

Informe de Tópicos III

N° 3: "Sistemas de Archivos"

Nombre: Juan Pablo Arancibia González

Carrera: Ingeniería en Computación e Informática

Fecha: 10/04/2014

Índice

<i>Introducción.....</i>	<i>3</i>
<i>Desarrollo.....</i>	<i>4</i>
<i>Sistema de Archivos FAT.....</i>	<i>7</i>
<i>Conclusiones.....</i>	<i>10</i>
<i>Anexo: Compatibilidad Sistemas de Archivos y Sistemas Operativos.....</i>	<i>11</i>
<i>Enlaces de Interés.....</i>	<i>12</i>

Introducción

El sistema de archivos es el componente de un sistema operativo encargado de administrar y facilitar el uso de las memorias periféricas. Entre sus funciones se encuentran asignar el espacio a los archivos, la administración del espacio libre y la administración del acceso a los datos existentes. Además estructuran la información en una unidad de almacenamiento (Disco Duro, Unidad Portable, Memoria SD), lo cual después será representado en un gestor de gráficos, finalmente destacar que la mayoría de los sistemas operativos utilizan su propio sistema de archivos. Además en este informe detallaremos acerca del Sistema de Archivos FAT, uno de los dos sistemas de archivos que se utilizan en el sistema operativo Windows.

Desarrollo

Los sistemas de archivos son estándares desarrollados por cada desarrollador de un sistema operativo, el cual indica la forma en que estos archivos van a ser almacenados y leídos (desde/hacia unidades de disco duro, unidades de disco óptico, unidades de disco sólido, memorias USB, etc.), además de la forma de arranque del sistema operativo, además la palabra formatear, además de borrar los datos existentes de una unidad de almacenamiento de archivos, en realidad prepara a la unidad de almacenamiento para guardar y utilizar la información a partir del sistema de archivos elegido.

Cuando se almacena un archivo, cualquiera que sea, este mantiene sus características y propiedades, pero se acopla al sistema de archivos existente en el sistema operativo donde se guarda el archivo, esta es una metáfora similar a la explicación de que una misma palabra suena y se pronuncia diferente en cada idioma, en este caso la palabra sería el archivo, y el idioma el sistema operativo.

Los tres sistemas operativos más importantes tienen sus propios sistemas de archivos y pueden ser compatibles o incompatibles entre sí, además las unidades de disco físico tienen también su propio sistema de archivos, a continuación se dará a detallar las características de estos sistemas de archivos.

El sistema operativo Linux utiliza diferentes sistemas de archivos a partir de sus diferentes distribuciones, VxFS, XFS, GPFS, GFS, ReiserFS, OCFS2, LTFs, por solo nombrar algunos, pero el más utilizado es el sistema ext, creado por Remy Card en 1992, y cuya actual versión ext4, comenzó a desarrollarse en 2006 y se convirtió en oficial a fines de 2008 (kernel Linux 2.6.28), su capacidad puede llegar hasta 2 Exbibytes (1 Exbibyte = 2^{60} de Bytes) y los ficheros hasta 16 Tebibytes (1 Tebibyte = 2^{40} de Bytes), además incorporó los extents, un conjunto de bloques físicos contiguos que mejoran el rendimiento cuando se trabaja con ficheros de gran tamaño y reduciendo así su fragmentación, ext4 es retrocompatible con ext3 con sólo ejecutar un par de comandos, permite la reserva de espacio para un fichero específico, ideal para transmisiones en streaming y bases de datos, supera los 32000 de límite de cantidad de subdirectorios impuesto por ext3 llegando a 64000 e incluso más, aplicación del asignador multibloque para un fichero, entre otras ventajas con sus antecesores, la gran desventaja de ext4 y ext3 es que sus archivos no pueden ser leídos en los S.O. Windows (salvo Windows NT mediante complementos) ni con Mac OS (mediante complementos se logra compatibilidad parcial), eso si los sistemas operativos de Linux si son capaces de leer cualquier sistema de archivos aunque con el HFS Plus de Mac tienen algunos problemas.

Otro sistema de archivos que se destaca es ZFS, Creado por Sun Microsystems, para su sistema operativo Solaris. ZFS se destaca por sus grandes capacidades, integra en un sola unidad la administración de volúmenes y el sistema de ficheros, sistemas de archivos ligero y una administración de espacio de almacenamiento sencilla pero muy eficiente, teniendo como cantidad máxima de archivos 2^{48} en un sistema de archivos de 128 bits. El sistema operativo Solaris (u Open Solaris) lee los sistemas de ficheros FAT, NTFS (vía complemento NTFS-3G) y ext3, mientras que el sistema de archivos ZFS es compatible con Free BSD Y Solaris, y mediante software externo en Mac OS y Linux.

El sistema operativo Mac de Apple, tiene el sistema de archivos HFS (Hierarchical File System), y su evolución HFS+, aparecido en 1998, sus grandes ventajas son el uso extendido de los nombres con un máximo de 255 caracteres y la codificación de texto hecha en Unicode para permitir sistemas de escritura mixtos sin problemas, en cuanto a compatibilidad los sistemas Mac OS son capaces de leer los sistemas de archivos de Windows, como los de disco físico, similarmente el Sistema Operativo Windows son capaces de leer archivos hechos en HFS, mientras que Linux y FreeBSD tienen compatibilidad parcial.

En el siguiente apartado se hablará del sistema de archivos FAT (File Allocation Table).

Los sistemas operativos de Windows tienen dos sistemas de archivos, FAT y NTFS (New Technology File System), este último hizo su aparición en el sistema operativo Windows NT y se convirtió en estándar a partir de Windows 2000, una de sus ventajas es que su tamaño específico del cluster se puede definir a partir de los 512 bytes independiente del tamaño de la partición, eso si el sistema no es recomendado su uso en particiones menores a 400 Megabytes de capacidad de espacio, su máxima dimensión de archivo es de 16 Tebibytes y el tamaño máximo de volumen es de 256 Tebibytes, en ambos casos puede llegar en futuras arquitecturas a los 16 Exbibytes por teoría. El sistema de archivo NTFS suele ser muy estable lo cual contribuyó notablemente en la evolución de las diferentes versiones de Windows NT, otra ventaja es que el uso de los ficheros está basado en el modelo de estructura del Árbol-B, lo cual mejora el acceso a los ficheros, y una mejora sustancial de la fragmentación de archivos respecto a FAT que era una de sus mayores desventajas. Los sistemas Operativos Windows que utilizan NTFS son Windows 2000, Windows XP, Windows Vista, 7 y 8, estos sistemas pueden reconocer los archivos FAT y de sistemas de archivos de disco físico y parcialmente HFS y ext mediante complementos. Mientras los archivos NTFS pueden leerse en Linux, Mac OS y Free BSD mediante el controlador NTFS-3G, que fue creado a base de ingeniería inversa debido a que Windows todavía no ha liberado el código fuente de NTFS.

Por último, cabe señalar la tecnología de sistemas de archivos que se utilizan en los formatos de discos físicos agrupadas en la ISO 9660, generalmente permite en una sola sesión (aunque existen discos regrabables), almacenar los datos a grabar y luego de eso, no volver a ser escritos, CDFS es el formato que se utiliza en los discos compactos y UDF (Universal Disk Format) es otro estándar que es común en los discos de DVD y Blu-Ray, generalmente tiene 3 funciones, la primera es la misma del sistema CDFS, que es escribir datos en una única sesión, la segunda, la llamada Tabla de Locación Virtual o VAT, que permite el uso de escribir datos en múltiples sesiones, y por último el acceso de escritura aleatoria limitado, el cual se utiliza en discos regrabables. UDF generalmente es compatible con los sistemas operativos Windows, Mac OS, Linux y BSD, aunque para Windows 95, Windows 98, y Windows ME soportan solo UDF 1.02, y Windows 2000 soportan UDF 1.02 y UDF 1.5. Así mismo, el sistema operativo Linux el Kernel 2.2.x soporta UDF 1.02 y el Kernel 2.4.x soportan UDF 1.02 y UDF 1.5, a partir del kernel 2.6.26 soporta UDF 2.5 y UDF 2.6, del cual es la versión actual del Protocolo.

Sistema de Archivos FAT

FAT (File Allocation Table, Tabla de Asignación de Archivos), es el sistema de archivos más utilizado debido a su fácil compatibilidad con muchos sistemas operativos (Windows, Linux, Free BSD, Mac OS, Solaris), y utilizarse en unidades de memoria portátil, además de tener una estructura sencilla, pero tiene variadas desventajas en este último apartado, como la fragmentación de archivos (es decir, cuando se borra y se escribe un archivo, quedan espacios de fragmentos en la unidad de disco duro, lo cual se soluciona desfragmentando la unidad frecuentemente), la limitación de que se utilizaba como máximo 11 caracteres para el nombre de archivo (8 como nombre propio y 3 de extensión), y la ausencia de permisos de seguridad, es decir, cualquier usuario puede acceder, modificar y eliminar archivos, a diferencia de otros sistemas de archivos que tienen permisos de seguridad en carpetas sensibles.

FAT fue creado en 1977, pero no sería hasta 1980 cuando apareció por primera vez en un sistema operativo, llamado QDOS, antecesora original de los sistemas MS-DOS de la década de los 80, en ese sistema existió la versión original de FAT (a quien a partir de ahora llamaremos FAT12, ya que las direcciones de bloque sólo podían ocupar 12 bits), además el máximo de espacio que podía manejar era de 32 Megabytes agrupados en clusters de 8 Kilobytes, en ese entonces se utilizaban disquetes de 5 ¼ pulgadas cuya capacidad máxima de espacio era de 160 kilobytes, y los discos duros para PC no superaban los 20 Megabytes de capacidad, lo cual FAT12 pudo estar vigente algunos años, un problema que surgió con la expansión de capacidad fue que al aparecer disquetes de 5 ¼ pulgadas de alta densidad (1,2 Megabytes de capacidad), forzó la optimización del cluster de 2 sectores de 16 bits a sólo 1, reduciendo así su rendimiento de tiempo en lectura y escritura de datos.

En 1988, Windows lanzó MS-DOS 4.0, que incorporó la segunda versión de FAT (que llamaremos en adelante FAT 16), el cual aumento la capacidad de clusters a 32 Kilobytes, el máximo de tamaño que puede tener un archivo es de 2 Gigabytes, y el límite máximo de espacio disponible es de 90 Gigabytes, a partir de Windows 95, se logró soportar nombres más largos de archivo mediante un truco de almacenar los índices de los directorios, a este proceso se le llamó LFN (Larger File Name), aunque el manual de Windows 95 lo denominó como VFAT (Virtual FAT), generando una confusión ya que VFAT se utilizaba en Windows 3.11, lo cual usaba un modo de protección de 32 bits (que lo usaba también Windows NT que utilizaba este modo de protección por defecto, solo que se llamó FASTFAT) que esquivaba el núcleo de MS-DOS, accediendo a la BIOS o a la unidad de hardware del disco, además de permitir el uso de caché, este proceso VFAT mejoró el acceso a los sistemas de datos.

En 1996, apareció la tercera y última versión de FAT (FAT32), que aumentó su capacidad potencial hasta 8 terabytes, aunque por limitaciones de los escaneos de los discos (por motivos de diseño), Sólo se podían manejar hasta los 32 Gigabytes de capacidad, aunque ese límite podía ser superado en otros sistemas operativos como Linux si estos son instalados como unidades maestras, en el mercado FAT32 apareció primero en Windows 95 OS2, aunque se necesitaba formateo para adoptar de FAT16 a FAT32, pero a partir de Windows 98, se podían convertirse sin necesidad de formateo, otro dato importante es que el tamaño máximo de un archivo alojado en FAT32 es de 4 Gigabytes, pero podía ser difícil de generar archivos a partir de captura o ediciones de video, los cuales pueden superar aquel límite de 4 Gigabytes.

Por último en la historia, en 2006 se creó exFAT, un sistema de archivos especificado para memorias flash y cuyo uso es generalmente presente en dispositivos móviles con sistema operativo Windows, exFAT incluye varias de las mejoras impuestas por NTFS, como el aumento de capacidad de un archivo hasta 16 Exbibytes, y el aumento de la capacidad del volumen de archivos hasta 64 Zebibytes (Potencial, 1 Zebibyte = $2^{70} = 1.18 * 10^{21}$) y 512 Tebibytes (Teórico), soporte de listas de control de acceso y a partir de la compatibilidad con Windows 7, elimina el tamaño máximo de 4 Gigabytes impuesto por FAT 32

Otro ítem importante que hay que señalar con FAT, es la ausencia de almacenamiento de metadatos, aunque hay sistemas operativos que si tienen sistema de archivos FAT pueden crearse bajo las reglas del propio sistema operativo, los cuales podrían ser destruidos en caso de desfragmentación perdiéndose así los metadatos, FAT solo soporta la última modificación al archivo, y soporta parcialmente las fechas de creación del archivo y la última vez que se accedió (o se leyó) el archivo, FAT carece de archivo de almacenamiento de metadatos (generalmente un archivo invisible presente en alguna carpeta), tampoco almacena quien fue el creador del archivo y tampoco tiene uso de permisos de archivo, los cuales si están presentes en NTFS, HFS+, ZFS y ext.

En cuanto a su estructura, el sistema de archivos FAT se puede dividir en cuatro secciones, la primera es el sector de arranque, siempre se ubica al comienzo de la partición y contiene información básica de la partición, punteros hacia otras secciones de datos y la dirección común de arranque del sistema operativo. La segunda sección es la región FAT, esta sección contiene dos copias de la tabla de asignación de archivos, es un mapa de la partición los cuales indican que clusters están ocupados por archivos. La Tercera sección es la región del directorio raíz, donde es el índice principal de las carpetas y archivos. Y la cuarta y última sección es la región de datos, en esta sección contiene el contenido de los archivos y carpetas, los cuales son ampliables si queda espacio disponible

en los clusters, y son unidos mediante punteros. Una partición se puede descomponer en clusters de idéntico tamaño, dependiendo de la versión de FAT, estos clusters varían entre 2 y 32 Kilobytes, cada archivo utiliza diferentes clusters de acuerdo a su tamaño, lo cual el archivo queda representado en una cadena secuencial de clusters (o una estructura de lista enlazada), esta cadena de clusters no necesariamente quedan en clusters consecutivos, lo cual como consecuencia eso provoca la fragmentación del archivo. El sistema de archivos FAT utilizan una serie de entradas en su estructura de datos, los cuales pueden contener: 1) Información del siguiente cluster de la cadena de archivos, 2) Un indicador de final de archivo, y por lo tanto el final de la cadena, 3) Un carácter especial que indica si el cluster es defectuoso, 4) Otro carácter especial que indica que el cluster a utilizar está reservado (es decir, ocupado para otro archivo), o 5) un 00 que indica que el cluster a utilizar está libre y que por lo tanto puede ser ocupado.

Conclusiones

Cada sistema operativo contiene su propio sistema de archivos, si el sistema operativo lanza al mercado nuevas versiones de su Sistema Operativo, generalmente en algún momento puede lanzar una actualización de su sistema de archivos e incluso lanzar un nuevo sistema de archivos con nuevas características y mejoras en su rendimiento respecto a su antecesor.

Otra característica importante es el uso de metadatos en los sistemas de archivos ya que en aquellos archivos generalmente contienen información de los archivos que están presentes en el sistema además de presentar diferentes propiedades como los permisos de usuarios y el creador del archivo.

El sistema de archivos FAT es uno de los más comunes por su fácil compatibilidad y su fácil estructura de uso, pero carece de características presentes en otros sistemas de archivos como la ausencia de permisos y su fácil fragmentación de los archivos.

Anexo: Compatibilidad entre Sistemas de Archivos y Sistemas Operativos

	<i>Win 9x</i>	<i>Win NT</i>	<i>Linux</i>	<i>Mac OS X</i>	<i>Free BSD</i>	<i>Solaris</i>
FAT16	Si	Si	Si	Si	Si	Si
FAT32	Si (Win 95 OSR2)	Si (Win 2000)	Si	Si	Si	Si
exFAT	Parcial	Si (W7, WXP SP2)	Si (S. Externo)	Si	No	Si
NTFS	SI (S. Externo)	Si	Si (NTFS3G)	Parcial	Si (NTFS3G)	Si (NTFS3G)
HFS+	SI (S. Externo)	SI (S. Externo)	Parcial	Si	Parcial	N/A
Ext3	N/A	SI (S. Externo)	Si	SI (S. Externo)	Si	Si
Ext4	No	SI (S. Externo)	Si (2.6.28)	SI (S. Externo)	No	N/A
ZFS	No	No	SI (S. Externo)	SI (S. Externo)	Si	Si
UDF	Parcial	Si	Si	Si	Si	Si

Enlaces de Interés

http://www.informaticamoderna.com/Sistema_arch.htm

<http://www.alegsa.com.ar/Dic/sistema%20de%20archivos.php>