

Datová úložiště CESNET

David Antoř

Přehled

- ▶ proč má smysl budovat národní datová úložiště pro vědecká data
- ▶ budovaná infrastruktura
- ▶ jak úložiště použít
- ▶ přístupové mechanismy
- ▶ správa uživatelů na úložištích
- ▶ použití pro náročnější uživatele
- ▶ speciální aplikace

Smysl externího úložiště

- ▶ uživatelé raději drží data na vlastním zařízení. . .
 - ▶ to ale nedává rozumné garance dlouhodobého uchování
 - ▶ používají se nevhodná média, chybí systematická péče
 - ▶ data je lépe uchovávat v živém systému
- ▶ konsolidace zdrojů – národní úložiště vědeckých dat
 - ▶ umožňuje pořídit zařízení dostatečné velikosti
 - ▶ spravované specializovaným týmem
 - ▶ s možností sdílení dat mezi skupinami uživatelů

Cíl: dlouhodobé uchování dat

- ▶ potřeba dlouhodobě uchovávat vědecká data vzrůstá
 - ▶ uchování primárních dat z experimentů a výpočtů
 - ▶ s ohledem na potenciál jejich dalšího využití
 - ▶ k návratu a revizi publikovaných výsledků
- ▶ uchování zajišťujeme na úrovni zachování binárních dat (bitstream)
 - ▶ správce úložiště nerozumí odborné povaze dat
 - ▶ ani nemůže rozhodovat o smysluplnosti jejich archivace
 - ▶ proto neděláme LTP (Long-Term Preservation)
 - ▶ bitstream je nicméně předpokladem pro uchování dat
 - ▶ další kopie, geografické repliky, redundance v systému

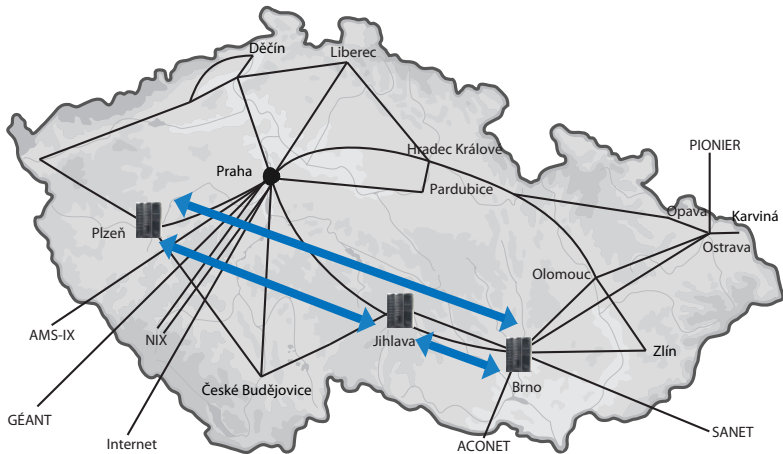
Možnosti využití datových úložišť

- ▶ 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 uchování
- ▶ 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
 - ▶ přímý provoz systémů s nutností 24/7 přístupu
 - ▶ „toto není úložiště pod váš SAP“
 - ▶ na to je obecně vzdálené úložiště nevhodné
- ▶ vhodné použití pro podobné scénáře
 - ▶ primární 24/7 data u uživatele, na CESNET zálohy/repliky
 - ▶ rozdělení životně důležitých dat vs. „nice to have“ dat

Infrastruktura



Infrastruktura

- ▶ trojice úložišť: Plzeň, Jihlava, Brno
- ▶ celková hrubá kapacita cca 21 PB
- ▶ Plzeň v rutinním provozu
 - ▶ v areálu ZČU
 - ▶ cca 500 TB disků, 4800 TB pásek, 3584 TB SGI COPAN
 - ▶ SGI Data Migration Facility, CXFS
- ▶ Jihlava a Brno v pilotním provozu
 - ▶ Krajský úřad Jihlava, rektorát VUT Brno
 - ▶ Jihlava 1041 TB disků, 3744 TB pásek, 2934 TB MAID
 - ▶ Brno 498 TB disků, 3500 TB pásek, 2116 TB MAID
 - ▶ řešení IBM založené na GPFS a TSM

Knihovna v Plzni



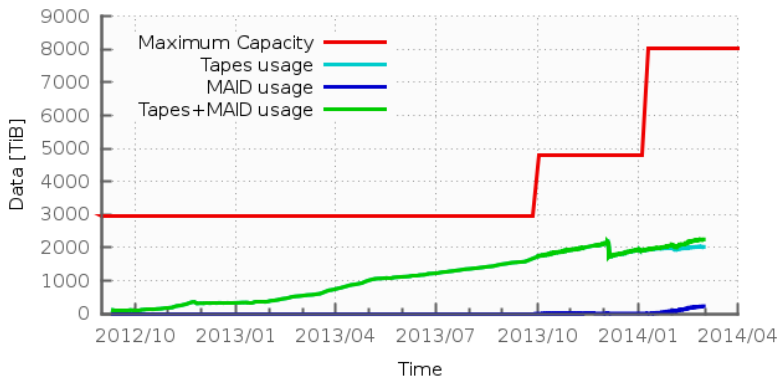
Úložiště jsou hierarchická

- ▶ vrstvy médií různé kapacity a rychlosti
 - ▶ rychlé disky/pomalejší disky/MAID/pásky
 - ▶ drahý provoz → levnější provoz
- ▶ a automatizovaný systém pro přesuny dat mezi nimi
 - ▶ déle nepoužívaná data do pomalejších vrstev
 - ▶ pro uživatele transparentní, resp. téměř transparentní
 - ▶ přístup k dlouho nepoužitému souboru trvá déle
 - ▶ uživatelům to obvykle nevadí

Zaplnění archivu v Plzni

Data stored on tapes/MAID

CESNET Pilsen storage, 2014-03-02 06:16:02



Jak na úložiště přenášet data – motivace

- ▶ úložiště musí podporovat řadu přístupových protokolů
 - ▶ možnosti a potřeby uživatelů nelze shrnout pod jedno univerzální
- ▶ od protokolů poskytujících souborový systém
 - ▶ dává aplikaci iluzi lokálního disku
 - ▶ ale platí se výkonem (za bohatost POSIX rozhraní)
- ▶ přes jednoduchou put/get sémantiku
 - ▶ použitelné pro stage-in/out práci nebo speciální aplikace
- ▶ speciální rozhraní
 - ▶ webové, synchronizační aplikace

Přenos dat na úložiště

- ▶ souborové
 - ▶ NFSv4 (známé uživatelům MetaCentra)
 - ▶ rsync, scp, 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 aplikace
 - ▶ gridové úložiště v systému dCache
 - ▶ FileSender
 - ▶ ownCloud
- ▶ bloková zařízení (není preferováno)

FileSender

- ▶ webová služba pro jednorázový přenos (velkých) souborů
 - ▶ „velkých“: aktuálně 500 GB
- ▶ 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>

ownCloud

- ▶ cloudové úložiště se synchronizací souborů
 - ▶ pro osobní počítače (Windows, Linux, Mac OS X)
 - ▶ mobilní Android, Apple (klient pod 1 Euro)
 - ▶ a webové rozhraní
- ▶ data se synchronizují přes úložiště
 - ▶ na počítači jsou i lokálně, na mobilní při otevření
- ▶ data lze sdílet
 - ▶ konkrétní osobě nebo „kdo zná odkaz“
- ▶ kalendář, kontakty
- ▶ registrace federací na <http://owncloud.cesnet.cz>
- ▶ standardní limit 100 GB na uživatele

Další speciální protokoly/aplikace

- ▶ GlobusOnline
 - ▶ „klikací FTP na steroidech“
 - ▶ kopírování velkého objemu dat
 - ▶ řízené pomocí webového rozhraní
 - ▶ mezi „koncovými body“ – úložišti podporujícími GlobusOnline nebo lokálním strojem
- ▶ dCache
 - ▶ jeden ze systémů správy dat vyvinutých pro CERN
 - ▶ pro ukládání PB objemů

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

Jak se stát uživatelem – snadno 2

- ▶ pokud
 - ▶ potřebujete (třeba okamžitě) i výpočetní zdroje
 - ▶ a vyhovují podmínky podobné jako ve VO Storage
- ▶ tak se zaregistrujte do MetaCentra na <http://metavo.metacentrum.cz/>
- ▶ jestli to nestačí, tak...

Jak se stát uživatelem – snadno 2

- ▶ pokud
 - ▶ potřebujete (třeba okamžitě) i výpočetní zdroje
 - ▶ a vyhovují podmínky podobné jako ve VO Storage
- ▶ tak se zaregistrujte do MetaCentra na <http://metavo.metacentrum.cz/>
- ▶ jestli to nestačí, tak... ale až za chvíli

Správa uživatelských účtů

- ▶ systém Perun
- ▶ uživatel (osoba) má (ideálně) jedinou identitu v infrastruktuře
- ▶ potřebujeme ověření identity a příslušnosti k oprávněné instituci
 - ▶ proč to? zodpovídáme v projektu za to, komu služby poskytujeme (akademické komunitě), řídíme mnoho různých služeb
- ▶ na identitu uživatele jsou navázány jednotlivé mechanismy přístupu ke zdrojům
- ▶ řízeno členstvím ve virtuálních organizacích
- ▶ uživatelé vše řídí přes webové rozhraní

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, ...

Aktuální stav

- ▶ úložiště v Plzni v běžném provozu
- ▶ přichází na něj uživatelé
 - ▶ individuální se mohou přihlásit do VO Storage
 - ▶ připojeno do MetaCentra
- ▶ úložiště v Jihlavě a Brně v provozu
- ▶ FileSender v provozu, ownCloud v pilotním provozu
- ▶ postupně obsluhujeme zájemce o náročnější využívání
 - ▶ vždy vyžaduje komunikaci
 - ▶ pro velká data (desítky TB+): nezbytná technická znalost uživatelů

Výhled

- ▶ cílem je udržet úložiště s otevřeným přístupem pro vědeckou komunitu
- ▶ ideové pilíře rozvoje
 1. technologie a její provoz
 2. organizace dat (Hadoop, databáze)
 3. aplikace – analýza velkých dat – včetně schopnosti fungovat jako testovací prostředí pro takové aplikace
 4. dlouhodobé uchování dat – poskytnutí nástrojů pro práci „aplikačního knihovníka“
- ▶ body 2 a 3 v úzké spolupráci s národním gridem MetaCentrum

Kontakty

- ▶ <http://du.cesnet.cz>
- ▶ uživatelská podpora: du-support@cesnet.cz