Contents: [Dobrica PavlinuÅ¡iÄ 's random unstructured stuff]

Have many disks. More disk spindles brings more than capacity alone! (Same as in databases)

Speed

Disk platter transfer speed

If you think that disk has constant transfer speed, ZCAV has interesting graphs

Individial disks

Slow laptop 2.5" 5400 disk

```
dpavlin@llin:~$ sudo hdparm -i /dev/sda

/dev/sda:

 Model=FUJITSU MHV2080BH                    , FwRev=00840028, SerialNo=      NW05T6B29HM5
 Config={ HardSect NotMFM HdSw>15uSec Fixed DTR>10Mbs }
 RawCHS=16383/16/63, TrkSize=0, SectSize=0, ECCbytes=4
 BuffType=DualPortCache, BuffSize=8192kB, MaxMultSect=16, MultSect=?16?
 CurCHS=16383/16/63, CurSects=16514064, LBA=yes, LBAsects=156301488
 IORDY=on/off, tPIO={min:240,w/IORDY:120}, tDMA={min:120,rec:120}
 PIO modes:  pio0 pio1 pio2 pio3 pio4
 DMA modes:  mdma0 mdma1 mdma2
 UDMA modes: udma0 udma1 udma2 udma3 udma4 *udma5
 AdvancedPM=yes: mode=0x80 (128) WriteCache=enabled
 Drive conforms to: unknown:  ATA/ATAPI-3,4,5,6,7

 * signifies the current active mode

dpavlin@llin:~$ sudo hdparm -tT /dev/sda
/dev/sda:
 Timing cached reads:   1566 MB in  2.00 seconds = 782.85 MB/sec
 Timing buffered disk reads:   66 MB in  3.03 seconds =  21.79 MB/sec
```

Interesting numbers are `BuffSize` (cache in disk) and `MaxMultSect` which we want to use for read-ahead param:

```
hdparm -m 16 -a 16 /dev/sda
```

This will **decrease** a bit speed of linerar buffer reads which `hdparm` uses, but we will pull from disk only blocks which are allready in cache, improving random read/write performance.

To find optimal readahead for your drive using hdparm access pattern you can use hdparm-readahead.pl which will try different combinations for you.

Faster (!) external 3.5 USB disk (no `hdparm -i` on USB), but just because it's **another** disk not loaded by system.

```
dpavlin@llin:~$ sudo hdparm -tT /dev/sdb

/dev/sdb:
 Timing cached reads:   1508 MB in  2.00 seconds = 753.72 MB/sec
 Timing buffered disk reads:   56 MB in  3.03 seconds =  18.48 MB/sec
```

Software RAID

Home-made RAID disk holder
Home-made software md RAID 5 array from SATA drives:

Content from RAID5 disk array

- dscf0002

  Dobrica Pavlinusic posted a photo:

- dscf0001

  Dobrica Pavlinusic posted a photo:

Final position in case, notice empty space above 4th disk which was occupied by disks before (which didn't have enough airflow because of that)
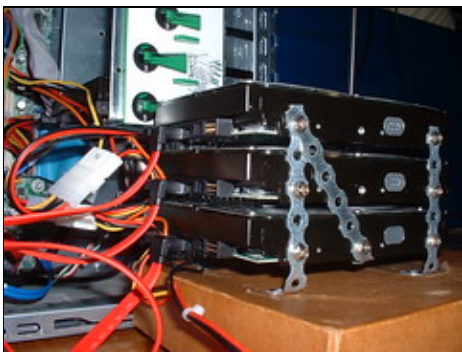
- dscf0012

Dobrica Pavlinusic posted a photo:

- dscf0007

Dobrica Pavlinusic posted a photo:

- dscf0006

Dobrica Pavlinusic posted a photo:

- dscf0005

Dobrica Pavlinusic posted a photo:

Note nice usage of construction metal stripes with holes which is usually used to hold fence. It has holes just the right size for screws to go through and hold disks nicely spaced (although a little bit more space would be ideal). It's soft enough to be bent at corners to produce nice and leveled space between it and case.

Blog post RAID5 for home describes setup in some details.

Drive info:

```
dpavlin@brr:~$ sudo hdparm -i /dev/sdd

/dev/sdd:

 Model=WDC WD5000AAKS-00YGA0                   , FwRev=12.01C02, SerialNo=    WD-WCAS80929678
 Config={ HardSect NotMFM HdSw>15uSec SpinMotCtl Fixed DTR>5Mbs FmtGapReq }
 RawCHS=16383/16/63, TrkSize=0, SectSize=0, ECCbytes=50
 BuffType=unknown, BuffSize=16384kB, MaxMultSect=16, MultSect=?16?
 CurCHS=16383/16/63, CurSects=16514064, LBA=yes, LBAsects=976773168
 IORDY=on/off, tPIO={min:120,w/IORDY:120}, tDMA={min:120,rec:120}
 PIO modes:  pio0 pio3 pio4
 DMA modes:  mdma0 mdma1 mdma2
 UDMA modes: udma0 udma1 udma2 udma3 udma4 udma5 *udma6
 AdvancedPM=no WriteCache=enabled
 Drive conforms to: Unspecified:  ATA/ATAPI-1,2,3,4,5,6,7

 * signifies the current active mode
```

Speed of individual drives in array:

```
dpavlin@brr:~$ sudo hdparm -tT /dev/sda /dev/sdb /dev/sdd

/dev/sda:
 Timing cached reads:    1982 MB in  2.00 seconds = 991.18 MB/sec
 Timing buffered disk reads:  232 MB in  3.03 seconds =  76.67 MB/sec

/dev/sdb:
 Timing cached reads:    2010 MB in  2.00 seconds = 1004.95 MB/sec
 Timing buffered disk reads:  228 MB in  3.01 seconds =  75.85 MB/sec

/dev/sdd:
 Timing cached reads:    2006 MB in  2.00 seconds = 1003.01 MB/sec
 Timing buffered disk reads:  230 MB in  3.01 seconds =  76.47 MB/sec
```

How are hey assembled into /dev/md0 RAID 5 array:

```
dpavlin@brr:~$ cat /proc/mdstat
Personalities : [raid6] [raid5] [raid4]
md0 : active raid5 sdd1[0] sda1[2] sdb1[1]
      976767872 blocks level 5, 64k chunk, algorithm 2 [3/3] [UUU]
```

Speed of array

```
dpavlin@brr:~$ sudo hdparm -tT /dev/md0

/dev/md0:
 Timing cached reads:    1986 MB in  2.00 seconds = 993.20 MB/sec
 Timing buffered disk reads:  434 MB in  3.01 seconds = 144.41 MB/sec
```

As expected RAID 5 speed is 75 + 75 + 0 (parity disk) ~ 144 MB/sec

Temperature

Disks don't like it hot!

```
root@brr:~# hddtemp /dev/sda /dev/sdb /dev/sdd
/dev/sda: WDC WD5000AAKS-00YGA0: 33Â°C
/dev/sdb: WDC WD5000AAKS-00YGA0: 32Â°C
/dev/sdd: WDC WD5000AAKS-00YGA0: 32Â°C
```

On output above, middle disk is /dev/sda so it's 1Â° hotter than other two. I could mitigate this with additional fan on front of case, but it's making enough noise already, so I'll leave it as is.

Data security

Smart

```
root@brr:~# smartctl --all /dev/sda | head -20
smartctl version 5.38 [i686-pc-linux-gnu] Copyright (C) 2002-8 Bruce Allen
Home page is http://smartmontools.sourceforge.net/

=== START OF INFORMATION SECTION ===
Model Family:     Western Digital Caviar Second Generation Serial ATA family
Device Model:     WDC WD5000AAKS-00YGA0
Serial Number:    WD-WCAS80815866
Firmware Version: 12.01C02
User Capacity:    500,107,862,016 bytes
Device is:        In smartctl database [for details use: -P show]
ATA Version is:   8
ATA Standard is:  Exact ATA specification draft version not indicated
Local Time is:    Sat Oct 11 00:27:01 2008 CEST
SMART support is: Available - device has SMART capability.
SMART support is: Enabled

=== START OF READ SMART DATA SECTION ===
SMART overall-health self-assessment test result: PASSED
```

Before you start to beleve in SMART as solution to all disk health problems, read Failure Trends in a Large Disk Drive Population

See also Bad block HOWTO for smartmontools if you ever get smart errors and don't just want to throw out your disk.

RAID

Also interesting is Some RAID Issues

Read also Why RAID 5 stops working in 2009