

# Infrastrukturális támogatás adatrepozitóriumok létrehozására

Pintér Ádám

Wigner Fizikai Kutatóközpont

ELKH ARP projekt

hardveres munkacsoport vezetője

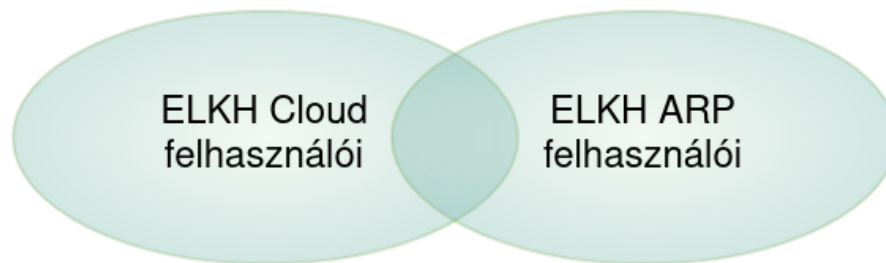


# Célok és kihívások

- Kutatástámogatás – ritkán változó nagy adatmennyiség megőrzése későbbi felhasználás céljából
- CONCORDA (Concentrated Cooperation on Research Data) példányok
- Infrastruktúra biztosítása (hardveres és szoftveres követelmények)
- Közös projekt (SZTAKI, Wigner FK, Társadalomtudományi Kutatóközpont)
- Két intézet (SZTAKI és Wigner FK) fog otthont adni hardvereknek, köztük adatok szinkronizálását biztosítani szükséges
- Üzemeltetési kihívások
- Kapacitás bővíthetősége fontos szempont a jövőre nézve
- Technológiai fejlődés és változások mellett adatok tárolásának biztosítása

# ELKH Cloud és ARP

- Költséghatékony megoldás
- ARP egy projekt mindkét felhős ágon
- Rugalmasság a virtualizált alkalmazásréteg miatt
- ARP felhasználói függetlenek a felhő felhasználótól
- Cél emellett az ELKH Cloud projektben futó virtuális gépről adott adatrezitórium projektek elérése is



# Hardver és szoftver

- SZTAKI: új központi Ceph klaszter ARP-s adatok számára
  - Tervezetten 1365 TB felhasználható tárterület
- Wigner: új központi Ceph klaszter ARP-s adatok számára
  - Tervezetten 512 TB felhasználható tárterület
- Mester példány: SZTAKI (adatfeltöltés itt)
- Adatrepozitórium projektre értelmezett tükrözési példányok: Wigner FK
  - Tervezetten csak olvasható
- 100G redundáns hálózati interfészek
- Ceph Pacific/Quincy verzió
- Monitorozó rendszer kialakítása
- Tartalék diszkek beszerzése üzemeltetési fázisra

# Adatredundancia és szinkronizálás

- Két telephely közötti szinkron
- Metaadatok szintjén a szinkronizáció teljes és folyamatos
- CONCORDA adatok írása SZTAKI oldalon
- Adatok átmásolása nem automatikus, kérvényezhető (szelektív georedundancia)
- Redundancia a Ceph klasztereken belül: tervezetten háromszoros

# Átviteli sebesség

- Becslések végezhetőek
- Tesztelés - tesztesetek
  - Ceph klaszter szerverei között
  - Két telephely közötti belső kommunikáció
  - Felhasználói feltöltés/letöltés adatrepozitóriumba
  - Alkalmazás réteg sebességi és mennyiségi korlátozása
  - Archiválás/visszatöltés

# Mentési stratégia és archiválás

- Metaadatok napi mentése - hetente archiválás
- Konfiguráció SQL adatbázisban található részének mentése hetente
- Alkalmazásszintű mentés ELKH Cloudba végezhető
- Visszaállítási teszt
- Archiválás tervezetten NFS-en keresztül valósul meg
- Szalagos rendszer (SL 3000) Wigner FK-ban üzemel

# Lokális tárolók

- Központi ARP mellett másik szint a lokális tárolóké
- Szerepe: kiválasztott intézetek számára ~60 TB (RAID 6 tömb) helyi kapacitás
- Egyszerű konfiguráció
- Cél: közvetlenül a kiválasztott intézeti hálózaton érhető el a tároló
- SZTAKI és Wigner FK is vásárol egy-egy tárolót, melyet egy-egy kiválasztott intézet tárolóként használhat helyben náluk
- Kiválasztott intézet elsajátítja az ARP-hez tartozó felhasználói folyamatokat
- Adatgyűjtés után adatcsomagok és mérési eredmények címkézésének elvégzése kutatók által (metaadatok)
- Interneten keresztül adatrepozitóriumba feltöltés
- Példaként szolgál jövőbeli tervezéshez



Köszönöm a figyelmet!

Wigner Fizikai Kutatóközpont - Adatközpont  
pinter.adam@wigner.hu