cesnet

ZÁLOHOVÁNÍ S DÚ

aneb sto způsobů jak (ne)přijít o data

Michal Strnad

CESNET, z. s. p. o.

30. 1. 2019

Agenda

- Datová úložiště a motivace pro workshop
- Popis prostředí pro hands-on
- Pohádka o sysadminovi

cesnet Datová úložiště jedním slidem

- Zajišťuje provoz a rozvoj národní infrastruktury pro ukládání dat pro vědu a výzkum
- Aktuálně provozujeme čtyři uložiště 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 uložiště pak přes S3/Swift, CephFS, RBD
- Další služby jako FileSender a ownCloud

Michal Strnad, CESNET,
CESNET, z. s. p. o.

Popis prostředí pro hands-on

- Virtuál s CentOS 7
- Login, heslo a IP adresu je na papírů
- Privilegovaný přístup na stroj přes sudo
- Přístup na úložiště CESNET přes servisní účet

V případě problémů s jednotlivými kroky Vám pomůžeme. Slidy, návody a další materiály

`https://du.cesnet.cz/cs/workshop`

Michal Strnad, CESNET,
CESNET, z. s. p. o.

Adřesářová struktura

```
/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

Pohádka o sysadminovi

- Jednorázová záloha či replika
  - Nástroj rsync
- Chceme spíše kontinuální zálohu
  - Použijeme nástroj duplicity
  - Provedeme obnovu a následnou kontrolu integrity dat
  - Pro uživatele i správce
- Potřebujeme mít snapshoty celé datové oblasti nebo systému
  - Využijeme Btrfs snapshoty a send/recieve
  - Jako cíl použijeme vzdálené RBD z Ceph clusteru
  - Primárně pro správce

Jednorázová záloha či replika

Rsync

- Prakticky na všech distribucích již předinstalován
- Pod sebou má SSH
- Používá ho velká část jiných aplikací (např. rsnapshot)
- Možnost skriptování
- Existuje grafická nadstavba Grsync

Plný návod

`https://du.cesnet.cz/cs/navody/rsync/start`

Rsync

- Základní použití nástroje
  - rsync -av –progress ./folder labX@ssh.du4.cesnet.cz: /VO_du_test-disk_only/
- Umožňuje zachovávat práva a informace o vlastníkovi souboru při přenosu na vzdálené úložiště v rozšířených atributech souborů
  - rsync -av –numeric-ids –rsync-path=”rsync –fake-super”./folder labX@ssh.du4.cesnet.cz: /VO_du_test-disk_only/

Rsync - háčky

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

Lab 1 - rsync

1. Z webu `https://du.cesnet.cz/cs/workshop` stáhněte testovací data a rozbalte archív
2. Připojte se 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 kernel
5. Smažte soubor "jedec_ddr.h"v adresáři ve svem VM (include/memory/jedec_ddr.h)
6. Zavolejte rsync s –delete a ověřte, zdali se smazal i na straně serveru (CESNET úložiště)

Lab 1 - rsync

Řešení:

1. mkdir data; cd ./data
2. wget https://goo.gl/YncDuX -O kernel.tar.xz
3. tar -xf kernel.tar.xz
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/
6. rm ./include/memory/jedec_ddr.h
7. rsync -av –progress –delete ./include labX@ssh.du4.cesnet.cz: /VO_du_test-disk_only/backup/

NFS s Duplicity

Chceme spíše kontinuální zálohu. My zde zvolíme NFS s Duplicity. Začněme s NFS.

- Umožnuje připojení vzdáleného svazku jako lokální disk
- Reconnect, nemá problémy s hardlinky
- Ve výchozím nastavení posílá data v clear textu, autentizace skrze /exports
- Ukážeme si případ s Kerberos autentizací

Plný návod
`https://du.cesnet.cz/cs/navody/nfs/start`

- Nainstalujeme základní nástroje a knihovny pro NFS + podporu Kerberos
  - sudo yum install nfs-utils nfs-utils-lib krb5-workstation
- Připravíme si adresář kam následně přimountujeme úložiště du4
  - mkdir /mnt/nfs
  - chown -R labX /mnt
- Stáhneme si konfigurační soubor pro Kerberos autentizaci
  - 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
  - sudo cp krb5.keytab /etc/krb5.keytab
  - sudo chmod 0600 /etc/krb5.keytab

- 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

Zkontrolujeme, že nám běží všechny služby, případně je zapneme

- systemctl enable nfs-idmap; systemctl enable nfs-idmap; systemctl status nfs-idmap
- systemctl enable nfs-secure; systemctl start nfs-secure; systemctl status nfs-secure

Nyní jsme hotovi a můžeme připojit NFS svazek

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

Lab 2 - NFS

1. Získejte keytab umístěte ho do správné cesty
2. Připojte NFS svazek do adresáře /mnt/nfs-du4
3. Nastavte statické mapování a ověřte jeho funkčnost

Michal Strnad, CESNET,
CESNET, z. s. p. o.

Lab 2 - NFS

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

Lab 2 -NFS

- 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
  - sudo cp krb5.keytab /etc/krb5.keytab
  - sudo chmod 0600 /etc/krb5.keytab

Lab 2 - NFS

- 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

Lab 2- NFS

Zkontrolujeme, že nám běží všechny služby, případně je zapneme

- systemctl enable nfs-idmap; systemctl enable nfs-idmap; systemctl status nfs-idmap
- systemctl enable nfs-secure; systemctl start nfs-secure; systemctl status nfs-secure

Lab 2 - NFS

Nyní jsme hotovi a můžeme připojit NFS svazek

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

Duplicity

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

Duplicity

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

Duplicity

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

Duplicity

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

Duplicity

- Inkrementální záloha
  - duplicity incr ./kernel file:///mnt/nfs
- Přehled o provedených zálohách
  - duplicity collection-status file:///mnt-nfs
- Kontrola zálohy proti lokálním souborům
  - duplicity verify file:///mnt-nfs ./folder
- Staré neaktuální zálohy je třeba smazat a nahradit novými
  - duplicity remove-older-than 12M –force file:///mnt/nfs

Lab 3 - Duplicity nad NFS

Předpokládáme splňený Lab 2 (připojené NFS).

1. Použijte duplicity na vytvoření plné zálohy adresáře include do cesty /mnt/nfs
2. Vytvořte nový soubor s názvem "jenicek"a vložte do něj text "marenka"
3. Proveďte přes duplicity inkrementální zálohu do stejné cesty (/mnt/nfs)
4. Zkontrolujte zálohu za pomoci duplicity

Lab 3 - Duplicity nad NFS

Řešení:

1. duplicity ./include file:///mnt/nfs
2. echo "marenka»> ./include/jenicek
3. duplicity incr ./include file:///mnt/nfs
4. duplicity verify file:///mnt-nfs ./include

Btrfs nad RBD z Cephu

- Preview - ukázka nového přístupu k zálohování
- Potřebujeme mít snapshoty celé datové oblasti nebo systému
- Využijeme Btrfs snapshoty a send/recieve
- Jako cíl použijeme vzdálené RBD z Ceph clusteru
- Instatní obnova

Michal Strnad, CESNET,
CESNET, z. s. p. o.

cesnet

Btrfs

- Copy-on-write file-systém
- RAID 0, 1, (pozor na 5 a 6), 10
- Online defragmentace
- Scrubing
- Komprese
- Subvolumes (oddíly), snapshoty
- Přidání/odebrání disku

Btrfs vs rsync

- rsync pracuje na souborové úrovni
- rsync nedetekuje přejmenovaný nebo přesunutý soubor
- Časy modifikací nebo velikost, případně hashe -> režie
- Výhoda v podobě exclude

- Btrfs pracuje na blokové úrovni
- Copy On Write
- Subvolume - vše nebo nic

Btrfs

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

Btrfs

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

Btrfs

- Vylistování subvolumů a snapshotů
  - btrfs subvolume show /
- Vytvoření snapshotu
  - btrfs subvolume snapshot [-r] <source> <dest>
- Smazání snapshotu
  - btrfs subvolume delete /path

Btrfs - tipy a triky

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

Lab 4 - Btrfs

1. Založte nový subvolume na cestě /btrfs
2. Nahrajte do něj rozbalený kernel viz Lab 1
3. Vytvořte pro tento subvolume snapshot připojený do cesty /btrfs/snapshot
4. Smažte soubor "/btrfs/include/xen/xen.h"a následně ho zkuste obnovit ze snapshotu
5. Po úspěšné obnově smažte snapshot

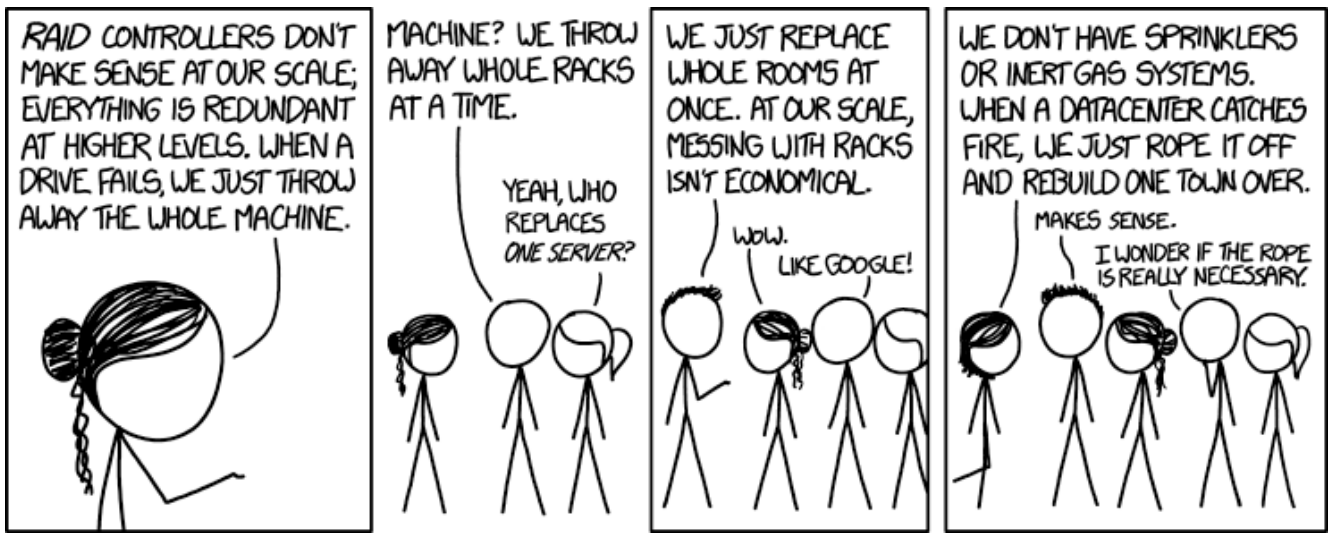
cesnet

Lab 4 - Btrfs

1. sudo btrfs subvolume create /btrfs
2. sudo cp -r ./include /btrfs/
3. sudo btrfs subvolume snapshot /btrfs/ /btrfs/snapshot
4. sudo rm /btrfs/include/xen/xen.h
5. sudo cp /btrfs/snapshot/include/xen/xen.h /btrfs/include/xen/xen.h
6. sudo btrfs subvolume delete /btrfs/snapshot/

Ceph

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



Ceph a DU CESNET

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

Ceph RBD

- 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

Btrfs nad RBD z Cephu

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

Příprava prostředí

- Nainstalujeme základní Ceph balík
  - sudo yum install ceph-common
- Přesuneme soubor ceph.conf do /etc/ceph/ (z balíku xyz)
- Následně i labX.keyring do cesty /etc/ceph/labX.keyrinig (z balíku xyz)

Ceph RBD

- Vytvoříme si RBD
  - rbd -n client.lab1 create DU-workshop/lab1 -s $((10*1024)) –image-format 2 –image-feature layering
- Zkontrolujeme zdárné vytvoření
  - rbd -n client.lab1 list DU-workshop
- Namapujeme RBD do systému
  - sudo rbd -n client.lab1 map DU-workshop/lab1
- Zkontrolujeme připojení ve zprávách od kernelu
  - dmesg

Ceph RBD, LUKS

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

Ceph RBD, btrfs

- Blokové zařízení dešifrujeme
  - sudo cryptsetup luksOpen /dev/rbd0 lusk_rbd
- Vytvoříme Btrfs na připojeném zařízení
  - sudo mkfs.btrfs /dev/lusk_rbd
  - comment: parted /dev/sdd
  - comment: mklabel gpt
  - comment: mkpart primary btrfs 1MiB 100%
- Připojíme
  - mkdir /mnt/rbd
  - sudo mount /dev/lusk_rbd /mnt/rbd

Ceph RBD, btrfs

- Cvičení - montovani btrfs - montování root btrfs
  - btrfs subvolume list
  - mkdir /mnt/localbtrfs
  - mount
- Výroba subvolume pro snapshoty
  - btrfs subvolume create /mnt/localbtrfs/@snapshots
  - btrfs subvolume list /mnt/localbtrfs/
  - mount /dev/sdaX /mnt/localbtrfs/ -o rw,relatime,seclabel,space_cache

Ceph RBD, btrfs

- Odpojíme a uzamkneme
  - umount /mnt/rbd/
  - sudo cryptsetup luksClose /dev/mapper/lusk_rbd
  - sudo rbd -n client.lab1 unmap DU-workshop/lab1

Lab 5 - Btrfs nad RBD

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 /btrfs
4. Udělejte nový snapshot a za pomoci send/recieve ho zapište do /btrfs

Michal Strnad, CESNET,
CESNET, z. s. p. o.

cesnet Lab 5 - Btrfs nad RBD, řešení

Řešení:

1. Opsat z predchozich slidu, jakmile overime jejich funkcnost

Michal Strnad, CESNET,
CESNET, z. s. p. o.

cesnet

btrbk

- 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


cesnet
Děkujeme za účast na workshopu!
www.du.cesnet.cz
du-support@cesnet.cz

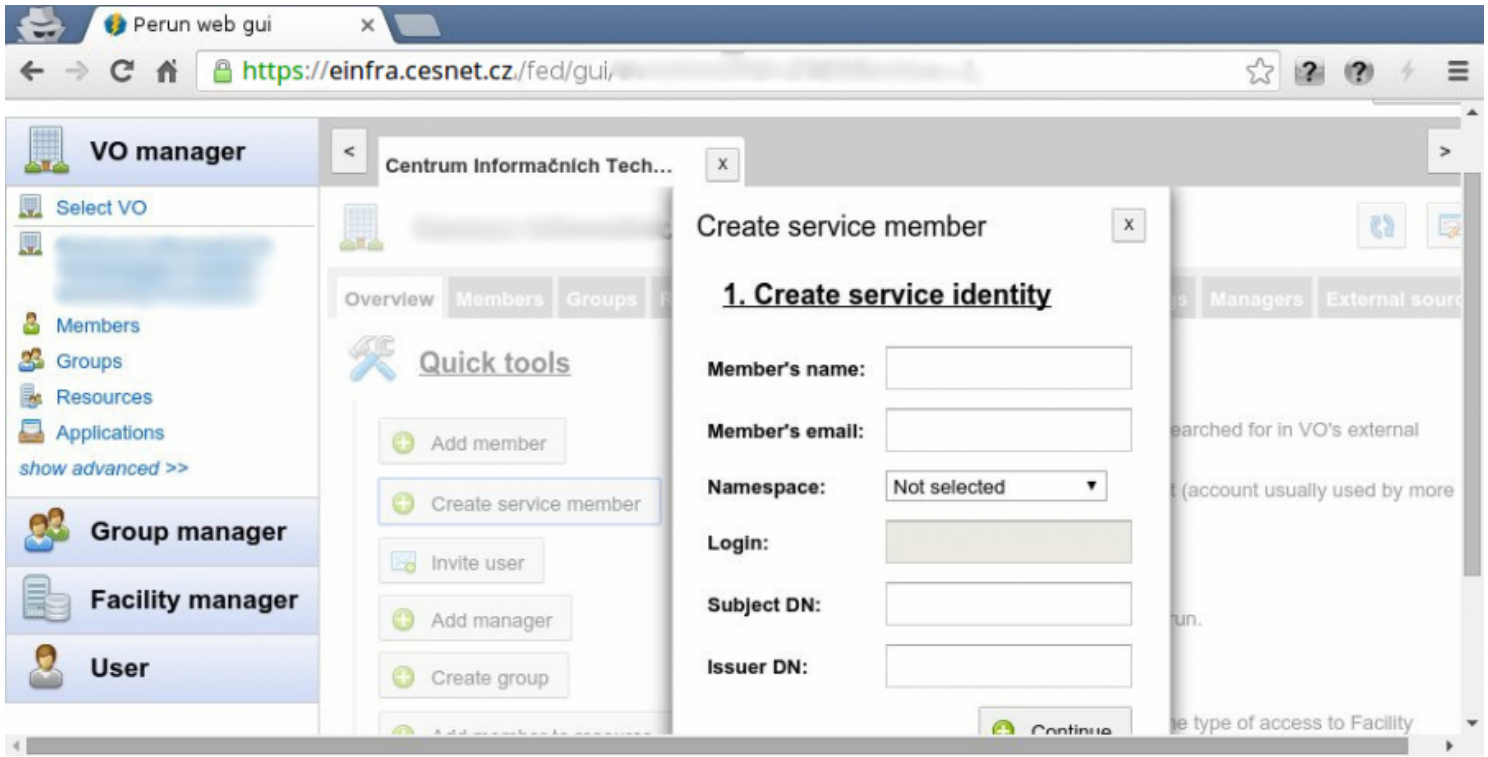
cesnet

Záložní slidy

Best practices

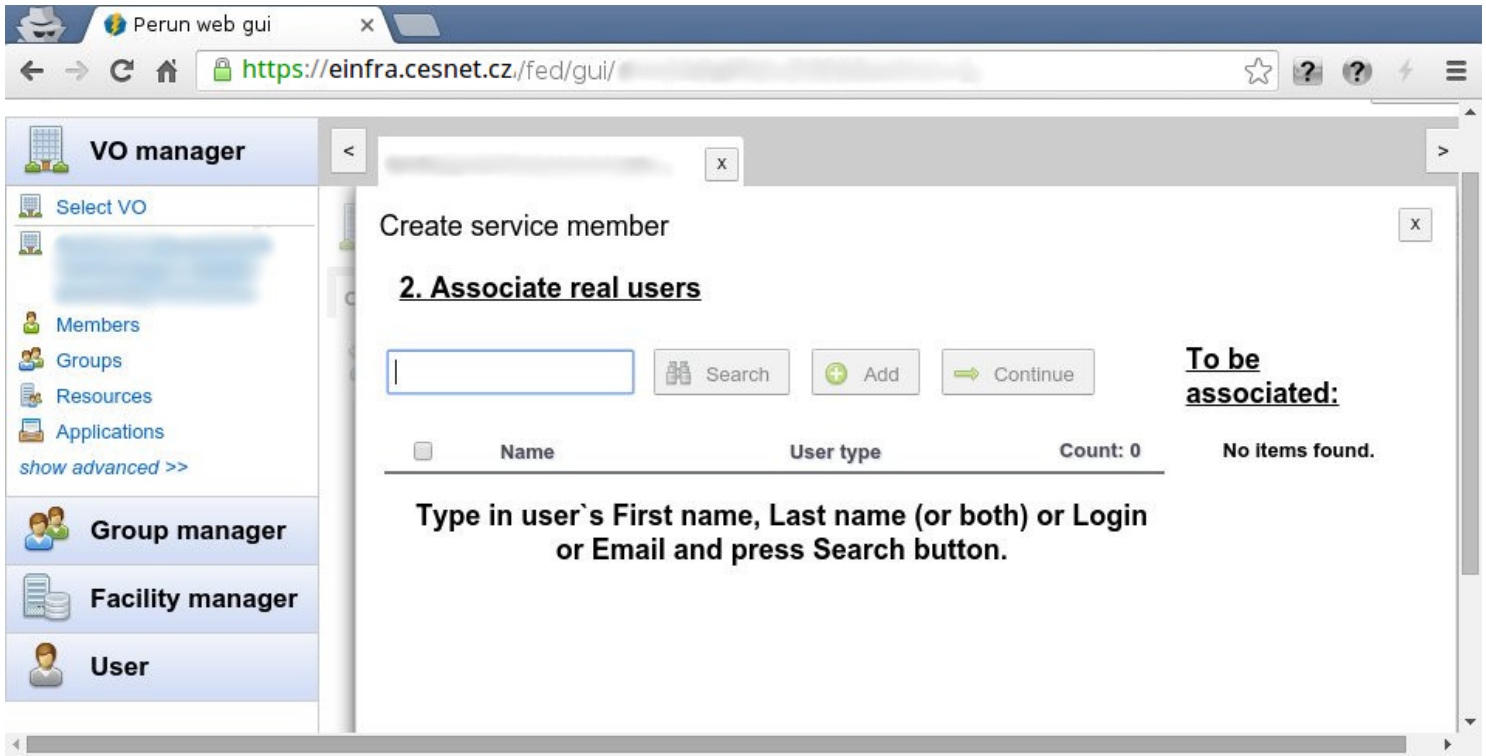
- Report provedených i neprovedených záloh (monitoring)
- Šifrujte zálohy pro zajištění důvěrnosti dat
- V případě přenosu dat po síti, kterou nemáte pod kontrolou, šifrujte i přenos dat
- Kontrolujte zálohy (verifikace)
- Označte si zálohy - co obsahují, datum vytvoření
- Ukládejte zálohy na různá místa. Pokud by došlo k lokální katastrofě (např. požár), tak pravděpodobně ztratíte zálohy zde uskladněné

Servisní účet



Michal Strnad, CESNET,
CESNET, z. s. p. o.

Servisní účet



Michal Strnad, CESNET,
CESNET, z. s. p. o.