

Analýza nastavení serveru

Obsah

1. Úvod	2
2. Konfigurace	2
3. Analýza stavu serveru	2
3.1. Práce s disky	2
3.2. Nastavení služeb	3
3.3. Využití paměti RAM	4
3.4. Komunikace na síti	4
4. Souhrn změn	4
5. Shrnutí	5
6. Závěr	5

1. Úvod

Na základě Vaší objednávky Vám předkládáme analýzu Linuxového serveru a návrh řešení Vámi popsaných problémových stavů.

Problémové stavy jsou pomalé odezvy poštovního serveru na akce klientů, zejména v uživatelských špičkách, doprovázené vysokou zátěží procesoru i paměti serveru. V následujících kapitolách naleznete detailní analýzu všech nalezených problémů, návrhy jejich řešení, a popis změn jenž byly provedeny v průběhu analýzy. Shrnutí nejdůležitějších problémů naleznete v kapitole 5..

2. Konfigurace

Procesor: Intel Pentium III 500 MHz

Paměť: 512 MB (+ 400 MB swap)

Disky: 2x HP 18 GB, 2x Seagate 18 GB, SCSI řadič Adaptec 7895, pásková mechanika HP

Síťová karta: Intel EtherExpress Pro 100 (100 Mbit/s)

Operační systém: Debian GNU/Linux 3.0 (woody)

Služby poskytované serverem: Email, News, LDAP, Intranet, DNS, DHCP, Proxy a tisk.

3. Analýza stavu serveru

3.1. Práce s disky

V serveru jsou 4 disky, dva starší disky HP (SCSI-2) původně dodané se serverem, dva novější disky Seagate (SCSI-3), instalované do serveru po poruše předchozích disků HP. Všechny jsou připojeny k jednomu ze dvou dostupných kanálů SCSI řadiče Adaptec 7895.

Doporučení: *Používání starších disků HP dodaných spolu se serverem je do budoucna rizikem. Pokud se již dva disky HP porouchaly, je velmi pravděpodobné, že zbývající dva disky téhož typu se porouchají také. Proto jsme poštovní schránky uživatelů přemístili ze staršího disku /dev/sdb na jeden z novějších disků /dev/sdd. Doporučujeme přenést i data z posledního disku HP /dev/sda na jeden z novějších disků, např. /dev/sdc a odstranit disky HP ze systému.*

Jistého zrychlení práce s poštou lze dosáhnout využitím druhého kanálu SCSI řadiče, t.j. připojením disku s poštou /dev/sdd na druhý kanál, na který by se už nic dalšího nepřipojovalo.

Disk /dev/sdb1 byl připojen do systému s parametry `sync, noatime`. `Noatime` sice může mírně zrychlit práci s diskem, tím že se nezapisuje čas přístupu k souborům (access time), avšak parametr `sync` způsobuje výrazné zpomalení logických zápisů na disk. Každý proces zapisující na disk (v tomto případě procesy IMAP/POP3 serveru) čeká na provedení fyzického zápisu dat, zatímco ve výchozím `async` režimu jsou data zapisována do vyrovnávací paměti OS a proces již nečeká na fyzické provedení zápisu.

Experimentálně jsme prověřili, že zápis 1MB dat na disk v synchronním režimu trval 40 sekund, zatímco v asynchronním režimu trval přibližně 20 milisekund (fyzický zápis pochopitelně probíhal déle, ale na pozadí v režii operačního systému). Stejně výsledky byly naměřeny na obou starších discích. Synchronní zápis 1 MB na nové disky trval pouze 4 sekundy. Asynchronní zápis byl stejně rychlý u všech disků. Čtení z disku není parametrem `sync` ovlivněno. (Naměřené hodnoty jsou pouze velmi přibližné neboť měření bylo prováděno za provozu.)

Doporučení: *Nastavené synchronní zapisování na disk výrazně zpomalovalo práci s poštou. Přenesením pošty na jiný disk byl problém vyřešen. Doporučujeme používat disky pouze v asynchronním režimu.*

Virtuální paměť (swap) je rozdělena na 3 disky, což nepřináší žádné zrychlení. Operační systém využívá pouze jeden z nich (má dostatečnou kapacitu), zbylé dva nejsou používány. (Alespoň po dobu naší analýzy tomu tak bylo.)

Doporučení: *Po odstranění starých disků bude virtuální paměť rozdělena na 2 disky, což ale není na závadu. Přerozdělování disků za účelem sjednocení virtuální paměti, či využití jedné její části na uložení dat by nepřineslo zlepšení, odpovídající námaze s tím spojené.*

SCSI řadič má nastavené různé přenosové rychlosti pro různé porty (disky). První port (/dev/sda) má nastavenou rychlost 40 MB/s, ostatní pouze 11.6 MB/s.

Doporučení: *Navrhujeme změnit nastavení přenosových rychlostí mezi řadičem a disky. Řadič i disky by měly být schopny přenosových rychlostí až 40 MB/s (SCSI Ultra Wide). Toto nastavení lze změnit v BIOSu SCSI řadiče, což vyžaduje restart počítače a s tím spojený krátkodobý výpadek služeb. Pokud by ani všechna výše uvedená řešení nezajistila dostatečné zrychlení, bylo by nutné uvažovat o diskovém poli RAID.*

3.2. Nastavení služeb

Na serveru běží následující služby:

- Mail server: SMTP – Postfix, IMAP/POP3 – Cyrus (celkem 370 mailboxů)
- News server: NNTP – innd
- Adresářový server: LDAP – slapd (OpenLDAP)
- Intranetový server: HTTP – Apache
- Name server: DNS – Bind
- DHCP server
- Tiskový server: IPP – Cups
- Proxy server: HTTP – Squid
- X Display Manager – gdm a X server
- NTP server
- Netbackup client
- Ident server
- Portmap – dracd (POP before SMTP), statd (součást NFS)
- Inetd – discard, daytime, time

Pravidelně jsou využívány služby Mail, News, LDAP, Intranet, DNS a DHCP, velmi zřídka potom (pro občasné administrativní úkony) služby Proxy, Tisk a X display. NTP (přesný čas) a Netbackup (zálohování po síti) jsou služby používané serverem samotným. Ostatní služby nejsou využívány vůbec a je možné jejich zastavení.

Drobné chyby v nastavení poštovního serveru byly opraveny a jsou zdokumentovány níže. Zásadní chyby v nastavení jednotlivých služeb nebyly nalezeny. IMAP/POP3 server Cyrus byl již před naší analýzou spouštěn s vyšší prioritou než normální procesy, což ale nezlepšilo výrazně dobu odezvy.

Doporučení: *Závažným problémem je, že na serveru je intenzivně využíván kromě pošty i intranetový server. Tento server má vysoké nároky na procesor, paměť i síťovou komunikaci a ve špičkách, kdy se uživatelé přihlašují jak do pošty, tak do intranetu, lze očekávat zátěž převyšující kapacitu serveru. Doporučujeme proto rozdělit služby serveru potato.csst.cz na dva počítače. Na jednom ponechat pouze Mail, News a LDAP, na druhém provozovat Intranet, DNS a DHCP server.*

Dále doporučujeme odstranění nepoužívaných, či málo používaných služeb a serverů. Tyto služby představují nejen zbytečnou zátěž serveru (zejména Proxy server a X server) ale také bezpečnostní problém. Každá služba může mít skrytou bezpečnostní díru a čím více služeb na serveru běží, tím více šancí má potenciální útočník na průnik. Proto doporučujeme všechny služby které nejsou nutné pro běh serveru a využívané všemi uživateli vypnout nebo přímo odinstalovat. Pokud administrátoři občas potřebují některé služby (např. tisk), doporučujeme jejich instalaci přímo na pracovní stanici některého z administrátorů.

3.3. Využití paměti RAM

Server je vybaven 512 MB RAM. Statistiky využití paměti ukazují, že systém je nucen využívat virtuální paměť (swap). Množství využití paměti postupně roste, až do dalšího restartu serveru. Nejvyšší hodnoty se pohybovaly kolem 60 MB. Systém virtuální paměti není možné nastavit tak, aby určité procesy (např. Cyrus a Postfix) ponechal v RAM a jiné (např. Squid) odložil na disk (swap). Tím se může stát, že některé často využívané procesy jsou odkládány na disk a vzápětí načteny do RAM, čímž se zpomalí odezva serveru. Nejnáročnější na RAM jsou procesy Apache (Intranet), Squid (Proxy), News server a X server.

Doporučení: *Problémy s pamětí a swapem lze částečně vyřešit rozšířením operační paměti (až na 1 GB), čímž by odpadla nutnost používat swap v běžném provozu a vypnutím nebo odinstalováním nepoužívaných či málo používaných služeb a serverů, zejména paměťově náročných: Proxy server (Squid), X server (včetně gdm). Seznam služeb a jejich využití je v předchozí podkapitole.*

3.4. Komunikace na síti

Některé služby serveru potato.csst.cz využívají služeb jiných serverů. Jedná se o Mail server, jenž využívá služeb NT serveru pro autentizaci uživatelů a Intranetový server, jenž využívá služeb databázového serveru (Oracle). Ke komunikaci s klienty i servery je využívána stejná síť a ve špičkách může docházet k jejímu přetížení. Parametry protokolu TCP v jádře OS byly již dříve upravovány, ale není pravděpodobné, že by tím bylo dosaženo výrazného zlepšení odezvy systému.

Doporučení: *S přihlédnutím k nárokům na komunikaci mezi servery (Intranet ? databáze, Mail ? autentizace NT) a s ohledem na dostupnost technologie gigabitového ethernetu, doporučujeme vytvoření samostatného ethernetového segmentu 1000Mbit/s, vyčleněného pro vzájemnou komunikaci mezi servery.*

4. Souhrn změn

- V souboru `/etc/imapd.conf` byla změněna hodnota parametru `popminpoll` z 1 na 0. Tento parametr určuje minimální přestávku (v minutách) mezi dvěma přístupy přes POP3 protokol.

Změnou se odstraní problém uživatelů z poboček, kteří při častější kontrole pošty, než jednou za minutu dostávali od poštovního serveru chybová hlášení.

- V souboru */etc/postfix/main.cf* byla změněna hodnota parametru **biff** z **yes** na **no**. Tento parametr určuje zda se má spouštět upozornění lokálních uživatelů na novou poštu. Vzhledem k tomu, že přímo na serveru žádní uživatelé (s výjimkou správce) nepracují, nemá tato služba smysl.
- V souboru */etc/squid.conf* byl povolen přístup na WWW z počítačů na nichž jsme pracovali.
- Systém balíčků byl v nekonzistentním stavu. Příkazem **apt-get -f install** byl uveden do konzistentního stavu, přičemž byly odstraněny balíčky: *kerberos4kth-dev* a *libssl-dev*. Odstranění těchto balíčků by nemělo narušit činnost serveru, neboť jsou zapotřebí pouze při kompilování programů, používajících knihovnu Kerberos nebo SSL.
- Nainstalovali jsme balíčky *rsync*, *hdparm* a *scsitools*.
- Poštovní schránky byly překopírovány z adresáře */var/cyrus* na disku */dev/sdb1* do adresáře */var/cyrus.new* na disku */dev/sdd2*. Disk */dev/sdb1* byl odmountován a adresář */var/cyrus.new* přejmenován na */var/cyrus*. Po dobu této operace byl IMAP/POP3 server dočasně nedostupný.
- V souboru */etc/fstab* byl změněn adresář do něhož se mountuje disk */dev/sdb1* z */var/cyrus* na */var/cyrus.old*. Poté byl disk */dev/sdb1* opětovně přimountován.

5. Shrnutí

V nastavení hardware serveru jsme objevili neoptimální nastavení řadiče disků (SCSI). Vzhledem ke skutečnosti, že z původních disků v serveru již dva musely být pro závadu vyměněny, lze předpokládat, že i zbývající dva disky se nacházejí na konci své životnosti. Poštovní schránky, jenž byly uloženy na jednom z původních disků, jsme proto preventivně přesunuli na jeden z nových disků.

V nastavení poštovního software jsme zjistili pouze drobné nepřesnosti, které byly ihned odstraněny. Závažnou komplikací bylo nastavení synchronního zápisu na disk s poštovními schránkami. Ta byla odstraněna změnou nastavení. Dalším problémem zvyšujícím zátěž serveru je přítomnost mnoha služeb, z nichž mnohé nejsou využívány. Softwarová nastavení používaných služeb a jádra operačního systému jsou v pořádku.

Množství služeb, poskytovaných serverem, je při současném počtu uživatelů na hranici jeho možnosti. Doporučujeme jej proto ponechat pouze jako server pro Intranet, DNS a DHCP. Email, News a LDAP doporučujeme přenést na samostatný stroj, dimenzovaný pro vyšší počet uživatelů tak, aby postačoval i Vaším budoucím potřebám. Dočasným řešením zátěže serveru je odstranění nevyužívaných služeb a rozšíření operační paměti.

6. Závěr

Chování serveru bylo po našich zásazích intenzivně sledováno. Bylo konstatováno, že problémové stavy serveru byly skutečně odstraněny. Věříme, že jsme tak naplnili Vaše původní očekávání.

Veškerá výše uvedená doporučení jsme schopni realizovat na základě další objednávky nebo smlouvy o správě serveru. V případě jakýchkoli dotazů nás prosím kontaktujte. Kontaktní údaje jsou uvedeny na první stránce nabídky či na přiložené vizitce. Velice se těšíme na další spolupráci.

S přátelským pozdravem,

Vladislav Kurz, M.Sc.
systémový administrátor

Andrej Pavelka
jednatel WebStep, s.r.o

V Brně, dne 11. prosince 2006

Analýzu provedl: Vladislav Kurz, M.Sc.
Odborný dohled: Ing. Jiří Kadeřávek
Zprávu vypracoval: Vladislav Kurz, M.Sc.
vladislav.kurz@webstep.net