

Výzkumná data

Yeast Colony Group

technické řešení

Potřeby

- převod 1-2 TB souborových dat z lokálního úložiště do centralizované infrastruktury
- velké množství malých a středně velkých souborů
- předpoklad intenzivní práce se soubory, dostupných formou “Windows share”
- řízení přístupu na základě IP adres ve vnitřní síti
- zálohování dat bez přímých finančních nákladů
- perspektivně přesun do infrastruktury pracoviště BIOCEV

Omezení

- přístupová práva dle IP adres: DÚ CESNET tento režim nepodporuje
- potřeba intenzivní práce s daty a přístup jen z vnitřní sítě fakulty - rychlejší ve vnitřní síti
- Windows share - DÚ CESNET podporuje jiné protokoly

Řešení sdílení dat

- virtuální server (Linux)
- Samba (SMB/CIFS ve virtualizované konfiguraci, určeno pro více projektů, nabízené služby se rozlišují na základě CNAME, každý může mít jiné nastavení serveru, zde konkrétně probíhá ověření ke konkrétním “shares” jednotnými přihlašovacími údaji jen z povolených IP adres)

Řešení zálohování

- vyhrazená složka v DÚ CESNET, připojená protokolem NFSv4 + Kerberos autentizace
- opensource zálohovací nástroj duply/duplicity jako prozatimní řešení vybrané na základě srovnání několika testovaných nástrojů
- plná záloha 1x měsíčně, přírůstky 1x denně, retence 3 měsíce

Zkušenosti z pohledu CIT

- k 11.5.2015 se využívá 1.03 TB uložených dat v 1.1 milionu produkčních souborů (2.34 TB v zálohách)
- bezproblémový chod na straně Samba serveru a fakulního centralizovaného úložiště - uživatelé vidí funkční řešení
- zálohování do DÚ CESNET není optimální: plná záloha trvá 30+ hodin (je třeba posílit hw a nalézt výkonnější zálohovací software, což je u opensource obtížné), obnova dat vyžaduje asistenci CIT - data jsou zálohována, řešení bude ale třeba zásadně vylepšit

Co dál ?

- vylepšit zálohování
- zaměřit se na nové potřeby týmu YCG

Zhodnocení

- máme modelové řešení pro další zájemce
- alternativa k přímému ukládání dat do DÚ CESNET v odůvodněných případech (nízkolatenční přístup k živým datům se zálohou do vzdálené lokality, složitější přístupové politiky)
- ověření NFS jakožto stabilního přístupového protokolu vůči DÚ CESNET