

Создание отказоустойчивого сетевого хранилища ГИС данных

[Обсудить в форуме](#) Комментариев — 47

Эта страница опубликована в основном списке статей сайта по адресу <http://gis-lab.info/qa/gis-dataserver.html>

Описание аппаратного и программного обеспечения для сервера геоданных

Содержание

- [1 Введение](#)
- [2 Железо](#)
- [3 Софт](#)
- [4 Установка](#)
- [5 Тестирование](#)

Введение

Встала задача создания хранилища ГИС данных большого объема (десятки Тб) "дешевое". Посмотрел имеющиеся на рынке NAS - цены порядка 70 тыс. р. за 8 ячеечную модель и это без дисков. Оказалось что дешевле собрать компьютер на базе Linux с программным RAID5.

Железо

Решил брать все в одном месте. Вот какой набор получился.

№	Наименование	Цена (на март 2012 г.)	Количество	Сумма
1.	Жесткий диск 3000ГБ Hitachi "Deskstar 5K3000 HDS5C3030ALA630" 5700об./мин., 32МБ (SATA III)	5544.82	3	16634.46
2.	Мат. плата Socket1155 ASUS "P8Z68-V LX" (iZ68, 4xDDR3, SATA III, SATA II, RAID, 2xPCI-E, D-Sub, DVI, HDMI, SB, 1Гбит LAN, USB2.0, USB3.0, ATX)	3115.2	1	3115.2
3.	Корпус Miditower Zalman "MS1000-HS1", ATX, черный (без БП)	4012	1	4012
4.	Кулер для процессора Socket775/115x Arctic Cooling "Alpine 11 Plus"	328.04	1	328.04
5.	Привод BD-ROM/DVD±RW 12xBD/16x8x16xDVD/48x32x40xCD Samsung "SH-B123L/BSBP", черный (SATA)	2185.36	1	2185.36
6.	Модуль памяти 2x2ГБ DDR3 SDRAM Kingston "ValueRAM" KVR1333D3S8N9K2/4G (PC10600, 1333МГц, CL9)	767	1	767
7.	Процессор Intel "Core i3-2100" (3.10ГГц, 2x256КБ+3МБ, EM64T, GPU) Socket1155	3314.62	1	3314.62
8.	Блок питания 750Вт Corsair "TX750 V2" CMPSU-750TXV2EU ATX12V V2.3 (20/24+4/8+6/8pin, вентилятор d140мм) + кабель питания EURO (1.8м)	3890.46	1	3890.46
9.	Жесткий диск 250ГБ Seagate "Barracuda ST250DM000" 7200об./мин., 16МБ (SATA III)	2153.5	1	2153.5

Итого: 36400.64 р.

Выбор комплектующих особо не принципиален и осуществлялся на основе имеющихся в одном магазине. Для

сетевого хранилища не нужен производительный процессор, но в перспективе предполагается развертывание на нем дополнительных сервисов (Веб сервер, база данных и т.д.), поэтому был выбран Core i3-2100.

Основа массива - 3 диска Hitachi Deskstar 5K3000 (для RAID 5 нужно минимум 3 диска) емкостью 3Тб. Это "зеленые" диски и не совсем подходят для производительных решений, но выбора не было, да и получилось "задешево". Кроме того, не стояла задача по скорости - главное **объем и надежность**. Предполагается получить объем хранения порядка 7-8Тб.

Для установки системы предусмотрим отдельный маленький жесткий диск 250Гб Seagate Barracuda.

Материнская плата имеет видеовыходы, которые работают в случае поддержки процессором (выбранный поддерживает). На материнской плате 4 разъема SATA3 и 2 SATA6. Я подключил их следующим образом: 1 SATA3 к приводу BD-ROM/DVD±RW, 3 к корзине из 3-х дисков с возможностью hot-swap (это особенность корпуса - в нем установлена корзина на 3 диска), 1 SATA6 к нашему системному диску и 1 SATA6 к внешнему eSATA.

В перспективе наш NAS может быть модернизирован добавлением второй сетевой карты для объединения их в единый пул. Для добавления дисков необходима отдельная плата с необходимым количеством портов SATA (PCI или PCI-1x,4x,16x), т.к. на данный момент портов SATA не осталось.

Выглядеть это будет так.



Дверца снизу закрывает корзину на 3 диска.



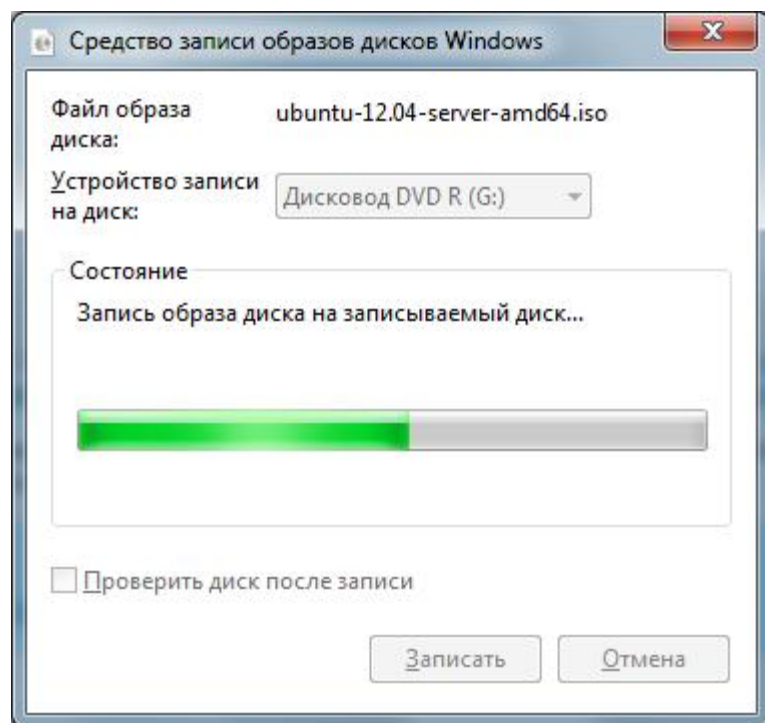
Софт

В качестве операционной системы для нашего сервера будем использовать Ubuntu Server 12.04 LTS. Скачиваем со [страницы загрузки](#) Ubuntu Server 12.04 LTS 64 bit. Будем ставить 64-битную операционную систему.

Полный адрес загрузки следующий: <http://mirror.bytemark.co.uk/ubuntu-releases//precise/ubuntu-12.04-server-amd64.iso>

Альтернативные образы: <http://mirror.bytemark.co.uk/ubuntu-releases//precise>

Скаченный образ записываем на диск.



Установка

Запускаем наш NAS и грузимся с только что записанного диска. Выполняем установку на отдельный маленький диск (250ГБ Seagate).



Выбор языка



Начальный экран - выбираем установка



Выбираем местоположение



•

Выбираем настройку клавиатуры



•

Задаем имя компьютера (я задал ubuntu-nas)



•

Папку не шифрую



•

Автоматически разбиваем диск



•

Для установки выбираю маленький (250ГБ Seagate) диск



Соглашаемся с автоматической разметкой



Ждем окончания процесса установки...



Ставим только OpenSSH - остальное будем делать через терминал

По окончании процесса установки перезагружаемся. Если все OK, то дальнейшую работу с NAS будем проводить удаленно. Для этого необходим DHCP сервер, который автоматически выдаст NAS адрес (можно и статически IP адрес назначить). Отключаем от NAS мышь, клавиатуру, монитор. Устанавливаем его в новом месте и обеспечиваем подключение сети питания и информационной сети (локалки). Дальнейшую работу будем проводить через [PuTTY](#).



Указываем IP адрес



Задаем кодировку UTF8

```
login as: bishop
bishop@192.168.2.86's password:
Welcome to Ubuntu 12.04 LTS (GNU/Linux 3.2.0-23-generic x86_64)
```

* Documentation: <https://help.ubuntu.com/>

System information as of Sun Apr 29 00:00:42 MSK 2012

System load:	0.0	Processes:	83
Usage of /:	1.9% of 228.70GB	Users logged in:	0
Memory usage:	0%	IP address for eth0:	192.168.2.86
Swap usage:	0%		

Graph this data and manage this system at <https://landscape.canonical.com/>

Last login: Sat Apr 28 22:40:57 2012 from bishop-hp.bishop-dom
bishop@ubuntu-nas:~\$

Устанавливаем mdadm (Mirrored Device Admin).

```
sudo apt-get install mdadm
```

Смотрим наши диски.

```
bishop@ubuntu-nas:~$ ls -l /dev/sd*
brw-rw---- 1 root disk 8,  0 app.  28 23:04 /dev/sda
brw-rw---- 1 root disk 8,  1 app.  28 23:04 /dev/sda1
brw-rw---- 1 root disk 8,  2 app.  28 23:04 /dev/sda2
brw-rw---- 1 root disk 8,  5 app.  28 23:04 /dev/sda5
brw-rw---- 1 root disk 8, 16 app.  28 23:04 /dev/sdb
brw-rw---- 1 root disk 8, 32 app.  28 23:04 /dev/sdc
brw-rw---- 1 root disk 8, 48 app.  28 23:04 /dev/sdd
```

Как видим диск /dev/sda разбит на /dev/sda1, sda2 и sda5 - это наш системный диск, а вот sdb, sdc, sdd нужно добавить в массив.

Выполняем команду:

```
sudo mdadm --create --verbose /dev/md0 --level=5 --raid-devices=3 /dev/sdb /dev/sdc
/dev/sdd
```

или более короткую запись для дисков с именами подряд

```
sudo mdadm --create --verbose /dev/md0 --level=5 --raid-devices=3 /dev/sdb-d
```

Получаем следующий вывод:

```
mdadm: layout defaults to left-symmetric
mdadm: layout defaults to left-symmetric
mdadm: chunk size defaults to 512K
mdadm: layout defaults to left-symmetric
mdadm: layout defaults to left-symmetric
mdadm: size set to 2930265088K
mdadm: Defaulting to version 1.2 metadata
mdadm: array /dev/md0 started.
```

Выполним команду:

```
bishop@ubuntu-nas:~$ watch cat /proc/mdstat
Every 2,0s: cat /proc/mdstat
Sun Apr 29 01:11:38 2012
```

```
Personalities : raid6 raid5 raid4
```

```
md0 : active raid5 sdd3 sdc1 sdb0
      5860530176 blocks super 1.2 level 5, 512k chunk, algorithm 2 3/2 UU_
      >..... recovery = 0.9% (26581700/2930265088) finish=656.1min
speed=73755K/sec

unused devices: <none>
```

Из которой узнаем, что массив создан и инициализируется. Как видно понадобится на это 10 часов. Пока идет процесс, посмотрим детали массива:

```
bishop@ubuntu-nas:~$ sudo mdadm --detail /dev/md0
/dev/md0:
    Version : 1.2
  Creation Time : Sun Apr 29 01:05:41 2012
    Raid Level : raid5
    Array Size : 5860530176 (5589.04 GiB 6001.18 GB)
  Used Dev Size : 2930265088 (2794.52 GiB 3000.59 GB)
    Raid Devices : 3
  Total Devices : 3
 Persistence : Superblock is persistent

    Update Time : Sun Apr 29 01:09:39 2012
      State : clean, degraded, recovering
  Active Devices : 2
Working Devices : 3
 Failed Devices : 0
   Spare Devices : 1


    Layout : left-symmetric
  Chunk Size : 512K

Rebuild Status : 2% complete

    Name : ubuntu-nas:0 (local to host ubuntu-nas)
   UUID : 5d16e4f0:04b6622a:d4649121:25fba69e
  Events : 5

   Number   Major   Minor   RaidDevice State
     0         8       16         0   active sync   /dev/sdb
     1         8       32         1   active sync   /dev/sdc
     3         8       48         2   spare rebuilding /dev/sdd
```

Запишем на всякий случай серийники дисков в RAID5.

```
bishop@ubuntu-nas:~$ sudo hdparm -I /dev/sdb-d | grep "Serial Number"
/dev/sdb      Serial Number:      MJ1311YNG6J5NA
/dev/sdc      Serial Number:      MJ1311YNG6LWZA
/dev/sdd      Serial Number:      MJ1311YNG6JMYA
```

Теперь глянем на загрузку процессора при этой операции.

```
top - 01:36:07 up 2:31, 1 user, load average: 1.00, 1.02, 0.89
Tasks: 88 total, 2 running, 86 sleeping, 0 stopped, 0 zombie
Cpu(s): 0.1%us, 1.3%sy, 0.0%ni, 98.2%id, 0.0%wa, 0.0%hi, 0.4%si, 0.0%st
Mem: 3959176k total, 343368k used, 3615808k free, 5132k buffers
Swap: 4102140k total, 0k used, 4102140k free, 210964k cached

  PID USER      PR  NI  VIRT  RES  SHR  S  %CPU  %MEM    TIME+  COMMAND
 8078 root        20   0    0    0    0  S   15   0.0   4:33.40 md0_raid5
 8082 root        20   0    0    0    0  D    8   0.0   2:17.74 md0_resync
 1008 root        20   0    0    0    0  S    0   0.0   0:02.05 kworker/0:1
```

Ну вот создание массива завершено. Итого имеем 6ТБ.

```
bishop@ubuntu-nas:~$ sudo mdadm --detail /dev/md0
sudo password for bishop:
/dev/md0:
    Version : 1.2
  Creation Time : Sun Apr 29 01:05:41 2012
    Raid Level : raid5
    Array Size : 5860530176 (5589.04 GiB 6001.18 GB)
  Used Dev Size : 2930265088 (2794.52 GiB 3000.59 GB)
    Raid Devices : 3
  Total Devices : 3
  Persistence : Superblock is persistent

    Update Time : Sun Apr 29 13:59:15 2012
      State : clean
Active Devices : 3
Working Devices : 3
Failed Devices : 0
Spare Devices : 0

    Layout : left-symmetric
  Chunk Size : 512K

        Name : ubuntu-nas:0 (local to host ubuntu-nas)
        UUID : 5d16e4f0:04b6622a:d4649121:25fba69e
        Events : 28
```

Number	Major	Minor	RaidDevice	State	
0	8	16	0	active sync	/dev/sdb
1	8	32	1	active sync	/dev/sdc
3	8	48	2	active sync	/dev/sdd

В качестве файловой системы будем использовать xfs. Я особо не выбирал исходя из текущей ситуации - взял проверенное ранее решение. Кроме того, я планирую расширять xfs на добавляемые в RAID диски. Эта файловая система позволяет расширять себя на добавленную емкость. Устанавливаем поддержку xfs.

```
sudo apt-get install xfsprogs
```

Создаем файловую систему.

```
bishop@ubuntu-nas:~$ sudo mkfs.xfs /dev/md0
log stripe unit (524288 bytes) is too large (maximum is 256KiB)
log stripe unit adjusted to 32KiB
meta-data=/dev/md0            isize=256    agcount=32, agsize=45785472 blks
                               sectsz=512    attr=2, projid32bit=0
data      =                   bsize=4096    blocks=1465132544, imaxpct=5
                               =             sunit=128   swidth=256 blks
naming    =version 2          bsize=4096    ascii-ci=0
log       =internal log      bsize=4096    blocks=521728, version=2
                               =             sectsz=512   sunit=8 blks, lazy-count=1
realtime  =none              extsz=4096    blocks=0, rtextents=0
```

Это происходит очень быстро. Монтируем созданную файловую систему.

```
sudo mkdir /mnt/fs0
sudo mount -t xfs /dev/md0 /mnt/fs0
```


Посмотрим, что получилось в результате:

```
bishop@ubuntu-nas:~$ df -H
Файл.система    Размер  Использовано  Дост  Использовано%  Смонтировано в
/dev/sda1        246G      4,8G      229G              3% /
udev             2,1G      4,1k      2,1G              1% /dev
tmpfs            811M      398k      811M              1% /run
none             5,3M        0      5,3M              0% /run/lock
none             2,1G        0      2,1G              0% /run/shm
/dev/md0         6,0T      35M      6,0T              1% /mnt/fs0
```

Проверим скорость записи:

```
root@ubuntu-nas:/etc# time dd if=/dev/zero of=/mnt/fs0/bigfile bs=4096k count=3072
3072+0 записей считано
3072+0 записей написано
скопировано 12884901888 байт (13 GB), 57,9814 с, 222 MB/c

real    0m58.028s
user    0m0.012s
sys     0m8.969s
```

Теперь чтения:

```
bishop@ubuntu-nas:~$ sudo time dd if=/mnt/fs0/bigfile of=/dev/null bs=4096k
3072+0 записей считано
3072+0 записей написано
скопировано 12884901888 байт (13 GB), 52,0242 с, 248 MB/c
0.00user 4.84system 0:52.11elapsed 9%CPU (0avgtext+0avgdata 20464maxresident)k
25166648inputs+0outputs (5major+1318minor)pagefaults 0swaps
```

С учетом теоретической скорости SATA2 в 300 MB/c - вполне неплохо. Пропишем автоматическое монтирование RAID. Выясним UUID RAID:

```
bishop@ubuntu-nas:~$ sudo blkid
/dev/sda1: UUID="6621a811-e7d5-495f-8995-c6aa2d5c0e5d" TYPE="ext4"
/dev/sda5: UUID="01a1bb84-8310-4044-85ea-d60eefba939d" TYPE="swap"
/dev/sdb: UUID="5d16e4f0-04b6-622a-d464-912125fba69e" UUID_SUB="24c097e7-ed25-b654-b0b8-5fdb8dbfd904" LABEL="ubuntu-nas:0" TYPE="linux_raid_member"
/dev/sdc: UUID="5d16e4f0-04b6-622a-d464-912125fba69e" UUID_SUB="c3cce69a-0865-777d-043c-0bfd9908e292" LABEL="ubuntu-nas:0" TYPE="linux_raid_member"
/dev/md127: UUID="a73aa5f6-c81b-4a3e-90a9-7c9bc5c1a43c" TYPE="xfs"
/dev/sdd: UUID="5d16e4f0-04b6-622a-d464-912125fba69e" UUID_SUB="bb6b6332-7fd0-c293-ebdf-c95cb1046529" LABEL="ubuntu-nas:0" TYPE="linux_raid_member"
```

И впишем в /etc/fstab в самый конец следующие строки:

```
# /dev/md0 RAID
UUID=a73aa5f6-c81b-4a3e-90a9-7c9bc5c1a43c          /mnt/fs0          xfs          defaults
0              0
```

Установим Samba для доступа к массиву по сети:

```
sudo apt-get install samba
```

Для доступа не требуется ввод в домен Windows или еще какие-то "особые" условия. Поэтому даем всем доступ. Отредактируем файл /etc/samba/smb.conf следующим образом:

```
#===== Global Settings =====

global
    workgroup = WORKGROUP
    server string = %h server (Samba, Ubuntu)
    dns proxy = no
    log file = /var/log/samba/log.%m
    max log size = 1000
    syslog = 0
    panic action = /usr/share/samba/panic-action %d

    socket options = TCP_NODELAY SO_KEEPAIVE SO_RCVBUF=8192 SO_SNDBUF=8192
IPTOW_LOWDELAY
    getwd cache = yes

# Allow users who've been granted usershare privileges to create
# public shares, not just authenticated ones
    usershare allow guests = yes

#===== NAS adds =====
preserve case = yes
guest ok = yes
guest only = yes
encrypt passwords = true
browseable = yes
os level = 33
security = SHARE
writeable = yes
short preserve case = yes
disable spoolss = Yes
load printers = No
show add printer wizard = No

#===== Share Definitions =====

# A sample share for sharing your CD-ROM with others.
cdrom
    comment = Samba server's CD-ROM
    read only = yes
    locking = no
    path = /media/cdrom
    guest ok = yes
    preexec = /bin/mount /cdrom
    postexec = /bin/umount /cdrom
share
    path = /mnt/fs0
    writable = yes
    force user = root
    guest ok = yes
```

Проверим правильность сделанных правок:

```
root@ubuntu-nas:/etc/samba# testparm
Load smb config files from /etc/samba/smb.conf
rlimit_max: increasing rlimit_max (1024) to minimum Windows limit (16384)
Processing section "cdrom"
Processing section "share"
Loaded services file OK.
Server role: ROLE_STANDALONE
Press enter to see a dump of your service definitions
```

```
global
server string = %h server (Samba, Ubuntu)
syslog = 0
log file = /var/log/samba/log.%m
max log size = 1000
load printers = No
disable spoolss = Yes
show add printer wizard = No
os level = 32
dns proxy = No
usershare allow guests = Yes
panic action = /usr/share/samba/panic-action %d
idmap config * : backend = tdb
guest only = Yes
guest ok = Yes

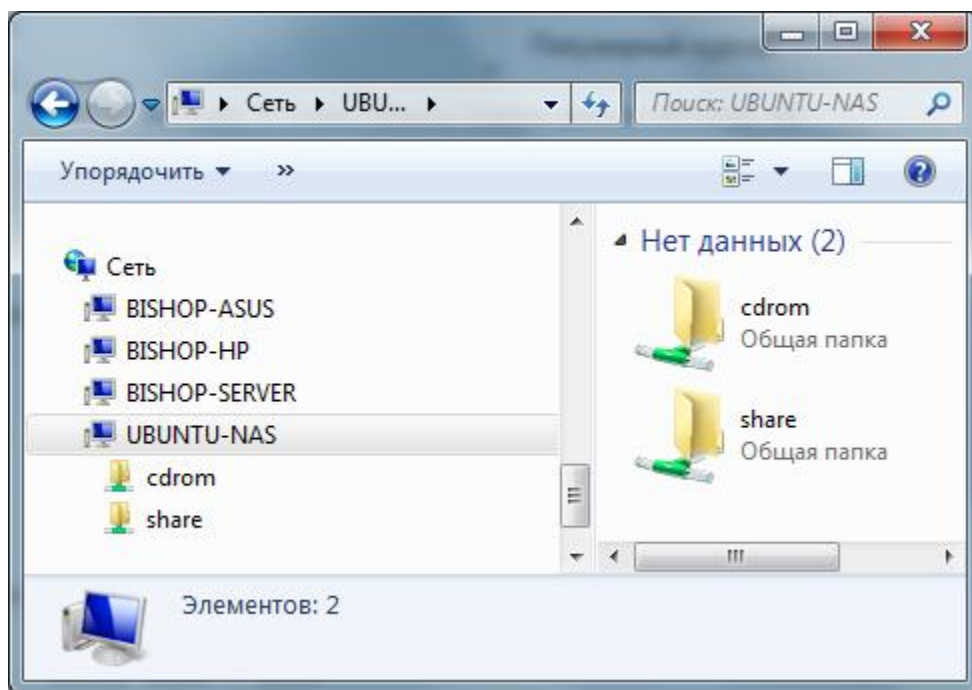
cdrom
comment = Samba server's CD-ROM
path = /media/cdrom
locking = No
preexec = /bin/mount /cdrom
postexec = /bin/umount /cdrom

share
path = /mnt/fs0
force user = root
read only = No
```

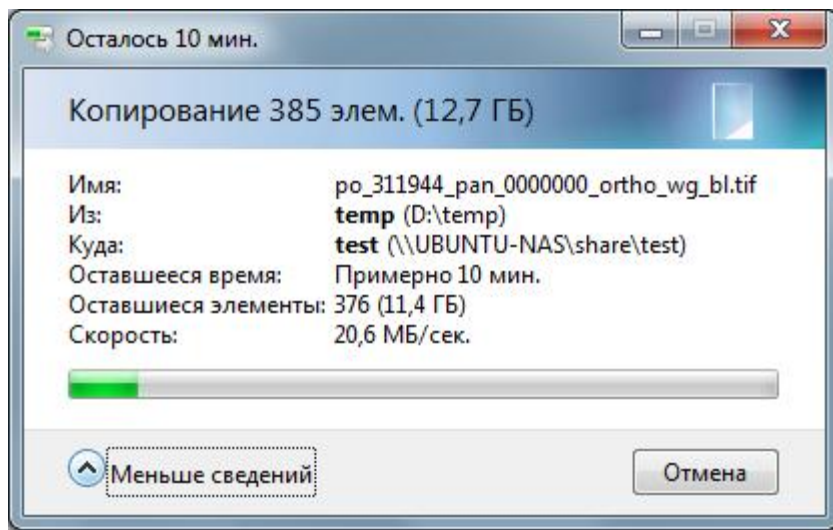
Перезапускаем Samba с новыми параметрами:

```
root@ubuntu-nas:/etc/samba# sudo service smbd restart
smbd stop/waiting
smbd start/running, process 2444
```

И входим через сетевое окружение.



Хранилище можно использовать по сети.



Для выполнения регулярной дефрагментации можно в cron добавить следующую строку:

```
30 21 * * 7 xfs_fsr -t 18000 /dev/md0
```

Обратите внимание в команде указывается именно устройство /dev/md*

Тестирование

Тестирование многопользовательского использования по сети (подключение Gigabit Ethernet).
Одновременное чтение/запись с 3-х компьютеров.

- Компьютер 1 - Ubuntu server из старого офисного компьютера (SATA2)
- Компьютер 2 - Windows 7 64 bit - ноутбук HP ProBook 4730s (SATA3 ?)
- Компьютер 3 - Windows Vista 32 bit - ноутбук Asus F3S (SATA1)

Операция	Компьютер 1	Компьютер 2	Компьютер 3
Чтение большого файла (4 Гб)	19,6 МБ/с	14,1 МБ/с	7,6 МБ/с
Чтение маленьких файлов (5-6 Мб)	17,5 МБ/с	13,9 МБ/с	7,5 МБ/с
Запись большого файла (4 Гб)	11 МБ/с	12 МБ/с	6,7 МБ/с
Запись маленьких файлов (5-6 Мб)	15,5 МБ/с	13 МБ/с	7,3 МБ/с

После всех манипуляций температура дисков:

```
bishop@ubuntu-nas:~$ sudo hddtemp /dev/sd?
/dev/sda: ST250DM000-1BD141: 33°C
/dev/sdb: Hitachi HDS5C3030ALA630: 34°C
/dev/sdc: Hitachi HDS5C3030ALA630: 33°C
/dev/sdd: Hitachi HDS5C3030ALA630: 35°C
```

[Обсудить в форуме](#) Комментариев — 47

Последнее обновление: 2014-05-15 00:08

Дата создания: 1.04.2012

Автор(ы): [Дмитрий Барышников](#)