

Xen na Západočeské univerzitě

Michal Švamberg

21. října 2006

- 1 Vlastnosti
 - Hledáme vhodné řešení
 - Výkonost

- 2 Xen na univerzitě
 - Migrace
 - Nastavení
 - Xen naživo

Hledáme vhodné řešení

Proč použít virtualizaci?

Výhody použití virtualizace:

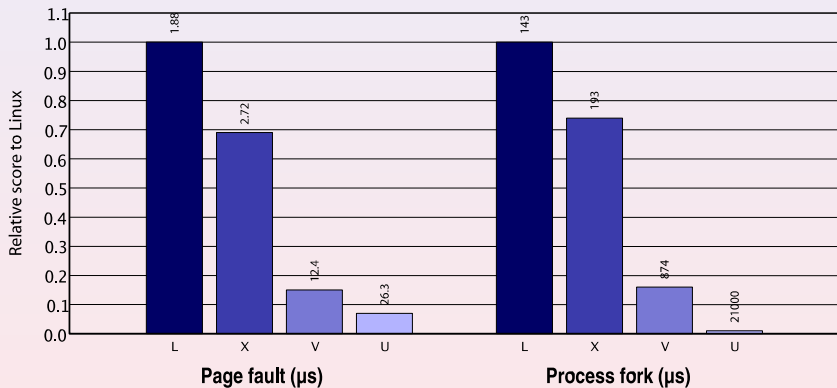
- snadno dostupný stroj
- snapshoty
- duplikace
- šetří náklady i prostor

Proč jsme vybrali Xen

- podpora Linuxu
- nízká režie virtualizace
- snadný management (CLI)
- lze navázat na existující prostředí
 - souborový distribuovaný systém AFS
 - instalační metoda FAI
 - síťová infrastruktura na VLAN (802.1q)
 - FibreChannel (FC) jako diskový subsystém
- cena

Výkonost

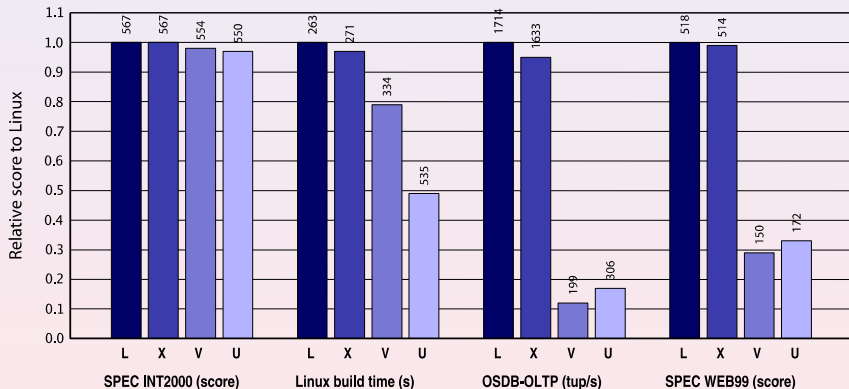
Paměť



Imbench results on Linux (L), Xen (X), VMWare Workstation (V), and UML (U)

zdroj: xen.sf.net

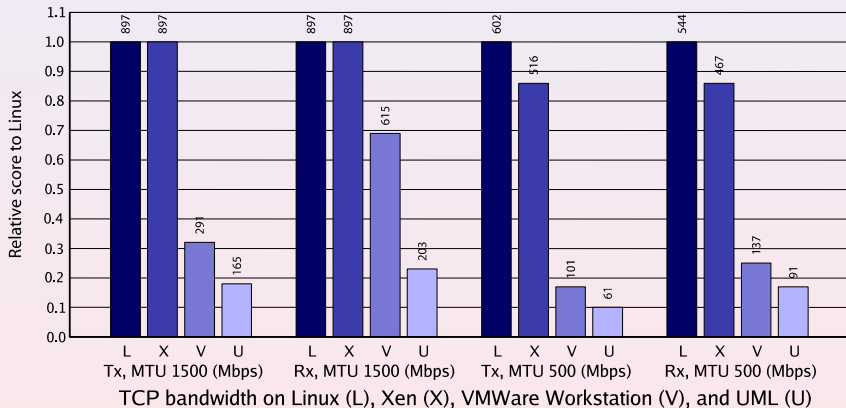
Výkonostní testy



Benchmark suite running on Linux (L), Xen (X), VMware Workstation (V), and UML (U)

zdroj: xen.sf.net

TCP testy



zdroj: xen.sf.net

Xen na univerzitě – Migrace

Migrace

Migrace je přesun virtuálního stroje (hosta) mezi jednotlivými hardwarovými nody (hostitely).

Migrací virtuálních strojů získáme:

- vyšší dostupnost při údržbě
- možnost vyrovnávání zátěže nodů

Migraci rozdělujeme na:

- off-line: rychlejší, stroj je ale suspendovaný
- on-line (live): pomalejší, stroj je v provozu

Prerekvizity pro migraci

- zachování síťového segmentu
- virtuální stroj umístěn na síťovém datovém prostoru:
 - NAS: NFS, CIFS
 - SAN: Fibre Channel
 - iSCSI, síťové blokové zařízení
 - drdb

Jak migrace probíhá

Migrace z virtuálního stroje X z nodu A na nod B :

pre-migration aktivace X na nodu A , výběr cíle na nodu B

reservation inicializace kontejneru pro X na nodu B

pre-copy cyklické kopírování "špinavých" paměťových stránek

stop-and-copy poslední fáze kopírování:

- pozastavení virtuálního stroje X na nodu A
- přesměrování síťového provozu
- synchronizace zůstávajícího stavu

commitment aktivace virtuálního stroje na nodu B , uvolnění
virtuálního stroje na nodu A

Jak migrace probíhá

Migrace z virtuálního stroje X z nodu A na nod B :

pre-migration aktivace X na nodu A , výběr cíle na nodu B

reservation inicializace kontejneru pro X na nodu B

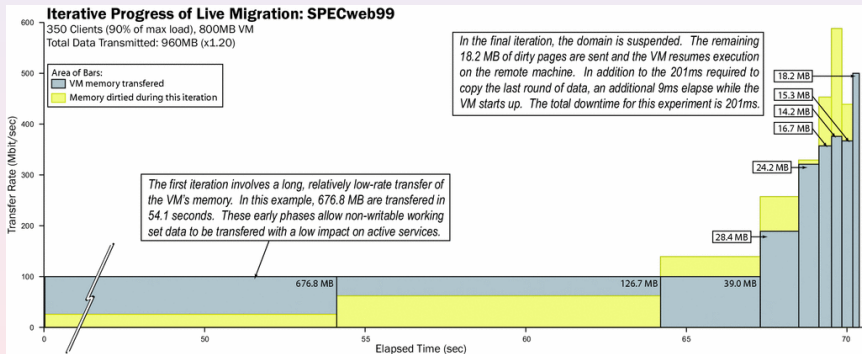
pre-copy cyklické kopírování "špinavých" paměťových stránek

stop-and-copy poslední fáze kopírování:

- pozastavení virtuálního stroje X na nodu A
- přesměrování síťového provozu
- synchronizace zůstávajícího stavu

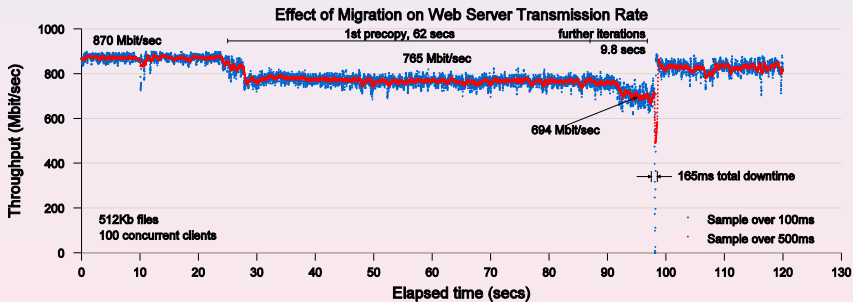
commitment aktivace virtuálního stroje na nodu B , uvolnění virtuálního stroje na nodu A

Fáze pre-copy



zdroj: xen.sf.net

Migrace webového serveru



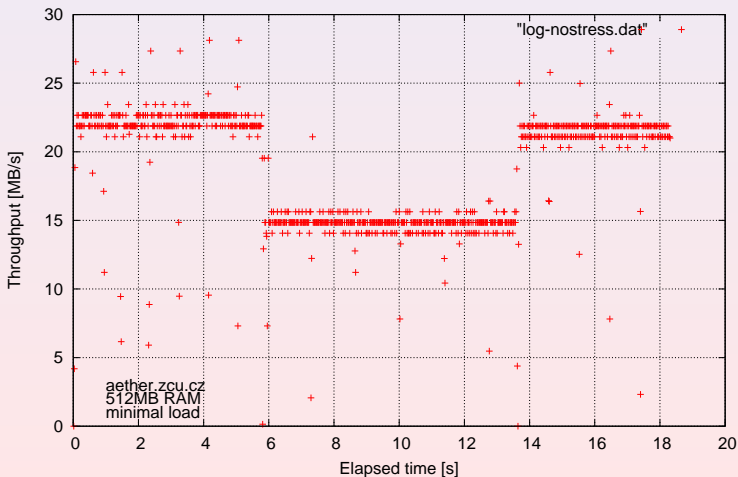
zdroj: xen.sf.net

Z vlastní zahrádky

- Konfigurace
 - migrovaný stroj o velikosti 512MB (*aether.zcu.cz*)
 - migrace probíhala ze stroje *xen2* na stroj *xen3*
 - měřicí stroj (*phoebe.zcu.cz*) umístěn na *xen3*
 - nutno nastavit forward delay: `brctl setfd br53 0`
- Měření
 - z *aether* se tahá dokola jeden soubor
 - z přírustku za 10ms se spočítá rychlost
 - použít Perl s `Time::HiRes` a `LWP::Parallel::UserAgent`
 - pro zatížení *aether* použít `stress --vm 3`

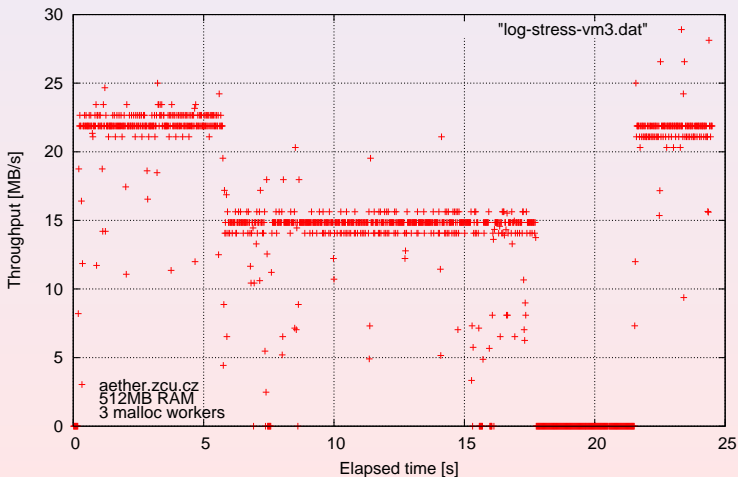
Migrace nezatížené *aether*

Live migration - transmission rate progress
aether.zcu.cz with minimal load



Migrace *aether* se zátěží

Live migration - transmission rate progress
aether.zcu.cz heavily loaded



Nastavení

Xen na ZČU

- 3x stroj hostující Xen (verze 2.0.7)
 - 2x DualCore Xeon na 3GHz
 - 4GB RAM, 2x Gbit ethernet, 2x 80GB SATA v SW RAIDu
 - FibreChannel karta
- celkem 35 virtálních strojů (9 v ostrém provozu)
 - 64-800MB RAM
 - 10-20GB HDD (včetně 2GB swap souboru) na FC
 - 1x virtualní CPU
 - 1x virtualní ethernet připojen na bridge s VLANou
- správa virtuálních strojů
 - instalace FAI
 - konfigurace na AFS

Nastavení sítě

- VLANy (správcovská je nativní jako záloha)
- bridge propojují 802.1q rozhraní s virtuálními interfacemi (vif)

Nastavení sítě – /etc/network/interfaces

```
auto br52
```

```
auto br4
```

```
iface br52 inet static
    address 147.228.52.25
    netmask 255.255.255.0
    broadcast 147.228.52.255
    gateway 147.228.52.1
    bridge_ports eth0.52
```

```
iface br4 inet loopback
    bridge_ports eth0.4
    up ip address del 127.0.0.1 dev br4
```

```
iface br53 inet loopback
    bridge_ports eth0.53
    up ip address del 127.0.0.1 dev br53
```

Napojení na bridge – brctl show

bridge name	bridge id	STP enabled	interfaces
br101	8000.003048683896	no	eth0.101
br4	8000.003048683896	no	eth0.4
br52	8000.003048683896	no	eth0.52 vif2.0 vif3.0 vif5.0 vif6.0 vif9.0
br53	8000.003048683896	no	eth0.53 vif7.0
br54	8000.003048683896	no	eth0.54
...

Diskový subsystém

Virtuální stroje:

- virtuální stroj má vyhrazenou partitionu na FC (RAID 5)
- swap do souboru z důvodu migrace
- partition na FC 10 nebo 20GB
- celkem 3 disky z FC po 200GB (každá 15 partitions)

Domain-0:

- 2x80GB SATA disk
- mirrorované partitions (mdadm)

Konfigurace na AFS

AFS (Andrew File System) nám poskytuje několik výhod:

- je třeba releasovat při změnách
- RW volume pro editace, přípravu a testování
- RO repliky na 4 serverech v různých lokalitách
- jedna konfigurace dostupná na všech Domain-0

<http://www.openafs.org/>

Instalace metodou FAI

FAI (Fully Automatic Installation) pro rychlou instalaci:

- instalační metoda pro Debian
- konfigurace uložena na AFS
- používáme vlastní mirror Debianu
- pro použití Xenu vlastní rozdělování disků (žádné)

`http://www.informatik.uni-koeln.de/fai/`

Konfigurace Xenu

boot:

```
kernel = "/boot/xen-linux-2.6.10-xenuboot"  
memory = 128  
name = "czosug"  
vif = [ 'mac=aa:bb:cc:05:30:61, bridge=br53' ]  
disk = [ 'phy:/dev/fc/fc-p3d1p8,hda1,rw' ]  
root = "/dev/hda1 ro"
```

install:

```
root = "/dev/nfs"  
extra = "ip=147.228.53.61:147.228.53.154:147.228.53.1:255.255.255.0:\nczosug:eth0:off FAI_LOCATION=147.228.53.154:/usr/local/share/fai.dev \nFAI_ACTION=install FLAGS=verbose,debug"
```

Další plány

- přechod na novou řadu 3.x.x
- uplatnit v praxi více migraci
- vyzkoušet Xen na 64bit architektuře
- vyzkoušet Xen s HW podporou virtualizace

Ukázka teorie v praxi

Jdeme na to

aneb jak se žije Xenu v Plzni.