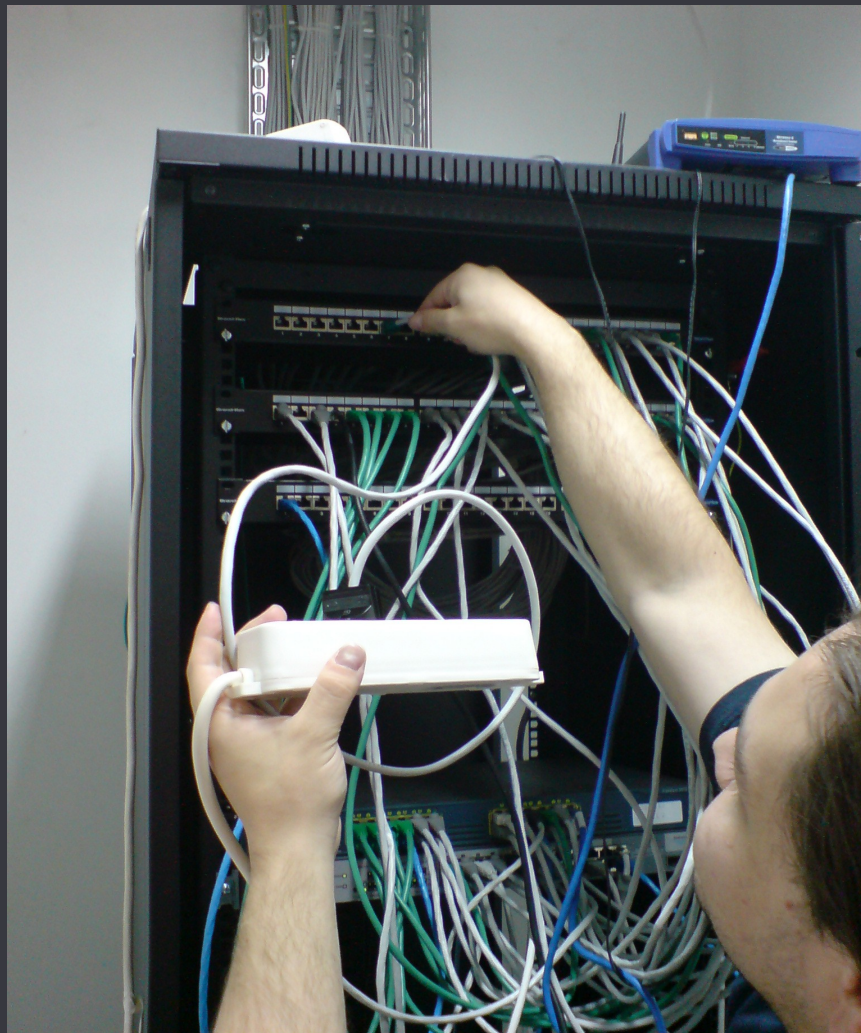


Защо (и как) поддържането на 700 машини е по-лесно от поддържането на 3

Васил Колев
vasil@ludost.net



Увод

- Играл съм си с много и малко сървъри
- Многого се оказаха по-лесни
- Аз съм ужасно мързелив...
- ... и не много умен
- Намирам го за лесно, значи трябва да е така за всички :)





Малко бележки

- Три типа операции
 - Малка фирма
 - Голяма, „Enterprise“
 - Голяма, „Porn company“
- Ще говоря за третия тип, другите не са интересни
- Защо тая лекция има смисъл при всичките тия облаци

Научени неща

- Всички големи системи ни принуждават да научим някои неща
- Изброил съм част от тях
- Те правят работата почти тривиална
- Някои ги има във всички книги, други не

Monitoring

- Двата типа хора
 - Тези, на които им е гърмял диск
 - Тези, на които ще им гръмне
- Виждате какво се случва
- Пази от downtime
- Дава данни за бъдещо планиране
- Полезно за всякаква система
 - Аз ползвам nagios за личните си машини, вкл. работната станция въщи.

на иб6

ко секунди от web

d LVM <http://ur1.ca/a>

utes ago от identi.ca

единени щати

et.soup.io/asset/384

ago от web

olev

ПРИЯТНО

ПОПУЛЯРНИ БЕЛЕЖКИ



<http://9.asset.soup.io>

Обект	Услуга	Положение	Продължителност	Информация
hfl (31.10.2012 (ср) 18,43,19 EET)				
u14	FTPP	critical	2м 27с	WARNING FTP conn 3
u12	FTPP	critical	2м 40с	WARNING FTP conn 3
s672	3WARE	warning	47м 56с	/c0/u0 is REBUILDING 8% /c0/u0/p3 is DEGRADED
s135	3WARE	warning	1x 17м 57с	/c0/u0 is REBUILDING 21% /c0/u0/p6 is DEGRADED
s264	3WARE	warning	1x 27м 15с	/c0/u0 is REBUILDING 19% /c0/u0/p6 is DEGRADED
s628	3WARE	warning	1x 37м 39с	/c0/u0 is REBUILDING 23% /c0/u0/p9 is DEGRADED
s333	3WARE	critical	1x 38м 11с	/c0/u0 is DEGRADED
s542	3WARE	critical	2x 22м 29с	/c0/u0 is DEGRADED /c0/u0/p3 is DEVICE-ERROR
s576-admin	-	down	1d 9x 13м 46с	CRITICAL - Host Unreachable (199.7.177.80)
s538	3WARE	warning	2d 22x 52м 48с	/c0/u0/p7 is DEVICE-ERROR

hfl 1

hfl 5

hfl 4

AS8262



Примерни monitoring-и

- <http://pnag.ludost.net/>
- <http://marla.ludost.net/munin/>
- <http://nagios.wikimedia.org/>

Single system description

- Дава възможност за мързелуване
- Всички конфигурации се генерират от централна база
- По-малко грешки
- Не особено използваемо в малки системи

```
#!/bin/sh
```

```
cmd="mysql -uUSER -pPASS -h HOST -r -s DBNAME"
```

```
$cmd -e "select concat('host ',s.name,'adm {\n\tdefault-lease-time 6000;\n\tmax-lease-time 72000;\n\thardware ethernet ',lower(s.admmac),';\n\tfixed-address ',(select ip from serveripaddr where hostname=concat(s.name,'-mgmt.domain.com') limit 1 ),';\n}\n') from server s where not s.admmac is null and switchid in (select id from switch where vlan=4)" > /tmp/mgmtmod-42.conf
```

```
if ! diff -q /tmp/mgmtmod-42.conf /etc/dhcp/mgmtmod-42.conf > /dev/null ; then
```

```
    cp /etc/dhcp/mgmtmod-42.conf /etc/dhcp/mgmtmod-42.conf.old
```

```
    cp /tmp/mgmtmod-42.conf /etc/dhcp/mgmtmod-42.conf
```

```
    if ! /usr/sbin/dhcpd -t /dev/null 2>/dev/null ; then
```

```
        /usr/sbin/dhcpd -t 2>&1 | mail -s "error in dhcpd configuration" vasil@ludost.net
```

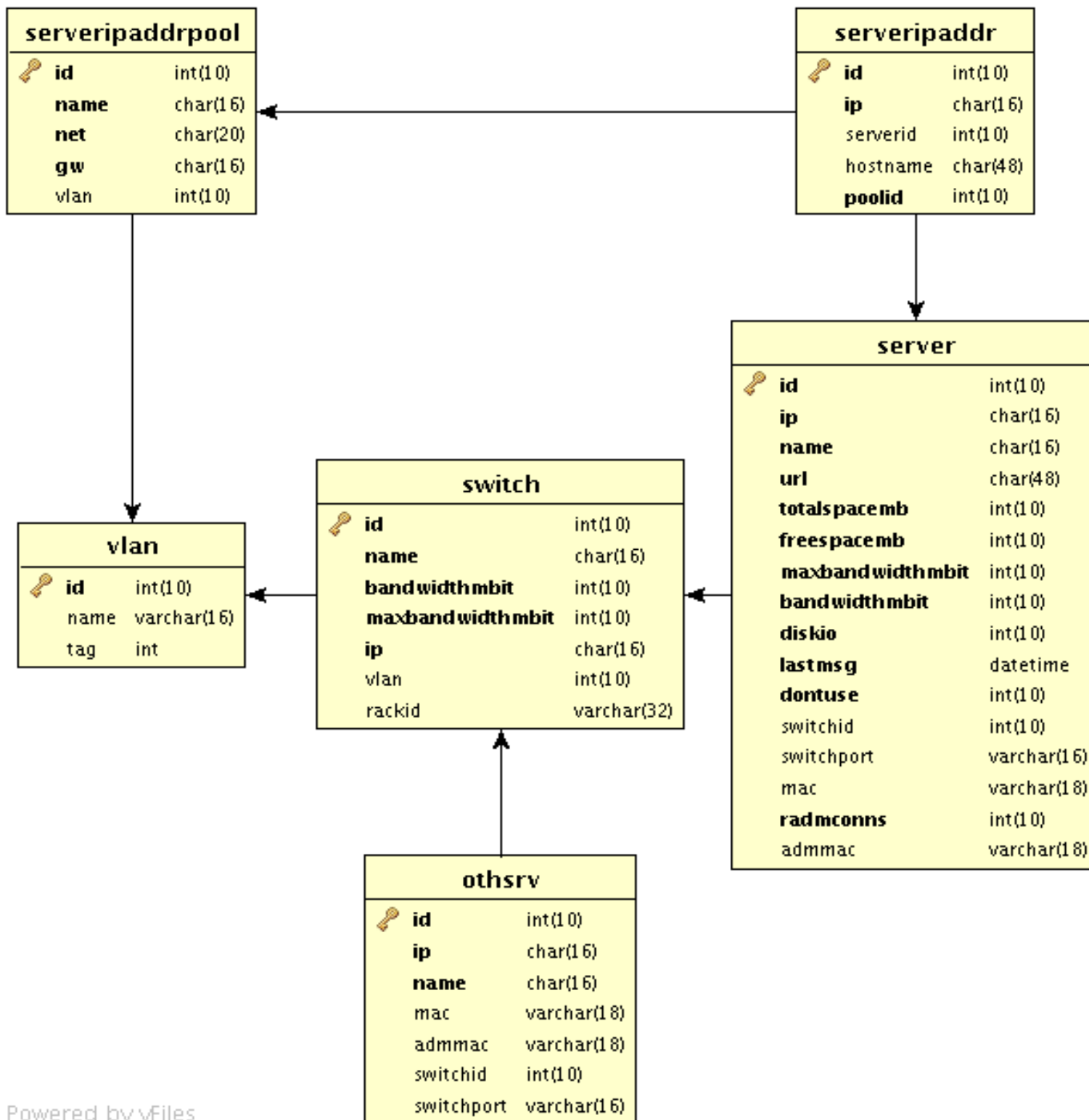
```
        cp /etc/dhcp/mgmtmod-42.conf.old /etc/dhcp/mgmtmod-42.conf
```

```
    else
```

```
        /etc/init.d/dhcpd restart >/dev/null 2>/dev/null
```

```
    fi
```

```
fi
```

(автоматични) Инсталации

- Началната инсталация е мястото, където може да направите най-много грешки
- Не трябва да има нужда от ръчна намеса
- Това е съвсем безсмислено в малки системи
 - Ако в малка система това ви се случва често, имате друг проблем
- При мен машините се инсталират съвсем сами, без никаква моя намеса.

бележка: автоматизация

- Системната администрация е основно автоматизация
- Админите сме мързеливи
 - ТРЯБВА да бъдем
- Всичко, което се прави често трябва да се автоматизира
- Научете се да пишете код
 - Не е задължително да е добре
- Един пример:

```
#!/bin/bash
echo -e 'rack/cage (id)\t\tserver\tport\tserial'
for i in `cat /var/nagios/status.dat |tr '\n' '|' | sed 's/}/\n/g'|grep '3WARE.*current_state=[12]' |
grep -vE 'REBUILDING|No passive checks' |sed 's/^.*host_name=(.*)|.*/\1/'|cut -d '|' -f 1`;
do

    rackid=`mysql -r -s -uUSER -pPASS -h HOST DB -e "select rackid from switch where
id=(select switchid from server where name='$i')"`
    echo -ne "$rackid\t$i\t"
    ssh root@$i.domain.com tw_cli /c0 show |grep '^p' |grep -v OK|awk '{print $1"\t"$7}'
done
```


Програмист, който тренира за админ



Naming conventions

- Религиозна тема
- Трябва да е разширяема
- Имената трябва да са лесни за произнасяне и разбиране през шумна връзка
- Моятa: `[a-z]\d+`

Резервни пътища

- Всички сме си рязали клона, на който седим
- `ifconfig eth0 down`
 - Неее, имах в предвид `eth1` ...
- Примери:
 - IPMI
 - резервни връзки
 - мрежов и локален boot

Мрежа

- Контрол
- Възможност за планиране и управление
- Възможност да знаете какво става
- ASNs, BGP, други трибуквени съкращения

Simplicity

- Трябва да можете да решавате проблеми в 4 сутринта
 - Все още пияни от предната вечер
- Това значи:
 - Да имате цялата картинка в главата си
 - Нещата да са максимално унифицирани
 - "Although there are no textbooks on simplicity, simple systems work and complex don't." (Jim Gray)

Проблеми в големи системи

- Разпределение на натоварвания и събития
- Премахване на SPOF-ове
- Стъпката от една машина към две е по-трудна от две до хиляда, но си струва

Всякакви редки неща са нещо нормално

Press <Alt 3> to access 3ware BIOS Manager

3ware SATA RAID Controller: 9650SE-12ML

BIOS: BE9X 4.00.00.001 Firmware: FE9X 4.00.00.006 Slot: 2

BBU Status: OK

Number of online units: 0, available drives: 0, hot spares: 0, offline units: 0

Incomplete Drives and Others:

SATA - SAMSUNG HE103UJ	931.51 GB	Phy	0	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	1	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	2	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	3	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	4	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	5	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	6	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	7	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	8	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	9	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	10	Drive error
SATA - SAMSUNG HE103UJ	931.51 GB	Phy	11	Drive error

Интересни неща

- Възможност за пълно дистрибутиране на натоварването
- Система за следене на трафика

Какво си струва да запомните

- Monitoring
- Правете нещата така, че да може да ги оправите в 4 сутринта
- Автоматизирайте
- Имайте резервни варианти

Въпроси?



Благодаря за вниманието!