



Subsistema de almacenamiento:

Controladoras e interfaces de disco y mantenimiento (Tema 6, puntos 4 y 5)

Mantenimiento de Equipos Informáticos
Curso 2008-2009
Universidad de Granada
Rubén Dugo Martín

Introducción

Como se podrá observar a continuación he decidido enfocar el trabajo en el apartado de mantenimiento, concretamente en el subapartado **Tecnología S.M.A.R.T.**, mis motivos son claros; por un lado es un sistema muy elaborado e interesante mediante el cual podemos obtener gran cantidad de información del disco duro así como prever posibles fallos y, por el otro, es una tecnología totalmente ignorada por la mayoría de los usuarios e incluso administradores de sistemas.

Debido a esto, expongo un estudio bastante extenso y detallado con casos prácticos concretos de dicha tecnología.

Controladoras e interfaces de disco

Funcionalmente la misión de una **controladora de disco** es la de más bajo nivel del subsistema de almacenamiento; se encarga del posicionamiento del cabezal y rotación del disco resolviendo las diferentes peticiones que le llegan del sistema.

En contraposición, el **interfaz de disco** simplemente adapta las funcionalidades de la controladora a un bus de expansión del ordenador.

Actualmente controladores e interfaces de disco se pueden encontrar conjuntamente o no y pueden estar situados tanto en la placa base, en el disco externo a ella o incluso en una tarjeta insertada en una ranura de expansión, antiguamente se encontraba de esta forma.

A continuación paso explicar brevemente las controladoras e interfaces de disco más relevantes la historia:

- ♦ ST-506: Diseñada por *Seagate* en 1980, fue la primera controladora para discos XT.
- ♦ ESDI (Enhanced Small Disk Interface): Fue diseñada originariamente por *Maxtor* en los años 80 como continuación a la ya citada ST-506. Podía soportar las velocidades de 10, 15 y 20 Mbits por segundo.
- ♦ IDE (Inteligent Drive Electronics): También llamada ATA (AT Attachment), diseñada por *Western Digital* en 1986. Originariamente soportaba velocidades de 16 Mbit/s entre sus limitaciones encontramos la exclusividad de conectar sólo discos duros, la velocidad de transferencia y la longitud de cable (<45 cms).
- ♦ EIDE (Enhanced IDE): Es una extensión de IDE creada por la misma compañía, con ella se consiguió un aumento considerable de la velocidad, así como la limitación en tamaño del disco (504 MB) y un nuevo concepto de conectar discos duros (esclavo-maestro). Incluye los siguientes estándares:
 - ATAPI (ATA Packet Interface): Hace posible conectar dispositivos de CD-ROM y cinta.
 - ATA-2: Mejora las velocidades hasta 16 MB/s. Incluye:
 - Modos PIO: 0, 1, 2, 3, 4.
 - Modos DMA: 0, 1, 2.
 - FAST ATA
 - FAST ATA-2
- ♦ ATA-3: Revisión de ATA-2 que incluye la tecnología SMART.
- ♦ ATA-4: También Ultra-ATA entre sus diversos nombres comerciales. Desarrollada por *Quantum* y seguido por las demás. Se consigue un aumento bastante alto de la velocidad (33 MB/s) mientras se reduce el consumo de CPU en E/S a disco y se mejora la integración de los datos (mediante CRC). Tiene una total compatibilidad hacia atrás.
- ♦ ATA-5: También ATA/66, soporta transferencias de 66 MB/s. El cable pasa a tener 80 hilos, el doble.
- ♦ ATA-6: También ATA-100 aumenta la velocidad de transferencia a 100 MB/s.
- ♦ ATA-7 (Serial ATA): Por fin se cambian los conectores. La nueva velocidad de transferencia del Serial ATA o SATA es de 1,2 Gbits/s

- ◆ SATA-II: Aumenta la velocidad hasta 3 Gbits/s. La especificación SATA-II para discos externos (eSATA) obtiene una velocidad de transferencia de 1,5 Gbits/s.
- ◆ SCSI (Small Computer System Interface): Es la alternativa a la archiconocida ATA. La primera versión (SCSI-1) fue aprobada en 1986. Permitía conectar hasta 7 dispositivos externos o internos e incluía su propia BIOS. La velocidad de transmisión de datos era 5 Mbps. Sus aplicaciones se situaban en entornos profesionales como servidores o estaciones de trabajo.
- ◆ SCSI-2: Diferentes mejoras con respecto a SCSI-1:
 - Fast SCSI: Dobla la velocidad hasta alcanzar los 10 MB/s.
 - Wide SCSI: Dobla la línea de datos y el número de dispositivos hasta 15.
- ◆ SCSI-3: Se crean los estándares Ultra-SCSI y Ultra Wide SCSI.
- ◆ SAS (Serial Attachmed SCSI): Implementación en serie del bus SCSI, aumenta la velocidad de transmisión, incorpora conexión/desconexión en caliente y utiliza el conector ATA.

Mantenimiento

MTBF

Generalmente los discos duros no necesitan mantenimiento preventivo, sin embargo, existen algunos parámetros relacionados entre sí:

MTBF: Valores típicos entre 500.000 y 1.000.000 horas.

Vida operativa: Entre 3 y 5 años, SCSI más que ATA.

Ciclos de arranque-parada: Situados entre 40.000 y 50.000.

Algunos factores que afectan a la tasa media de fallos (MTBF) son:

Temperatura: Aunque según los investigadores menos de lo que cabe esperar, ya que a temperatura media el disco duro se encuentra lubricado y funcionando en perfecto estado, sin embargo, es a altas temperaturas cuando realmente afecta al funcionamiento del mismo, para evitarlo, simplemente instalar ventiladores que reduzcan estas temperaturas críticas.

Tiempo encendido (*Power-On-Hours*): El tiempo de encendido del disco duro reduce el MTBF sobre todo en las 2.500 primeras horas, más adelante este factor apenas afecta al funcionamiento.

Uso: El "*Duty Cicle*" o tiempo que el disco duro está leyendo o escribiendo, normalmente supuesto del 100%, algo bastante raro, afecta directamente al valor MTBF del mismo.

Este valor depende directamente del propósito del ordenador del que estamos hablando, por ejemplo, no es lo mismo un servidor que, generalmente, realizará más E/S a disco que un ordenador doméstico.

Otros factores a tener en cuenta para el mantenimiento y ergonomía relacionados con el disco duro:

Altitud y humedad: También son factores a tener en cuenta en el funcionamiento de un disco; la altitud traducida en presión atmosférica afecta al funcionamiento, ya que este se basa en ciertos principios aerodinámicos y la humedad (y mucho peor la condensación) puede dañar el funcionamiento eléctrico y mecánico del mismo.

Impactos: Los discos duros están diseñados para soportar impactos, especialmente los destinados a ordenadores portátiles.

Vibraciones: Generalmente un buen atornillado prácticamente las elimina.

Ruido: Actualmente hay una tendencia a reducir el ruido producido por los discos duros, éste aumenta con el aumento de la velocidad de giro.

Síntomas de deterioro

Síntomas deterioro de un disco duro son, por ejemplo:

- Errores de lectura en sectores concretos
- Corrupciones de ficheros (registro de Windows, sector de arranque, etc)
- Ruidos mecánicos:
 - Golpeteos del cabezal
 - Chirridos (es posible que el cabezal esté roto y además rallando la superficie del disco)

```
[ 4666.836843] __ratelimit: 71 messages suppressed
[ 4666.836846] Buffer I/O error on device sda, logical block 0
[ 4668.477837] sd 4:0:0:0: [sda] Result: hostbyte=DID_OK
driverbyte=DRIVER_SENSE,SUGGEST_OK
[ 4668.477848] sd 4:0:0:0: [sda] Sense Key : Hardware Error
[current]
[ 4668.477853] sd 4:0:0:0: [sda] Add. Sense: Data phase error
[ 4668.477860] end_request: I/O error, dev sda, sector 1
```

Registros del kernel de Linux tras encontrarse con un sector defectuoso

Aunque hay casos en los que no todo está perdido; existe una utilidad llamada *HDD Regenerator* (ya comentada en la actividad opcional *Software para mantenimiento*), según los desarrolladores de esta herramienta es posible recuperar ciertos sectores defectuosos mediante un proceso de magnetización, aseguran que hasta el 60%, claro que, en discos duros donde existen gran cantidad de estos sectores es mejor que vayamos pensando en comprar otro.

Tecnología S.M.A.R.T.

Introducción

Son las siglas de *Self Monitoring Analysis and Reporting Technology*, muchos somos los que hemos visto estas siglas, sobre todo en el arranque de nuestros ordenadores, pero ¿de qué se trata? ¿qué técnicas y mecanismos hay detrás de este sistema?

Me parece un tema bastante interesante y curioso por lo que he decidido centrar el trabajo en este punto haciendo más breve los anteriores.

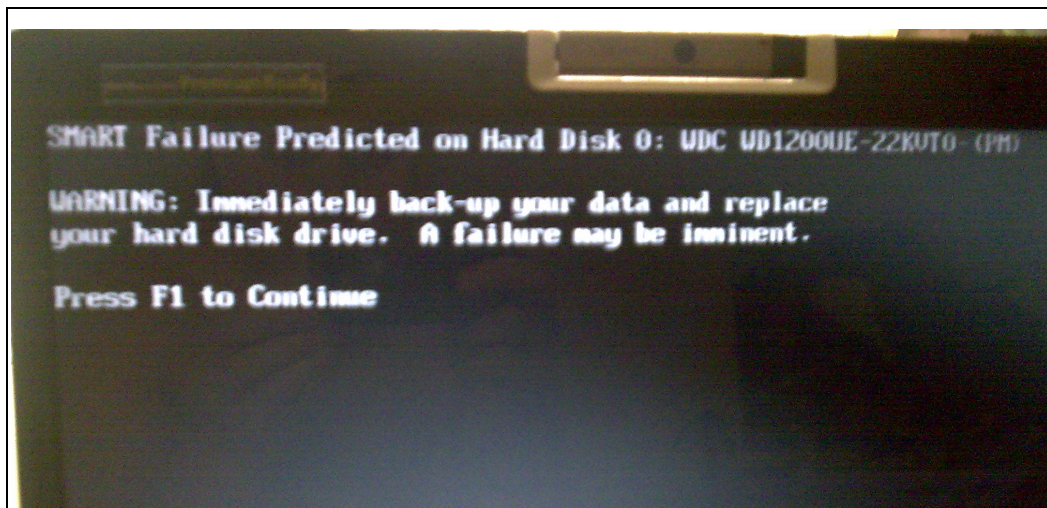
La tecnología S.M.A.R.T. Consiste un conjunto de métodos y técnicas mediante los cuales es posible la detección a corto y medio plazo de los conocidos como **fallos predecibles** de disco duro.

Estos fallos son producidos por la degradación progresiva principalmente mecánica debida del uso y se trata ni más ni menos que del 60% de los fallos de disco.

En contraposición, los **fallos impredecibles** son aquellos provocados por un sobrevoltaje puntual, golpes, mal uso de disco, etc, cuya propia naturaleza los hace imposible pronosticar.

S.M.A.R.T. forma parte del estándar ATA (a partir de la versión 2, como ya mencionamos en el apartado **Controladores e interfaces de disco**) por lo que su integración entre los fabricantes se encuentra muy extendida.

El propósito principal de esta tecnología es avisar al usuario o administrador del sistema de un posible fallo del disco a corto o medio plazo con la finalidad de que tomen medidas preventivas (reemplazar unidad, realizar copia de seguridad, etc).



Esta foto la he tomado en el portátil de un amigo, después de aproximadamente dos semanas apareciendo el error al encender el portátil, éste dejó de funcionar; el pronóstico de SMART fue bastante acertado, el disco duro estaba prácticamente inutilizable.

El funcionamiento básico de S.M.A.R.T. se basa en monitorear y analizar una serie de parámetros o atributos del disco y ver si estos están dentro de un umbral preestablecido, de forma que, si estos parámetros se salen de su rango "normal" la propia BIOS o un programa externo instalado para tal fin lanzará el mensaje de aviso pertinente.

Los atributos más básicos son los correspondientes al *SMART status* y "casi" podríamos decir que están estandarizados ya que se encuentran en la mayoría de los discos.

Sin embargo, la mayoría de la implementación de SMART depende de los fabricantes; cada uno define un conjunto de estos atributos y unos valores de umbral bajo los cuales no debería pasar en estado normal.

Cada atributo tiene un valor bruto (o *raw value*), cuyo significado depende totalmente del fabricante (aunque a menudo corresponde a contadores o a una unidad física, como grados Celsius o segundos), y un valor estandarizado cuyo rango va de 1 a 253 (con 1 representando el peor caso y 253 representando el mejor). Dependiendo del fabricante, un valor de 100 o 200 será elegido como el valor "normal".

Fabricantes que integran SMART en varios de sus productos son: Samsung, Seagate, IBM (Hitachi), Fujitsu, Maxtor, Toshiba, Western Digital y ExcelStor Technology.

Mejor que exponer uno a uno los atributos más frecuentes que nos encontramos en los discos duros, he preferido analizar la salida de la utilidad *smartctl* (incluida en la suite *smartmontools*) en mi equipo y de esta forma explicar los atributos propios de mi disco duro a la vez que doy un caso práctico, entre otras cosas.

Atributos SMART: Análisis de la salida de *smartctl*

Ejecuto la utilidad *smartctl* como *smartctl --all /dev/hda* y guardo la salida en un fichero para un posterior análisis.

La suite *smartmontools* (de la que forma parte *smartctl*) la podemos encontrar en cualquier repositorio de GNU/Linux que se precie.

Ahora paso a analizar los datos más relevantes de la salida de dicha utilidad, he tenido que pasar por alto alguna información ya que hay demasiados datos estadísticos que no nos son interesantes.

El disco del que hablamos es un Hitachi Travelstar 80GN de 60 GB de capacidad.

```
SMART capabilities:                (0x0003)  Saves SMART data before entering
                                     power-saving mode.
                                     Supports SMART auto save timer.
Error logging capability:          (0x01)  Error logging supported.
                                     General Purpose Logging supported.
Short self-test routine
recommended polling time:          (  2) minutes.
Extended self-test routine
recommended polling time:          ( 53) minutes.
```

Extracto de la información SMART general

- SMART capabilities: Este apartado se refiere a las capacidades S.M.A.R.T generales que tiene el disco duro, según su valor (3) tiene la capacidad de guardar los datos antes de entrar en modo ahorro de energía y también posee un temporizador de autoguardado de las mismas
- Error logging capability: El disco soporta registros de errores y además registros de propósito general.
- Short self-test routine recommended polling time: El tiempo recomendado para realizar la rutina corta de autoprueba del disco duro es de 2 minutos.
- Extended self-test routine recommended polling time: El tiempo recomendado para realizar la rutina extendida de autoprueba del disco duro es de 53 minutos.

ID#	ATTRIBUTE_NAME	FLAG	VALUE	WORST	THRESH	TYPE	UPDATED	WHEN_FAILED	RAW_VALUE
1	Raw_Read_Error_Rate	0x000b	100	098	062	Pre-fail	Always	-	1
2	Throughput_Performance	0x0005	100	100	040	Pre-fail	Offline	-	437
3	Spin_Up_Time	0x0007	190	100	033	Pre-fail	Always	-	1
4	Start_Stop_Count	0x0012	096	096	000	Old_age	Always	-	7058
5	Reallocated_Sector_Ct	0x0033	100	100	005	Pre-fail	Always	-	786481
7	Seek_Error_Rate	0x000b	100	100	067	Pre-fail	Always	-	0
8	Seek_Time_Performance	0x0005	100	100	040	Pre-fail	Offline	-	0
9	Power_On_Hours	0x0012	070	070	000	Old_age	Always	-	13342
10	Spin_Retry_Count	0x0013	100	100	060	Pre-fail	Always	-	0
12	Power_Cycle_Count	0x0032	098	098	000	Old_age	Always	-	3253
191	G-Sense_Error_Rate	0x000a	099	085	000	Old_age	Always	-	131072
192	Power-Off_Retract_Count	0x0032	100	100	000	Old_age	Always	-	187
193	Load_Cycle_Count	0x0012	052	052	000	Old_age	Always	-	484323
194	Temperature_Celsius (Lifetime Min/Max 11/56)	0x0002	152	100	000	Old_age	Always	-	36
196	Reallocated_Event_Count	0x0032	100	100	000	Old_age	Always	-	13
197	Current_Pending_Sector	0x0022	100	100	000	Old_age	Always	-	2
198	Offline_Uncorrectable	0x0008	100	100	000	Old_age	Offline	-	0
199	UDMA_CRC_Error_Count	0x000a	200	200	000	Old_age	Always	-	0
<i>Información de los atributos SMART que soporta la unidad</i>									

Las columnas de la izquierda a derecha muestran: n° identificación del atributo; banderas; valor actual; peor valor alcanzado; umbral; tipo, puede ser *Pre-fail* (sirve para pronosticar posibles fallos) o *Old_age* (para cálculos estimados de tiempos de uso del disco duro); actualizado, puede ser *Always* (constantemente) u *Offline* (no se actualiza); última vez que falló y el ya comentado *raw value*.

A continuación una breve descripción y análisis de los atributos anteriores.

NOTA: *Resalto en negrita los atributos que pueden ser indicativos de un posible fallo electromecánico inminente.*

- **Raw Read Error Rate**: Indica la tasa de errores de lectura hardware. Un valor diferente de cero indica problemas en el cabezal de lectura/escritura o en la superficie del disco. Como podemos observar el *raw value* en mi disco es 1, muy próximo a cero por lo que tenemos que alarmarnos.
- **Throughput Performance**: Comportamiento con carga, en mi equipo no se actualiza por lo que no lo comento.
- **Spin Up Time**: Tiempo medio en milisegundos en alcanzar una velocidad de giro adecuada (desde reposo hasta estar totalmente operativo), como vemos, en mi disco duro es de un 1 ms, muy poco por lo que puede haber un error en la medida.
- **Start Stop Count**: Contador de ciclos de inicio-parada, un valor normal para mi disco duro.
- **Reallocated Sector Ct**: Contador de sectores reasignados, aumenta cuando se encuentran sectores dañados. Estoy en el umbral "bueno".
- **Seek Error Rate**: Tasa de errores de búsqueda, el valor *raw* 0 indica que no existen problemas.
- **Seek Time Performance**: Media de errores de las operaciones de búsqueda de los cabezales magnéticos. *Offline* así que no lo comento.

- Power_On_Hours: Horas encendido, valor *raw* de 13342 horas (algo más de un año y medio, altísimo), aunque son muchas horas no representan un riesgo.
- Spin_Retry_Count: Contador de los intentos de iniciar el giro, de nuevo un valor *raw* de 0 califica a mi disco como "normal".
- Power_Cycle_Count: Contador de ciclos completos de encendido/apagado, otra vez mi disco duro es calificado como "normal".
- G-Sense_Error_Rate: Tasa de errores debidos a impactos, valor dentro de lo normal.
- Power-Off_Retract_Count: Número de veces que han salido del disco, valor dentro del umbral "normal".
- Load_Cycle_Count: Contador de ciclos de carga, tenemos un valor *raw* alto para este atributo, pese a encontrarse dentro de lo normal.
- Temperature_Celsius: Temperatura en grados Celsius, debe estar en el rango de 11° a 56° nada fuera de lo normal.
- Reallocated_Event_Count: Contador de las operaciones de reasignación, fallidas y exitosas son contadas sin distinción. Valor bastante normal.
- Current_Pending_Sector: Número actual de sectores "pendientes" de ser reasignados tras un fallo. Valor dentro de lo normal.
- Offline_Uncorrectable: Errores en sectores irrecuperables. No se actualiza en mi equipo.
- UDMA_CRC_Error_Count: Número de errores en la transferencia de datos. Valor normal pasa este caso.

Aparte de estos valores podemos encontrar información relacionada con estimaciones de horas de uso, posibles fallos en el futuro, registro de fallos, etc, información realmente completa.

Podemos concluir que a pesar de la cantidad de horas de uso de el disco duro de mi portátil personal este goza de *muy buena salud electromecánica*, ciertamente, nunca he tenido problemas con el disco duro aunque no creía que fuera a obtener tan buenos resultados.

Otro caso práctico:
¿Una configuración incorrecta de la administración de energía puede *envejecer* mi disco duro?

La respuesta es un rotundo si, ya que al configurar un ahorro energético muy alto el sistema intentará ahorrar energía en todos sus dispositivos, entre ellos el disco duro. Esto es aplicable principalmente a portátiles.

Esto hace que el disco duro sitúe tras un periodo de tiempo de inactividad (no hay lectura ni escritura) su cabezal en una posición de reposo fuera del disco magnético. En esta posición el cabezal no sólo ahorra energía si no que esta a salvo de golpes e impactos que lo dañarían tanto a él como a la superficie del disco.

Claro que a valores de ahorro energético demasiados altos el cabezal tras apenas unos segundos de inactividad regresa a su posición de reposo y, curiosamente, este es el valor por defecto que tienen los discos duros de portátiles.

Esta acción precisamente es la que aumenta el valor *raw* del atributo *Load_Cycle_Count*, haciendo cada vez más viejo el disco duro y no sólo eso, también ralentiza las entradas y salidas a disco.

Aún mayor es el problema cuando el sistema operativo que usas es GNU/Linux que realiza escrituras constantes en diversos registros, como podemos observar, comparado con los demás valores de tipo *Old_age*, mi disco duro tiene un alto valor *raw* para este atributo.

Externamente podemos observar cuando el disco duro realiza esta acción escuchando un leve *tic*.

Para solucionar este problema en mi portátil, utilizaré la utilidad *hdparm* que viene en cualquier sistema GNU/Linux, para el parámetro *-B* tenemos la descripción:

-B Set Advanced Power Management feature, if the drive supports it. A low value means aggressive power management and a high value means better performance. Possible settings range from values 1 through 127 (which permit spin-down), and values 128 through 254 (which do not permit spin-down). The highest degree of power management is attained with a setting of 1, and the highest I/O performance with a setting of 254. A value of 255 tells hdparm to disable Advanced Power Management altogether on the drive (not all drives support disabling it, but most do).

De forma que elijo ejecutar la utilidad como *hdparm -B 254 /dev/hda*, cuya salida es satisfactoria.

Escojo el valor 254, el cual no llega a desactivarme del todo la administración de energía, de forma que si dejo el equipo encendido por un largo periodo de tiempo sin estar trabajando con él quiero que me sitúe el cabezal en la posición de reposo.

Referencias

1. Apuntes de la asignatura
2. http://en.wikipedia.org/wiki/Enhanced_Small_Disk_Interface
3. <http://www.virtualforos.com/archive/index.php/t-115947.html>
4. <http://es.wikipedia.org/wiki/ATA>
5. <http://es.wikipedia.org/wiki/EIDE>
6. <http://es.wikipedia.org/wiki/SCSI>
- 7.
8. <http://pinux.info/fitxers/discosduros.html>
9. <http://jorge.huerga.org/2008/07/prevenir-errores-de-disco-duro-gracias-a-smart/>
10. <http://es.wikipedia.org/wiki/S.M.A.R.T.>
11. http://www.recuperaciondedatos.com.es/archivos/vp-tid:3-pid:6-sintomas_comunes_que_muestra_una_unidad_corrupta_o_defectuosa.html
12. <http://www.kabytes.com/tutoriales/guia-como-utilizar-el-hdd-regenerator/>
13. <http://en.wikipedia.org/wiki/S.M.A.R.T.>
14. <http://es.wikipedia.org/wiki/S.M.A.R.T.>
15. <http://smartmontools.sourceforge.net/>
16. <http://www.ariolic.com/activesmart/smart-attributes/g-sense-error-rate.html>
17. <http://www.linuxquestions.org/questions/ubuntu-63/hdparm-apm-hard-drive-clicking-sound-588770/>
18. Y mi experiencia personal y/o profesional