

Storage Over IP III

Implementace iSCSI v komerčních zařízeních

(Cisco SN 5428 Storage Router)

Jan Haluza, Filip Staněk
Technická zpráva CESNETu č. 15/2002
30. listopadu 2002

Úvod

Tato zpráva navazuje na dvě předchozí technické zprávy (viz. [1] a [2]), kde se nachází i stručný úvod problematiky iSCSI a pojmy potřebné k bližšímu pochopení následujícího textu.

Zpráva popisuje zkušenosti s prací se zařízením Cisco SN 5428 Storage Router s připojeným diskovým polem Fortra v roli *targetu* spolu s *initiatorem*, který tvoří standardní PC s operačním systémem Linux.

Konfigurace systému

Veškerá zde popsaná měření byla provedena na následujících zařízeních.

Konfigurace hardware

target:

- 2 x Cisco SN 5428 Storage Router
- Diskové pole Fortra 10 Bay Fibre JBOD Rackmount
- 4 x SCSI disk Seagate ST 318452FC

initiator:

- Cisco MS 7800, procesor Intel Pentium III 1133MHz s 512 kB cache, 1 GB paměti RAM, NetGear GA620 Gigabit Ethernet, IDE řadič CMD649 rev. 2, pevný disk MAXTOR 6L040J2 (IDE, Ultra DMA 100, cache 1819 kB), 2 x Intel EtherExpress 100

Připojení iSCSI routeru k diskovému poli bylo realizováno přes rozhraní Fibre Channel. iSCSI router má osm Fibre Channel, dva gigabitové, jedno *high availability (HA)* a jedno *management* rozhraní. Testy jednoho klienta byly prováděny pomocí přímého propojení gigabitovými ethernet rozhraními bez použití dalších síťových prvků (routery switchy), které by mohly ovlivnit celkový výkon.

Konfigurace software

target:

firmware verze 2.3.1.3-K9

initiator:

OS Linux Redhat 7.3 (Valhala), jádra:

- ac kernel: 2.4.19-ac4 (Alan Cox)
- redhat kernel: 2.4.18-10 (Redhat)
- vanilla kernel: 2.4.19 (Marcello Tossati)

linux-iscsi verze 2.1.2

Firmware SN 5428 byl povýšen na verzi 2.3.1.3-K9, protože původně dodávaný firmware verze 2.3.1 vykazoval neshodné výsledky při opakovaných testech čtení.

Jádro dále označováno jako „ac kernel“ bylo standardní linuxové jádro s aplikovanými záplatami od Alana Coxe. Tyto je možno získat na [3].

Jádro dále označováno jako „redhat kernel“, je standardní linuxové jádro se speciálními úpravami firmy RedHat. Je možno jej získat na [4].

Jádro dále označováno jako „vanilla kernel“, je standardní linuxové jádro bez jakýchkoliv úprav, tak jak jej spravuje Marcello Tossati. Získat jej lze na [5]. Jak je vidět z označení všechna jádra jsou z větve 2.4.

V roli *initiatora* byl použitý balík *linux-iscsi* firmy Cisco [6]. Firma Cisco vyžaduje aby se s momentálními verzemi firmware iSCSI routeru počítal balík *linux-iscsi* řady 2 (tedy vyhovující iSCSI draftu verze 8).

Popis vlastností

iSCSI router SN 5428 se dá spravovat jak přes sériovou konzoli, tak i přes WWW rozhraní implementované Java appletem. WWW rozhraní nabízí dva módy přístupu. První (*monitor*) je určen pouze pro sledování, zatímco druhý (*admin*) i administraci routeru.

Dalšími zajímavými vlastnostmi routeru jsou funkce *failover* a možnost bootování po síti, tzv. *network boot*.

Funkce *failover* funguje tak, že dva routery jsou vzájemně propojeny přes *HA porty* a v případě výpadku jednoho z routerů, převezme automaticky ten druhý veškeré požadavky. Převezme samozřejmě i nastavení síťových rozhraní a pro *initiatory* se tedy nic nemění. Takto je možno *zabezpečit* pouze dva routery. Propojení routerů je realizováno kříženým *UTP* kabelem.

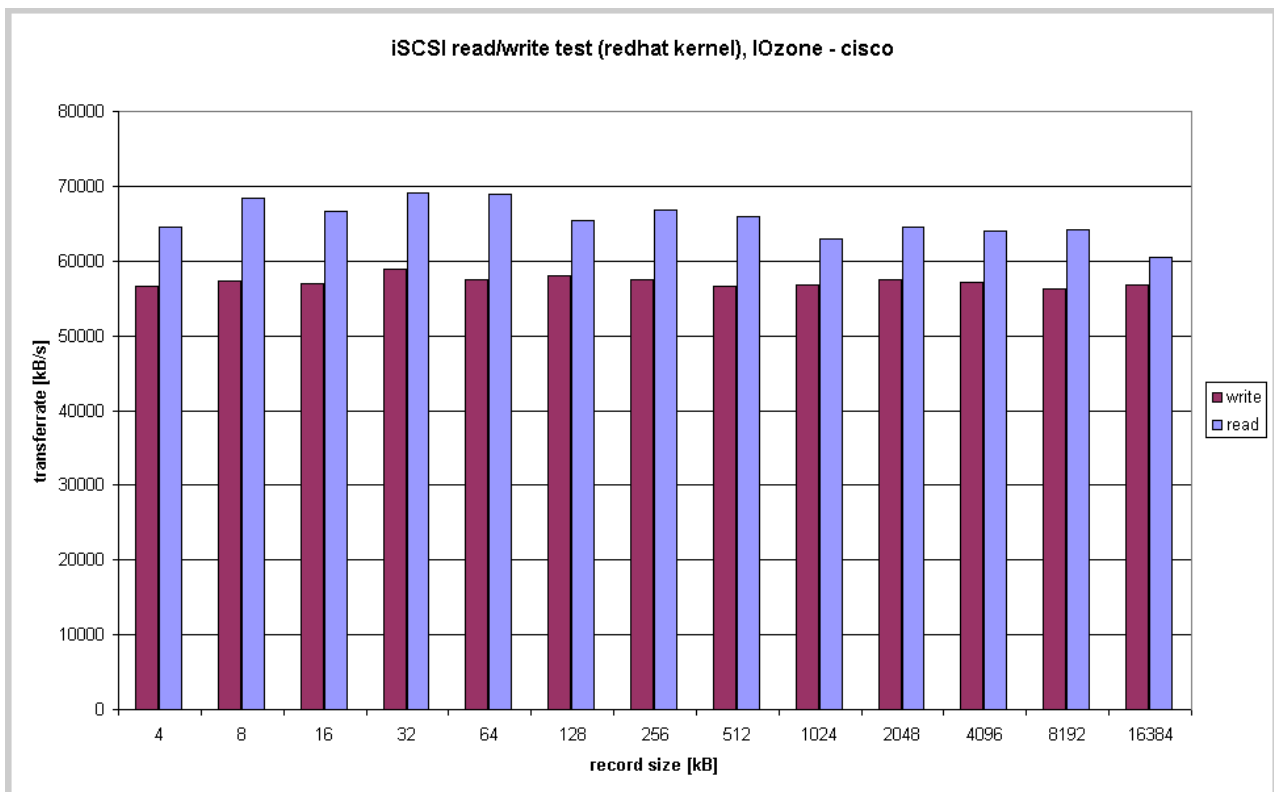
Funkce *network boot* by měla umožnit bootování po síti pro karty vyhovující standardu PEX 2.1. Současně je firmou Cisco podporováno pouze bootování platformy MS Windows 2000.

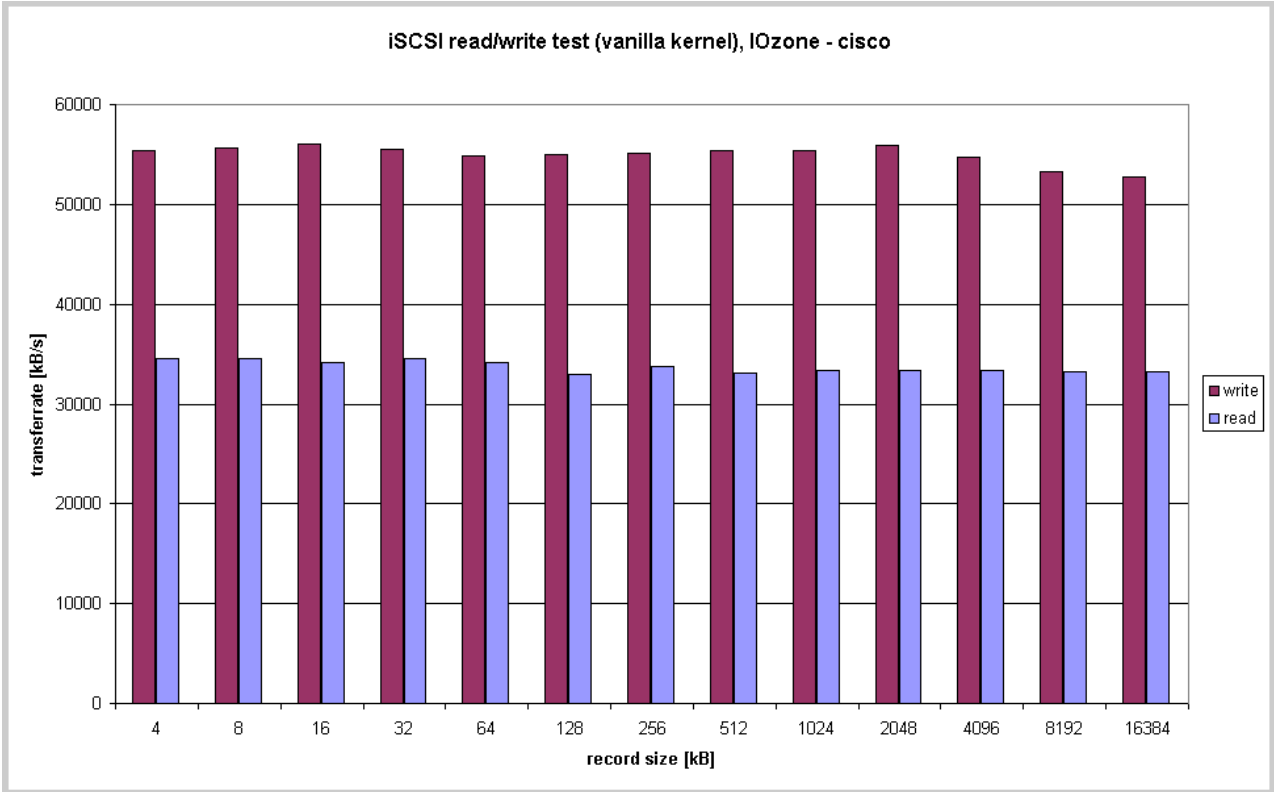
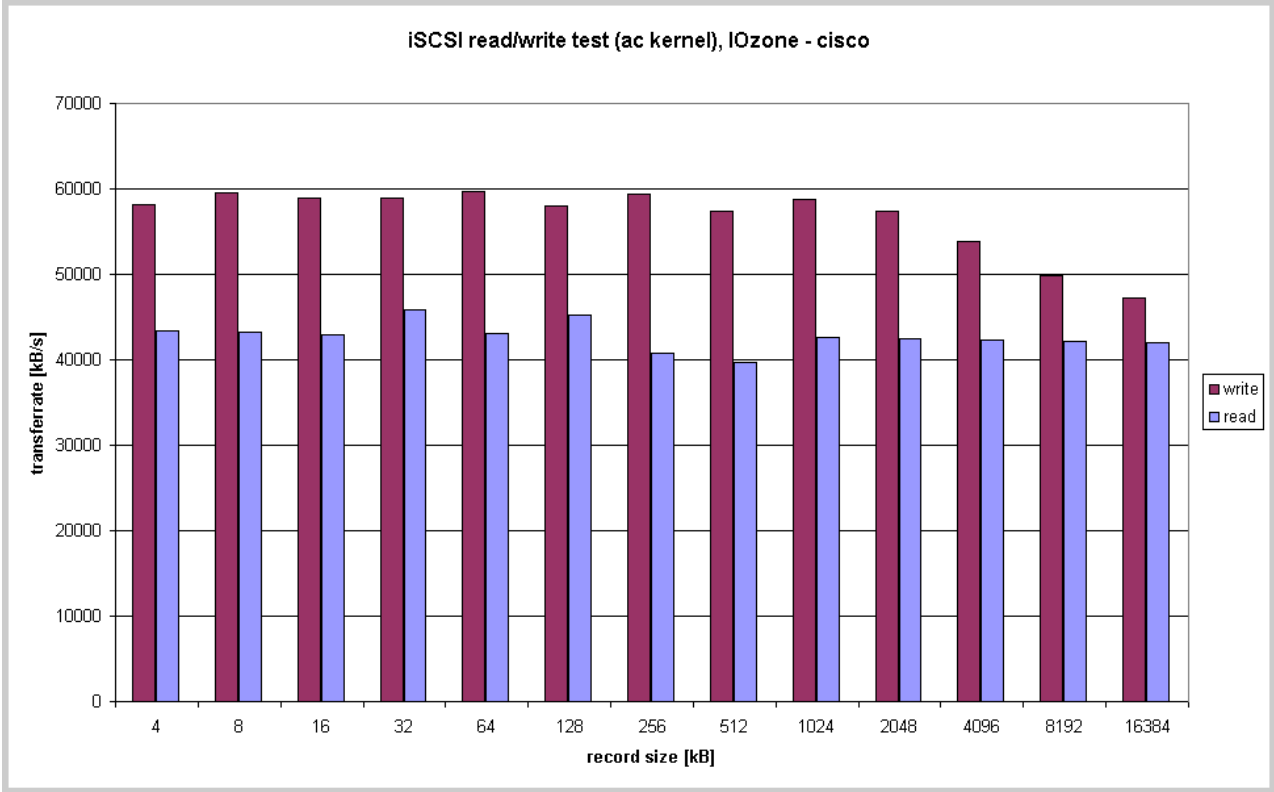
Výkon

Měření probíhala pomocí benchmarkovacího programu IOzone [7] s následujícími parametry:

```
iozone -Rb iscsi.wks -n 3g -g 3g -z -c -a -i 0 -i 1
```

Tedy s výstupem do binárního souboru ve formátu MS Excel, s minimální a maximální velikostí souboru 3 GB, s použitím funkce close() a automatickým měřením, s velikostmi přenášených bloků od 4 kB do 16 MB.





redhat kernel

record size	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384
write	56557	57249	56914	58873	57495	58063	57568	56687	56804	57485	57150	56273	56829
read	64588	68430	66605	69162	68907	65422	66816	65937	62976	64572	64066	64209	60491

ac kernel

record size	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384
write	58140	59488	58930	58870	59735	58038	59342	57349	58756	57318	53892	49842	47296
read	43431	43231	42967	45832	43139	45260	40843	39641	42555	42391	42376	42161	42069

vanilla kernel

record size	4	8	16	32	64	128	256	512	1024	2048	4096	8192	16384
write	55350	55619	56026	55496	54887	54967	55174	55397	55346	55867	54701	53262	52692
read	34565	34595	34188	34583	34104	32904	33714	33160	33338	33324	33372	33252	33233

Závěr

Z grafů jasně vyplývá, že nejvyšší výkon byl docílen se speciálním jádrem firmy Redhat. Bohužel počet všech záplat na standardní jádro je 215 (velikost 35 MB) a zatím není zcela jasné, co je příčinou tohoto nárůstu. Součástí je i soubor záplat Alana Coxe, proto bylo měření provedeno i s jeho verzí jádra. Výsledky byly lepší než u standardního jádra, avšak ne tak dobré jako u jádra firmy Redhat.

V porovnání s dříve testovaným zařízením Nishan IPS3300 [2] je SN 5428 rozhodně rychlejší volba. Taktéž funkce *failover* se jeví jako pozitivní. Bohužel funkci *network boot*, ač by byla dozajista také přínosem, se nepodařilo uvést do fungujícího stavu.

Literatura a odkazy

[1] Cesnet „Storage over IP“, <<http://www.cesnet.cz/doc/techzpravy/2002/iscsi/iscsi.pdf>>, June 18, 2002

[2] Cesnet „Storage over IP implementace iSCSI v komerčních zařízeních“, <<http://www.cesnet.cz/doc/techzpravy/2002/ipstorage/ipstorage.pdf>>, October 28, 2002

[3] Alan Cox, linux 2.4 kernel patches <<ftp://ftp.kernel.org/pub/linux/kernel/people/alan/linux-2.4>>

[4] Redhat, linux kernel 2.4 sources <<ftp://ftp.redhat.com/pub/redhat/linux/updates/7.3/en/os/i386>>

[5] Marcello Tossati, linux kernel 2.4 <<ftp://ftp.kernel.org/pub/linux/kernel/v2.4>>

[6] Cisco, linux-iscsi <http://sourceforge.net/project/showfiles.php?group_id=26396>

[7] IOzone, IOzone Filesystem Benchmark <<http://www.iozone.org>>

Přílohy

výpis konfigurace Cisco SN 5428

```
*[prglab-sn5428a]$ sh runningconfig
!
! CLUSTER
!
! cluster Cluster28
!
! ACCESSLIST
!
accesslist termit_access
accesslist termit_access 158.196.32.0/255.255.255.0
!
! VTP DOMAIN
!
vtp domain none
!
! VTP MODE
!
vtp mode transparent
!
! VLAN
!
! (no vlan(s) found)
!
! SCSIROUTER
!
scsirouter win_router
scsirouter win_router authenticate "no"
scsirouter win_router primary "none"
scsirouter win_router reserve proxy disable
scsirouter win_router failover primary none
scsirouter win_router failover secondary none
scsirouter win_router lun reset yes
scsirouter win_router serverIf gel 158.196.32.25/255.255.255.0
scsirouter win_router target winbootdisk lun 0 lunwnn
"20:00:00:04:cf:92:7b:3c"
scsirouter win_router target winbootdisk enabled
scsirouter win_router target winbootdisk accesslist "all"
!
scsirouter unix_router
scsirouter unix_router authenticate "none"
scsirouter unix_router primary "none"
scsirouter unix_router reserve proxy disable
scsirouter unix_router failover primary none
scsirouter unix_router failover secondary none
scsirouter unix_router lun reset no
scsirouter unix_router serverIf ge2 10.1.116.1/255.255.255.0
scsirouter unix_router target unixdisk wwpn "21:00:00:04:cf:92:7e:e9"
scsirouter unix_router target unixdisk enabled
scsirouter unix_router target unixdisk accesslist "all"
!
! SYSTEM
!
hostname prglab-sn5428a
!
! Mgmt Port
!
interface mgmt ip-address 10.62.8.185/255.255.255.0
!
! HA Port
!
interface ha ip-address 10.1.115.17/255.255.255.248
!
! GE
!
```

```
interface ge2 mtusize 1500
interface ge2 autonegotiation autodetect
no interface ge2 vlan enable
!
! GE
!
!
! ROUTES
!
ip default-gateway 158.196.32.254
!
! ADMIN
!
admin contactinfo name "Jarry Pilar"
admin contactinfo email "jpilar@cisco.com"
admin contactinfo phone "+420602339424"
!
! ADMIN LOGIN
!
admin password <password>
!
! MONITOR LOGIN
!
monitor password <password>
!
! Sntp
!
clock timezone Europe/Prague
!
! SNMP
!
snmp-server community public ro
snmp-server community private rw
no snmp-server host all traps
no snmp-server sendauthtraps
snmp-server linkupdown mgmt
snmp-server linkupdown ha
snmp-server linkupdown ge2
snmp-server linkupdown fc1
snmp-server linkupdown fc2
snmp-server linkupdown fc3
snmp-server linkupdown fc4
snmp-server linkupdown fc5
snmp-server linkupdown fc6
snmp-server linkupdown fc7
snmp-server linkupdown fc8
!
! DNS
!
ip name-server 144.254.71.184 144.254.10.123
ip domain-name cisco.com
!
! SOFTWARE
!
software http url "http://144.254.88.201/~jpilar"
software tftp hostname "10.62.8.165"
software tftp directory "/tftpboot"
software http username none
software proxy username none
!
! CDP
!
cdp enable
cdp timer 60
cdp holdtime 180
cdp interface mgmt enable
cdp interface ha enable
cdp interface ge2 enable
!
! HA
```

```
!  
! ha configuration standalone  
!  
! REMOTE LOG  
!  
logging 158.196.32.1  
!  
! LOGGING ROUTE FACILITY  
!  
logging level debug from all to rslog  
!  
! RESTRICT  
!  
restrict mgmt ftp  
no restrict mgmt telnet  
no restrict mgmt http  
no restrict mgmt snmp  
no restrict mgmt rlogin  
no restrict mgmt ssl  
!  
restrict ha ftp  
restrict ha telnet  
no restrict ha http  
restrict ha snmp  
no restrict ha rlogin  
no restrict ha ssl  
!  
restrict ge2 ftp  
restrict ge2 telnet  
restrict ge2 http  
restrict ge2 snmp  
no restrict ge2 rlogin  
no restrict ge2 ssl  
!  
!  
! FC  
!  
interface fc zoning default all  
interface fc zoning autosave enable  
interface fc domainid 1  
no interface fc domainid lock enable  
interface fc interop-credit 12  
!  
! FC Port(s)  
!  
interface fc1 enable  
no interface fc1 al-fairness enable  
interface fc1 fan-enable enable  
interface fc1 mfs-bundle enable timeout 10  
interface fc1 linkspeed auto  
interface fc1 type gl-port  
!  
interface fc2 enable  
no interface fc2 al-fairness enable  
interface fc2 fan-enable enable  
interface fc2 mfs-bundle enable timeout 10  
interface fc2 linkspeed auto  
interface fc2 type fl-port  
!  
interface fc3 enable  
no interface fc3 al-fairness enable  
interface fc3 fan-enable enable  
interface fc3 mfs-bundle enable timeout 10  
interface fc3 linkspeed auto  
interface fc3 type fl-port  
!  
interface fc4 enable  
no interface fc4 al-fairness enable  
interface fc4 fan-enable enable  
interface fc4 mfs-bundle enable timeout 10
```

```
interface fc4 linkspeed auto
interface fc4 type fl-port
!
interface fc5 enable
no interface fc5 al-fairness enable
interface fc5 fan-enable enable
interface fc5 mfs-bundle enable timeout 10
interface fc5 linkspeed auto
interface fc5 type fl-port
!
interface fc6 enable
no interface fc6 al-fairness enable
interface fc6 fan-enable enable
interface fc6 mfs-bundle enable timeout 10
interface fc6 linkspeed auto
interface fc6 type fl-port
!
interface fc7 enable
no interface fc7 al-fairness enable
interface fc7 fan-enable enable
interface fc7 mfs-bundle enable timeout 10
interface fc7 linkspeed auto
interface fc7 type fl-port
!
interface fc8 enable
no interface fc8 al-fairness enable
interface fc8 fan-enable enable
interface fc8 mfs-bundle enable timeout 10
interface fc8 linkspeed auto
interface fc8 type fl-port
!
! AAA
!
aaa new-model
aaa authentication iscsi default local group tacacs+ group radius
username "gry73" password "9 da28171b333baacabfebf7d5a519676c42"
tacacs-server key "9 08711eb2d0cd0f9266ed5a164e30d897e0"
tacacs-server host 158.196.149.9 auth-port 49
*[prglab-sn5428a]$
```