

Správa serverů a počítačových sítí

2020/2021

Přednáška 8
(ver. 2021-04-13-01)



Souborové servery a systémy

- Souborové systémy
 - Význam a nedostatky
 - Journal
 - Ext2-4, XFS, Fat, NTFS
- Síťové souborové systémy
 - Samba, NFS2-3, WinShares
- Distribuované souborové systémy
 - AFS, CODA, NFS4
- Blokované síťové systémy
 - DRBD / Linstor
 - Fibre Channel
 - iSCSI



1. Souborové systémy

- Systém uložení adresářů a souborů
- Detaily na ZOS
- Vždy velice důležitá část systému
- Důležité vlastnosti při výběru
 - Stabilita
 - Výkon
 - Obnovitelnost
 - Journal
 - Dostupnost ovladačů



1.1 Stabilita

- Schopnost dlouhodobého běhu bez nutnosti restartu či následné obnovy
- Schopnost využívat celou kapacitu disku
- Duplikace důležitých dat
 - Superblok
- Schopnost dalšího provozu po tvrdém ukončení systému
 - Např Ext2 po tvrdém restartu či delším chodu bez restartu vyžaduje kontrolu FS
- Rozumná reakce na chybové stavy
 - XFS při tvrdém restartu nuluje otevřené soubory



1.2 Obnovitelnost

- Dostupnost nástrojů na údržbu a obnovitelnost systému
 - Ne jen v běžícím OS, ale i z LiveCD např
- Schopnost pracovat s journalem
 - Atomicita operací
- Ideální kombinovat s kvalitním HW
 - Např řadiče a baterií na zápis cache
- Rychlost rebuildu FS, například při havárii disku
- Možnosti zálohování
 - Např možnost vytvořit kompletní dump FS a jeho obnova



1.3 Journal

- Dnes téměř u všech standart
 - Aktuálně není u FAT a Ext2
- Jedná se o transakční zpracování úloh
 - Data jsou na disk zaznamenána kompletně nebo vůbec
 - Atomické zpracování s možností návratu
- V případě havárie disku je možné přehrát poslední nedokončené operace ze journalu
- Občas problém s obnovou
 - například u ext3 vazba superblock a journal



1.4 Dostupnost ovladačů

- U Windows typicky není problém
 - Drtivá většina je NTFS nebo Fat, obojí nativně dostupné
- V Linuxu je velké množství FS
 - ext2, ext3 , ext4, xfs, reiserfs, jfs, btrfs....
- Aby systém mohl nastartovat musí jádro znát FS při startu
 - 1. jako součást jádra OS
 - 2. jako modul, pak musí být součástí ramdisku - initrd
- Při použití exotického FS, je dobré mít recovery CD/USB, které tento FS podporuje
 - Ideálně ne doma šuplíku, ale přímo u serveru



2.1 Používané souborové systémy na MS Window

- FAT32 či FAT16
 - Starší a jednodušší systém
 - Nepodporuje ACL
 - Snadné čtení i zápis z jiných OS
- NTFS
 - Původně jen pro serverové systémy - poprvé Windows NT
 - Podporuje
 - ACL a práva spravovaná prostřednictvím domény
 - Journal
 - Snapshoty FS
 - Dříve velmi komplikované čtení z jiných OS
 - V linuxu implementované jako Fuse
 - Dnes už je čtení standart
 - Zápis je možný, ale přepis většího množství dat nemusí být bezpečný – nedoporučuje se



2.2 Používané souborové systémy na Linux

- Ext2
 - Snad nejstarší FS pro Linux
 - Velice jednoduchý, umí ho snad každý systém
 - Nepodporuje journal
 - Velmi citlivý na dlouhý chod – nedovolí start bez kontrol
- Ext3, Ext4
 - Nástupci ext2
 - Podporují journal
 - Zpětně kompatibilní mezi sebou
 - V jednotlivých verzích se řeší
 - Rychlost / Stabilita / dostupná maximální velikost
 - Nutné například pro AFS-cache



2.2 Používané souborové systémy na Linux

- XFS

- V posledních letech na ústupu, problém se stabilitou
- Velmi dobrá schopnost obnovy po pádu
 - Za cenu někdy drastických opatření – nulování otevřených souborů
- Nemá problém s dlouhodobým během systémů
- Má zásadní problém při fungování na téměř zaplněných discích – nad 95 procent – dokáže poškodit sám sebe

- ReiserFS

- Dříve často nasazován jako alternativa k ext3
- Vlastní opravné utility, ale pokud dojde ke skutečné havárii, většinou jsou data ztracena
- Dlouho nebyl součástí jádra, kvůli neshodě vývojářů



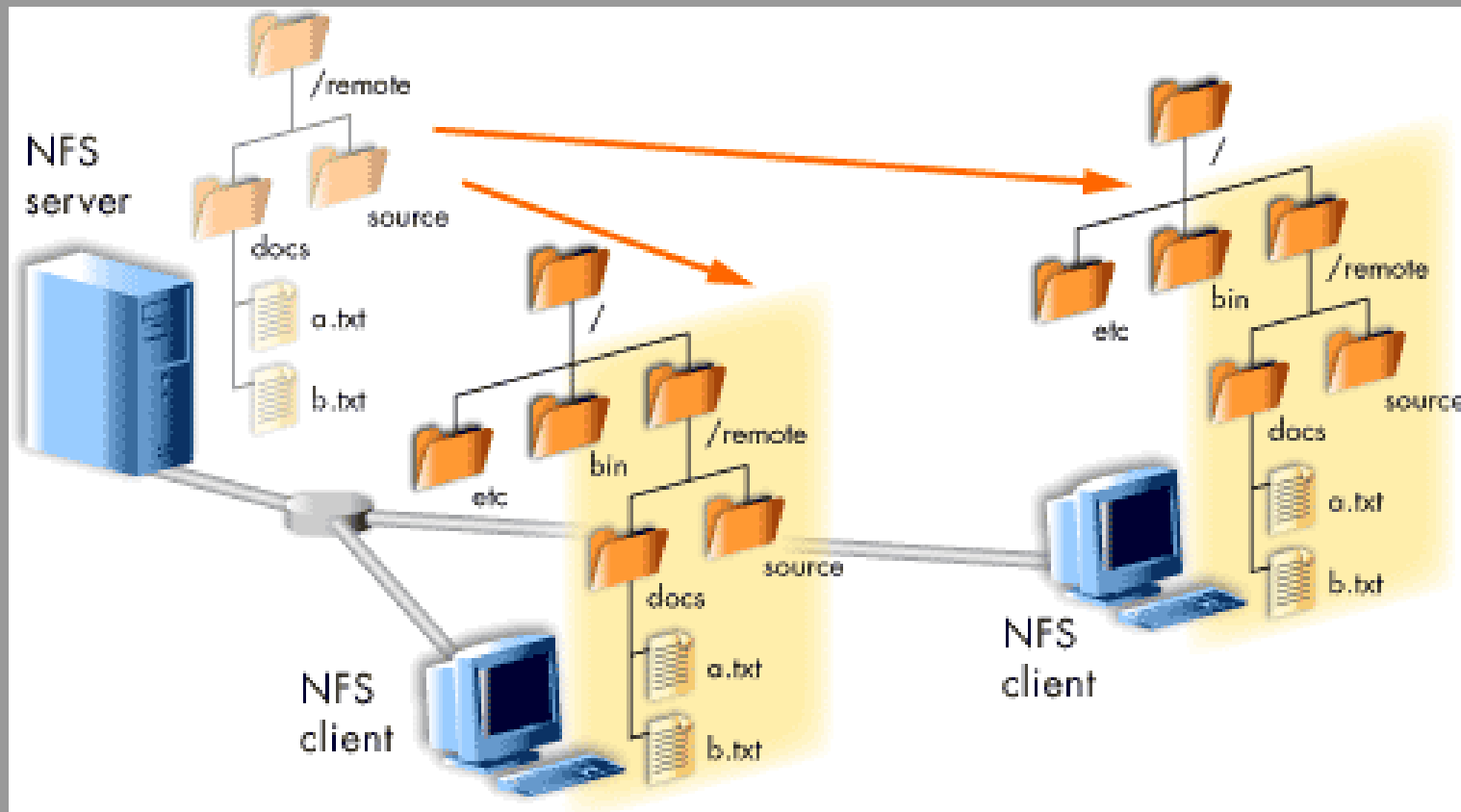
3. Síťové souborové systémy

- Data nejsou uložena na lokálním stroji, ale jsou přístupná pomocí počítačové sítě
- Veškerá data jsou typicky v rámci systému dostupná z jednoho zdroje
- Při výpadku zdroje jsou data nedostupná v jakékoliv podobě
- Běžně používané v lokálních sítích
- Zástupci :
 - NFS2, NFS3, Samba, WinShares, Fuse-FS



3. Síťové souborové systémy

příklad

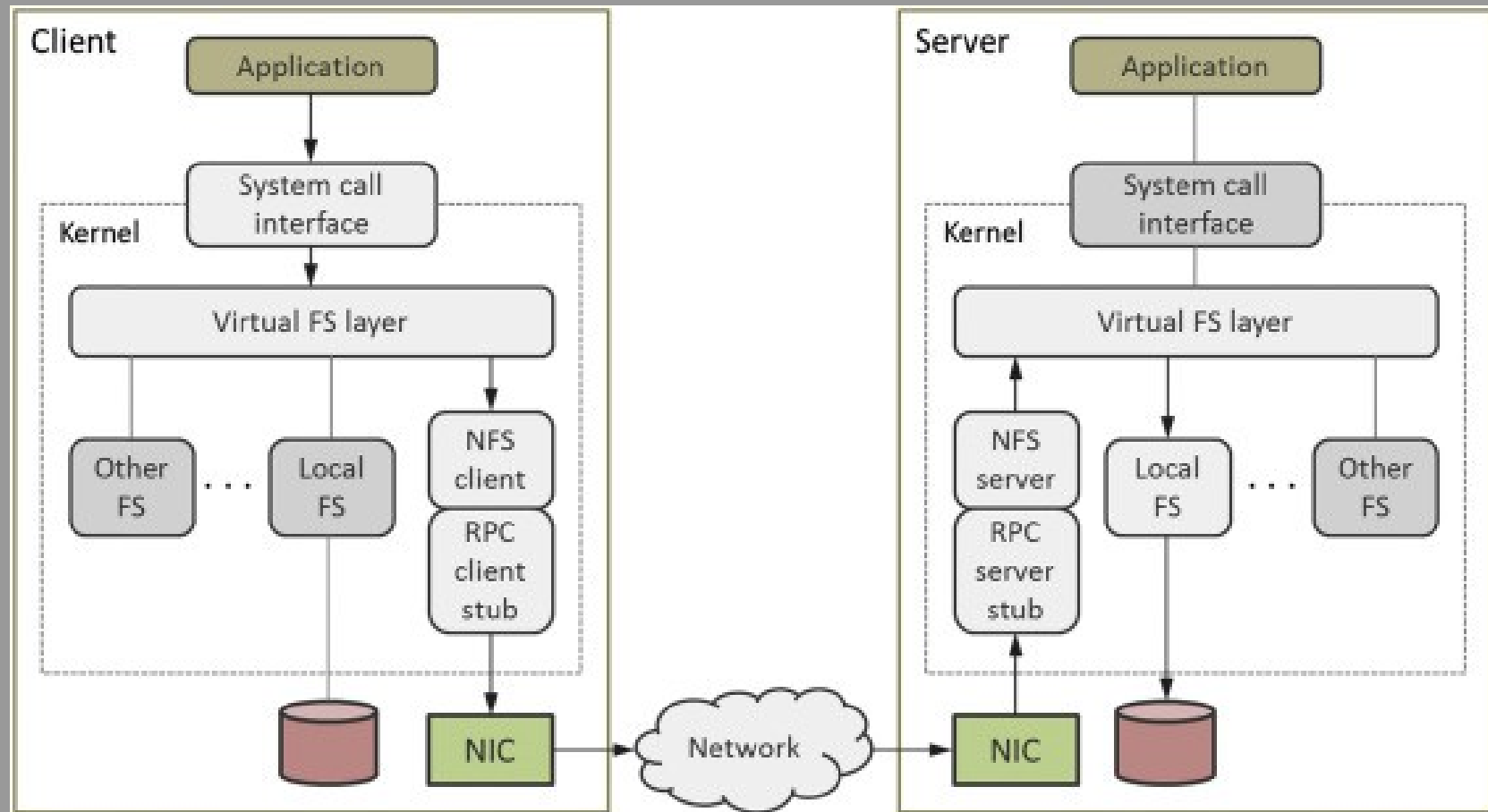


The NFS server exports the `/remote` filesystem to two NFS clients.

<https://cs.puntomariner.com/what-is-nfs-network-file/>



3. Síťové souborové systémy příklad II



<https://www.sciencedirect.com/topics/computer-science/network-file-system>



3.1 NFS2 a NFS3

- Patrně nejstarší síťový souborový systém
- Využívají RPC
- NFS3 už umí fungovat na TCP, NFS2 funguje jen nad UDP, což není v lokálních sítích výhodné
- Neumí ACL, ověřování jen na úrovni stroj a share
- Práva se přenáší s exportovaného systému na hostitelský
- Musí souhlasit uid uživatelů



3.1 NFS2 a NFS3

- V lokálních sítích patrně nejrychlejší způsob sdílení dat
 - Na instalaci
 - Na rychlost přenosu
- Velmi citlivý na výpadky paketů, dokáže tak zastavit celý hostitelský systém
- Dostupné ve všech Linuxech
- Dostupný klient do Windows
- Je třeba zabezpečit aby se provoz nedostal mimo lokální síť, není šifrovaný
 - Nutnost nasazení Firewallu a omezení sdílení



3.1 NFS2 a NFS3

- Musí běžet portmap / rpcbind

- Konfigurace klienta/etc/fstab

```
192.168.1.3:/data /mnt/ nfs rsize=8192,wsiz=8192,noac
```

- Konfigurace serveru v /etc/exports

```
/data 192.168.1.1/255.255.255.0(rw,no_root_squash,async )
```



3.2 Samba

- Odvozeno od WinShares
- Neumožňuje vše co sdílení na Windows, například dědění práv
- Stále se vyvíjí, dnes už umí základní ACL či ověřování přes Kerberos
- Nižší propustnost než NFS
 - Ale vyšší úroveň zabezpečení a transparentnosti
- Často používán v kombinaci Win-klienti a Linuxový server jako fileservr či řadič AD
- Existují dva moduly v rámci Linuxu
 - SMB a CIFS



3.2 Samba

- Konfigurace /etc/samba/smb.conf
- Uživatelé mohou být ve více systémech :

- Files
 - Ldap
 - Mysql
- [faktury]

```
comment = Moje data
read only = No
create mask = 0770
directory mask = 0770
browseable = Yes
valid users = pepa, honza
path = /home/data
```

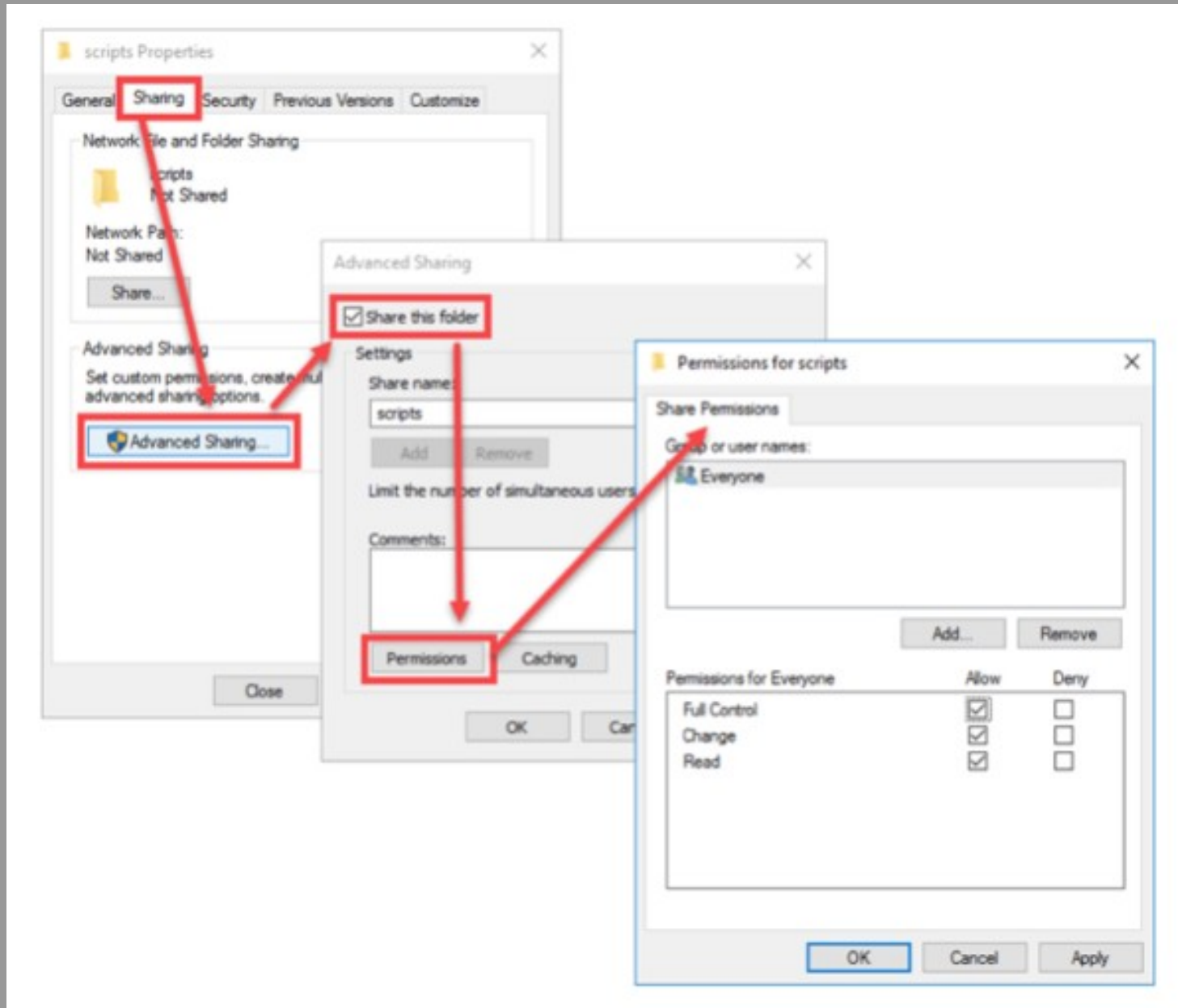


3.3 WinShare

- Dostupné na všech verzích MS Windows
- V Linuxu je dostupný pomocí SMB/CIFS
- Nejrozšířenější systém ve win-sítích
- Velmi snadné nastavení přes GUI
- Umožňuje pracovat s ACL nad NTFS
- Může být kombinováno s doménou



3.3 WinShare - příklad



<https://www.businessnewsdaily.com/11020-create-file-share-windows-server-2016.html>

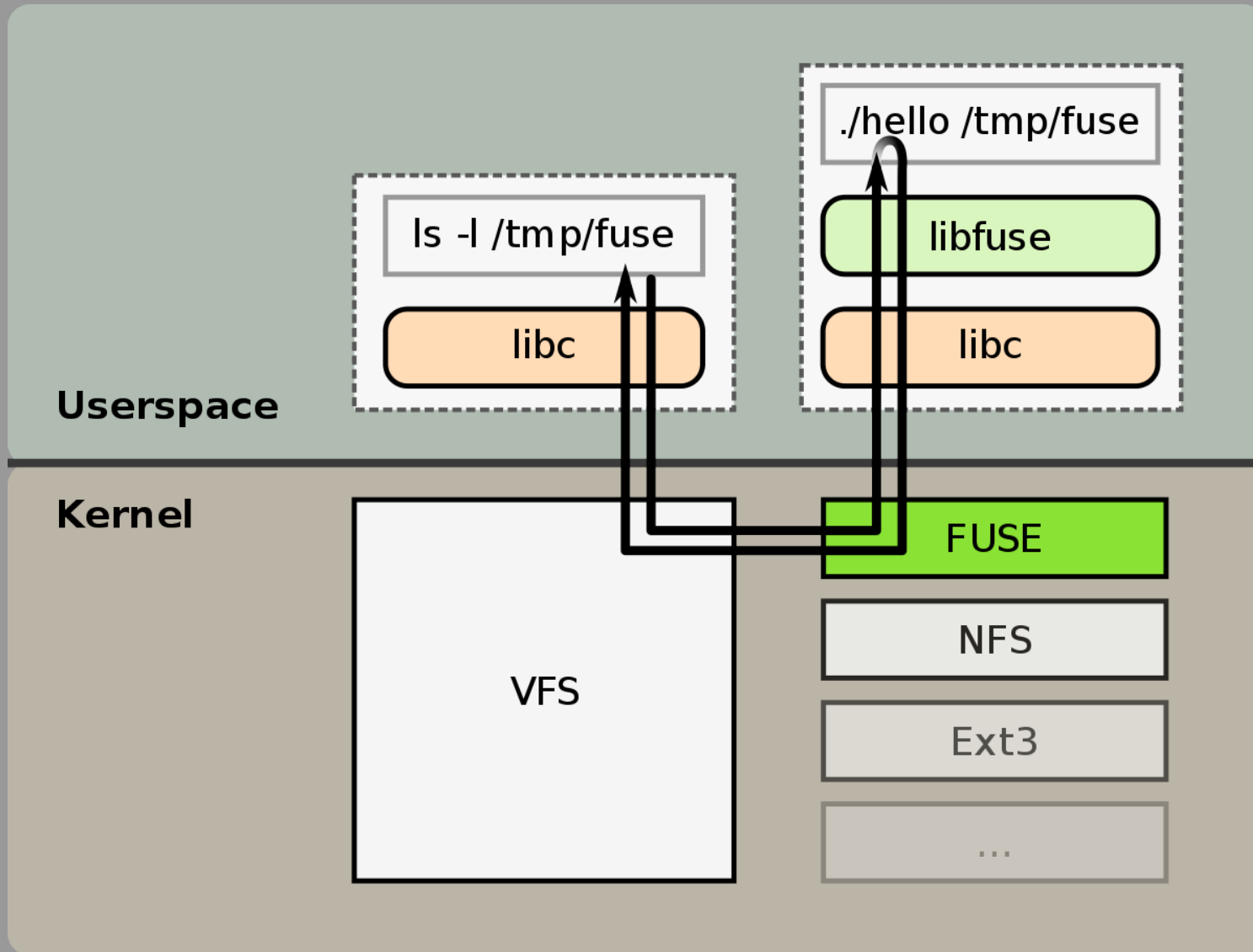


3.4 Fuse-FS

- Nejedná se skutečný souborový systém
- Je to API v jádře, které dovoluje dopisovat vlastní moduly
- Je pak možné snadno připojit různé systémy :
 - SSH, FTP, NTFS, ZFS, IMAP, atd ...
- Velmi pomalé proti nativním ovladačům



3.4 Fuse-FS - příklad



https://cs.wikipedia.org/wiki/Filesystem_in_Userspace

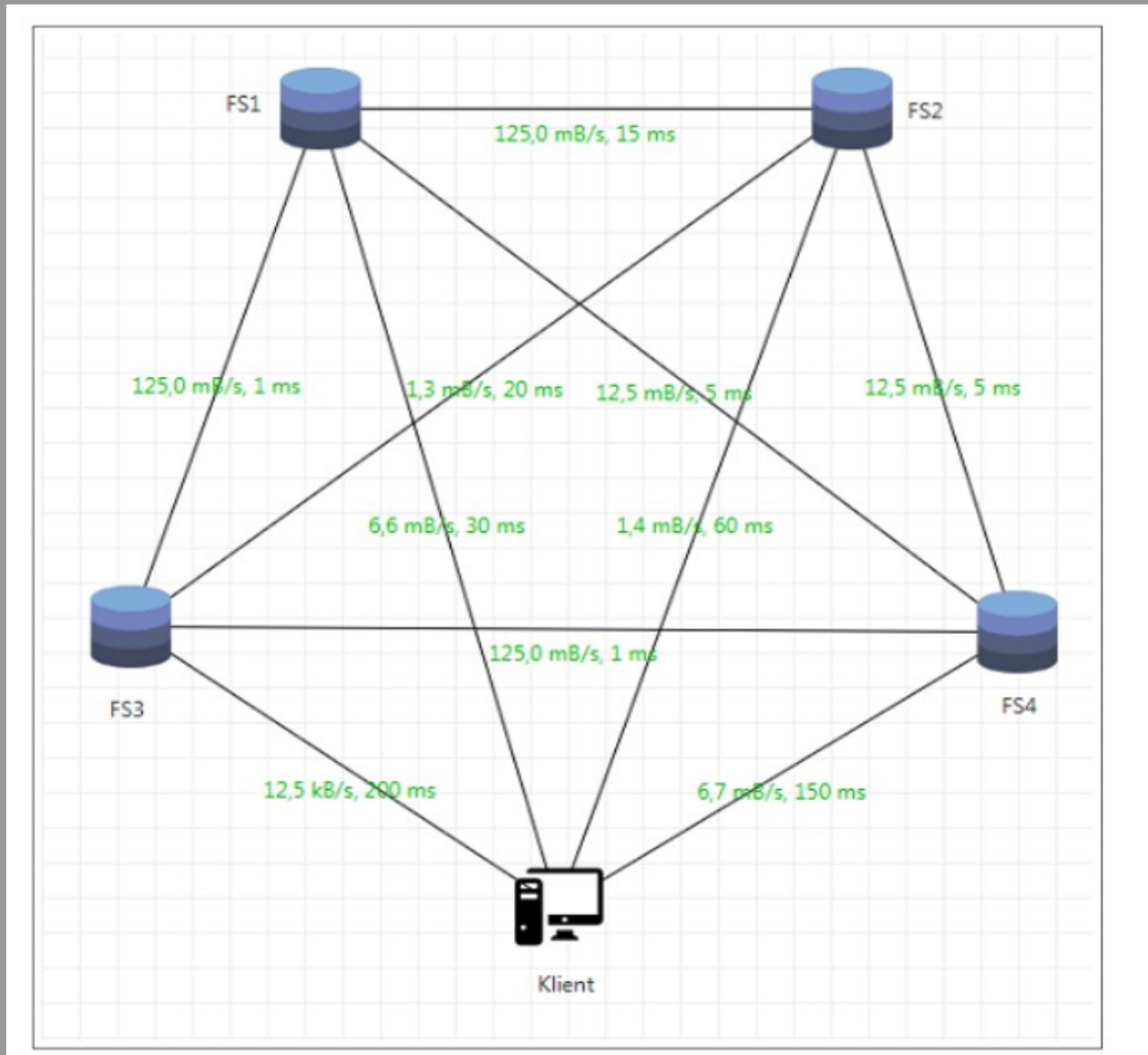


4. Distribuované souborové systémy

- Podobně jako síťové systémy neposkytují typicky lokální data, ale data sdílejí prostřednictvím počítačové sítě
- ALE - umožňují definovat pro jedna data více zdrojů
- Podporují transakce, replikace, ACL, šifrované přenosy, transparentnost z pohledu uživatele
- Většina systému nedotažena
- S přidanou hodnotou přináší i další režii



4. Distribuované souborové systémy - příklad



<https://docplayer.cz/105228756-Diplomova-prace-simulator-distribuovaneho-souboroveho-systemu.html>



4.1 DFS - Transakce

- Stejně jak v databázích, transakce je ucelený blok operací, který se musí provést celý nebo vůbec
 - Atomické provedení
 - Možnost návratu k původnímu stav
- Transakce řeší možné souběhy operací v distribuovaném prostředí
 - DFS jsou velmi citlivé na přesný čas
 - Může docházet k souběžným zápisům a je třeba JEDNO-ZNAČNĚ rozhodnout o jejich pořadí



4.2 DFS - Replikace

- Stejná data mohou být dostupná z více zdrojů
 - Mohou – může to tak být, ale nemusí
- Ne všechny zdroje umožňují data zapisovat, některé jsou jen pro čtení
 - Typicky Master-Write replikace
- Replikace nemusí probíhat okamžitě, ale může být záměrně zpožděná
- Z pohledu uživatele by měla být existence replik transparentní
 - Měla by být, ale ne vždy to tak je
 - Např AFS / afs / .zcu.cz a / afs / zcu.cz



4.3 Typické použití DFS

- Rozsáhlé sítě jedné organizace
- Více poboček, které nejsou spojené bezpečným spojením, tedy typicky obecné internetové spoje
- Velké množství uživatelů s různými úrovněmi oprávnění
- Nevhodné na jednotlivé lokální sítě – příliš velká režie



4.4 Dostupné systémy

- DFS splňující všechny požadavky patrně neexistuje
- Každá implementace má svá omezení
- Nejpoužívanější představitelé :
 - AFS,Coda
 - NFS4
 - LustreFS, GFS, GoogleFS, CEPS,.....
 - ParalelFS, pNFS
 - „Win DFS“



4.5 AFS

- Původně vyvíjen v IBM, následně portován na open source jako OpenAFS
- V dnešní době patrně nejpoužitelnější DFS
 - Pro klasické použití jako souborový systém
- Podporuje replikace s master-write kopií
- Replikace nejsou automatické, musí se ručně nastavit a také ručně spustit jejich synchronizace
- Nejmenší oddíl je volume
- Má dvě větve, RO a RW, pokud neexistuje replikace volumu jsou v RO i RW stejná data



4.5 AFS

- Existují klienti i servery na různé platformy a OS: Linux, Windows, Solaris
- Typicky servery běžící na Linuxech
- Neumí oddíl větší jak 2TB, pro cache nesmí XFS
- Autorizace vlastní nebo přes Kerberos
- Funguje na více portech nad UDP, TCP se připravuje
- Velká režie, reálná rychlost kolem 7MB/s



4.6 Coda

- Jedná se o akademický projekt s cílem vytvořit DFS, který splňuje všechny požadavky na DFS
- V základu vychází z AFS
- Podporuje offline operace
- V zásadě nikdy nedotažená do použitelné podoby
- Už dlouho se uvažuje o kompletním přepsání
- Zdrojový kód jsou uloženy na AFS;)



4.7 PVFS a pNFS

- Neúplně typické, ale velmi zajímavé řešení, především pro prostředí s nutností rychlého odbavování homogenních dat
- Typicky ve výpočetních centrech
- V zásadě jsou na všech serverech stejná data a čtou se z několika zdrojů najednou, čímž se zvyšuje jak rychlost tak dostupnost
- PVFS samostatný projekt, pNFS nadstavba nad NFS



4.9 Win DFS

- Opět ne zcela klasický DFS
- Existuje v rámci MS Windows serverů
- Umožňuje skrýt několik sdílených zdrojů pod jeden share
- Zjednodušuje údržbu sdílení
- Používá se zřídka
- Oficiální zkratka je DFS, ačkoliv neřeší ani transakce ani replikace

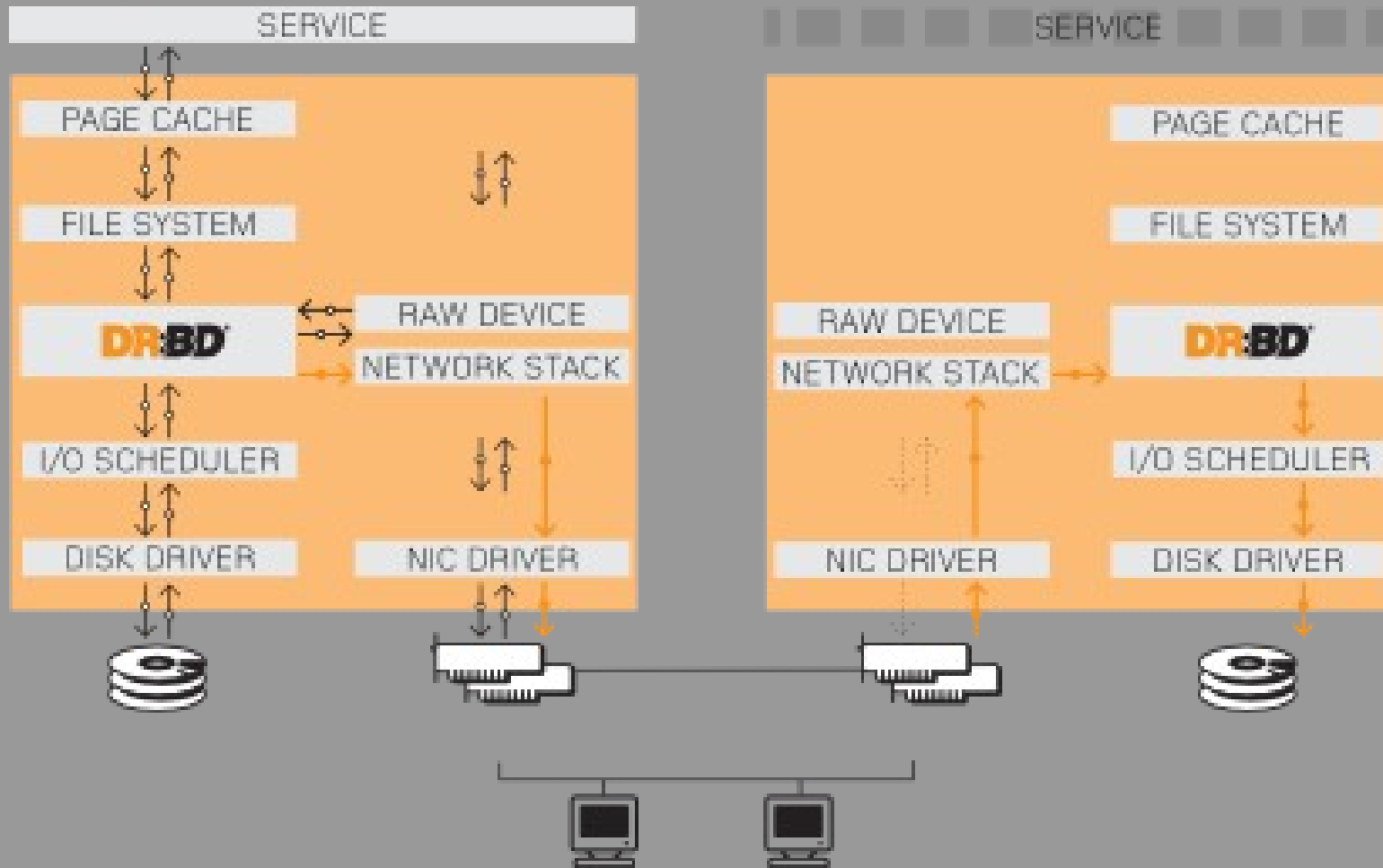


5.0 Bloková síťové systémy

- DRBD / Linstor – síťová verze SW Raidu
 - Konfigurace Master x Master, i Master x Slaves
 - Typické použití pro HA clustery
 - Nutná kombinace s Heartbeat
 - Do verze 8, „raid1 po síti“
 - Od verze 9 s podporou Diskless node
- Fibre Channel – HW SAN systém
 - HW implementace, typicky využívá optické spoje
 - Výkonné, ale velice nákladné, vše musí být extra pro FB
- ISCSI / AoE – SW SAN systém
 - Softwarová implementace
 - Funguje nad IP
 - Podporované v Unixu i Window, dnes už i v HW



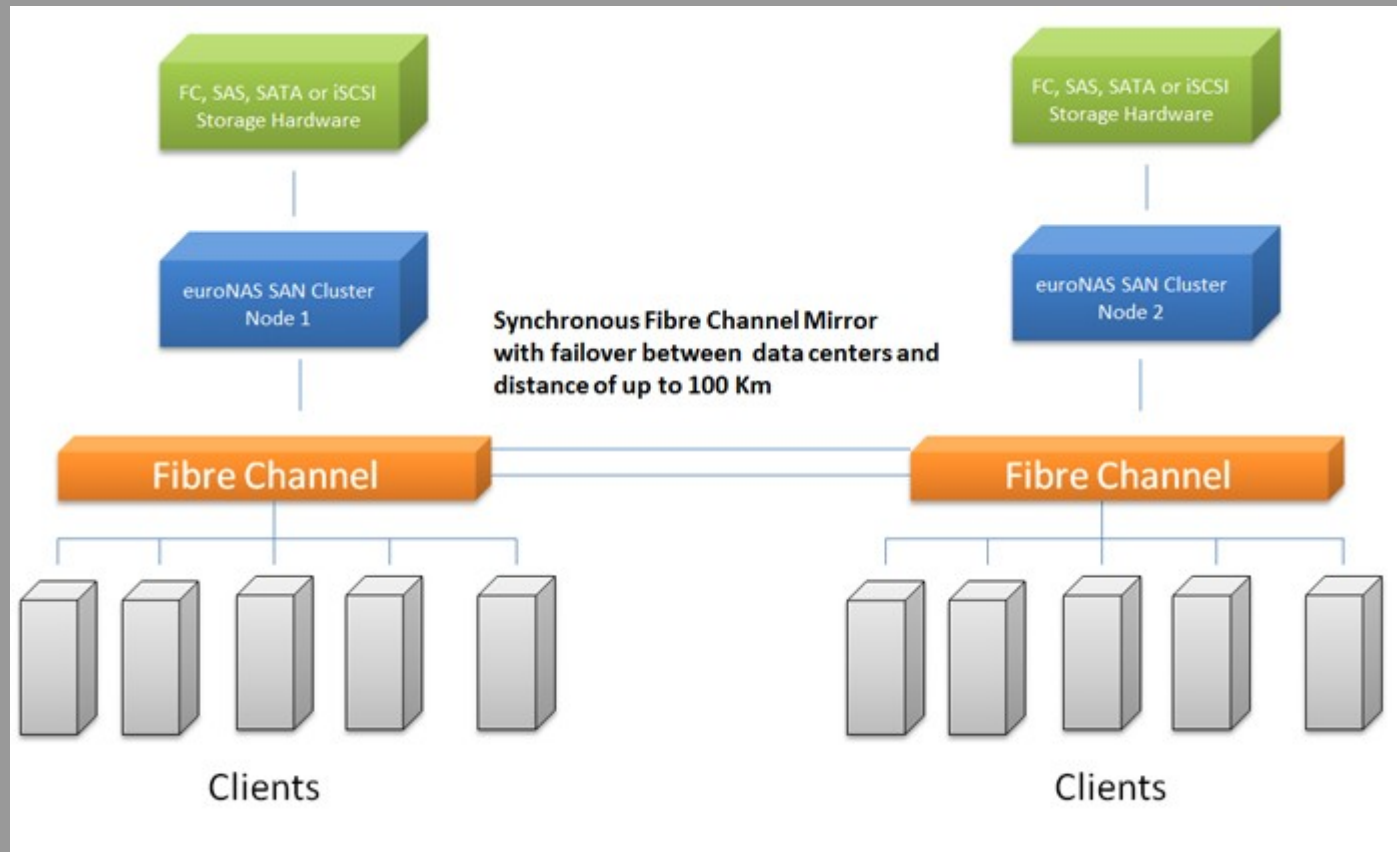
5.0 Bloková síťové systémy - DRBD - příklad



https://pandorafms.com/docs/index.php?title=Pandora:Documentation_en:DRBD1



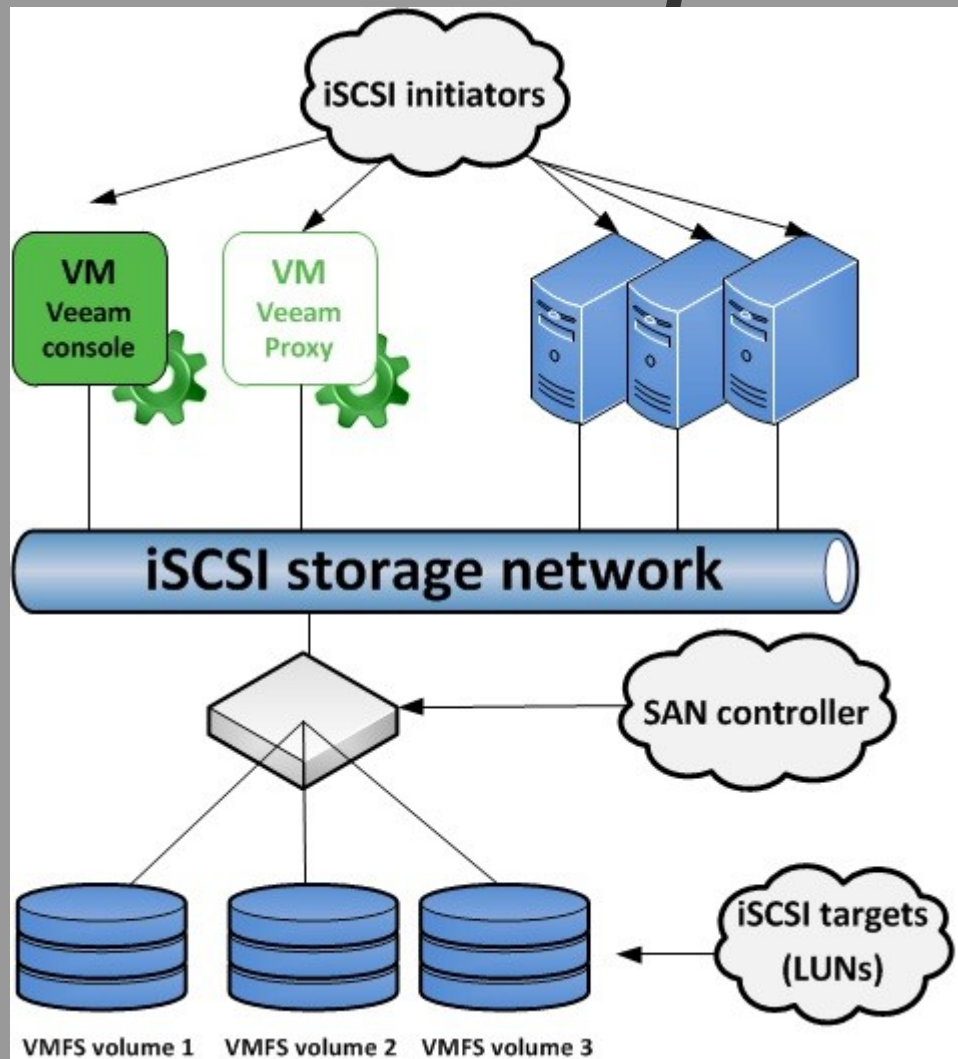
5.0 Bloková síťové systémy - FibreChannel - příklad



<https://euronas.com/product/fiber-channel-cluster-2/>



5.0 Bloková síťové systémy - iSCSI - příklad



<https://www.veeam.com/blog/direct-san-access-tips-for-iscsi-vmfs-volumes-and-backup-proxies.html>

