



# ZÁLOHOVÁNÍ S DÚ

## aneb sto způsobů jak (ne)přijít o data

**Michal Strnad**

CESNET, z. s. p. o.

---

30. 1. 2019



- Datová úložiště a motivace pro workshop
- Popis prostředí pro hands-on
- Teorie a laborky

- Zajišťuje provoz a rozvoj národní infrastruktury pro ukládání dat pro výzkum a vývoj
- Aktuálně provozujeme čtyři úložiště založené na HSM
- Další HSM a nové servery pro object storage se právě instalují
- HSM úložiště jsou přístupná přes NFSv4, FTP, rsync, SCP, Globus ...
- Objektové úložiště pak přes S3/Swift, CephFS, RBD
- Další služby jako FileSender a ownCloud

- Virtuál s CentOS 7
- Přihlašovací jméno, heslo a IP adresu máte na papíru
- Privilegovaný přístup na stroj přes sudo
- Přístup na úložiště CESNET (ostravské du4) přes servisní účet
- V případě problémů s jednotlivými kroky Vám pomůžeme
- Slidy, návody a další materiály  
<https://du.cesnet.cz/cs/workshop>
- Raději manuálové stránky než copy-and-paste ze slidů

- Jednorázová záloha či replika
  - Nástroj rsync
- Chceme spíše kontinuální zálohu
  - Použijeme nástroj duplicity
  - Provedeme obnovu a následnou kontrolu integrity dat
  - Aby to bylo zajímavější, budeme zapisovat do připojeného NFS
  - Pro uživatele i správce
- Potřebujeme mít snapshoty celé datové oblasti nebo systému
  - Využijeme Btrfs snapshoty a send/receive
  - Jako cíl použijeme vzdálené RBD z Ceph clusteru
  - Primárně pro správce

## Jednorázová záloha či replika

- Prakticky na všech distribucích je možné rsync binárku nainstalovat z repa
- Pod sebou má SSH
- Delta přenos
- Používá ho velká část jiných aplikací (např. rsnapshot)
- Možnost skriptování
- Existuje grafická nadstavba Grsync

## Plný návod

<https://du.cesnet.cz/cs/navody/rsync/start>

## ■ Základní použití nástroje

- `rsync -av --progress ./directory`  
`labX@ssh.du4.cesnet.cz:~/VO_du_test-disk_only/`

## ■ Umožňuje zachovávat práva a informace o vlastníkovi souboru (UID a GID) při přenosu na vzdálené úložiště v rozšířených attributech souborů

- `rsync -av --numeric-ids`  
`--rsync-path="rsync --fake-super" ./directory`  
`labX@ssh.du4.cesnet.cz:~/VO_du_test-disk_only/`

- Pozor na lomeno na konci zdrojové cesty
  - `rsync -av ./directory/  
labX@ssh.du4.cesnet.cz:~/VO_du_test-disk_only/`
- Zvláštní obezřetnost je na místě při použití přepínače `--delete`
  - `rsync -av --delete ./directory  
labX@ssh.du4.cesnet.cz:~/VO_du_test-disk_only/`



```
/home/labX:  
|-- .bash_history  
|-- .ssh  
|  \-- authorized_keys  
|-- VO_du_test-disk_only  
|-- VO_du_test-tape_tape  
\-- VO_du_test-tape_tape-shared
```

Budeme operovat jen v politice VO\_du\_test-disk\_only

1. Z webu <https://du.cesnet.cz/cs/workshop> stáhněte testovací data do VM a rozbalte archiv
2. Z VM se připojte za pomoci SSH na `ssh.du4.cesnet.cz`
3. Vytvořte adresář backup v politice `VO_du_test-disk_only`
4. Nahrajte do vytvořeného adresáře backup přes rsync rozbalený archiv
5. Upravte dle svého uvážení soubor `"include/keys/user-type.h"` na svém VM
6. Smažte soubor `"include/memory/jedec_ddr.h"` na svém VM
7. Zavolejte rsync tak, aby provedl změny i na straně serveru (CESNET úložiště)

Řešení:

1. `yum install rsync`
2. `wget`  
`https://du.cesnet.cz/\_media/cs/kernel.tar.gz`
3. `tar -xf kernel.tar.gz`
4. `ssh labX@ssh.du4.cesnet.cz "mkdir`  
`./VO_du_test-disk_only/backup"`
5. `rsync -av --progress ~/include`  
`labX@ssh.du4.cesnet.cz:~/VO_du_test-disk_only/backup/`

Řešení:

1. `vim ~/include/keys/user-type.h`
2. `rm ~/include/memory/jedec_ddr.h`
3. `rsync -av --progress --delete ~/include  
labX@ssh.du4.cesnet.cz:~/VO_du_test-disk_only/backup/`

Chceme spíše kontinuální zálohu. My zde vyzkoušíme NFS a Duplicity. Začneme s NFS.

- Umožňuje připojení vzdáleného svazku jako lokální disk
- Reconnect
- Nemá problémy s hardlinky
- Ve výchozím nastavení posílá data v clear textu
- Ukážeme si případ s Kerberos autentizací a šifrování přenosu

Plný návod

<https://du.cesnet.cz/cs/navody/nfs/start>

- Nainstalujeme základní nástroje a knihovny pro NFS + podporu Kerberos
  - `sudo yum install nfs-utils libnfs-utils krb5-workstation`
- Připravíme si adresář kam následně přimountujeme úložiště du4
  - `sudo mkdir /mnt/nfs`
- Stáhneme si konfigurační soubor pro Kerberos autentizaci
  - `sudo wget https://du.cesnet.cz/_media/cs/navody/nfs/krb5.conf -O /etc/krb5.conf`

- Vygenerujeme si `krb.keytab`, který použijeme k přístupu na úložiště
  - `ssh -o PubkeyAuthentication=no -o GSSAPIAuthentication=no labX@ssh.du4.cesnet.cz "remctl kdccesnet.ics.muni.cz accounts nfskeytab" > krb5.keytab`
- Přesuneme vygenerovaný `krb5.keytab` do `/etc` a nastavíme příslušná práva
  - Zkopírujeme `krb5.keytab` do `/etc`
  - Zajistíme, aby `/etc/krb5.keytab` vlastnil `root:root` a měl práva `600`

- Vygenerujeme tiket
  - kinit labX@EINFRA
- Nastavení statického mapování pro našeho uživatele - uživatel na úložišti pak bude mapován na lokálním stroji jako lokální uživatel
  - `sudo vim /etc/idmapd.conf`
  - do sekce [General] doplníme `Domain = EINFRA`
  - do sekce [Translation] doplníme `Method = static, nsswitch`
  - do sekce [Static] doplníme `labX@EINFRA = labX` a `du_test@EINFRA = labX`



Zkontrolujeme, že nám běží všechny služby, případně je zapneme

- `systemctl is-enabled nfs-idmap nfs-secure`
- `sudo systemctl enable nfs-idmap nfs-secure`
- `sudo systemctl restart nfs-idmap nfs-secure`
- `systemctl status nfs-idmap nfs-secure`

Nyní jsme hotovi a můžeme připojit NFS svazek

- `sudo mount -vvv -t nfs nfs.du4.cesnet.cz:/mnt/nfs -o vers=4,sec=krb5p,rsiz=1048576,wsiz=1048576`
- `nfs.du4.cesnet.cz:/mnt/nfs nfs4 _netdev,sec=krb5p,rsiz=1048576,wsiz=1048576 0 0`

1. Nastavte Kerberos autentitazaci
2. Získejte keytab umístěte ho do správné cesty
3. Vygenerujte tiket pro uživatele labX
4. Připojte NFS svazek do adresáře /mnt/nfs
5. Nastavte statické mapování a ověřte jeho funkčnost pod uživatelem

- Nainstalujeme základní nástroje a knihovny pro NFS + podporu Kerberos
  - `sudo yum install nfs-utils libnfs-utils krb5-workstation`
- Připravíme si adresář kam následně přimountujeme úložiště du4
  - `sudo mkdir /mnt/nfs`
- Stáhneme si konfigurační soubor pro Kerberos autentizaci
  - `sudo wget https://du.cesnet.cz/_media/cs/navody/nfs/krb5.conf -O /etc/krb5.conf`

- Vygenerujeme si `krb5.keytab`, který použijeme k přístupu na úložiště
  - `ssh -o PubkeyAuthentication=no -o GSSAPIAuthentication=no labX@ssh.du4.cesnet.cz "remctl kdccesnet.ics.muni.cz accounts nfskeytab" > krb5.keytab`
- Přesuneme vygenerovaný `krb5.keytab` do `/etc` a nastavíme příslušná práva
  - `sudo cp krb5.keytab /etc/`
  - `sudo chmod 0600 /etc/krb5.keytab`

- Vygenerujeme tiket z keytabu
  - kinit labX@EINFRA
- Nastavení statického mapování pro našeho uživatele - uživatel na úložišti pak bude mapován na lokálním stroji jako lokální uživatel
  - sudo vim /etc/idmapd.conf
  - do sekce [General] doplníme Domain = EINFRA
  - do sekce [Translation] doplníme Method = static, nsswitch
  - do sekce [Static] doplníme labX@EINFRA = labX a du\_test@EINFRA = labX

Zkontrolujeme, že nám běží všechny služby, případně je zapneme

- `systemctl is-enabled nfs-idmap nfs-secure`
- `sudo systemctl enable nfs-idmap nfs-secure`
- `sudo systemctl restart nfs-idmap nfs-secure`
- `systemctl status nfs-idmap nfs-secure`

Nyní jsme hotovi a můžeme připojit NFS svazek

- `sudo mount -vvv -t nfs nfs.du4.cesnet.cz:/mnt/nfs -o vers=4,sec=krb5p,rsiz=1048576,wsiz=1048576`



- Zalohovací nástroj napsán v Python
- Plné a inkrementální zálohy
- Používá standardní Unix nástroje (rsync, tar, GnuPG)

- Jednoduchá obnova z libovolné bodu v čase
- Nepodporuje hardlinky
- Zabudované šifrování (GPG)
- Na Windows pod Cygwin
- Duply
- GUI nadstavba Déjà Dup

- SSH/SCP
- FTP
- Lokální disky (NFS, Samba ...)
- Podpora pro Amazon S3, Google Cloud Storage, Rackspace Cloud Files
- mnoho dalších

- Nainstalujeme z repozitáře balík
  - `sudo yum install duplicity`
- Běh nanečisto
  - `duplicity --dry-run ./directory file:///mnt/nfs`
- Plná záloha
  - `duplicity ./directory file:///mnt/nfs`

## ■ Inkrementální záloha

- `duplicity incr ./directory file:///mnt/nfs`

## ■ Přehled o provedených zálohách

- `duplicity collection-status file:///mnt/nfs`

## ■ Přehled odzálohovaných souborů

- `duplicity list-current-files`  
`file:///mnt/nfs/duplicity/ | sort -n -k 4`

## ■ Kontrola zálohy proti lokálním souborům

- `duplicity verify file:///mnt/nfs ./directory`

## ■ Staré neaktuální zálohy je třeba smazat a nahradit novými

- `duplicity remove-older-than 12M --force`  
`file:///mnt/nfs`

Předpokládáme splnění Lab 2 (připojené NFS).

1. Odpojit NFS a připojit místo / konkrétní politiku (/home/labX/VO\_du\_test-disk\_only)
2. Použijte duplicity na vytvoření plné zálohy adresáře include do cesty /mnt/nfs
3. Vytvořte nový soubor s názvem "marenka" a vložte do něj text "jenicek"
4. Provedte přes duplicity inkrementální zálohu do stejné cesty (/mnt/nfs)
5. Zkontrolujte zálohu za pomoci duplicity
6. Provedte obnovu do adresáře ~/restore

## Řešení:

1. `umount /mnt/nfs; mount -vvv -t nfs  
nfs.du4.cesnet.cz:/home/labX/VO_du_test-disk_only  
/mnt/nfs -o  
vers=4,sec=krb5p,rsiz=1048576,wsiz=1048576`
2. `duplicity ~ /include file:///mnt/nfs`
3. `echo "marenka" > ~/include/jenicek`
4. `duplicity incr ~ /include file:///mnt/nfs`
5. `duplicity verify file:///mnt/nfs ~ /include`
6. `mkdir ~/restore`
7. `duplicity file:///mnt/nfs ~/restore`

- Na velký počet souborů a velká disková pole již tradiční způsoby nestačí
- Řešením jsou snapshoty celé datové oblasti nebo systému
- Využijeme Btrfs snapshoty a send/recieve
- Jako cíl použijeme vzdálené RBD z Ceph clusteru
- Instatní obnova
- Preview - ukázka nového přístupu k zálohování



- Copy-on-write file-systém
- Subvolumes (oddíly), snapshoty
- Checksumy dat i metadat
- RAID 0, 1, (pozor na 5 a 6), 10
- Defragmentace
- Scrubing
- Kompresce
- Přidání/odebrání disku

- send/recvie stabilní od kernelu 3.6
- btrfs-send - stream instrukcí, diff mezi dvěma subvolume
- Full a inkremental mod
- Stream instrukcí umí zpracovat btrfs-recvie (user-space)

- Vylistování subvolumů a snapshotů
  - `btrfs subvolume show /`
- Zjištění výchozího subvolume
  - `btrfs subvolume get-default /`
- Zjištění obsazeného místa, ale ...
  - `btrfs filesystem df /`

## ■ Vytvoření snapshotu

- `btrfs subvolume snapshot [-r] <source> <dest>`

## ■ Smazání snapshotu

- `btrfs subvolume delete /path`

## ■ Připojení / subvolume do cesty

- `sudo mount /dev/sdaX /mnt/btrfs_test -o rw,relatime,space_cache,subvolid=5`

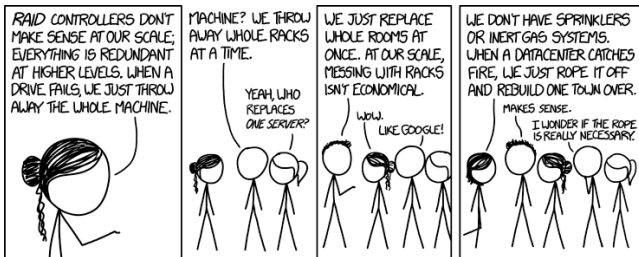
- Copy-on-write na souboru
  - `cp --reflink`
- Automatické vytváření snapshotů před instalací balíků
  - `yum-plugin-fs-snapshot.noarch`

1. Namontujte si / subvolume do cesty /mnt/btrfs
2. Vytvořte subvolume /mnt/btrfs/@snapshots
3. Vytvořte pro /home snapshot připojený do cesty /mnt/btrfs/@snapshots
4. Smažte soubor "~/.include/xen/xen.h" a následně ho zkuste obnovit ze snapshotu
5. Po úspěšné obnově smažte snapshot

Řešení:

1. `sudo mkdir /mnt/btrfs`
2. `sudo mount /dev/vda3 /mnt/btrfs -o  
rw,relatime,space_cache,subvolid=5`
3. `sudo btrfs subvolume /mnt/btrfs/@snapshots`
4. `sudo btrfs subvolume snapshot /home  
/mnt/btrfs/@snapshots`
5. `sudo rm ~/include/xen/xen.h`
6. `sudo cp  
/mnt/btrfs/@snapshots/home/labX/include/xen/xen.h  
~/include/xen/xen.h`
7. `sudo btrfs subvolume delete  
/mnt/btrfs/@snapshots/home/`

- Objektově orientované uložení
- Cluster sám udržuje minimální počet nastavených replik
- Protokoly S3/Swift, CephFS a RBD
- OSD, MON ...





- Aktuálně pilotní provoz, testy s OpenStack
- Probíhá instalace nových serverů do Jihlavy

- Ceph RADOS Block Device (RBD)
- Striping a replikace napříč clusterem
- Read-only snapshoty, revertování snapshotů
- Možnost připojit do Linuxu nebo QEMU KVM klientů
- RBD mirroring
- 10 Gbps jeden klient proti šesti serverům

- Nyní spojíme zmíněné technologie (RBD a Btrfs)
- Dostaneme možnost vytvořit zálohu subvolume/snapshotu (send/recv)
- Cílem však bude vzdálené blokové zařízení
- Zápis a případné čtení (obnova) dosahují rychlostí dostupné linky
- Provést to můžeme ručně nebo za pomoci skriptů (btrfs, snapper)

- Nainstalujeme základní Ceph balík
  - `sudo yum install ceph-common`
- Stáhněte `~/ceph.conf` z  
[https://du.cesnet.cz/\\_media/cs/ceph.conf](https://du.cesnet.cz/_media/cs/ceph.conf)  
do `/etc/ceph/`
- Zkopírujte `~/ceph.client.labX.keyring` do cesty  
`/etc/ceph/`

- Vytvoříme si image
  - `rbd -n client.labX create DU-workshop/labX -s $((100*1024)) --image-format 2 --image-feature layering`
- Zkontrolujeme zdárné vytvoření
  - `rbd -n client.labX list DU-workshop`
- Namapujeme RBD do systému
  - `sudo rbd -n client.labX map DU-workshop/labX`
- Zkontrolujeme připojení ve zprávách od kernelu
  - `dmesg`

- Blokové zařízení nejdříve zašifrujeme přes dm-crypt/LUKS
  - `sudo yum install cryptsetup-luks`
  - `sudo cryptsetup -s 512 luksFormat --type luks2 /dev/rbdX`
- Kontrola nastavení
  - `sudo cryptsetup luksDump /dev/rbdX`

## ■ Blokové zařízení dešifrujeme

- `sudo cryptsetup luksOpen /dev/rbd0 luks_rbd`

## ■ Vytvoříme Btrfs na připojeném zařízení

- `sudo parted /dev/mapper/luks_rbd`
- `mklabel gpt`
- `mkpart primary btrfs 1MiB 100%`
- `sudo mkfs.btrfs /dev/mapper/luks_rbdX`

## ■ Připojíme

- `sudo mkdir /mnt/rbd`
- `sudo mount /dev/mapper/luks_rbdX /mnt/rbd`

## ■ Odpojíme a uzamkneme

- `sudo umount /mnt/rbd/`
- `sudo cryptsetup luksClose  
/dev/mapper/luks_rbd`
- `sudo rbd -n client.labX unmap  
DU-workshop/labX`



1. Vytvořte RBD o velikosti 100GB a připojte ho do svého VM
2. Blokové zařízení zašifrujte za pomoci `dm-crypt/LUKS`
3. Vytvořte na zařízení Btrfs file-systém a připojte ho do cesty `/mnt/rbd`

Řešení:

- Nainstalujeme základní Ceph balík
  - `sudo yum install ceph-common`
- Stáhněte `~/ceph.conf` z  
`https://du.cesnet.cz/\_media/cs/ceph.conf`  
do `/etc/ceph/`
- Zkopírujte `~/ceph.client.labX.keyring` do cesty  
`/etc/ceph/`

- Vytvoříme si image
  - `rbd -n client.labX create DU-workshop/labX -s $((100*1024)) --image-format 2 --image-feature layering`
- Zkontrolujeme zdárné vytvoření
  - `rbd -n client.labX list DU-workshop`
- Namapujeme RBD do systému
  - `sudo rbd -n client.labX map DU-workshop/labX`
- Zkontrolujeme připojení ve zprávách od kernelu
  - `dmesg`

- Blokové zařízení nejdříve zašifrujeme přes dm-crypt/LUKS
  - `sudo yum install cryptsetup-luks`
  - `sudo cryptsetup -s 512 luksFormat --type luks2 /dev/rbd0`
- Kontrola nastavení
  - `sudo cryptsetup luksDump /dev/rbd0`

## ■ Blokové zařízení dešifrujeme

- `sudo cryptsetup luksOpen /dev/rbdX luks_rbd`

## ■ Vytvoříme Btrfs na připojeném zařízení

- `sudo parted /dev/mapper/luks_rbd`
- `mklabel gpt`
- `mkpart primary btrfs 1MiB 100%`
- `sudo mkfs.btrfs /dev/mapper/luks_rbdX`

## ■ Připojíme

- `sudo mkdir /mnt/rbd`
- `sudo mount /dev/mapper/luks_rbdX /mnt/rbd`

- Nástroj pro práci se subvolumes
- Vytváří inkrementální snapshoty na zadané cestě
- Možnost definovat retention policy
- Přenos na více cílů i skrze SSH
- V repu nebo na <https://github.com/digint/btrbk>

- `sudo yum install btrbk`
- Ukázka konfiguračního souboru
- `btrbk -c /etc/btrbk/btrbk.conf -v dryrun`

## Legend:

=== up-to-date subvolume (source snapshot)  
+++ created subvolume (source snapshot)  
--- deleted subvolume  
\*\*\* received subvolume (non-incremental)  
>>> received subvolume (incremental)



Předpokládáme splněný lab 5.

1. Nainstalujte btrbk a projděte jeho nastavení
2. Opět namapujte, dešifrujte a připojte RBD (viz lab 5)
3. Upravte konfigurační soubor btrbk, aby
  - target\_preserve\_min byl na all
  - zálohoval se subvolume home
  - datovou retenci zvolte dle svého uvážení
4. Zavolejte btrbk s vaším konfiguračním soubor a volbou `--dry-run run`
5. Pokud je vše v pořádku, aplikujte ho
6. Zkuste smazat souboru z `/home` a následně ho obnovit ze snapshotu

- `sudo yum install btrbk`

```
target_preserve_min    all

snapshot_dir    @snapshots
volume /mnt/btrfs
    subvolume    home
        snapshot_create    ondemand
        target send-receive    /mnt/rbd/
```

- Ostatní data retention pravidla upravte dle svého uvážení nebo ponechte ve výchozím stavu

- `btrbk -c /etc/btrbk/btrbk.conf -v dryrun`
- `btrbk -c /etc/btrbk/btrbk.conf -v run`



Děkujeme za účast na workshopu!

[www.du.cesnet.cz](http://www.du.cesnet.cz)

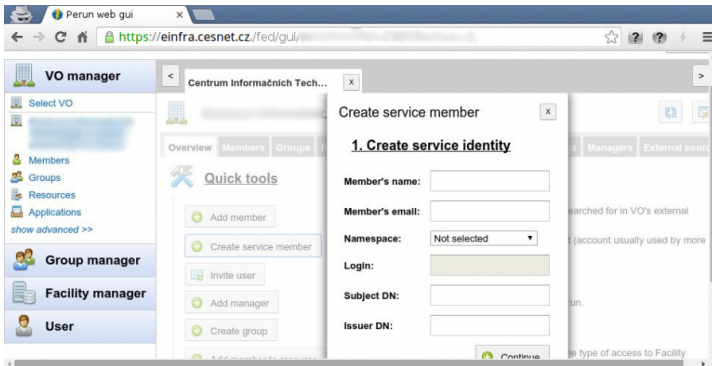
[du-support@cesnet.cz](mailto:du-support@cesnet.cz)





**Záložní slidy**





Perun web gui

<https://einfra.cesnet.cz/fed/gul/>

**VO manager**

- Select VO
- Members
- Groups
- Resources
- Applications
- show advanced >>
- Group manager**
- Facility manager
- User

Create service member

**2. Associate real users**

**To be associated:**

Name	User type	Count: 0
No items found.		

Type in user's First name, Last name (or both) or Login or Email and press Search button.