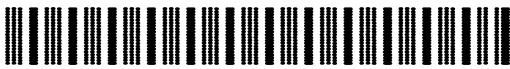


CLASSIFICATION: pro interní použití		IDENTIFICATION NUMBER:  0 0 0 0 0 0 0 0 0 0 0 0 0 2	
TITLE: Instalace GIS ve virtualizovaném prostředí		AUTHOR:	Jan Forman
		TYPE OF DOCUMENT:	Popis instalace
		VERSION:	1
		DATE:	25.9.2011
Public URL: http://janforman.org/files/instalace.pdf		ORIGINAL DIGITAL LOCATION: network attached storage	
OFFICE NAME AND ADDRESS: ---			
DISTRIBUTION STATEMENT: modifications restricted (c) / modifikace nejsou dovoleny, copyright			
KEY WORDS: kvm openvz proxmox arcgis mysql nginx apache			
ABSTRACT: Popis krok za krokem instalace hypervisoru a komponent pro poskytování webových mapových služeb. Dokument již není aktuální!			
PLACE OF ORIGIN: Plzeň, Česká Republika			



Instalace komponent pro platformu GIS

jedná se o backend poskytující služby / funkcionality SW klientům
dostupné přes otevřené rozhraní. jeho virtualizování, cluster a cloud computing

Obsah

Úvod - filosofie systému.....	4
Instalace hypervizoru.....	5
Stavební prvky (plně autonomní).....	6
Propojení jednotlivých stavebních prvků.....	7
Konfigurace zálohování.....	10
Migrace strojů v rámci clusteru.....	10
Instalace ArcSDE (D).....	11
Instalace Mapové komponenty (B).....	12
Instalace ArcGIS Enterprise (A).....	24
Instalace ORACLE DB (E).....	26
Poznámky k návrhu.....	28

Úvod - filosofie systému

Rozložení systému na jednotlivé komponenty z důvodu jednoduchosti a bezpečnosti. Každá komponenta je plně soběstačná a vyžaduje minimálně jednu IP adresu v síti.

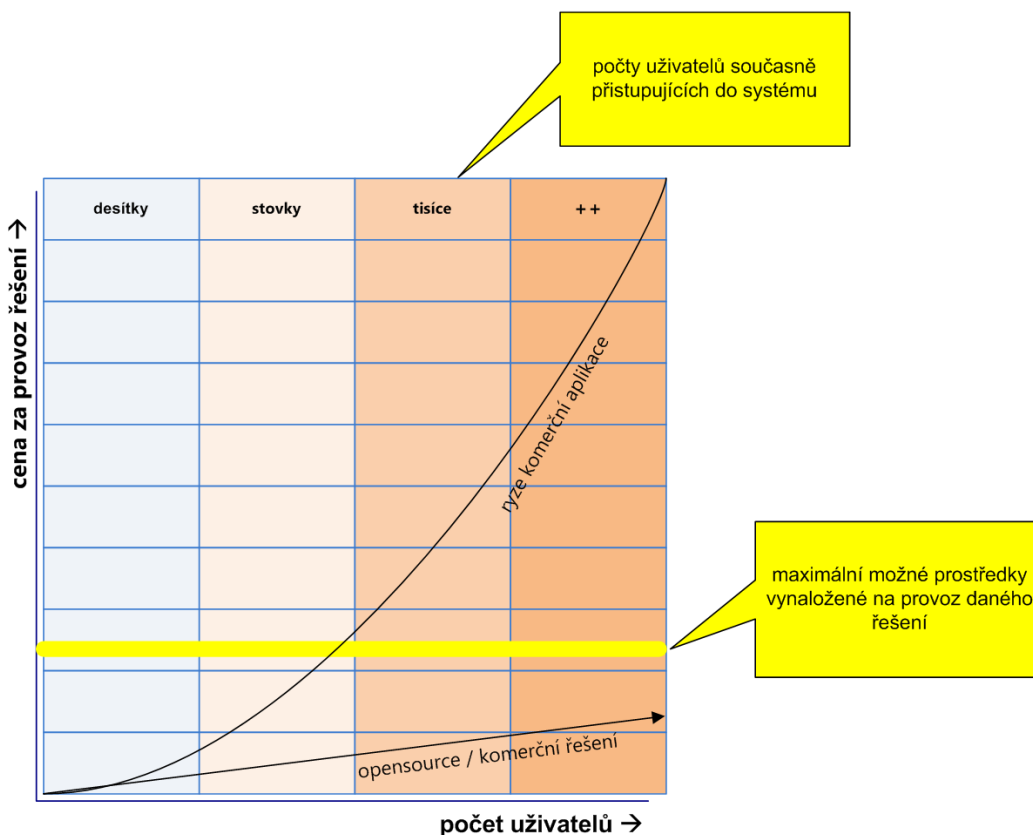
Systém je postaven na SOA platformě s REST komunikací mezi prvky která zajišťuje abstraktní přístup k jednotlivým prvkům. Struktura dat pro komunikaci je ve formátech XML a JSON. Zároveň zde je využito horizontální škálování výkonu, replikace, je podporován SNMP monitoring a SNAP-shot zálohy s použitím funkcionalit LVM.

Komerční komponenty jsou separovány a jejich využití je z velké části dobrovolné. Využití opensource komponent je možné pouze za podmínek že nijak záporně neovlivní cílové řešení (nejde primárně o úsporu finančních prostředků, ale o maximální výkon za co nejnížší cenu a možnost jednotlivé komponenty zaměňovat mezi sebou dle potřeb a aktuálních finančních prostředků).

V maximální míře je vyžadována kompatibilita na úrovni rozhraní (abstrakce + API) tak, aby bylo možné jednotlivé komponenty zaměňovat a to i za chodu v případě, že se dodavatel komponenty začne jevit jako neperspektivní.

Do budoucna se počítá s připojením dalších speciálních funkcionalit s modelováním v reálném čase přes tato rozhraní => cokoliv uzavřeného znemožňuje další vývoj a v systému všeobecně nemá co dělat.

Uživatelský SOFTWARE = nejdražší a nejhodnotnější komponenta



Instalace hypervizoru

První krok – instalace hypervisoru:

Pro podrobný popis prosím navštivte poslední část dokumentu „Poznámky k návrhu“.

Stáhněte si hypervisor a vypalte ho na CD

(instalace je možná i z USB klíče, nicméně není uživatelsky přívětivá).

Vyberte si přijatelný hardware (můžete použít téměř jakýkoliv HW systém) procesor ale musí podporovat **64bit rozšíření** a **bez podpory HW virtualizace nelze použít KVM**.

Plně podporované jsou následující CPU

AMD – téměř všechny 64bit CPU (vyjma nejstarších pro socket 939), **Opterony druhé generace**
INTEL od Core i3 všechny – u přechozích verzí to byla „nadstandartní funkce“ za příplatek

Balík je jednoúčelově orientovaná distribuce obsahující fúzi KVM a OpenVZ technologie

Po instalaci se připojte na <https://{IPSERVERU}>, použijte jméno root a heslo jaké jste zadali v průběhu instalace.

Doporučený počet serverů jsou 3 minimální počet 1

(další servery lze připojovat za provozu přemigrováním, rovněž je možné je i zahazovat bez ovlivnění provozu nebo s minimálními výpadky – řádově sekundy až minuty)

Na následujících dvou stranách je popis jednotlivých komponent a jejich provázání mezi sebou.
Obojí je volné a lze to kdykoliv měnit →

Dostupné autonomní komponenty (předpřipravené stavební prvky)

každá komponenta je plně soběstačná a nevyžaduje přítomnost jiné komponenty (vyjma hypervizoru)



otevřené dokumentované rozhraní

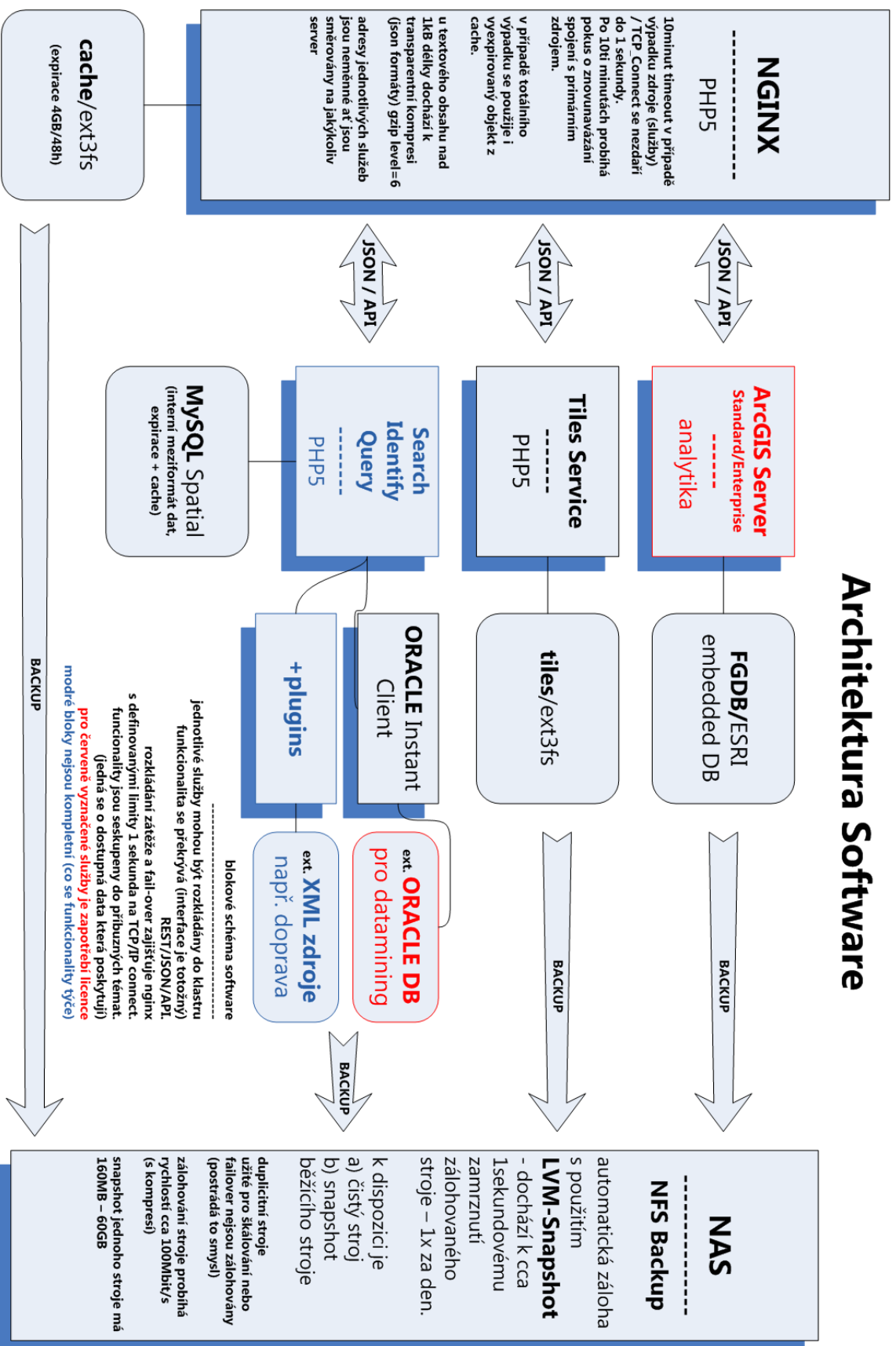
uzavřené proprietární rozhraní

* typ stroje – KVM (běžný virtualizovaný obraz s virtuálními ovladači), OpenVZ (speciálně upravený kontejner s přímým přístupem na síť a disk)

* červené prvky vyžadují odpovídající licenci pro provoz (pokud jsou aktivované)

* k zeleně označeným prvkům jsou dostupné zdrojové kódy

Architektura Software



Druhý krok - konfigurace vzdáleného úložiště:

pro jednoduchou zálohu a obnovení virtuálních strojů

You are logged in as 'root'

proxmox

Home | Logout

Proxmox Virtual Environment 1.8

www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Storage

Add NFS Share

Storage Name: backup

Server: 10.160.150.99

Export: /mnt/storage/backup

Enable: ☒

Content: VZDump Backups

scan

save

Doporučuji připojit NFS storage pro zálohu který bude nasdílený zároveň jako SMB/CIFS tak, aby byl dostupný i z produktů Microsoft Windows.

Možná je i varianta připojení vzdáleného NAS přes fibre-channel (mapování do lokálního fs) nebo přes iSCSI.

Configuration → Storage → červená šipka → Add NFS Share

pro zjednodušení pojmenujte sdílený disk jako backup (Storage Name)

Server (vlozte IP serveru s běžící NFS službou) a klikněte na scan (v nabídce se objeví sdílené adresáře)

Content → VZDump Backups (na toto místo se bude ukládat záloha běžících strojů snapshotem) a uložte konfiguraci

You are logged in as 'root'

proxmox

Home | Logout

Proxmox Virtual Environment 1.8

www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Storage

Storage List

Add iSCSI Target

Add NFS Share

Add LVM Group

Add Directory

ssd

Type	Enabled	Active	Shared	Used (GB)	Capacity (GB)	
Directory	Yes	Yes	No	191.11	813.05	23.51%
NFS	Yes	Yes	Yes	70.10	1105.63	6.34%
Directory	Yes	Yes	No	110.29	214.89	51.32%

Verze 1.3 – © 25.9.2011 Jan Forman – dokument nesmí být použit ke komerčním účelům bez souhlasu

8/24

Pokud instalujete na více fyzických serverů vytvářejte virtuální stroje pouze na primárním serveru a další servery připojte do clusteru.

Je důležité aby jednotlivé virtuální servery měli unikátní ID a tudíž je šlo jednoduchým způsobem migrovat mezi fyzickým hardwarem. Nezávislé instance se hůře migrují a nedochází k synchronizaci. viz obrázek – dva servery jednotné číslování, jednotná správa →

You are logged in as 'root'

proxmox

Home | Logout

Proxmox Virtual Environment 1.8

www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

ListCreateMigrate

Running Maintenance Tasks

Nothing to view

Cluster Node 'gis-i7-1'Online

VMID	Status	Name	Uptime	Disk	Memory	CPU
101	running	arcgis1	9d	10.00GB	3.17GB	0.00%
102	running	tiles1	14d	0KB	9MB	0.00%
103	running	search	14d	0KB	274MB	0.00%

Cluster Node 'gis-i7-2'Online

VMID	Status	Name	Uptime	Disk	Memory	CPU
104	running	tiles2	14d	0KB	9MB	0.00%
105	running	arcgis2	10d	10.00GB	3.17GB	0.00%

Oživení stroje:

Systém podporuje následující možnosti importu

- 1) **OpenVZ obraz** (stroj má přímý přístup k HW, obsahuje jednu vrstvu - Linux)
- 2) **KVM snapshot** (plná virtualizace, emulace HW, passthru - Windows)

--- méně používané varianty

- 3) **OpenVZ snapshot** (záloha stroje i s daty)
- 4) **ISO obraz – standartní instalace** (jsou dostupné Windows XP s VirtIO ovladači a většinou aktualizací do poloviny roku 2011) bez omezení síťových připojení, velikost cca. 300MB

Vysvětlivky:

{VMID} => číslo přiřazené virtuálnímu počítači, systém začíná číslovat od 101
{NÁZEV_SLUŽBY} => název REST služby / alfanumerické znaky bez mezer
{IPSERVERU} => ip adresa serveru

Verze 1.3 – © 25.9.2011 Jan Forman – dokument nesmí být použit ke komerčním účelům bez souhlasu

9/24

Konfigurace zálohování

You are logged in as 'root'

Proxmox Virtual Environment 1.8

Home | Logout

www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Backup

Edit backup job

Cluster Node: gis-i7-2 (10.160.150.37) Send email to:

Storage: backup (nfs) Include all VMs: ☐

Day of Week: Mon Tue Wed Thu Fri Sat Sun

Start time: 02 : 00

Mode: snapshot

Compress files: snapshot

VMID	Status	VM Name	VM Type	Disk size
☐ 104	running	tiles2._____	openvz	200.00 GB
☐ 105	running	arcgis2._____	kvm	10.00 GB

➔ save

Vyberte fyzický stroj, storage ponechte na backup (nfs) zaškrtněte požadované dny. Následně označte virtuální stroje určené pro zálohu. **Nevytvářejte zálohu duplicitních strojů.**

Migrace strojů v rámci clusteru

You are logged in as 'root'

Proxmox Virtual Environment 1.8

Home | Logout

www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

List Create Migrate

Define migration task

Source Node: gis-i7-1 (10.160.150.36) VMID: VM 101 (arcgis1._____)

Target Node: gis-i7-2 (10.160.150.37) Online migration

➔ migrate

VMID	Name	Status
101	arcgis1._____	running
102	tiles1._____	running
103	search._____	running

Zvolte zdrojový a cílový fyzický server, následně si vyberte virtuální stroj. Online migrace v některých případech nemusí být úspěšná. Tudiž **v případě potřeby stroj normálně zastavte.**

Proces migrace je časově závislý na rychlosti sítě. Live migrace strojů v OpenVZ kontejneru je možná i bez externího NAS úložiště, nicméně vyžaduje jádro 2.6.18. KVM pro live migraci vyžaduje storage na externím NAS poli nebo replikovaný souborový systém. V případě použití externího pole připojeného přes fibre-channel nebo iSCSI je migrace v podstatě okamžitá – bez výpadku běžících spojení. Stroj se „zmrazí“ na jednom hardware a následně „rozmrazí“ na druhém - doba je závislá na přesunutí obsahu RAM.

Instalace ArcSDE (D)

Vložte OpenVZ obraz přes webové rozhraní:

Appliance Templates → Upload file → vyberte obraz a stiskněte na upload.

Název obrazu je:

centos-4-arcsde_0-0_i386.tar.gz

(obraz se automaticky zreplicuje do podřízených fyzických serverů)

Vytvořte stroj přes webové rozhraní

Virtual machines → create

Type: Container (OpenVZ)

Template: centos-4-arcsde_0-0_i386.tar.gz

Hostname: zvolte hostname dle uvážení

Memory: 1000MB (stroj konzumuje 70-256MB dle konfigurace postgresql)

Swap: 0MB

Start at boot: zaškrtnout

Disk space: 128GB nebo více

Network type: virtual network (venet)

Zvolte IP adresu a názvy DNS serverů

klikněte na create

(dojde k vytvoření stroje z obrazu)

Proxmox Virtual Environment 1.8

Home | Logout

www.proxmox.com

You are logged in as 'root'

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

List Create Migrate

Configuration

Type: Container (OpenVZ) VMID: 107

Template: centos-4-arcsde_0-0_i386.t Cluster Node: gis-i7-1 (10.160.150.36)

Hostname: arcsde Start at boot: ☒

Memory (MB): 3000 Disk space (GB): 100

Swap (MB): 0

Password: Password Confirm Password:

Network

Network Type: Virtual Network (venet) DNS Domain:

IP Address: 10.160.150.222 First DNS Server: 10.160.150.41

Second DNS Server: 0.0.0.0

create

klikněte na záložku list a vybalením nabídky na červené šipce můžete virtuální stroj spustit.

ArcSDE běží a výchozí parametry jsou →

Jméno: sde **Heslo:** <bude sděleno ústně>

Port: 5151

IP adresa: zadali jste při vytváření virtuálního stroje.

V systému jsou účty: sde (správa sde) a postgres (správa DB)

Instalace je kompletní.

Instalace Mapové komponenty (B)

Vložte OpenVZ obraz přes webové rozhraní:

Appliance Templates → Upload file → vyberte obraz a stiskněte na upload.

Název obrazu je:

gentoo-2011-tiles_0-0_i386.tar.gz

(obraz se automaticky zreplicuje do podřízených fyzických serverů)

Vytvořte stroj přes webové rozhraní

Virtual machines → create

Type: Container (OpenVZ)

Template: gentoo-2011-tiles_0-0_i386.tar.gz

Hostname: zvolte hostname dle uvážení

Memory: 512MB (stroj konzumuje 9-256MB dle počtu připojených klientů typicky 9MB)

Swap: 0MB

Start at boot: zaškrtnout

Disk space: 128GB nebo více

Network type: virtual network (venet)

Zvolte IP adresu a názvy DNS serverů

klikněte na create

(dojde k vytvoření stroje z obrazu)

You are logged in as 'root'

PROXMOX

Home | Logout Proxmox Virtual Environment 1.8 www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

List Create Migrate

ConfigurationType: Container (OpenVZ) VMID: 106Template: gentoo-2011-tiles_0-0_i386 Cluster Node: gis-i7-1 (10.160.150.36)Hostname: tiles3 Start at boot: ☒Memory (MB): 512 Disk space (GB): 128Swap (MB): 0Password:Confirm Password:

NetworkNetwork Type: Virtual Network (venet) DNS Domain: IP Address: 10 . 160 . 150 . 128 First DNS Server: 10 . 160 . 150 . 41Second DNS Server: 0 . 0 . 0 . 0

➔ create

klikněte na záložku list
kliknutí na červenou šipku můžete virtuální stroj spustit

You are logged in as 'root'

proxmox

Home | Logout Proxmox Virtual Environment 1.8 www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

List
Create
Migrate

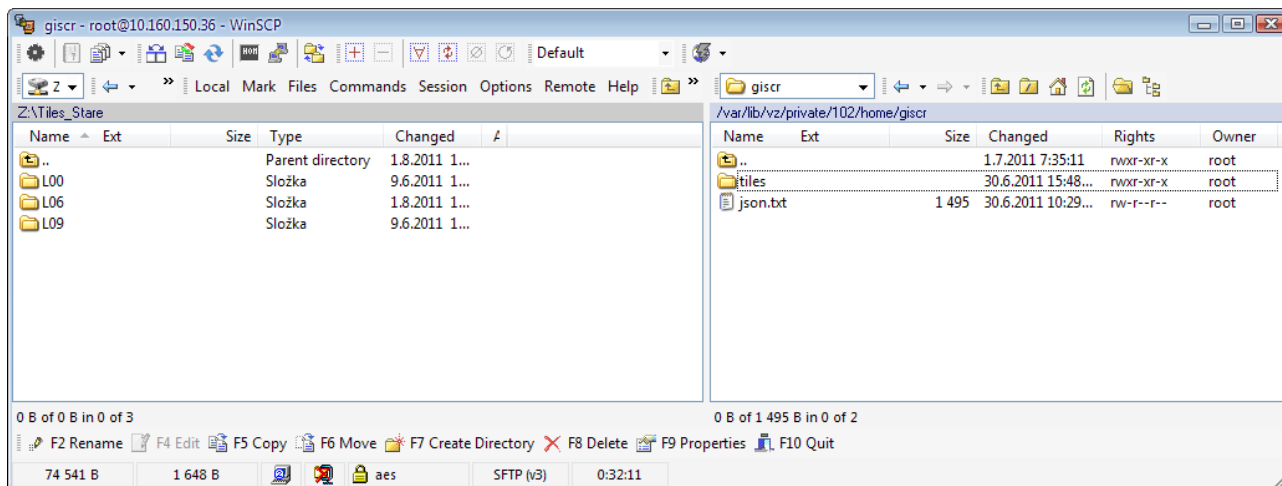
Running Maintenance Tasks

Nothing to view

Cluster Node 'gis-i7-1' Online							
VMID	Status	Name	Uptime	Disk	Memory	CPU	
↓ 101	running	arcgis1	9d	10.00GB	3.17GB	0.00%	
↓ 102	running	tiles1	14d	0KB	9MB	0.00%	
↓ 103	running	search	14d	0KB	274MB	0.00%	

Cluster Node 'gis-i7-2' Online							
VMID	Status	Name	Uptime	Disk	Memory	CPU	
↓ 104	running	tiles2	14d	0KB	9MB	0.00%	
↓ 105	running	arcgis2	10d	10.00GB	3.17GB	0.00%	

Nakopírujte dlaždice do virtuálního stroje + JSON popis služby.
Nejjednodušší bude použít WinSCP a kopírovat data do následujícího umístění



(v obrázku je VMID 102 a název služby gis-cr)

složka /var/lib/vz/private/{VMID}/home

vytvořte adresář (název služby)

dovnitř nakopírujte soubor json.txt (konfigurace služby) a dlaždice do dalšího podadresáře tiles

struktura by tedy měla být následovná - cesta k souboru jako příklad:

/var/lib/vz/private/{VMID}/home/{NÁZEV_SLUŽBY}/json.txt

/var/lib/vz/private/{VMID}/home/{NÁZEV_SLUŽBY}/tiles/L00/R0000007c/C00000074.jpg

..... a další milióny souborů :-)

Ověřte dostupnost REST webové služby z internetového prohlížeče

`http://{IPSERVERU}/arcgis/rest/services/{NÁZEVSLUŽBY}/MapServer?f=json`

Server musí vrátit json popis mapy

`http://{IPSERVERU}/arcgis/rest/services/{NÁZEVSLUŽBY}/MapServer/tile/0/128/121`

Server vrátí dlaždici v měřítku 1:2M

Při pokusu o načtení neexistující služby nebo obsahu server vrátí 404 – nenalezeno.

Pokud dlaždice neexistuje vrací se kód 200 a dlaždice s nápisem „mapa v tomto místě neexistuje“

Instalace je kompletní.

Instalace ArcGIS Enterprise (A)

qmrestore arcgisserver.tar.gz {VMID} (http://{IPSERVERU}:8099/ jméno: arcgismanager heslo: admin)

You are logged in as 'root'

proxmox

Home | Logout Proxmox Virtual Environment 1.8 www.proxmox.com

VM Manager

- Virtual Machines
- Appliance Templates
- ISO Images

Configuration

- System
- Storage
- Backup

Administration

- Server
- Logs
- Cluster

Virtual Machines

List **Create** **Migrate**

Configuration

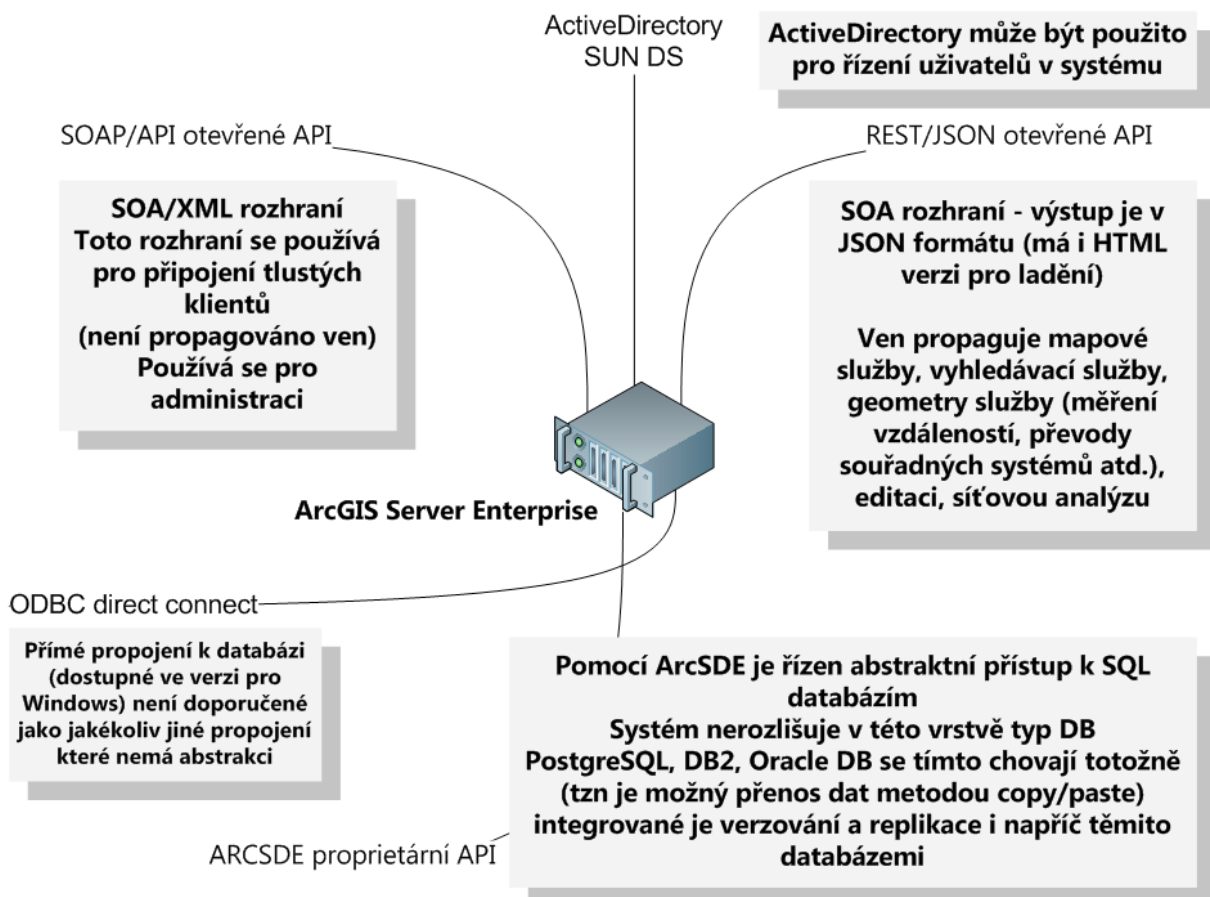
Type:	Fully virtualized (KVM)	VMID:	106
ISO Storage:	local (dir)	Cluster Node:	gis-i7-1 (10.160.150.36)
Installation Media:	WinXPSP3-VE-052011.iso	Start at boot:	☒
Disk Storage:	local (dir)	Image Format:	raw
Disk space (GB):	10	Disk type:	VIRTIO
Name:	kvm	Guest Type:	Windows XP
Memory (MB):	512	CPU Sockets:	1

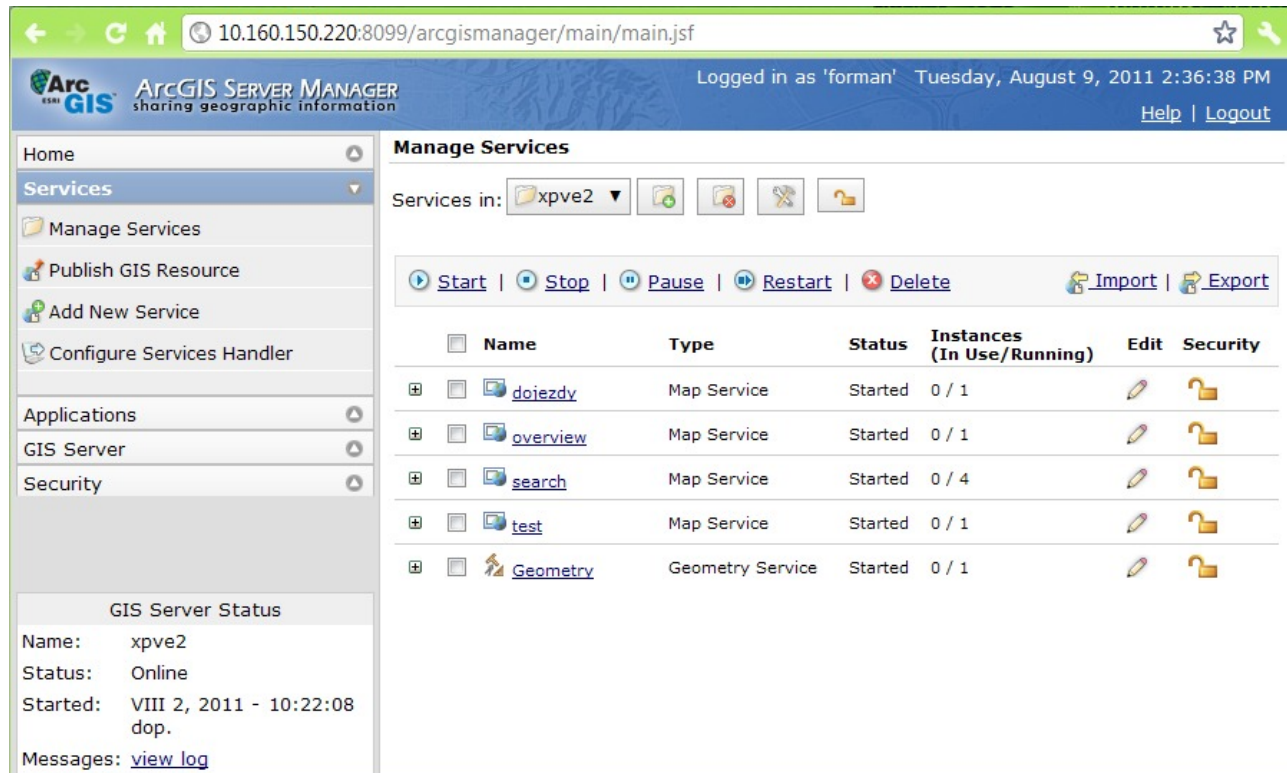
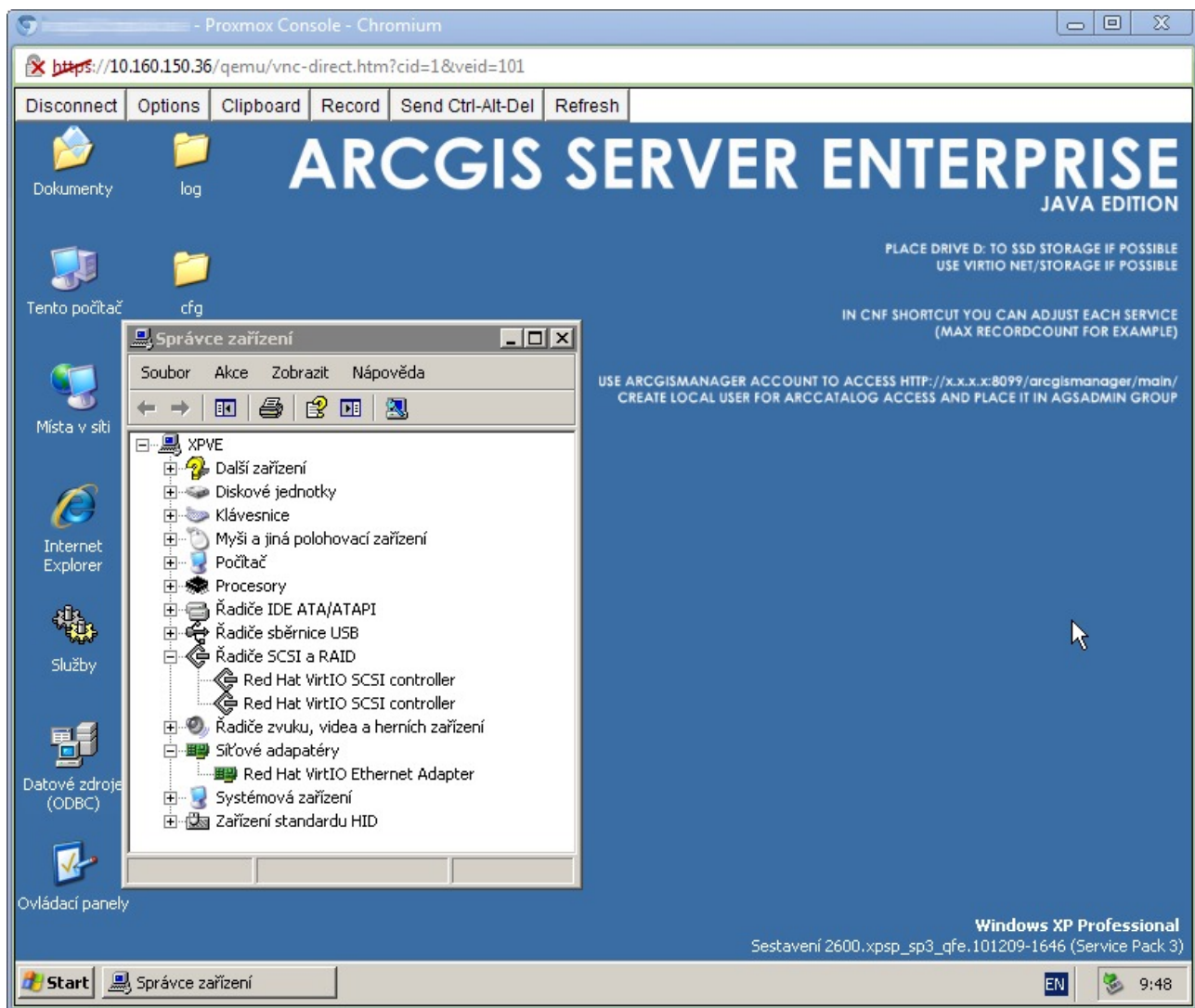
Network

Bridge:	vbr0	Network Card:	virtio
		MAC Address:	BE:34:98:08:65:52

create

Vnější komunikační schéma ArcGIS Serveru Enterprise





Instalace ORACLE DB (E)

Vložte OpenVZ obraz přes webové rozhraní:

Appliance Templates → Upload file → vyberte obraz a stiskněte na upload.

Název obrazu je:

centos-4-oracle10gr2_4.8_i386.tar.gz

(obraz se automaticky zreplicuje do podřízených fyzických serverů)

Vytvořte stroj přes webové rozhraní

Virtual machines → create

Type: Container (OpenVZ)

Template: centos-4-oracle10gr2_4.8_i386.tar.gz

Hostname: ve-oracle

Memory: 2048MB

Swap: 0MB

Start at boot: zaškrtnout

Disk space: 64GB nebo více

Network type: virtual network (venet)

Zvolte IP adresu a názvy DNS serverů

klikněte na create

Zatím stroj nespouštějte!!!

(JRE obsažený v tomto balíku si klade enormní nároky na paměťové struktury a je třeba upravit limit tak, aby nedošlo k jejímu omezení hypervizor by odmítl přidělit potřebnou RAM virtuálnímu stroji)

Přes WinSCP upravte konfiguraci v adresáři:

/etc/vz/conf/{VMID}.conf – najděte položku PRIVVMPAGES a radikálně ji navyšte.

Databáze je předinstalovaná a má SID: **orclve**
Storage má nakonfigurované ukládání dat v **UTF8**,
standartně se **řetězce počítají na znaky** a ne bajty
(což je nelogické výchozí nastavení)

Nyní můžete stroj spustit a přihlásit se přes webové rozhraní na adrese:

<http://{IPSERVERU}:1158/em/>

ORACLE Enterprise Manager 10g Database Control

Setup Preferences Help Logout

Database

Logged in As SYS

Error
java.lang.Exception: Could not open /proc/partitions

Database Instance: orclve

Home Performance Administration Maintenance

Page Refreshed 4.8.2011 3:39:59 Refresh View Data Automatically (60 sec) ▼

General

Shutdown

↑

Status Up
Up Since 6.6.2011 3:41:29 MDT
Instance Name orclve
Version 10.2.0.1.0
Host ve-oracle
Listener LISTENER_ve-oracle

[View All Properties](#)

Host CPU

Load Unavailable Paging Unavailable Maximum CPU 2

Active Sessions

Maximum CPU 2

SQL Response Time

Baseline is empty.

[Reset Baseline](#)

Diagnostic Summary

ADDM Findings	0
All Policy Violations	7
Alert Log	6.6.2011 3:39:52

Space Summary

Database Size (GB)	21.968
Problem Tablespaces	0
Segment Advisor Recommendations	0
Space Violations	0
Dump Area Used (%)	94

High Availability

Instance Recovery Time (sec)	11
Last Backup	n/a
Usable Flash Recovery Area (%)	100
Flashback Logging	Disabled

Paradoxně nedojde téměř k žádnému zpomalení takto nainstalované ORACLE DB - systém má cca 3% režii na abstrakci souborového a síťového subsystému. Abstrakce probíhá pouze v jedné vrstvě.

Toto pravidlo se vztahuje na všechny stroje v OpenVZ

Instalace je kompletní



Poznámky k návrhu infrastruktury **IS**

popis otevřené servisně orientované architektury, odůvodnění volby jednotlivých komponent

OpenVZ

Jedná se o opensource variantu komerčního software Virtuozzo firmy Parallels.

Firma Parallels je známá svými automatizačními nástroji pro správu větších množství serverů, B2C/B2B aplikace a PLESK, jejich programátoři z Ruska, Asie a Ameriky se podíleli i na dalších produktech - např. pro vyhledávání v miliónech stran textu, automatizovaném indexování dokumentů apod. Seznam.cz udává že cca. 50% strojů používá OpenVZ.

<http://www.parallels.com/>

KVM

Kernel-based Virtual Machine (opensource)

Umožňuje plnou virtualizaci pomocí funkcí v novějších CPU VT-X / AMD-V. Jedná se o plnou virtualizaci na rozdíl od OpenVZ. Zrychlení umožňují VirtIO ovladače pro síťovou kartu a disk.

Výhody: lze spouštět stroje s OS-Windows (někdy je to bohužel opravdu nutné)

Nevýhody: plná virtualizace emuluje hardware počítače, je znát výrazné zpomalení obzvláště v oblasti I/O a z tohoto důvodu není vhodná pro databázové systémy a zatížené web-servery.

<http://www.linux-kvm.org>

V novém jádru jsou integrovány obě technologie vč. KSM Kernel Samepage Merging (nižší spotřeba RAM)

Dvě dostupné filosofie ukládání dat

Centralizování datových úložišť (jejich přimapování virtuálním strojům)

x

Distribuovaný souborový systém (LUSTRE, GFS2, HDFS)

Distribuovaný paměťový systém (MEMCACHE NoSQL databáze – vhodné pro přenos real time informací)

Dostupné zdrojové kódy

K většině komponent jsou dostupné zdrojové kódy a kompilační skripty i s některými zdrojovými kódy jsou součástí jednotlivých komponent.

Hardware

Pro provoz je vyžadována virtualizace a nepočítá se s jiným způsobem provozu z důvodů jednoduché správy systému.

Doporučené jsou tři stroje

- a) master
- b) slave1
- c) slave2

Optimální je oddělit jednotlivé komponenty (i když to není nutné). Pro běžný provoz dostačují stroje s 4 jaderným CPU, běžným diskem a jedním SSD diskem. V rackovém provedení lze jeden HW prvek pořídit v hodnotě cca 30tisíc (spotřeba do 100W).

Pro rozsáhlé a náročné analytické úlohy je další varianta použití 24 jaderného systému s procesory Opteron (spotřeba cca 300W na stroj) které pro meziprocessorovou komunikaci používají hypertransport. cena stroje je cca 50 tisíc vč DPH.

(nicméně tak daleko nejsme jednalo by se o sofistikované aplikace jejichž vývoj zatím nenastal)

Je třeba si uvědomit že komerční software má problémy s využíváním více-vláknových aplikací (technicky nebo cenovou politikou – analytický software **většinou použije na jednu úlohu jedno jádro**)

Opensource komponenty lze spouštět na libovolném počtu jader. Nicméně je třeba upravit konfiguraci. (většina aplikací, ale na jednu úlohu taktéž použije jedno jádro)

S využitím rozsáhlých RAID polí na úrovni serverů se nepočítá, už z principu návrhu nejsou nutné - systém zbytečně prodražují a komplikují. **Prvek se bere jako celek, tudíž se počítá s výpadkem celého serveru.** (proč umřel je nepodstatné - podstatné je co převeźme jeho funkcionalitu či jak ho okamžitě obnovit jinde)

Single Point of Failure musí být minimalistický (aktuálně cca 180MB) aby bylo možné v případě výpadku ho okamžitě odstartovat v jiném fyzickém stroji.

Využití externích NAS systémů je vhodné k live-migraci virtuálních serverů ovšem systém **MUSÍ** být schopen provozu i na desktopu pracovníka podatelny nebo sekretářky klidně i notebooku.

To znamená obnovy na zelené louce kdekoli v řádu maximálně hodin ze snapshotu. V případě jakýchkoli problémů např. blesk, výbuch, zamoření, zaplavení či útoku mimozemšťanů. Systém jako takový je odizolován od hardware.

Využití SSD disků u dat bez častých změn je velmi doporučeno! (číselníky, mapové dlaždice, vyhledávací data). Lze zde pozorovat až **10x nárůst** výkonu, pokud data dosahují větších objemů.

Způsob aktualizací

Doporučený způsob je aktualizace **celé komponenty**.

Individuální update je nevyhovující a není doporučován.

Odpovědná osoba aktualizuje balík (celý serverový systém) a prověří funkcionalitu, poté se do jednoho bodu nahraje a provede se záměna za aktuální prvek. Vyčká se definovaný čas pokud je vše v pořádku distribuuje se do celého systému kopie celého virtuálního stroje - **komponenty mají separované úložiště pro data.**

Bezpečnost

Systém je dekomponován na jednotlivé moduly, ty v podstatě běží v sandboxu (v našem případě virtualizace). Pokud dojde ke kompromitování systému útočník získává kontrolu nad komponentou a je schopen modifikovat komunikaci s ní - od ostatních komponent je však izolován do úrovně rozhraní. Jednotlivé komponenty lze oklasifikovat stupněm rizika napadení dle její funkcionality a komplexnosti.

Funkcionalita je v mnoha případech záměrně omezovaná z bezpečnostních důvodů.

V případě vytváření nových modulů se **nepoužívají výchozí konfigurace operačních systémů**, všechny komponenty jsou odladěny zvlášť a minimálně se dbá na doporučení NSA (National Security Agency – USA). K většině komponent jsou dostupné zdrojové kódy potřebné pro případný audit.

Vše je minimalistické a servery nemusí mít aktivovány ani administrátorské účty.

Primárně se očekává že síť je kompromitována útočníkem, nepřipadá v úvahu aby se spoléhalo na firewall. Je nutné změnit výchozí hesla do serverů (která jsou pouze pro případ aktivování virtuálního obrazu serveru). **Se snapshoty serverů je třeba pracovat rozumným způsobem** (v případě jejich oživení někde mimo je možné z nich dešifrovat hesla pomocí dostupných nástrojů podporujících OpenCL a CUDA - dnes lze běžně docílit rychlosti miliónů kombinací za sekundu)

Rizika lze kategorizovat stupňi

a) velmi nízké riziko

neexistuje administrátorský účet, systém nepracuje aktivně se soubory. běží jedna triviální služba

b) střední riziko

běží databáze + skripty, komponenty jsou ověřené a neobjevují se časté záplaty.

c) vysoké riziko

virtualizační hypervizor // běh aplikací v systému **Windows / IIS / JAVA**
(v těchto případech je třeba dbát zvýšené opatrnosti)

NGINX

Reverzní proxy, IMAP proxy, Webserver v jednom. Software původně vytvořený na zakázku poskytovatele webových/emailových služeb v Rusku, následně uvolněný jako opensource. Jeho podíl na trhu stále stoupá a vykazuje velmi vysoký výkon pravděpodobně téměř shodný jako nejrychlejší Web Server na světě firmy ZEUS Technology. Primárně cíleno na rozsáhlé cluster – statisíce až milióny uživatelů současně. Jeho velkým uživatelem jsou Facebook.com, Wordpress.org, Seznam.cz

Nativně podporuje loadbalancing s váhou jednotlivých podřízených serverů a fail-over.

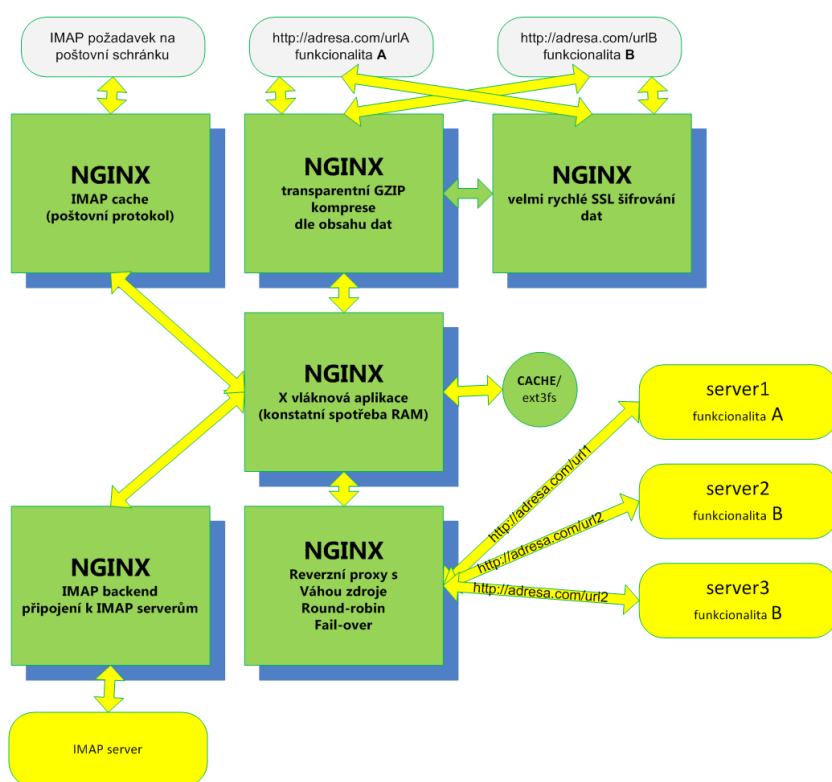
Binárka může mít cca 300kB (64bit O3) konfigurace i instalace je celkem triviální. Platformě téměř nezávislá tudíž bych doporučoval použít Gentoo VPS (bezpečnost, rychlost)

(udává se že běží na 7.3 miliónu serverů a obsluhuje 12.7 miliónu domén – zdroj. Netcraft 2010).

Jeho velká výhoda je velikost (do 1MB v paměti) a architektura – nevyžaduje vlákna.

Podporuje z těch klíčových funkcí např. limitování provozu, transparentní kompresi dat, reverzní proxy jak pro web tak poštu (imap-proxy – využíváno www.mail.ru). Online transformaci XML přes šablonu XSLT.

nginx core functionality (zjednodušené schéma)



server dle URL nebo portu zjistí co má s požadavkem dělat

- A) obslouží lokálně sám
- B) odešle ho do pluginu
- C) předá ho reverzní proxy
- D) načte ho z cache

Zároveň může být provedena nějaká transformace

- A) komprese dat (u textu)
- B) zašifrování SSL

V případě vstupní URL

http://adresa.com/urlA pokud výsledek není v cache dojde k odeslání požadavku modulu reverzní proxy která kontaktuje **server1** a následně se výsledek uloží do cache. Pokud je výsledek v cache použije se odpověď z něj stejně tak pokud cílový stroj **server1** není dostupný.

V případě vstupní URL

http://adresa.com/urlB pokud výsledek není v cache dojde k odeslání požadavku modulu reverzní proxy která kontaktuje server v listu, který je nyní v pořadí a vynechá ten, který aktuálně není dostupný. **server2** nebo **server3** výsledek uloží do cache.

Obsah server2 a server3 by měl být totožný!

- obsluhují stejnou funkcionalitu
- v nejhorším případě může poskytovat starší data

poznámka: podpora IMAP protokolu vychází z původního zadání objednavatele který vyžadoval cache pro svojí poštovní službu s milióny uživatelů. v našem systému je funkcionality vyjmuta. rovněž další desítky rozšíření nejsou obsaženy – mnoho z nich je rovněž určeno pro velmi stresující provoz na úrovni stovek tisíc až miliónů uživatelů současně.

Rozkládání zátěže, detekování vypadlého uplinku, váhy upstreamů a mnoho mnoho dalších funkcí.

<http://www.nginx.com>

CentOS Linux

Redhat linux kompilovaný ze zdrojových kódů RedHat (je téměř 100% kompatibilní)

je nasazen z důvodu používání **binárních uzavřených balíků komerčních dodavatelů** aplikací

typicky ESRI, ORACLE, IBM

Výhody – je zdarma, tudíž jeho licenci ve virtuálním prostředí netřeba řešit.

Gentoo Linux

Distribuce velmi vhodná na vytváření embedded systémů (od nuly)

Výhody

- a) minimalistická – základ cca 160MB s vývojovým prostředím (C, C++, Python, PHP)
- b) všechny aplikace vznikají přímou kompilací proti sobě (100% stabilita)
- c) maximální rychlost (volba optimalizace na jakýkoliv systém)
- d) vysoká bezpečnost (v systému v optimálním případě nic neběží a není skoro nic nainstalováno)
- e) je zdarma

MySQL + Spatial

Dle společnosti Oracle je to velmi rychlá databáze vhodná **pro real-time aplikace** využívaná převážně v telekomunikačním sektoru. Podporuje clustering (grid), vysokou dostupnost (99.999%), hot-backup, sub-second failover, pohledy, trigger, uložené procedury, fyzické rozdělení rozsáhlých tabulek a základní geografické funkce. V placené verzi je navíc grafické rozhraní viz níže.



Oficiálně ji používají následující organizace:

Google, Facebook (části systému migruje na Hadoop a NoSQL)

PostgreSQL

Vysoce sofistikovaná opensource databáze - funkcionality se blíží ORACLE DB ze které se na ní často migruje z důvodu úspory finančních prostředků. (žijí se na tom spousty firem, které si dodělali i vlastní rozšíření tak, aby kompatibilita byla ještě vyšší)

Oficiálně ji používají následující organizace:

Yahoo! (prý petabajty dat), Myspace, SONY, Skype, OpenStreetMap

ORACLE DB

Největší míra nasazení je v bankách. Velké portfolio funkcí v podstatě cokoliv (ovšem ne vždy funkční)

Výhody: clustering/grid na úrovni TNSLISTENERu (od verze 10g), velmi dobré vlastnosti při potřebě vrátit data k určitému časovému úseku. slušná stabilita.

Nevýhody: cena, uzavřené řešení (obtížně migrovatelné), většinu funkcionalit nikdo nevyužije (složitost), z pohledu zpětné kompatibility lze většinu funkcí vyvolat z různých míst nicméně nechovají se shodně.

Memcache

NoSQL distribuované paměťové úložiště v operačních pamětech serverů, jeho kapacita se množstvím připojených strojů sčítá. Běžně tedy jeho kapacita je od desítek GB po TB. V různých programovacích jazycích je řešeno jako plug-in. Jedná se o nejrychlejší způsob jak přistupovat k velkému objemu dat velmi rychle. Popř. jak přenést stavové informace na větší vzdálenost systémově - bez režie způsobené replikací dat a diskových systémů.

Hadoop

Zatím není žádná vize jak využívat plnohodnotné řešení clusteringu s NoSQL databází a automatickým replikováním souborového systému z důvodu omezení aplikací které jsou schopné na něm běžet. Nicméně jeho budoucí potenciál je nezpochybnitelný – omezení jsou zatím ty neexistující aplikace které je třeba doprogramovat. V analytických úlohách bude naprostou nezbytností.