

Tru64: Upravljanje performansama

Dinko Korunić
v1.1, studeni 2005.



O predavaču

- višegodišnji vanjski suradnik časopisa Mrež@, vlastita kolumna "Digitalna radionica - Linux", itd.
- vanjski suradnik SRCE-a: forenzike provaljenih sustava, izgradnja sistemskih paketa, helpdesk za sistemce, sigurnost Unix baziranih sustava, predavač, itd.
- sigurnosni ekspert pri InfoMAR d.o.o.

Tijekom prezentacije

- **ako što nije jasno - pitajte i tražite objašnjenje!**
- **ako što nije točno - ispravite!** greške su moguće i česte, posebice za prvu verziju
- **diskusija** je poželjna i produktivna
- **ako je prebrzo - tražite da se uspori!**
- **ako je pak presporo i uspavljuje vas - lako se ubrza** sa sadržajem
- **podijelimo** zajedno vlastita iskustva

Dio I: Ponavljanje



Repetitorij

- dodavanje korisnika:
 - naredbe **useradd**, **adduser**
 - primjer: `useradd -c "Dinko Korunic" -d /usr/users/kreator -g users -m -p -s /usr/bin/bash kreator`
- brisanje korisnika:
 - naredba **userdel**
 - primjer: `userdel -r kreator`
- promjena lozinke:
 - naredba **passwd**
 - primjer: `passwd korisnik`

Repetitorij (2)

- gašenje sustava, resetiranje sustava:
 - naredbe **shutdown**, **init**
 - primjer: `shutdown -h now`
 - primjer: `init 1`
- izbjegavati:
 - naredbe **halt**, **reboot**
- editiranje datoteka:
 - naredbe **vi**, **sed**, **ex**, **ed**, **emacs**
- kopiranje/pomicanje datoteka:
 - naredbe **cp**, **mv**

Repetitorij (3)

- brisanje datoteka i direktorija:
 - naredbe **rm**, **rmdir**
- pregled procesa:
 - naredba **ps**
- gašenje procesa:
 - naredba **kill**
- općenito administriranje sustava:
 - naredba **sysman**

Dio I: Teorija održavanja



Uvod

- uspješni sistemac - stalni nadzor, iskustvo:
 - prepoznavanje normalnog i abnormalnog ponašanja sustava i korisnika
 - predviđanje rezultata promjena aplikacija, promjena dijelova sustava ili načina korištenja
 - prepoznavanje tipičnih uzoraka korištenja
- moguće automatizirati - ali nikad u potpunosti!

Uvod (2)

- tipični nadzor:
 - korištenje CPU
 - korištenje memorije
 - korištenje diskovnog prostora
 - brzina izmjena stranica i keširajući podsustav
 - mrežne U/I statistike
 - diskovne U/I statistike
 - broj procesa i otvorenih datoteka
- nužan vremenski nadzor, grafovi, statistike!

Uvod (3)

- ponašanje:
 - različito o profilu korisnika
 - različito o broju korisnika
 - različito o namjeni poslužitelja
 - različito o dobu dana
 - itd.
- behavioralni sustavi - obično nemogući!

Logovi

- uvijek u `/var/adm`
 - spremaju se kroz `syslogd` i `binlogd`
 - konfiguracija `/etc/syslogd.conf`, `/etc/binlogd`
 - nije uvijek lako pronaći tražene informacije
 - `grep`, `awk`, itd
- naredba `uerf`:
 - pretražuje sistemske zapisnike/logove
 - hardverske greške, OS problemi, aplikacije, itd
 - primjer: `uerf -R`

Logovi (2)

- specijalizirani alati:
 - Event Manager - EVM
 - DECevent - naredba **dia**
 - Compaq Analyze i WEBES alati - naredba **ca**

Parametri poslužitelja

- brzina rada i efikasnost uvjetovana hardverskim karakteristikama:
 - brzina procesora
 - količina radne memorije
 - komunikacijski medij
 - ulazno-izlazni uređaji
 - moguće nadogradnje
- softver - mora inicijalno biti prilagođen hardveru

Softverski parametri

- komercijalna aplikacija - teško utjecati
- mogućnosti promjene:
 - odabir podatkovnih struktura (memorija)
 - odabir algoritama i načina obrade (CPU, I/O)
 - korištenje sistemskih resursa (datotečni sustav), distribuirani rad (brzina rada, I/O..)
- moguće vrlo velike oscilacije - podrazumijeva se da je softver unaprijed dobro podešen
- primjer: zaključavanje, veliki I/O - usporenja!

Savjeti

- upoznajte se sa sustavom!
- jesu li svi servisi nužni?
- počnite nadgledati prije problema!
- ostvarite grafički vid nadgledanja cijelog poslužitelja - resursa, servisa, itd.
- definirajte procedure detektiranja problema, rješavanja problema

Pojmovnik

- resursi:
 - memorija
 - I/O uređaji
 - CPU
- performanse resursa ograničene:
 - kapacitet
 - potraživanje
- značajke performansi:
 - propusnost
 - vrijeme odgovora, vrijeme obrade

Pojmovnik (2)

- moguće poboljšati performanse odgovora i obrade:
 - pravedno dijeljenje resursa
 - smanjenje potrošnje resursa
 - dijeljenje opterećenja
 - prebacivanje dijela/dijelova zadaće
- moguće poboljšati kapacitet resursa:
 - horizontalnim ili vertikalnim nadogradnjama

Memorija

- kritični resurs
- operacijski sustav i svi korisnički procesi
- memorija - skupa
- disk - jeftin
- razlika latencija memorije i diska - 1-2 tisuće puta
- međuspremnici - poboljšavaju performanse sporih I/O i inih uređaja; najčešće blokovski orijentirani (za stranice i datoteke)

Memorija (2)

- virtualna memorija - swap
- page fault - stranica nije u fizičkoj memoriji, vadi se iz međuspremnik ili virtualne
- straničenje - spremanje blokova u virtualnu memoriju - skup za I/O, spor, opterećuje
- memorije - nikad dovoljno
- više algoritama spremanja (ovisi o OS):
 - konzervativni
 - agresivni

Memorija (2)

- na procesoru (L1, L2, L3) integrirano ili vanjski - brzina uglavnom identična procesoru
- glavna memorija - velika, sporija, nužno pristupati preko sabirnice
- međuspremnicu za disk (UBC) - dio glavne memorije, dinamički rastu ili padaju s obzirom na diskovne aktivnosti i količinu slobodne/zauzete memorije

Memorija (3)

- dijelovi upravljačkih programa - također spremnici
- fizički I/O uređaj - dodatna interna memorija
- virtualni prostor: korisnički 4TB, neiskorišteno, kernel 4TB
- virtualni prostor - mapiranje prema fizičkom!

Memorija (4)

- page fault - traži se virtualna stranica, nužno kopirati u fizičku stranicu:
 - zero-filled-on-demand: prvi put, alocira se, prazni se, itd
 - short: lokalno u memoriji
 - page-in: u swapu
 - copy-on-write: r/o a nužna r/w modifikacija

Diskovni I/O

- DMA ili ne
- inteligentni I/O uređaji, RAID, međuspremnicu, baterije
- datotečni sustavi - zapisnički, sinkroni, itd.
 - ovisno o tipu, mijenja se upotreba. brzina
- "sirovi" pristup - nestrukturirani, znakovni
- blokovski pristup - strukturirani, međuspremnicu
- visok I/O - "jede" performanse

Metode rješavanja problema

- povećati kapacitet resursa:
 - kroz parametre jezgre
 - dodatni/alternativni hardver
- "smanjiti" kapacitet resursa:
 - pokretati manje istovremenih poslova/korisnika
 - redizajnirati aplikacije
 - podesiti dobar kompromis između konkurentnih procesa

Metode rješavanja problema (2)

- odterećenje:
 - izvršavati aplikacije i sl. sa udaljenog računala
 - distribuirati datoteke i datotečne sustave po višestrukim diskovima
 - smanjiti maksimalnu količinu programa po korisniku
 - itd.

Dio II: Alati za nadzor



Naredba uptime

- **uptime** - ispis vremena, koliko je poslužitelj aktivan, broj korisnika i opterećenje (srednji broj poslova koji čekaju na izvršenje) u zadnjih 5, 30 i 60 sekundi
- primjer:
 - `uptime`
- poželjno: što manje opterećenje
- broj poslova na čekanju - maksimalno treba porasti do broja procesora (naredba **psrinfo**)

Naredba swapon

- **swapon** - prikazuje statistike potrošnje virtualnog memorijskog prostora
- parametri:
 - -s - statistike
- primjer:
 - swapon -s
- poželjno: što manje aktivnih(!) procesa u virtualnoj memoriji, što više slobodne radne memorije

Naredba iostat

- **iostat** - napredne ulazno/izlazne statistike za terminale i CPU statistike
- sintaksa:
 - `iostat [uredjaj...] [interval_prikaza [broj_upita]]`
- primjer:
 - `iostat 1 5`
- izlazne vrijednosti:
 - `tin, tout` - ulazno i izlazno znakova iz terminala u sekundi

Naredba iostat (2)

- bps - blokova (1024 bajtova) po sekundi
- tps - transakcija po sekundi
- us - % CPU proveden u korisničkim aplikacijama
- ni - % CPU proveden izvršavajući aplikacije niskog prioriteta
- sy - % CPU proveden izvršavajući jezgrine procese
- id - % CPU proveden ne radeći ništa
- veliki skokovi u us, ni, sy - uglavnom pokazuju jače sistemske aktivnosti

Naredba vmstat

- **vmstat** - statistike procesa, virtualne memorije, diskovlja, prekida i CPU aktivnosti
- sintaksa:
 - `vmstat [interval_prikaza [broj_upita]]`
 - `vmstat -f`
 - `vmstat -s`
- parametri:
 - `-f` - broj stvorene djece od početka rada sustava
 - `-s` - nakupljene statistike od početka rada

Naredba vmstat (2)

- ispis:
 - procs - informacije o procesima
 - r, w, u - dretve koje se izvršavaju, čekaju s mogućim prekidima, čekaju bez mogućih prekida
 - memory - informacije o memoriji
 - act, free, wire - stranice koje se koriste i moguće ih je prebaciti u virtualnu memoriju, slobodne stranice, zaključane stranice
 - pages - informacije o stranicama
 - fault, cow, zero - greške u pronalaženju stranice, greške u kopiranju stranice radi pisanja, greška u punjenju stranice nulom

Naredba vmstat (3)

- react, pin, pout - greške u stranicama koje treba izbaciti dok su neaktivne, broj zahtjeva za stranicom sa diska, broj zahtjeva za izbacivanjem stranice
- intr - informacije o prekidima
 - in, sy, cs - broj prekida od uređaja po sekundi, sistemskih poziva po sekundi, broj izmjena konteksta zadataka i dretvi po sekundi
- cpu - informacije o procesoru
 - us, sy, id - provedeno procesorsko korisničko vrijeme, provedeno sistemsko vrijeme, provedeno vrijeme ne radeći ništa

Naredba ps

- **ps** - osnovni prikaz liste procesa
- parametri:
 - -e - ispisuje sve procese
 - -f - generira puni ispis
 - -l - generira dugi ispis
- primjer:
 - standardna sintaksa: `ps -elf`
 - BSD sintaksa: `ps xuaw`

Naredba ps (2)

- ispis:
 - USER - korisnik, UID
 - PID - jedinstveni identifikator procesa
 - PPID - jedinstveni identifikator roditelja procesa
 - %CPU - prosječna CPU potrošnja u zadnjoj minuti
 - STARTED - kada je naredba pokrenuta
 - TT - osnovni/roditeljski terminal
 - TIME - koliko je ukupno provedeno CPU vrijeme u procesu
 - COMMAND - naredba i njeni argumenti

Naredba df

- **df** - nadzor potrošnje diskovnog prostora, u 512-bajtnim blokovima (potrošeno, na raspolaganju), postotak potrošnje, točka montiranja
- parametri:
 - -i - prikazuje informacije s obzirom na inode

Naredba `dumpfs`

- `dumpfs` - prikaz naprednih informacija o diskovnom podsustavu i uređajima
- sintaksa:
 - `dumpfs diskovni_uredjaj`
 - `dumpfs datotecni_sustav`
- prilično napredno, zahtijeva mnogo predznanja!

Naredba nfsstat

- **nfsstat** - statistike NFS i RPC sustava
- parametri:
 - -i - prikaz informacija u nekom intervalu
 - -c - klijentske informacije
 - -s - serverske informacije
 - -n - NFS informacije
 - -r - RPC informacije
 - -z - obriše statistike na 0
- napredno, očekuje mrežno predznanje!

Naredba netstat

- **netstat** - prikazuje sve mrežne informacije: o uređajima, sustavu, konekcijama, itd
- sintaksa:
 - netstat [parametri] [-f adresna obitelj] [-p protokol] [interval]
 - ... kompliciranija sintaksa, vidjeti man stranice ...
- parametri:
 - -a - ispisuje stanja socketa

Naredba netstat (2)

- parametri:
 - -f - adresne obitelji: inet, inet6, unix, all, any
 - -g - statistike od podizanja sustava do trenutnog vremena
 - -H - ARP tablica, imitira **arp** naredbu (arp -a)
 - -i - stanja trenutno konfiguriranih uređaja, MAC adrese, itd.
 - -l - informacije o željenom uređaju
 - -M - ispisuje multicast usmjerivačke informacije
 - -n - ispisuje informacije čisto numerički

Naredba netstat (3)

- parametri:
 - -p - statistike po traženom protokolu (datoteka **/etc/protocols**)
 - -r - tablice usmjerivanja (kao naredba **route**)
 - -s - statistike za tablicu usmjerivanja
 - -d - ispisuje broj ispuštenih/odbačenih paketa, nužno je specificirati i uređaj
- izlazne informacije:
 - Iface, MTU, addr, Ipkts, Ierrs, Opkts, Oerrs, col, drop, timer

Naredba sar

- **sar** - sistemski nadzorni softver - SystemV
- prikuplja informacije o
 - korisničkom potrošenom vremenu
 - sistemskom potrošenom vremenu
 - itd.
- parametri:
 - -u - CPU potrošnja
 - -b - aktivnost međuspremnika
 - -d - aktivnost za diskove

Naredba sar (2)

- -y - aktivnost na terminalima
- -c - sistemski pozivi
- -a - korištenje poziva za pristupanje datotekama
- -q - potrošnja sistemskih redova
- -v - status procesa, inodea i datotečnih tablica
- -m - aktivnost IPC resursa
- -p - aktivnost straničenja
- -r - neiskorištene stranice i blokovi diska
- -k - resursi koje nije moguće straničiti
- -A - svi podaci!

Naredba prdaily

- **prdaily** - ispisuje skupljene statistike i omogućava pregled koji procesi koriste najviše sistemskih resursa - SystemV
- nema argumenata
- tek informativne vrijednosti

Naredba ipcs

- **ipcs** - ispisuje podatke o IPC resursima (SystemV shared memory)
 - -a - najkorisnije, sve mogućnosti
 - -b - ispisuje veličine u bajtovima
 - -o, -p, -t, itd. - broj poruka, redova, itd

Različiti alati

- naramak raznih:
 - **monitor** (sprema u datoteku kao sar)
 - **top**
 - **tcpdump**
 - **xload** (histogram)
 - **volstat** (diskovi unutar LSM, nalik iostat)
 - **volwatch** (e-mail u slučaju greške diskova i sl)

Što nismo spomenuli

- AdvFS podešavanja:
 - advscan, showfdmn, showfile, defragment, migrate, chvol, advfsstat, switchvol
- praćenja aplikacija:
 - atom, third, hiprof, pixie, prof, gprof, uprofile, dbx, ladebug, lsof

Dio III: Alati za promjenu jezgre



Početak i kraj

- nužno veliko predznanje; krivi parametri - loš rad, gubitak podataka, problemi!
- podešavanje parametara jezgre:
 - datoteka **/etc/sysconfigtab**
 - naredba **sysconfigdb**
 - naredba **sysconfig** (-s, -q)
- izgradnja jezgre:
 - naredba **doconfig**
 - direktorij **/sys/conf**

Sysconfigtab - IPC resursi

- Informix, Oracle - IPC resursi!
- ipc:
 - shm-max = 2147483647
 - shm-mni = 2048
 - shm-seg = 1024
 - sem-mni = 400
 - sem-msl = 1600
 - sem-opm = 400
 - sem-ume = 400
 - ssm-threshold = 0

Sysconfigtab - procesi i stog

- proc:
 - # dovoljno threadova i procesa
 - max-proc-per-user = 1024
 - # maxfd i sl
 - maxusers = 2048
 - # mnogo threadova za bazu
 - max-threads-per-user = 4096
 - # 8GB za procesni stog
 - per-proc-stack-size = 8589934592
 - max-per-proc-stack-size = 8589934592

Sysconfigtab - procesi i stog

- # 1T
- per-proc-data-size = 1099511627776
- max-per-proc-data-size = 1099511627776
- per-proc-address-space = 1099511627776
- max-per-proc-address-size = 1099511627776

Sysconfigtab - asinkroni rad i pristup disku

- "raw" pristup disku
- rt:
 - aio_max_percent = 3
 - aio_task_max_num = 60108
 - aio_listio_max_num = 256
 - sigqueue_max_num = 256

Sysconfigtab - mreža, socketi

- socket:
 - # TCP connection listen queue
 - somaxconn = 65535
 - sominconn = 65535
 - # 4MB socket buffer
 - sb_max = 4194304

Sysconfigtab - mreža, IP

- inet:
 - # IP input queues
 - ipqs = 16
 - # TCP hash tables
 - tcbhashnum = 16
 - # max TCP segment size
 - tcp_mssdflt = 1460
 - # send and receive spaces
 - tcp_recvspace = 61440
 - tcp_sendspace = 61440

Poboljšanje performansi

- naredba **sys_check**:
 - HTML izlaz
 - sugestije o poboljšanjima - generički!
 - -all - svi podsustavi, sigurnost, itd
 - -perf - samo poboljšanja za performanse
 - -nohtml - čisti tekst

Kraj i diskusija

