

Zákulisí provozu Seznam.cz

Martin Doleček

Manažer provozního oddělení

21. 10. 2013, MUNI



SEZNAM.CZ

Obsah

- Datacentra
 - aktuální stav
 - vlastní DC
- Hardware
 - servery
 - storage
 - network
- Software
- Aplikace podpory provozu
- Bezpečnost
- Kdo se o provoz stará
- Statistiky
- Na čem pracujeme

Datacentra

- Dvě datacentra = srdce Seznamu
- 2007 active-pasive → 2008 active-active
- 2013
 - 170 racků + rezervace
 - roční spotřeba cca 5 GWh (hodinová 570 kWh)
 - 2.400 fyzických serverů
 - 8.100 operačních systémů
 - 105 TB RAM
 - 17.500 HDD (13 PB)

Datacentra – vlastní DC



HW - servery

- Dual / triad vendor strategie
- Důkladný výběr a testování
 - vlastní benchmark systém
 - syntetické i provozní testování



HW - servery

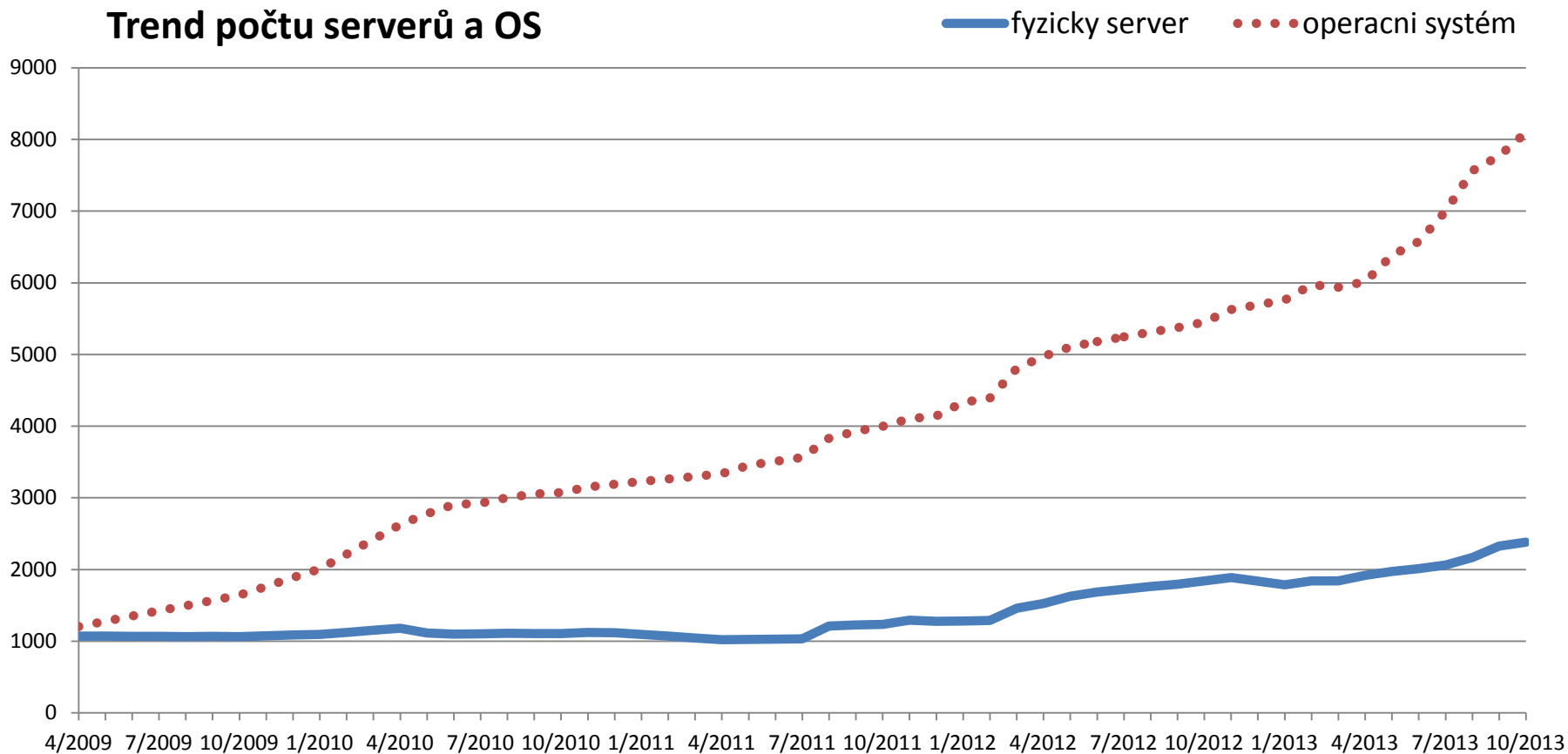
Běžná konfigurace

- CPU: 2x E5-2630L
- RAM: 64-128 GB
- NET: 1-4x 1Gb
- MNG: IPMI, dedikovaný eth port
- DISK:
 - vždy HW RAID + baterka
 - 1U: 8x 140G 15k SAS 2,5“ / 8x 300G 10k SAS 2,5“
 - 2U: 24x 1T SATA 2,5“ / 12x 4T SATA 3,5“
- Redundantní zdroje, větráky



HW - trend

Trend počtu serverů a OS

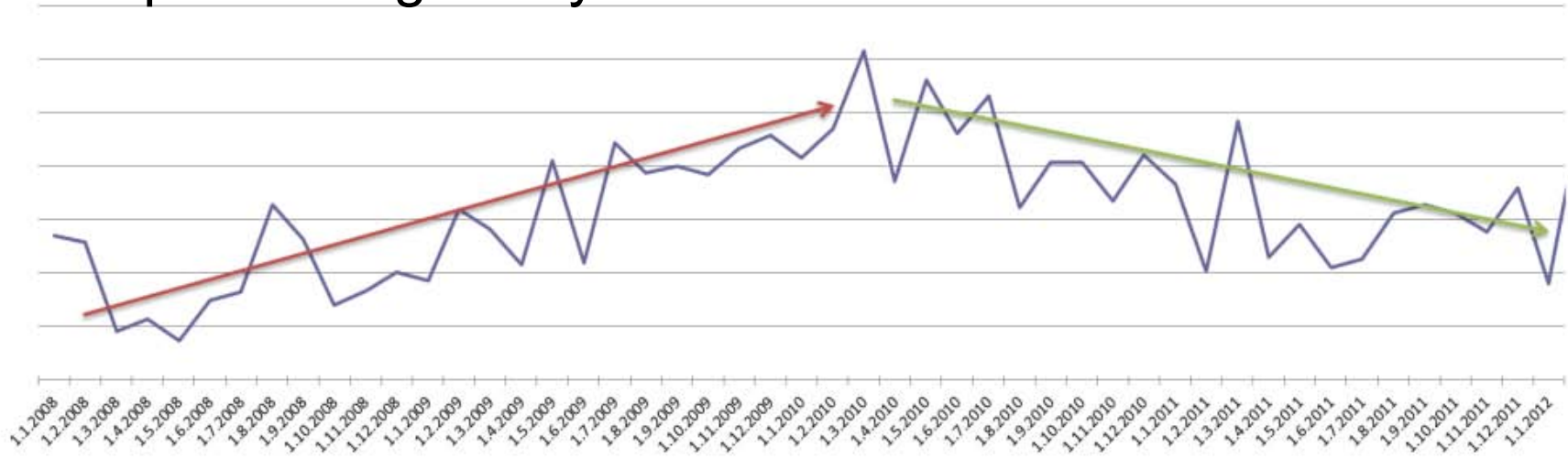


HW – pravidelná obnova

- Po záruce a odpisech server nahradíme novým

rok	konverze obnovy	
	serverů	spotřeby
2010	5:1	7:1
2011	3:1	4:1
2012	3:1	4:1
2013	3:1	3:1

- Úspora energie díky obnově serverů



HW - storage

- Email.cz uložště (cca 6PB, 8000 SATA disků)

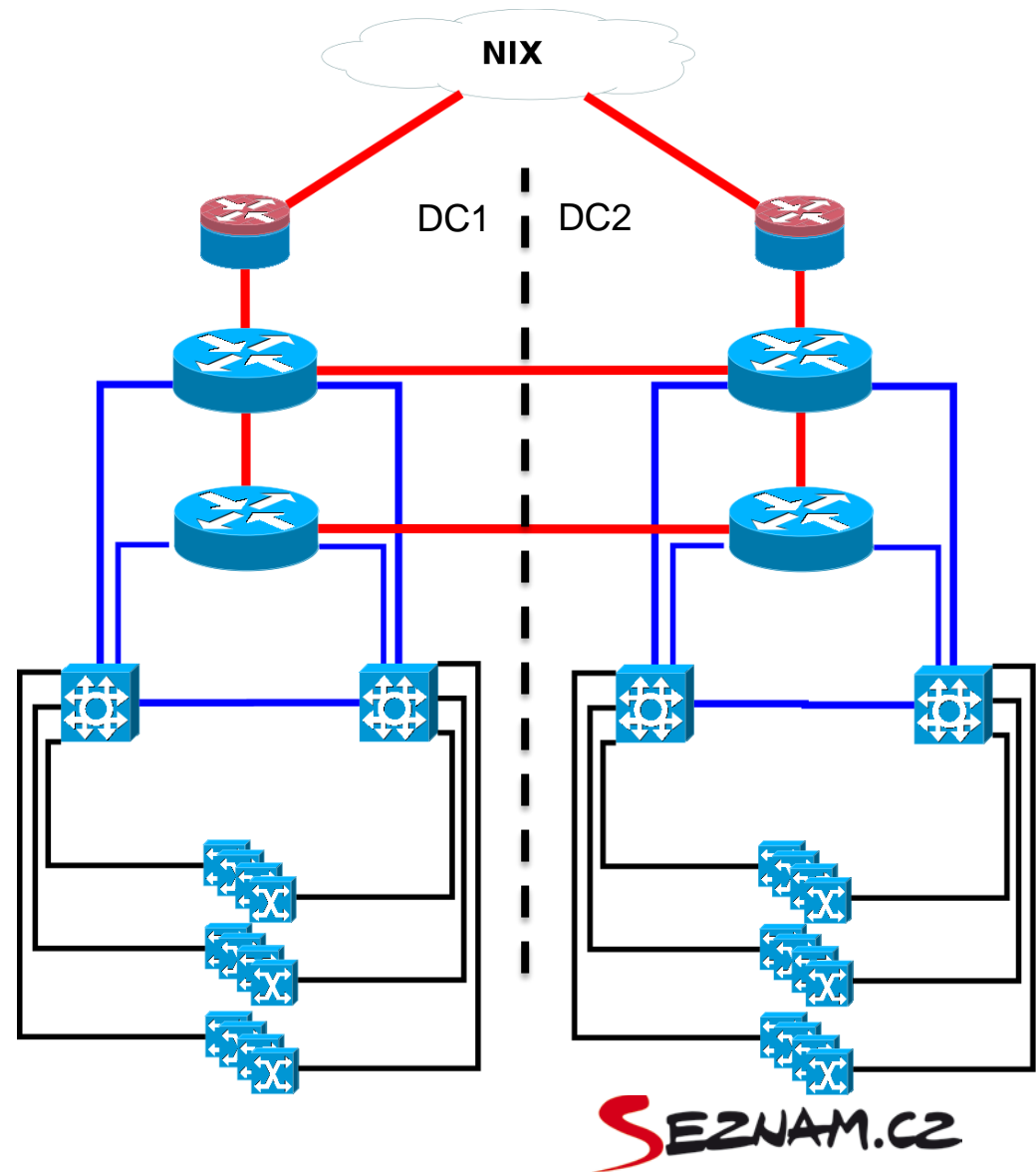


- Disková pole pro servery

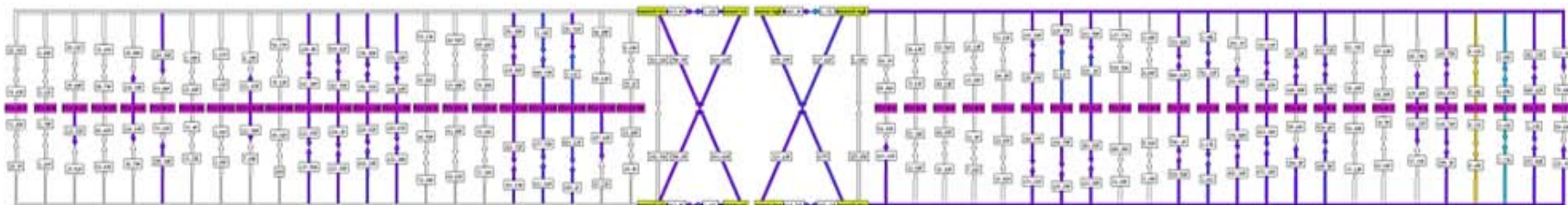
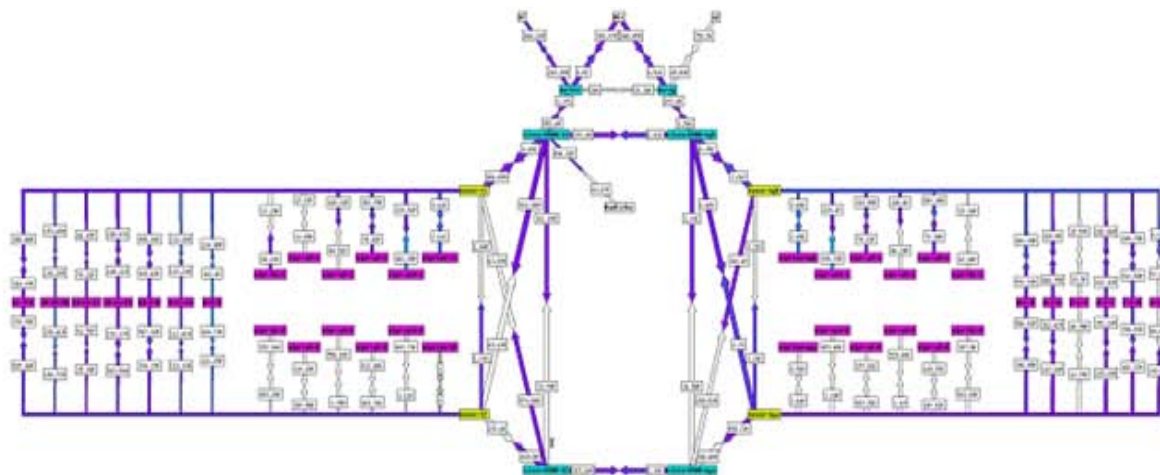


HW - network

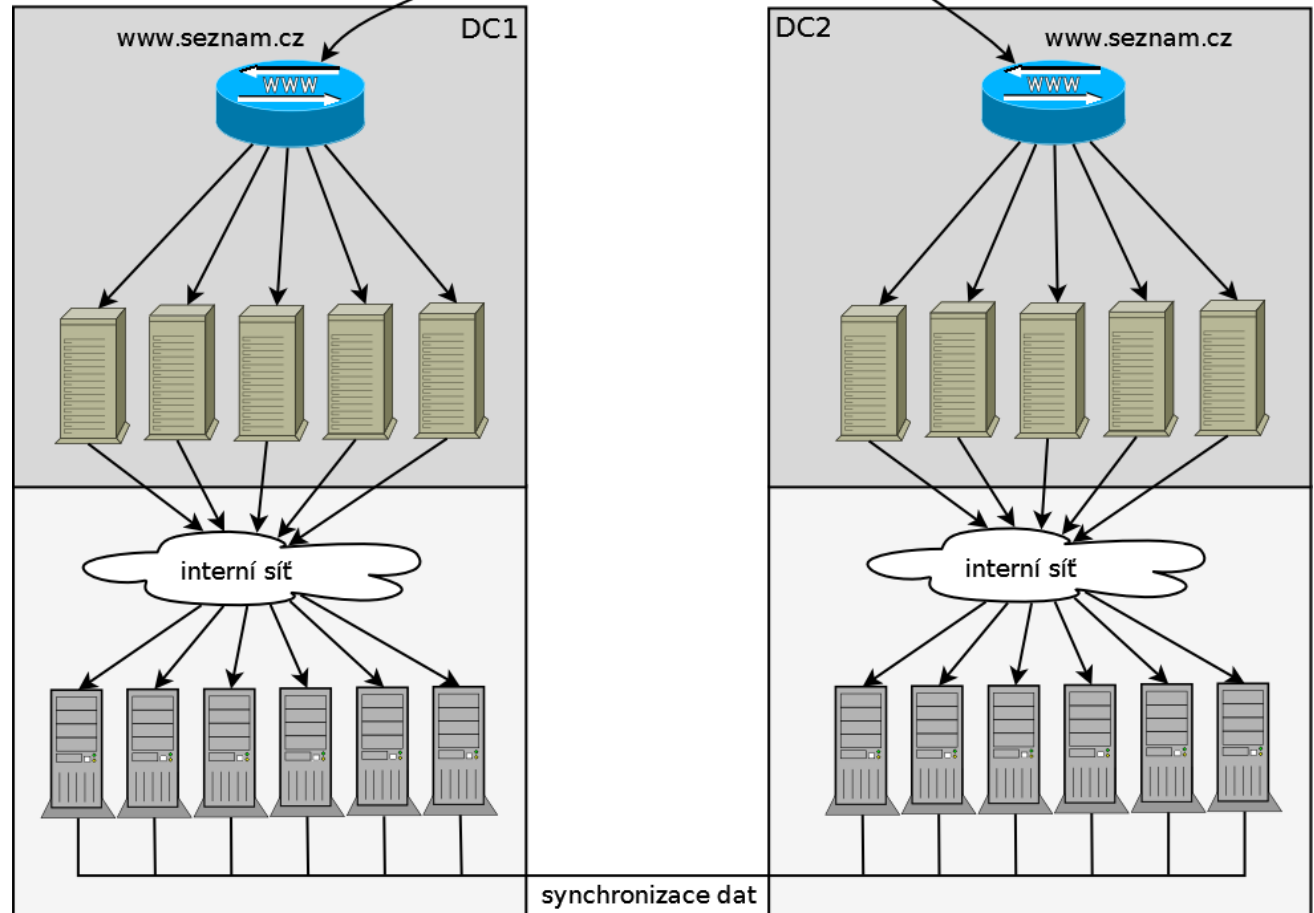
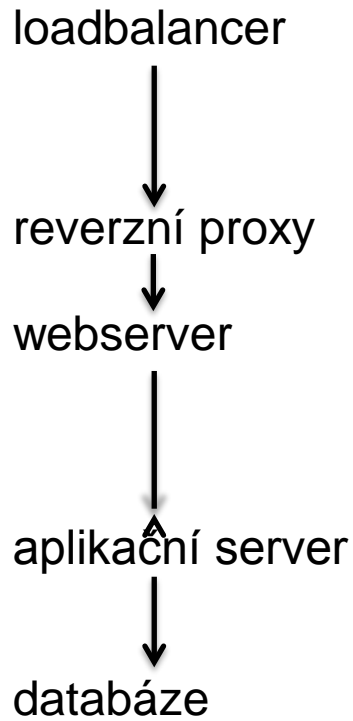
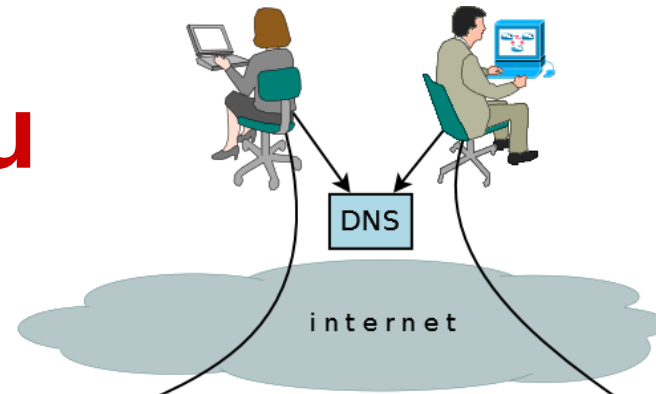
- Core routery 2x
- Core L3 SW 4x
- Aggregate SW 12x
- Access SW 228x
- Mng SW 147x



HW – mapa sítě



Cesta k serveru



SW - přehled

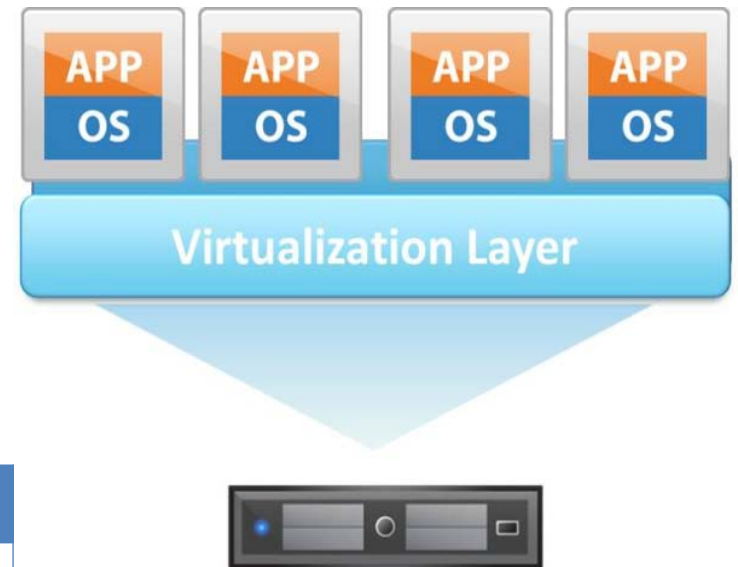
- virtualizace – openVZ, LXC, KVM
- stejná distribuce – Debian, Solaris
- jednotný monitoring a API
- instalace aplikací - vše přes balíky
- for smyčky, salt, puppet, chef
- síťová instalace
- selfcheck / supervisor / puzzle

SW - virtualizace serverů

Důvody:

- oddělitelnost aplikací
- manipulace s virtuály
- využití HW

Typ operačního systému	host	guest	%
Nevirtualizovaný Solaris	363	0	4%
Nevirtualizovaný Debian	480	0	6%
OVZ - Debian	1387	5268	82%
LXC - Debian	137	485	8%
KVM - Debian	16	17	0,4%
Celkem OS	2383	5770	8153



SW - aplikace

Open-source

- OS: 96 % Debian + 4 % Solaris
- Kernel: RHEL / UEK + vybrané patche
- Web: Nginx + Apache2
- DB: MySQL, no-SQL
- HPC: Hadoop, Cassandra

Proprietární

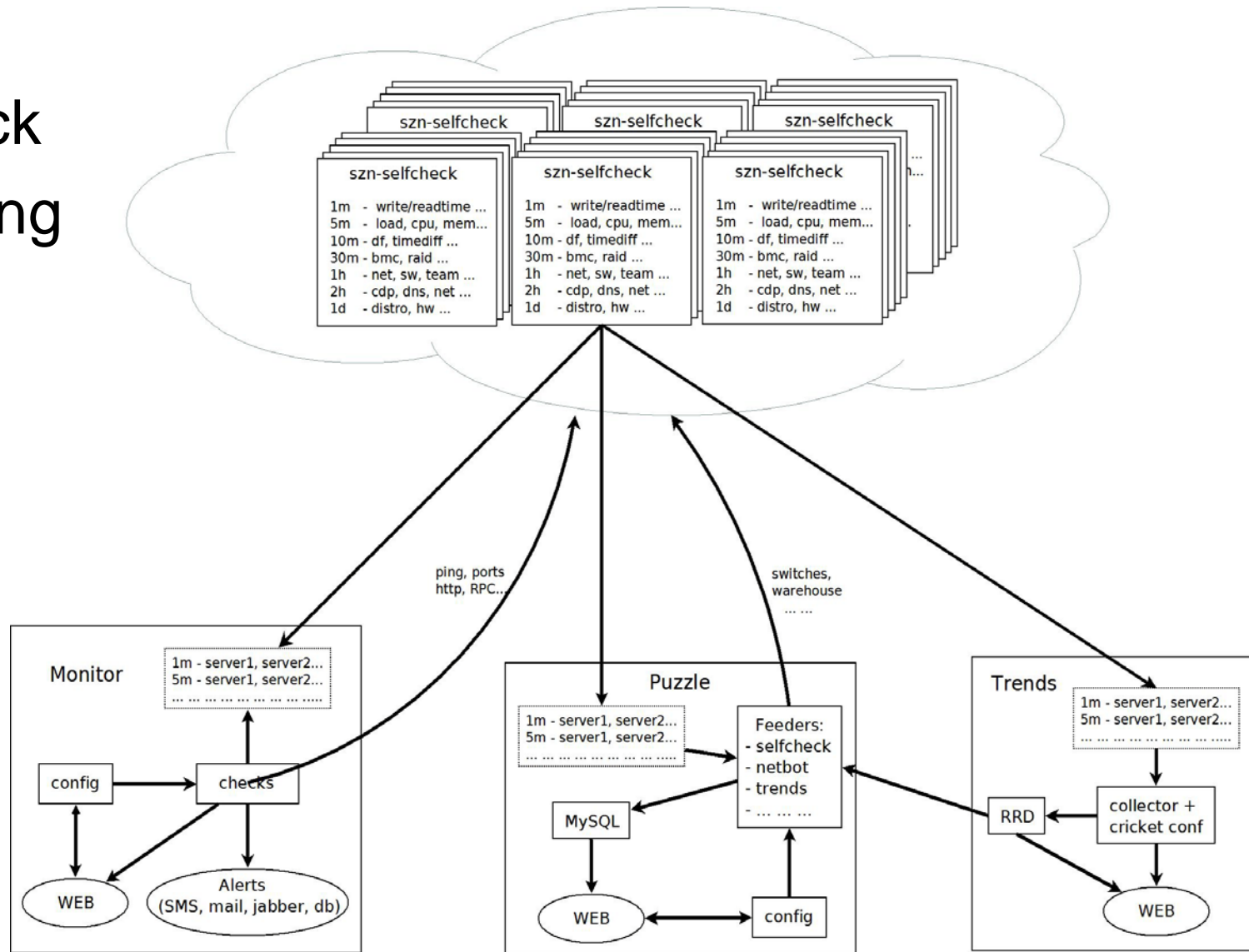
- python: webovky, scripty
- C++: backendy (Metaserver)
- js, java, ruby, django, teng ...

SW – aplikace podpory provozu

- Monitoring
 - externí - Nagios (vlastní scripty)
 - interní – Mon (vlastní téměř vše)
- Grafy
- DNS – PowerDNS
- Puzzle – webový přehled a správa serverů
 - sběr informací, filtrace, přehledy
 - ovládání serverů, init scriptů, portů na switchi
 - instalace systémů, reporty, SMS zprávy
 - využití datacenter atd.

Aplikace pro podporu provozu

Selfcheck
Monitoring
Puzzle
Trends

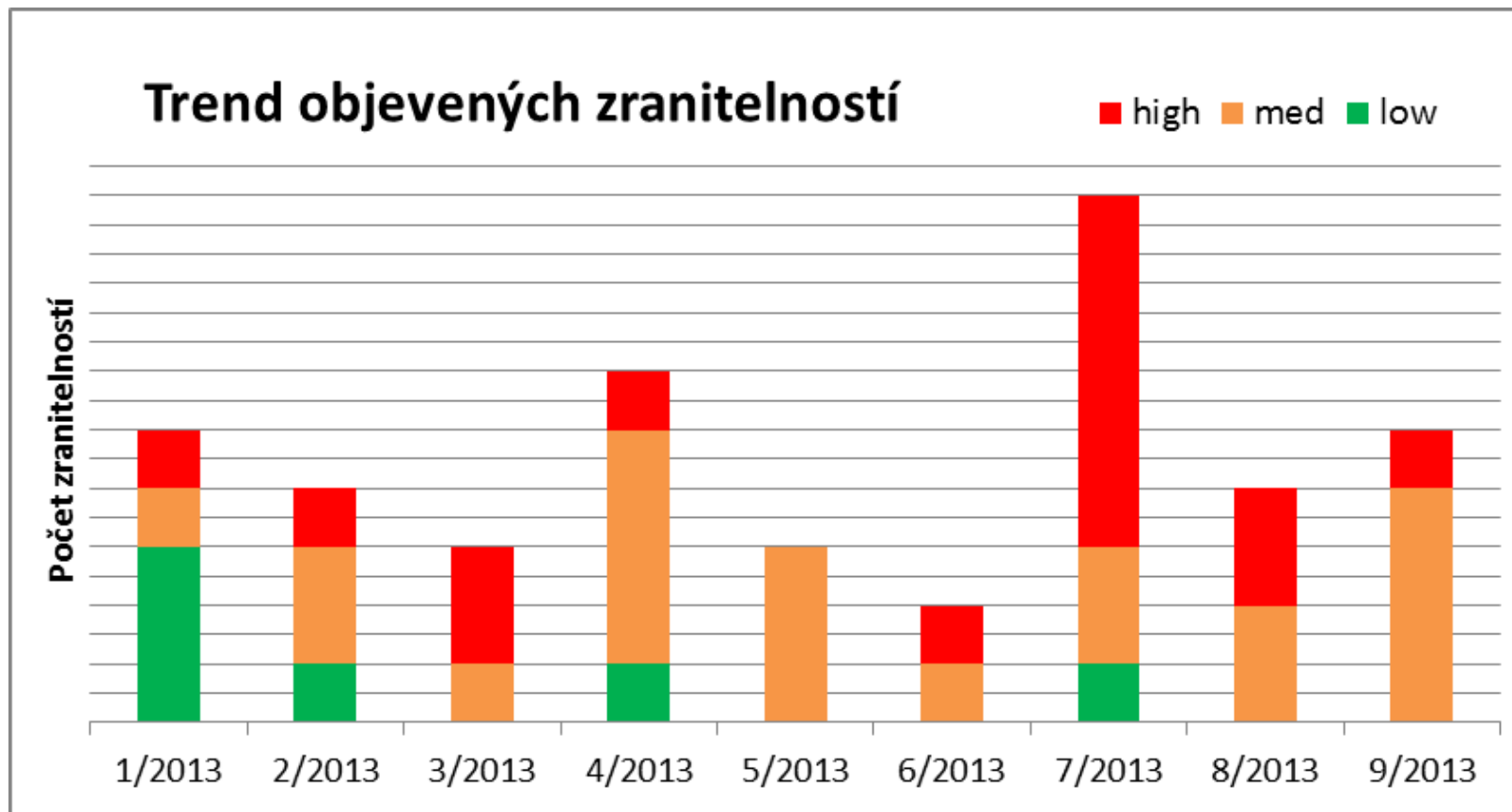


Bezpečnost

- Externí komunikace
 - SEZNAM.CZ-CSIRT
 - vlastní technické kontakty
- Interní bezpečnost
 - skenery zranitelností
 - anti-ddos, IDS
 - flow monitoring



Bezpečnost



Kdo se o provoz stará?

Administrátoři (31x)

Admins1

Email.cz, Login,
AntiSpam
Peněženka
Interní email

Admins2

Novinky.cz, Super
Sport.cz, Lide.cz

DNS, Backup

Admins3

Homepage,
Mapy.cz Zbozi.cz,
Stream.cz, CDN,
Firmy.cz

Admins4

Email.cz cluster

(Solaris)

Admins5

Fulltext,
Encyklopedie

Kernel, OVZ

Admins6

Hadoop clustery

Adm-Brno

Sklik, Reklama

Cassandra

Netadmins

Network
LoadBalancers

Security

Provozní vývoj (6x)

Vývoj

Puzzle, MySQL, DNS
Preinstall, Selfcheck
Monitoring, Trends

Support (15x)

Operators

Onsite 24x7
Monitoring
Installing

Consultants

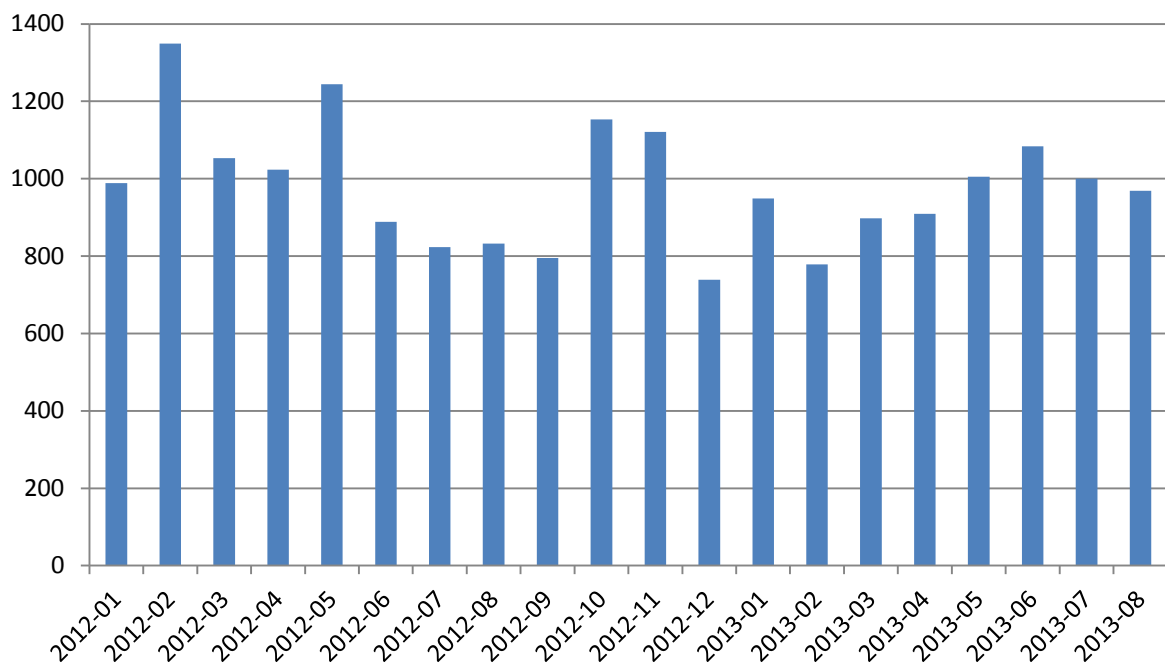
Výrobci,
objednávky
logistika, servis,
datacentra

Admini

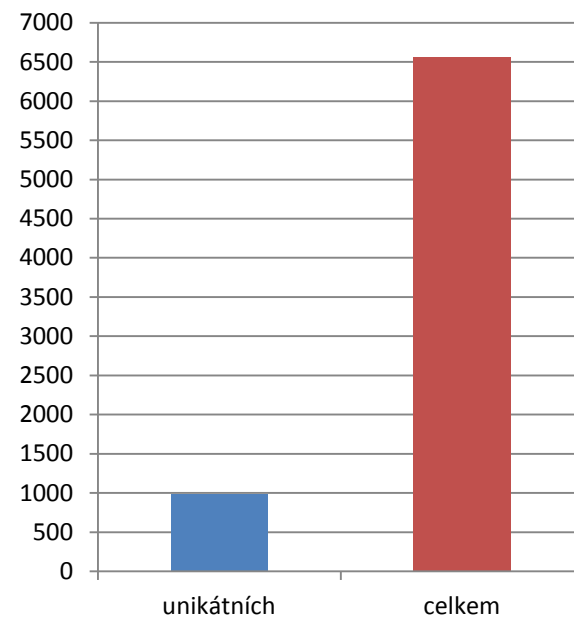


Statistiky

Počet nových aplikací a úprav



Počet komponent

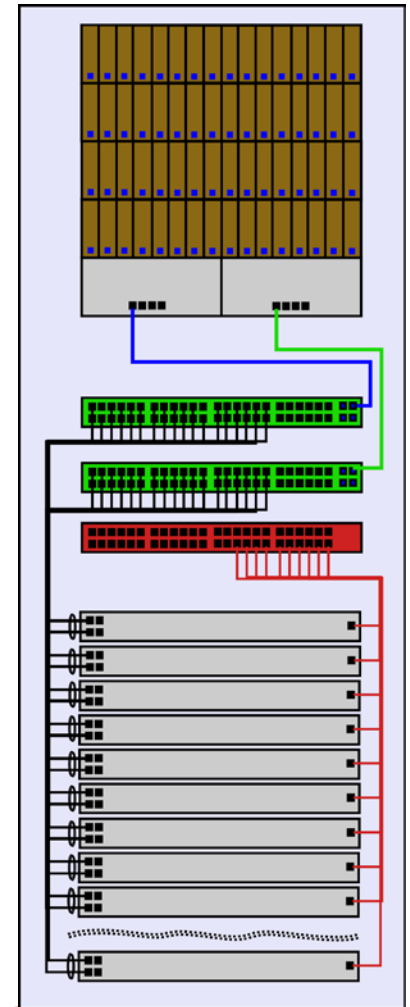


Kudy dál... 1/3

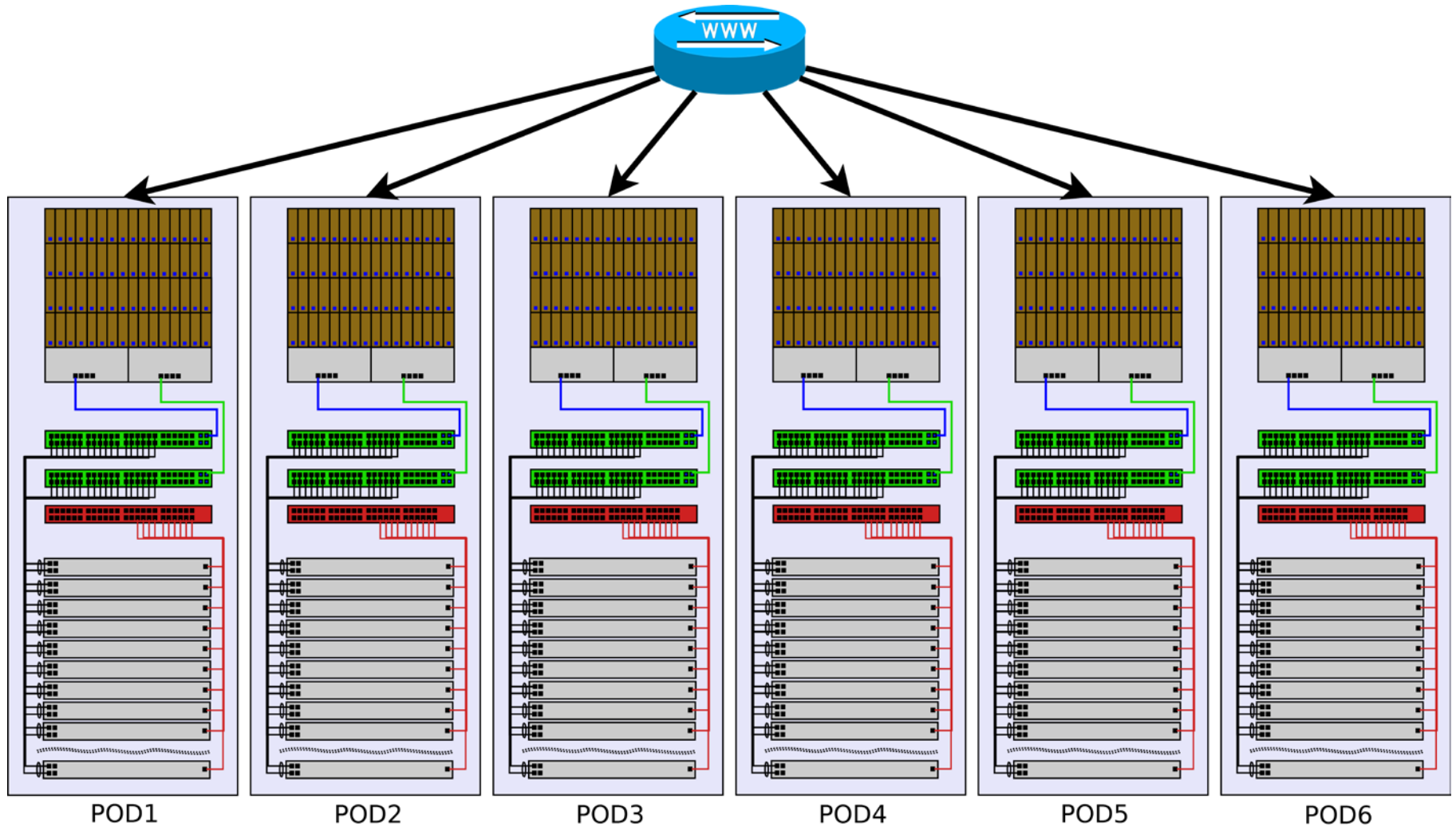
POD – uzavřený dimenzovaný blok HW

- sdílená iSCSI storage
- levné bezdiskové servery
- společná síť LAN-SAN
- nadstavbový SW
 - failover serveru
 - automatický balancing
 - auto deployment

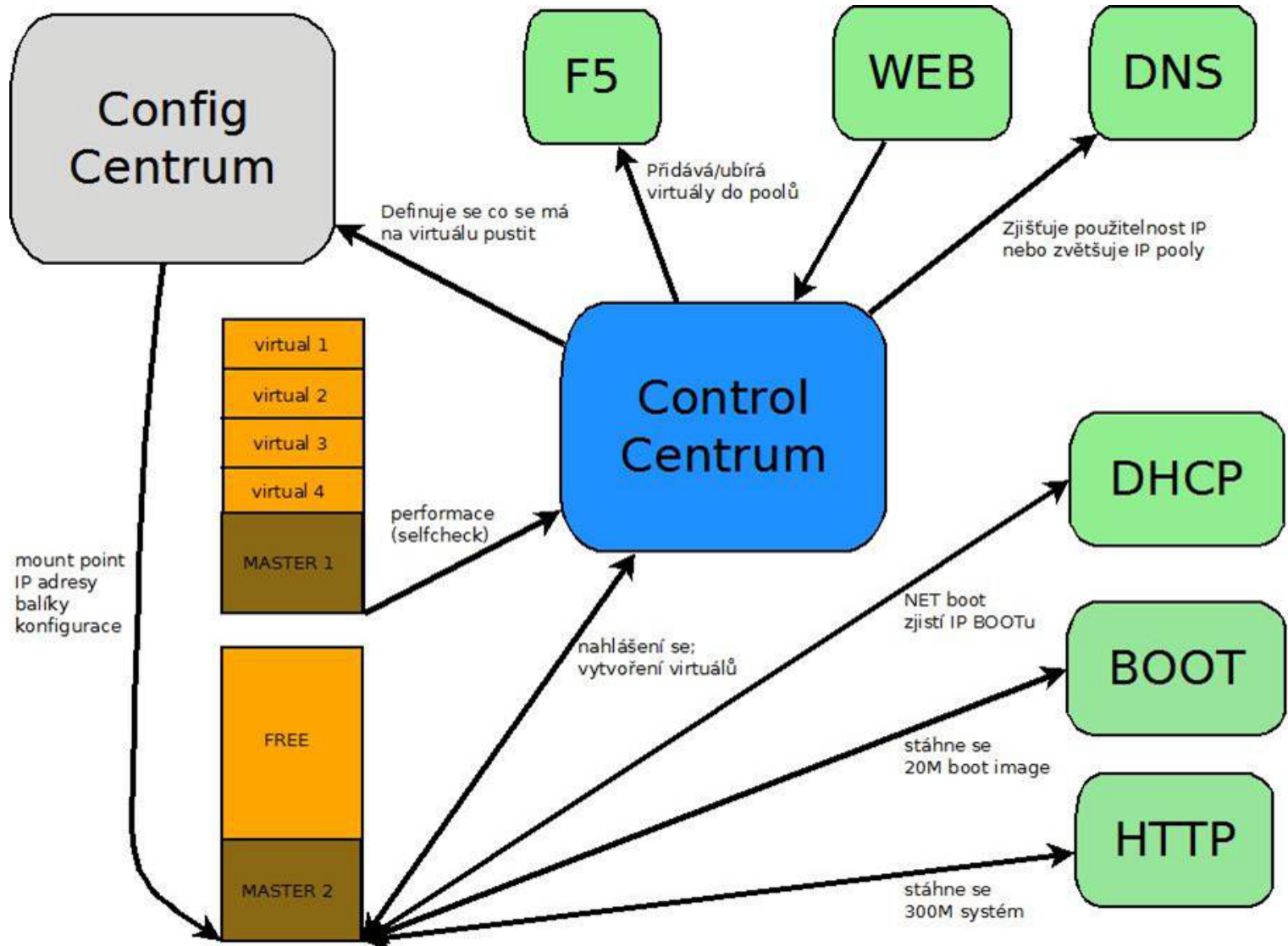
... ale co při pádu storage?



Kudy dál... 2/3



Kudy dál... 3/3



SEZNAM.CZ
...najdu tam, co neznám!

Děkuji za pozornost

Martin Doleček
martin.dolecek@firma.seznam.cz, @MatesDolda