Datová úložiště CESNET

Michal Strnad

5. 10. 2014

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

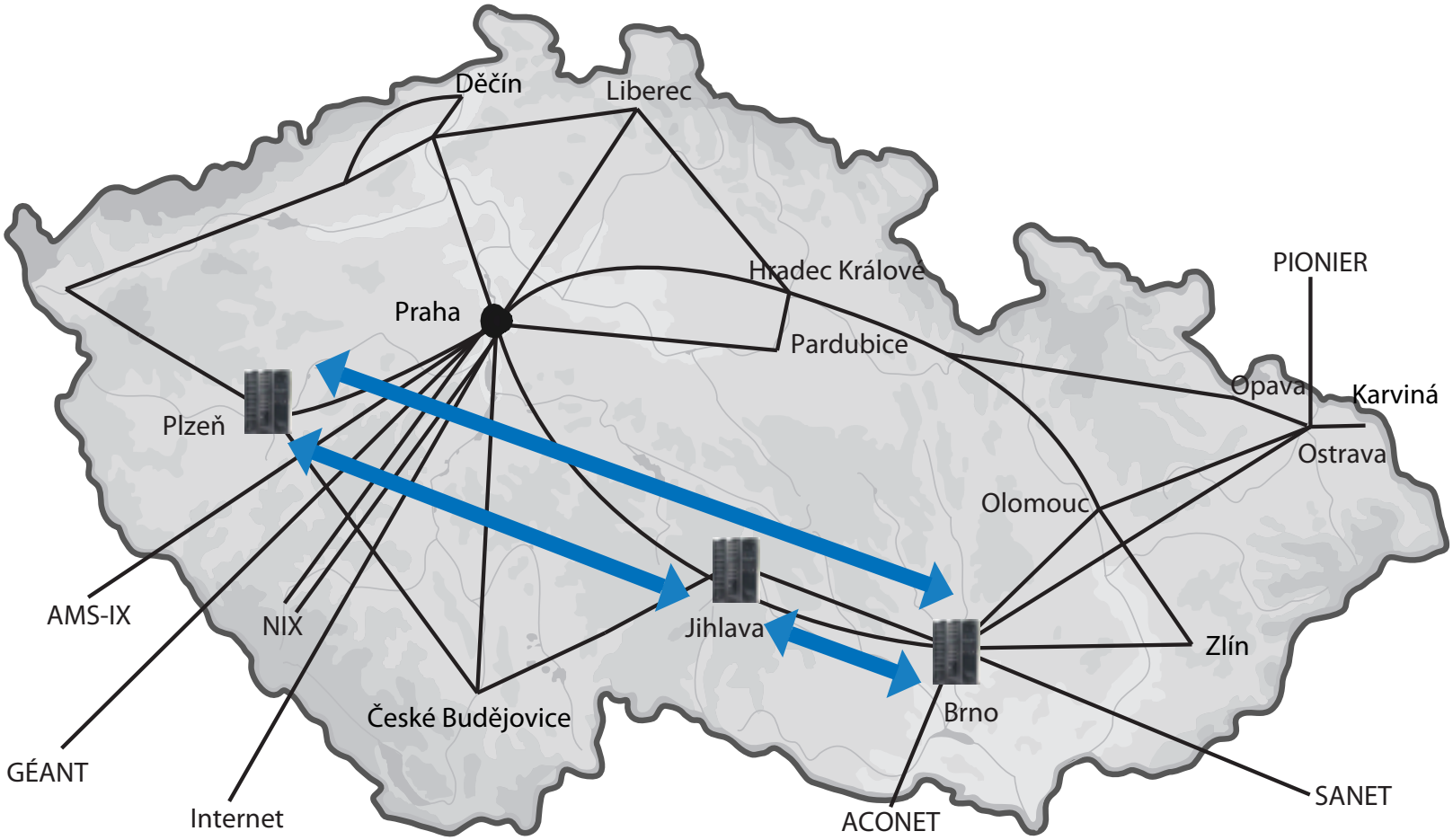
EVROPSKÁ UNIE
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
INVESTICE DO VAŠÍ BUDOUCNOSTI

2007–13
OP Výzkum a vývoj pro inovace

Agenda

- Infrastruktura
- Jak úložiště použít
- Přístupové protokoly
- ownCloud

Infrastruktura



Možnosti využití datových úložišť

- zálohy
  - uživatelé mají primární data u sebe
  - na úložiště odkládají zálohu pro případ havárie
- archivace
  - uživatelé na úložiště odkládají cenná primární data
  - data nejsou často využívána
  - uživatelé nemají prostředky pro jejich uchovávání
- sdílení dat
  - distribuovaný tým potřebuje společně pracovat nad většími objemy dat, případně je zveřejňovat
- „něco jiného“
  - distribuce obsahu, jiné speciální aplikace

Možnosti využití datových úložišť

- a naopak: na co se vzdálené úložiště příliš nehodí
  - interaktivní práce zejména s velkým množstvím malých souborů
  - ukládání dat s potřebou přístupu v reálném čase
    - prioritou je spolehlivost uložení

Přenos dat na úložiště

- souborové
  - NFSv4 (známé uživatelům MetaCentra)
  - rsync, scp, SFTP
  - FTPS
  - obvykle autentizované systémem Kerberos
  - existují GUI klienti i nástroje pro příkazovou řádku pro hlavní platformy
  - Globus Online – vysokorychlostní přenosy
  - výhledově CIFS (známý „síťový disk“ z Windows)
- speciální
  - gridové úložiště v systému dCache
  - FileSender
  - ownCloud

FileSender

- webová služba pro přenos (velkých) souborů
  - velkých: aktuálně 500 GB (pouze s HTML5)
- přispíváme i k jeho vývoji
- `http://filesender.cesnet.cz`
- alespoň jedna strana komunikace musí být oprávněný uživatel infrastruktury
  - autentizace federací eduID.cz
- oprávněný uživatel může nahrát soubor a poslat mu oznámení
- lze poslat komukoli pozvánku
- testovací verze na `http://filesender2.cesnet.cz`

Jak se stát uživatelem – snadno

- pokud
  - chcete využívat souborově orientované protokoly
  - dostačuje prefabrikovaná politika migrací dat
  - ukládáte jen individuální data
  - potřebujete ukládat nejvýše jednotky TB
- tak se jen zaregistrujte na `http://du.cesnet.cz`
  - vyžaduje to ověření uživatele z akademické instituce
- členství se po roce prodlužuje
  - jako indikace, že uživatel „ještě ví o svých datech“
- kódové označení: VO Storage

Úložiště pro náročnější

- nestačí VO Storage? založíme vám virtuální organizaci!
- VO je skupina uživatelů se společným zájmem, kteří vystupují jako celek
- VO má správce, který
  - jedná s poskytovatelem zdrojů
  - rozhoduje o podmínkách členství jednotlivých uživatelů
- dohoda o poskytování zdrojů mezi VO a správcem zdroje
  - popisuje nastavení technických parametrů
    - velikost poskytovaného prostoru
    - použité přístupové protokoly
    - politika ukládání dat, počty kopií/replik
  - VO může využívat různé typy zdrojů

Založení VO prakticky

- napište nám na `du-support@cesnet.cz`
- domluvíme se, co potřebujete
  - opravdu to potřebujeme pochopit
  - kolik dat máte, jak s nimi pracujete, ...
- připravíme konfiguraci úložiště
- založíme virtuální organizaci
- pozveme vás do ní a povýšíme na administrátora
- dál si uživatele spravujete sami, vytváříte skupiny, ...
- ukládáte data přes FTP/rsync/NFSv4/...
  - my se staráme, aby na se vytvářely jejich repliky, ...

Kam se po registraci připojit?

Aktuální adresy serverů (hostů) pro jednotlivá datacentra:
ssh.duX.cesnet.cz
ftp.duX.cesnet.cz
nfs.duX.cesnet.cz

X nahraďte podle následujího klíče:
Plzeň -> 1
Jihlava -> 2
Brno -> 3

Adřesářová struktura

Příklad pro uživatele „user“ s členstvím ve dvou VO (např. VO_cern a VO_storage), kde každá z nich používá dvě migrační politiky:

```
/home/user:
|-- .bash_history
|-- .ssh
|   \-- authorized_keys
|-- VO_cern-tape_tape -> /exports/tape_tape/VO_cern/home/user
|-- VO_cern-disk_only -> /exports/disk_only/VO_cern/home/user
|-- VO_storage-tape_tape -> /exports/tape_tape/VO_storage/home/user
\-- VO_storage-cache_tape -> /exports/cache_tape/VO_storage/home/user
```

Migrační politiky

- tape - SATA -> MAID
- tape_tape -> tape + záloha na pásku
- disk_only - data jsou jen na discích, nikdy se nepřesouvají
- cache_tape - FC disky, jedna kopie na páskách a SATA discích

ownCloud

- Nejen webový přístup k vašim datům
- Cloudové úložiště se synchronizací souborů

K čemu je dobrá další cloudová služba?

Use-cases

- Distribuovaný tým potřebuje společně pracovat s daty, případně je zveřejňovat (sdílení formou "kdo zná link")
- Automatická synchronizace mezi počítači a mobilními zařízeními. Na počítači jsou i lokálně, na mobilní při otevření
- Poskytuje platformu pro sdílení kontaktů, kalendářů a úkolů
- Vhodnější pro menší objemy dat

Under the Hood

- PHP + JavaScript
- WebDAV, CardDAV, CalDAV
- PostgreSQL + Apache

Klientské aplikace

- Pro vybrané platformy (aktuálně Windows, OS X, Linux, Android, iOS) je k dispozici nativní klient, který zpřístupní vaše data z ownCloud na daném zařízení
- Pro použití desktopového klienta je nutné si nastavit pro něj heslo

ownCloud u CESNETu

- Uložení dat organizované standardním způsobem do složek a souborů s celkovou kapacitou 100 GB
- `https://owncloud.cesnet.cz`
  - zdarma, jen je nutné se autentizovat pomocí federace eduID.cz
  - není nutné se registrovat, účet je vám automaticky založen při prvním přihlášení
- Služba byla spuštěna 26. 2. 2014
- Data jsou denně zálohována a služba běží v režimu high-availability

Známé problémy

- Přihlásil jsem se přes jiného poskytovatele identit, než při svém prvním přihlášení a nevidím svá data
- Nahrával jsem soubor přes webové rozhraní a po skončení uploadu ho nikde nevidím
- V Internet Explorer 11 nefunguje správně JavaScript

ownCloud - demo

Quick demo

Sumarizace

- Single účet vs. separátní VO
- FileSender
- OwnCloud

Q and A

- `http://du.cesnet.cz`
- uživatelská podpora: `du-support@cesnet.cz` (RT systém)