesgo seria para todo el sistema de cómputo.

Preocupados por esta situación, los especialistas de IBM han creado un concepto que viene a revolucionar la confiabilidad de los sistemas de cómputo, así como la disponibilidad, no solo de la información almacenada en disco, sino también del sistema de cómputo en general, logrando así un mayor porcentaje de disponibilidad en sistemas de cómputo.

5.6.1 Conceptos y funcionamiento de DUAL COPY.

El concepto de DUAL COPY, significa poder contar con una copia idéntica de cualquier unidad de disco (principalmente los que contengan información vital), en cualquier momento, esto es, que Dual Copy nos permite crear una copia de los datos de un volumen, y almacenarlos en un dispositivo diferente dentro del mismo subsistema de discos en tiempo real.

5.6.1.1 Ventajas de DUAL COPY.

Entre las principales ventajas de esta facilidad de los controladores cache podemos mencionar las siguientes:

- Proteje a volúmenes críticos de fallas en unidades de disco simples ya que su información es respaldada en tiempo real durante cada actualización en un disco alterno.
- Permite la concurrencia de operaciones de escritura de información en dispositivos separados.
- Las operaciones de transferencia de información, son automáticamente switcheadas al volumen secundario, si una falla imposibilitara la utilización del volumen primario, el switcheo es transparente para la aplicación.
- Provee de una estrecha operación de ambas unidades de disco, creando una copia respaldo, por si se perdiera el acceso a cualquier dispositivo conectado en dual copy.
- Asegura que la segunda copia sea idéntica a la primera.

Dual Copy mejora considerablemente la disponibilidad de los datos. El estado que guardan los dispositivos de Dual Copy es almacenado en la memoria no-volátil. Dual Copy es activada y desactivada por comandos de utilería del sistema. El utilizar Dual Copy no requiere modificaciones a las aplicaciones ni a los métodos de acceso.

Para funcionar con esta facilidad, los dispositivos requieren estar conectados en PAR DUPLEX. (En inglés, Duplex Pair), es decir, un dispositivo secundario con acceso a uno primario; las operaciones de Dual Copy son administradas por el subsistema de entrada/salida. Todas las operaciones de E/S son aplicadas sobre el volumen primario y la unidad de control se encarga de actualizar la copia secundaria automáticamente. Así pues los datos son accedidos en el volumen secundario en caso de que el primario presentara alguna falla o no pudiera ser accedido.

Debido a que el dispositivo secundario esta fuera de linea, (no esta montado para el sistema), el procesador solo reconoce a un dispositivo: el dispositivo primario; después de que la información es escrita en el volumen primario, son enviados al sistema el "fin de canal" y "fin de dispositivo", y posteriormente la unidad de control completa la actualización de memoria cache a disco; esta última operación es transparente para el sistema. La figura 5.3 muestra el flujo de las operaciones de Dual Copy.

Con esto se concluye el presente trabajo de investigación sobre diversas técnicas encaminadas a controlar y hacer óptimo el uso de espacios magnéticos en un ambiente de cómputo complejo como los que existen hoy día, no sin antes recalcar que al pretender implantar una administración de recursos de almacenamiento centralizada, la empresa deberá realizar un análisis exhaustivo de las condiciones actuales de su ambiente de cómputo para determinar cuáles de las propuestas aquí planteadas son las que mejor encajarían en su muy particular situación, podría inclusive darse el caso de que en algún aspecto ninguna de las alternativas aquí planteadas fuera factible de instalarse, ein embargo a nivel conceptual se tendría ya un buen pivote para ser tomado como punto de referencia.

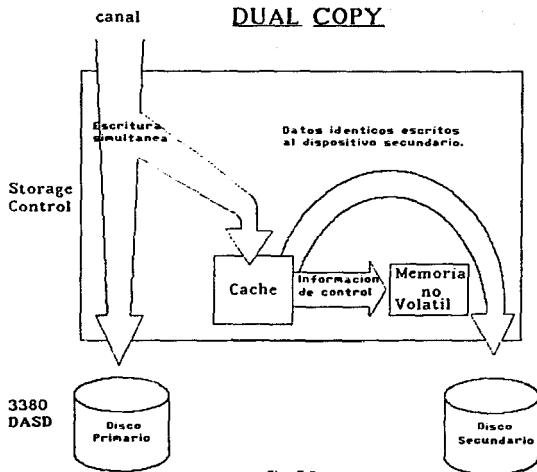


Fig. 5.3

B I B L I O G R A F I A

- 1.- Williams. Davis. Sistemas operativos de la computación. Fondo Educativo Interamericano, U.S.A. (1985).
- 2.- Xephon. C. The DASD handbook. Technology Transfer. England. (1989).
- 3.- IBM. The MVS/XA storage management library. IBM Corporation. U.S.A. (1990).
- 4.- IBM. Using QSAM, BSAM, ISAM and VSAM to manage data. IBM. Corporation. U.S.A. (1990).
- 5.- Proaks. J. Configuring storage subsystems. Mc.Graw Hill. U.S.A. (1984).
- 6.- Cadow H. OS/360 Job Control Language. Englewood Cliffs. U.S.A. (1986).
- 7.- Cockrell D. Data Management for mainframes. Scientific American. U.S.A. (1987).
- 8.- Hoffman K.L. The MVS enviroment. Basic Books Inc. U.S.A. (1986).
- 9.- Arvind J. Introduction to CACHE processing. Prentice Hall. U.S.A. (1990).

IMPRESA "MARTINEZ"

Tesis Directas y Mecanografiadas en I. B. M.
URGENTES EN 24 HORAS

Rodolfo Martinez Cerezo

Portal de Sto. Domingo 12 Altos 11
06010 México, D. F. Tel. 510-25-24