



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
ÚSTAV POČÍTAČOVÝCH SYSTÉMŮ

FACULTY OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER SYSTEMS

TESTOVACIA SADA SLUŽIACA NA ANALÝZU TUNED PROFILOV

TESTSUITE FOR ANALYSIS OF PROPERTIES OF TUNED PROFILES

BAKALÁRSKA PRÁCA

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

BRANISLAV BLAŠKOVIČ

VEDÚCI PRÁCE

SUPERVISOR

Ing. ALEŠ SMRČKA, Ph.D.

BRNO 2012

Abstrakt

Výtah (abstrakt) práce v slovenskom jazyku.

Abstract

Výtah (abstrakt) práce v anglickom jazyku.

Klíčová slova

tuned, testovanie, linux, fedora

Keywords

tuned, testing, linux, fedora

Citácia

Branislav Blaškovič: Testovacia sada slúžiaca na analýzu
Tuned profilov, bakalárska práca, Brno, FIT VUT v Brně, 2012

Testovacia sada slúžiaca na analýzu Tuned profilov

Prehlásenie

Prehlasujem, že som túto bakalársku prácu vypracoval sám pod vedením pána...

.....
Branislav Blaškovič
27. apríla 2013

PodĎakovanie

PodĎakovanie.

© Branislav Blaškovič, 2012.

Táto práca vznikla ako školské dielo na Vysokém učení technickém v Brně, Fakultě informačních technologií. Práca je chránená autorským zákonom a jej použitie bez udelenia oprávnenia autorom je nezákonné, s výnimkou zákonom definovaných prípadov.

Obsah

1	Úvod	3
2	Popis komponenty tuned	4
2.0.1	História <i>tuned</i>	4
2.1	Profily	4
2.1.1	Prehľad profilov	5
3	Plán testovania pre Fedora Linux	6
3.1	Úvod	6
3.2	Testovacie položky	6
3.3	Softvérové riziká	6
3.4	Čo sa bude testovať	6
3.5	Čo sa nebude testovať	7
3.6	Kritéria pre splnenie testov	7
3.7	Kritéria na prerušenie testovania	7
3.8	Čo obsahuje plán testovania a jeho výsledky	7
3.9	Remaining Test Tasks	8
3.10	Environmental Needs	8
3.11	Staffing and Training Needs	8
3.12	Responsibilities	8
3.13	Schedule	8
3.14	Planning Risks and Contingencies	8
3.15	Approvals	8
3.16	Glossary	8
4	Testovanie	9
4.1	Príprava systému	9
4.2	Použitie virtuálneho stroja	10
4.2.1	Výhody	10
4.2.2	Nevýhody	10
4.2.3	Zhrnutie	10
4.3	Implementácia testov	10
4.4	Vyhodnotenie testov	11
5	Bočné produkty práce	12
5.1	Problém s <i>libvirt</i>	12
5.2	Power Management Test Day	12
5.3	Nájdené chyby	13

Kapitola 1

Úvod

Každý linuxový server alebo osobný počítač môže slúžiť na niečo iné. Preto je veľmi náročné vytvoriť linuxovú distribúciu, ktorá by pokrývala požiadavky každého a bola optimalizovaná pre všetky operácie. Preto je potrebné systém nastaviť tak, aby presne vyhovoval našim potrebám a získali sme maximálny výkon pre naše potreby. Keďže sa jedná o množstvo druhov nastavení, vznikol balíček *tuned*[8], ktorý ich zahrňuje.

Cieľom tejto práce je priblížiť *tuned*, zhrnúť jeho hlavné funkcie, popísať profily a na záver implementovať sadu testov pre I/O operácie nad najpoužívanejšími súborovými systémami a diskovými zariadeniami. Na záver budú testy spustené a výsledky vyhodnotené.

Kapitola 2

Popis komponenty tuned

Balíček *tuned* je primárne napísaný pre linuxovú distribúciu Fedora[4] a Red Hat Enterprise Linux. Démon *tuned* neustále beží, skenuje systém a upravuje nastavenia podľa potreby. Napríklad najväčšia záťaž na disk je pri štarte systému alebo pri ukladaní dát na disk (napríklad filmov). Inak je disk skoro nečinný. *tuned* dokáže optimalizovať zápis práve v tej dobe, keď je to potreba. Rovnako je to aj pri sieťových operáciach.

Niektoré profily zamedzujú aj prepínaniu Cx stavov¹.

Súčasťou *tuned* je aj program *tuned-adm*, ktorý nám dovoľuje prepínať medzi profilmi. Každý z profilov slúži na iné zameranie a napriamo podľa toho upravuje systém, čím dosahujeme ešte lepšie výsledky.

2.0.1 História *tuned*

Komponenta *tuned* je vyvíjaná od roku 2008. Prvými autormi boli *Philip Knirsh* a *Thomas Woerner*. Dnes sú najväčšími prispievateľmi *Ján Včelák*, *Jaroslav Škarvada* a *Ján Kaluža*. Dnes je *tuned* vo verzii 2. Medzi verziami 1 a 2 je veľký rozdiel, pretože bol celý kód od základov prepísaný. Jedna z najväčších zmien je v používaní *D-BUS*². Ďalšia zmena je v profiloch. Niektoré profily boli zmenené alebo odobraté. Pre zachovanie spätnej kompatibility vznikol preto balík *tuned-profiles-compatible*, ktorý obsahuje všetky profily z verzie *tuned* 1.

2.1 Profily

Profily su hlavne zamerané na CPU, disky, sieť a FSB³. Samotný balíček obsahuje niekoľko predvolených profilov a ako základný profil je po spustení *tuned* profil *balanced*.

Profily si môžeme aj samy vytvárať. Ak si nie sme istý, čo je potrebné upraviť, môžeme využiť odporúčania z programu *powertop*[5] a za pomoci skriptu *powertop2tuned* si nechať profil vytvoriť automaticky na základe výstupu z *powertop*. Bližšiemu popisu profilov sa venuje sekcia 2.1.1.

¹Cx sú stavy, v ktorých sa môže vyskytovať procesor, typicky firmy Intel. Tieto stavy sa volajú Spiacie stavy (ang. Sleep states) [7]. Spiacie stavy procesoru slúžia na šetrenie energie.

²Systém pre jednoduché zasielanie správ a komunikáciu medzi aplikáciami

³Front-Side Bus - datová zbernica, ktorá zaisťuje komunikáciu medzi CPU a hardvérom. Využíva sa v procesoroch *Intel*

2.1.1 Prehľad profilov

Profily *tuned* sa nachádzajú v adresári */usr/lib/tuned*. V tomto adresári sa taktiež nachádza súbor s funkciami, ktoré tieto profily využívajú. Práve tieto súbory sú najviac vyvíjané a menené.

Prehľad profilov zo základného balíčka *tuned*. Zoznam je platný pre verziu **tuned-2.2.2-1** na *Fedora 18*:

- **balanced** - predvolený profil pre väčšinu systémov s výnimkou virtuálnych
- **latency-performance** -
- **powersave** - na zníženie odberu
- **throughput-performance**
- **virtual-guest** - predvolený profil pre virtuálne systémy
- **virtual-host**

Balíček *tuned-profiles-compat* rozširuje zoznam o tieto ďalšie profily:

- **default**
- **desktop-powersave**
- **enterprise-storage**
- **laptop-ac-powersave**
- **laptop-battery-powersave**
- **server-powersave**
- **spindown-disk**

Medzi najčastejšie operácie profilov patrí menenie governoru⁴ procesoru medzi **ondeman**⁵ a **performance**⁶. Viac o tejto vlastnosti je možné dočítať sa na wiki stránkach Arch Linuxu[1].

Ďalej je to nastavovanie plánovačov diskov. Toto nastavenie sa mení v súbore */sys/block/<dev>/queue/scheduler*. Niektoré profily ho prepínajú z predvoleného plánovača na *deadline*⁷. Viac o plánovačoch diskov je možné sa dočítať na stránkach dokumentácie *OpenSuse*[2] alebo v článku [9].

Niektoré profily taktiež vypínajú bariéry⁸ pri pripojovaní diskov.

⁴Rýchlosť procesora je možné meniť a tým šetriť energiu v čase, keď ho naplno nevyužívame. Túto rýchlosť ovplyvňujú rôzne druhy plánovačov.

⁵Nastavuje rýchlosť procesora podľa využitia.

⁶Nastaví rýchlosť procesora na najvyššiu hodnotu bez ohľadu na využitie.

⁷Tento plánovač sa používa pre zníženie latencie. Obsahuje dve fronty a každá požiadavka má deadline

⁸Vypnutím bariér hrozí poškodenie dát pri odpojení napájania diskov, pokiaľ disk nemá záložnú batériu.

Kapitola 3

Plán testovania pre Fedora Linux

3.1 Úvod

Na testovanie *tuned* využijeme pomocnú knižnicu *beakerlib* [3] pre jednoduchšie písanie testov a prehľadnejšiu interpretáciu dosiahnutých výsledkov. Cieľom testovania je zistiť, nakoľko *tuned* ovplyvňuje rýchlosť diskových operácií.

3.2 Testovacie položky

Napísané testy budú overovať správnu funkcionálnosť *tuned* démona a taktiež profilov v zameraní na diskové operácie. Všetky testy budú pripravené pre linuxovú distribúciu Fedora 18[4].

Overí sa rýchlosť zápisu na bežných aj RAID diskoch a s najpoužívanějšími súborovými systémami.

3.3 Softvérové riziká

V prípade zlyhania niektorých testov môže prísť k poškodeniu už pripojených diskov. Preto je vhodné spúšťať sadu testov na virtuálnom stroji. V prípade vydania novej verzie *tuned* alebo inej použitej komponenty je tu riziko, že testy nebudú stabilné a môžu sa správať nepredvídateľne. V tomto prípade ale môžeme uvažovať o nájdení chyby (dokonca regresie) v *tuned*.

3.4 Čo sa bude testovať

Hostiteľský systém bude spúšťať predpripravené obrazy virtualizovaného systému Fedora 18. K virtualizovanému systému sa budú pripájať ďalšie disky. Tieto nové disky budú formátované na najpoužívanějšíe súborové systémy a testované ich rýchlosti pri rôznych profiloch *tuned*.

Na testovacie účely použijeme najnovšiu verziu *tuned* z repozitára.

3.5 Čo sa nebude testovať

Pretože testy bežia na virtualizovanom hardvéri, výsledky môžu byť trochu skreslené. Viac o testovaní na virtualizovanom systéme v kapitole doplniť

3.6 Kritéria pre splnenie testov

Počas testovania so zapnutým démonom *tuned* by všetky I/O operácie diskov mali byť rýchlejšie alebo aspoň tak rýchle ako s vypnutým *tuned*. Žiadna operácia by nemala skončiť s chybou a disky by sa nemali poškodiť. Zápis so zapnutým *tuned* musí mať rovnaké výsledky ako s vypnutým.

3.7 Kritéria na prerušenie testovania

Ak zlyhá operácia pripájania disku k virtuálnemu stroju, ďalšie testovanie stráca význam. Preto je potrebné počas testovanie kontrolovať, či táto kľúčová operácia dopadla úspešne. Rovnaká situácia môže nastať, ak zlyhá nahratie uloženého obrazu disku systému pred testovaním.

Prvý prípad (pripájanie obrazu disku) sa rieši opakovaním testu. Takto dosiahneme vždy požadované výsledky, aj keď operácia zlyhá. Problematiku pripájania diskov je bližšie popísaná v sekcii ??.

3.8 Čo obsahuje plán testovania a jeho výsledky

Celé testovania sa skladá z niekoľkých súčastí:

- Plán testovania.
- Zdrojové kódy jednotlivých testov.
- Knižnica na spracovanie výsledkov.
- Zoznam chýb, ktoré nastali.
- Vyhodnotenie výsledkov.

- 3.9 Remaining Test Tasks**
- 3.10 Environmental Needs**
- 3.11 Staffing and Training Needs**
- 3.12 Responsibilities**
- 3.13 Schedule**
- 3.14 Planning Risks and Contingencies**
- 3.15 Approvals**
- 3.16 Glossary**

Kapitola 4

Testovanie

4.1 Príprava systému

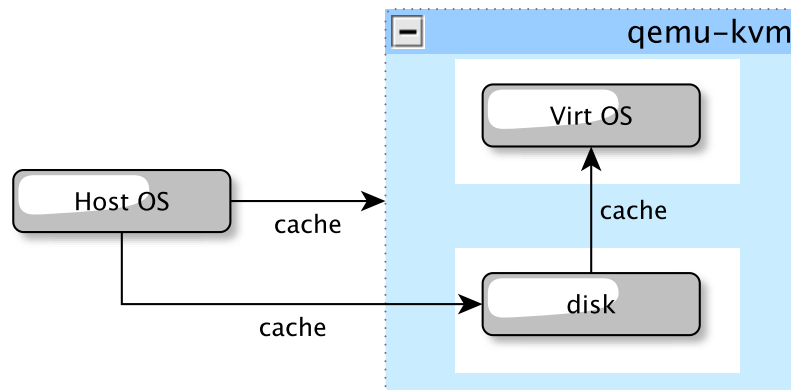
Pred testovaním je potrebné pripraviť si nainštalovaný systém *Fedora 18* ako obraz disku. Tento obraz sa bude spúšťať cez *qemu-kvm*. Disk, na ktorom sa bude testovať rýchlosť zápisu a čítania by mal byť v ideálnom prípade nekešovaný. Pred testovaním aj po testovaní je potrebné hostiteľský aj virtualizovaný systém synchronizovať a zmazať nakešované stránky (Algoritmus 4.1).

Algoritmus 1 Synchronizácia systému

```
/bin/sync
```

```
echo 3 >/proc/sys/vm/drop_caches
```

Každé miesto, kde je možnosť, že by systém si uchovával nejaké nakešované data, ktoré by mohli ovplyvniť výsledky testovania je potrebné poznať (Obrázok 4.1).



Obrázok 4.1: Vyznačenie miest, kde môže nastať kešovanie

Virtualizovaný systém musí obsahovať všetky potrebné balíčky, ktorých programy sa používajú v testoch. Ďalej by mali byť vypnuté všetky nepotrebné služby na pozadí (napríklad *abrt* alebo *ntp*). Treba si dať pozor aj na zoznam *cron*¹ úloh, ktoré môžeme nájsť

¹Časovo závislý plánovač úloh. Každý užívateľ si môže naplánovať vykonávanie príkazov v daných časových intervaloch.

v adresároch `/etc/cron.*`.

4.2 Použitie virtuálneho stroja

Testovanie diskových operácií prebieha vo virtuálnom stroji za použitia *qemu-kvm* pod *libvirt*. Tento spôsob testovania som zvolil pre minimalizáciu problémov, ktoré môžu nastať pri testovaní (popísané v sekcii 3.3) a ochránenie hostiteľského operačného systému.

4.2.1 Výhody

Medzi hlavné výhody testovania na virtualizovanom systéme patrí:

- Jednoduchšia správa obrazov celého systému.
- Pohodlné pripájanie a odpájanie diskov.
- Dokonalé prispôbenie systému pre požiadavky testov - na hostiteľskom systéme moc nezáleží a dá sa využiť skoro akákoľvek Linuxová distribúcia.

4.2.2 Nevýhody

Toto testovanie prináša ale aj množstvo nevýhod.

- Hostiteľský systém môže obmedzovať virtualizovaný systém.
- Medzi systémami môže nastať kešovanie (ako je zobrazené aj na obrázku 4.1). Toto sa ale dá do veľkej miery eliminovať a nemalo by mať vplyv na výsledky testov.
- Virtualizovaný systém a taktiež aj virtuálne disky budú vždy odlišné od fyzických. Aj keď virtualizácia beží na úrovni jadra, stále su v systémoch malé rozdiely. Preto je celé testovanie experimentálne a výsledky sa budú zbierať a priemerovať z viacerých iterácií testovania.

4.2.3 Zhrnutie

Aj napriek všetkým nevýhodám je testovanie na virtualizovanom systéme vhodné. Výsledky testov sú porovnávacie (so zapnutým a vypnutým *tuned*) - to znamená, že aj keď by v skutočnosti časy mohli byť odlišné, rozdiel medzi *tuned* a bez *tuned* verziou ostáva rovnaký.

Taktiež treba poznamenať, že je dnes virtualizácia využívaná veľmi často a preto virtualizácia pri testovaní nie je prekážkou.

4.3 Implementácia testov

Na implementáciu testov je využitý prevažne jazyk *bash*. Testovanie riadi hlavný súbor *test_disk.sh*, v ktorom sa nastavujú parametre testov. Tieto testy su popísane v tabuľke 4.1. Konkrétne podtesty začínajú prefixom *inc_**. Virtuálny systém sa pred každým testom vypne, obnoví zo zálohy disku a znova zapne.

Názov premennej	Popis
TEST_COMMAND	Sada príkazov pre diskové operácie
FS	Asociatívne pole, obsahujúce názov súborového systému a príkazu na jeho vytvorenie
TO_TEST	Zoznam testov, ktoré sa prevedú

Tabuľka 4.1: Popis premenných pre parametrizáciu testov.

Disky sa pripájajú podľa toho, ako to vyžaduje aktuálny test. Kešovanie diskov medzi hostiteľským a virtualizovaným systémom je vypnuté na miestach, ktoré su uvedené na obrázku 4.1.

4.4 Vyhodnotenie testov

Test	bez tuned	s tuned	rozdiel	rozdiel [%]
simple.disk virtio ext3	112s	104s	8s	8%
raid0 virtio ext3	120s	111s	9s	8%
raid1 virtio ext3	131s	119s	12s	10%
Priemery	121.00s	111.33s	9.66s	8.66%

Tabuľka 4.2: Výsledky testov pre súborový systém ext3

Test	bez tuned	s tuned	rozdiel	rozdiel [%]
simple.disk virtio ext4	116s	101s	15s	13%
raid0 virtio ext4	121s	107s	14s	12%
raid1 virtio ext4	121s	109s	12s	10%
Priemery	119.33s	105.66s	13.66s	11.66%

Tabuľka 4.3: Výsledky testov pre súborový systém ext4

Kapitola 5

Bočné produkty práce

5.1 Problém s *libvirt*

Pri pripájaní diskov k virtuálnemu systému za použitia príkazu `virsh attach-device`, sa systém občas náhodne vypne. Nepodarilo sa mi spoľahlivo zreprodukovať túto chybu do takého stavu, aby som ju mohol nahlásiť vývojárom, pretože nastávala nepravidelne.

O páde systému som nenašiel žiadnu zmienku v logoch virtualizovaného ani hostovacieho systému.

5.2 Power Management Test Day

Dňa 17. apríla 2013 sa konal deň otvorených dverí Brnenskej pobočky firmy *Red Hat*. Súčasťou programu bol aj *Power Management Test Day* najnovšej verzie operačného systému *Fedora 19*, ktorého som bol spoluorganizátor.

Mimo iné sa testovalo aj *tuned* samotné. Najväčším prekvapením bolo testovanie *tuned* profilu *powersave*, ktorý by mal znižovať elektrický odber. Na testovanie prišlo množstvo ľudí s rôznymi notebookmi. Konkrétne výsledky sú uvedené v tabuľke 5.1.

Typ notebooku	Odber bez tuned	Odber s tuned	Zlepšenie
Lenovo ThinkPad X230	4.540 Wh	4.110 Wh	9.47%
HP Elitebook 8540w	7.566 Wh	7.493 Wh	0.97%
Lenovo T60 laptop	4.720 Wh	4.500 Wh	4.66%
Dell Optiplex 990	8.500 Wh	7.700 Wh	9.41%
Lenovo ThinkPad T61	4.880 Wh	4.340 Wh	11.07%
Lenovo ThinkPad T520	5.387 Wh	4.815 Wh	10.62%
Lenovo T510	3.310 Wh	2.820 Wh	14.80%
ThinkPad T430	5.690 Wh	5.370 Wh	5.62%
Samsung N210 (Intel Atom)	1.402 Wh	1.394 Wh	0.58%
Priemerné zlepšenie	7.47%		

Tabuľka 5.1: Odber rôznych notebookov s vypnutým a zapnutým tuned.

5.3 Nájdene chyby

Počas písania tejto bakalárskej práce, narábanie s *tuned* a testovaní som narazil na niekoľko chýb, ktoré som nahlasoval do oficiálnej *Red Hat* a *Fedora* bugzilly [6]. Chyby sa nachádzali väčšinou priamo v balíku *tuned*, ale taktiež v balíku *selinux-policy*.

Tieto chyby boli väčšinou opravené do pár dní. Ich prehľad je v tabuľke 5.2.

Číslo chyby	Názov	Status
BZ#907065	tuned-adm can crash because of unknown variable log	opravené
BZ#953128	tuned does not start if active profile is missing	opravené
BZ#953132	AVCs after starting tuned daemon	opravené
BZ#901689	Typo in tuned-adm.py	opravené
BZ#911138	tuned-adm profile virtual-guest/host AVC denied	zatiaľ neopravené

Tabuľka 5.2: Zoznam nájdených chýb

Kapitola 6

Záver

Z výsledkov je jasné, že *tuned* naozaj optimalizuje I/O operácie na diskových poliach.

Literatúra

- [1] *CPU Frequency Scaling*, [online]. 2013 [cit. 2013-04-20].
URL https://wiki.archlinux.org/index.php/CPU_Frequency_Scaling
- [2] *Tuning I/O Performance*, [online]. 2013 [cit. 2013-04-20].
URL http://doc.opensuse.org/products/draft/SLES/SLES-tuning_sd_draft/cha.tuning.io.html
- [3] BeakerLib - domovská stránka. <https://fedorahosted.org/beakerlib>.
- [4] Fedora - domovská stránka. <http://fedoraproject.org>.
- [5] PowerTOP - domovská stránka. <https://01.org/powertop>.
- [6] Red Hat bugzilla. <https://bugzilla.redhat.com>.
- [7] Spiace stavy - sleep states.
<http://www.intel.com/support/processors/sb/CS-028739.htm>.
- [8] Tuned - domovská stránka. <https://fedorahosted.org/tuned>.
- [9] Stanovich, M.; Baker, T.; Wang, A.-I.: Throttling On-Disk Schedulers to Meet Soft-Real-Time Requirements. In *Real-Time and Embedded Technology and Applications Symposium, 2008. RTAS '08. IEEE*, 2008, ISSN 1545-3421, s. 331–341, doi:10.1109/RTAS.2008.30.