

AZ ELKH ADATREPOZITÓRIUM PLATFORM INFRASTRUKTURÁLIS ÉS TERVEZÉSI MEGOLDÁSAI

INFRASTRUCTURE AND DESIGN SOLUTIONS OF THE ELKH DATA REPOSITORY PLATFORM

Kacsuk Péter¹, Pintér Ádám², Tenczer Szabolcs³, Hajnal Ákos⁴

¹az MTA doktora, tudományos tanácsadó, tudományos laborvezető
ELKH Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI), Budapest
peter.kacsuk@sztaki.hu

²IT-mérnök, adatközpont IT-vezető, ELKH Wigner Fizikai Kutatóközpont, Budapest
pinter.adam@wigner.hu

³IT-mérnök, DevOps-csoportvezető, ELKH Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI), Budapest
szabolcs.tenczer@sztaki.hu

⁴PhD, tudományos munkatárs, kutató, ELKH Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI), Budapest
akos.hajnal@sztaki.hu

ÖSSZEFOGLALÁS

A tudományos eredményeket megalapozó adathalmazokhoz való hozzáférhetőség, ezek nyilvánossá tétele, kereshetősége további kutatások számára nemcsak egy adott tudományterület fejlődését segítheti elő, de kaput nyit interdiszciplináris és nemzetközi kutatások számára is, ami napjainkban már egyre inkább alapelv, elvárás, esetenként követelmény. A publikációkhoz kapcsolódó adatok nyilvánossá tétele és tárolása sok esetben ad hoc módon történik, amelynek azonban számos hátránya van. A hosszú távú elérhetőség biztosítása nem egyszerű feladat, továbbá az adatok kereshetősége, elérhetősége nehézségekbe ütközhet a heterogenitás és széttagoltság miatt.

Az ELKH Adatrepozitórium Platform (ELKH ARP) projekt célja egy központi magyar kutatási adatrepozitórium létrehozása, amely lehetővé teszi a kutatási adatok FAIR*-elvé tárolását és kezelését, illetve a publikációhoz társított kutatási adathalmazok hosszú távú tárolását és ezek (nyílt vagy zárt) megosztását adott tudományos közösség számára, tudományágak között, vagy akár nemzetközi szinten is, ezzel biztosítva egy későbbi nemzetközi kutatási infrastruktúrába való bekapcsolódás lehetőségét. Az adatrepozitórium kiépítéséhez szükséges digitális infrastruktúrát az Eötvös Loránd Kutatási Hálózat (ELKH) két kutatóhelye: a Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI) és a Wigner Fizikai Kutatóközpont (Wigner FK) készíti az ELKH Titkárság támogatásával, az ELKH Adatrepozitórium projekt keretein belül.

Egy ilyen rendszerrel szemben elvárás a magas rendelkezésre állás, nagy fokú adatbiztonság, nagy sávszélességű adatkapcsolat, a hosszú távú tárolás biztosítása, hatalmas tömegű és mé-

*FAIR – Findable, Accessible, Interoperable, Reusable; az adatok megtalálhatók, hozzáférhetők, együttműködők, interoperábilisak.

retű adattartalmak befogadása, és nem utolsósorban az adatok közötti hatékony keresetőség biztosítása.

Ebben a cikkben az adatrepozitórium hardver és szoftver infrastruktúrájának tervezési szempontjait és megvalósítási lépéseit, illetve ezek fő aspektusait tekintjük át, kitérve műszaki részletekre, fontosabb tervezési döntésekre és a választott megoldásokra. Bemutatjuk, milyen szerepet játszik, és milyen előnyökkel jár az ELKH Cloud rendszeréhez való kapcsolódás, az adatbiztonság hogyan garantálható a redundancia, adatreplikáció különböző szintjeinek alkalmazásával, illetve hogyan tehetők még kényelmesebbé és hatékonyabbá a kutatóintézeti primer adatgyűjtési folyamatok, alkalmazkodva az esetleges egyedi, helyi munkafolyamatokhoz.

ABSTRACT

Access to the data sets that underpin scientific results, making them publicly available and searchable for further research can not only promote the development of a given field of science but also open the door to interdisciplinary and international research, which is nowadays increasingly becoming a basic principle, an expectation, and sometimes a requirement. The publication and storage of data related to scientific publications is often done in an ad-hoc manner but this has a number of drawbacks. Ensuring long-term accessibility is not easy and the searchability and availability of data can be difficult due to heterogeneity and fragmentation.

The ELKH Data Repository Platform (ELKH ARP) project aims to create a central Hungarian research data repository, which will allow for the FAIR storage and management of research data, the long-term storage of research datasets associated with publications and their sharing (open or closed) with the scientific community, between disciplines or even internationally, thus ensuring the possibility of future integration into an international research infrastructure. The digital infrastructure required for the construction of the data repository is being developed by the two research centres of the Eötvös Loránd Research Network (ELKH): the Institute for Computer Science and Automation (SZTAKI) and the Wigner Research Centre for Physics (Wigner FK), with the support of the ELKH Secretariat, within the framework of the ELKH Data Repository Platform project.

Such a system is expected to provide high availability, high data security, high bandwidth data connectivity, long-term storage, the ability to accommodate huge data volumes, and last but not least, the ability to efficiently search across data.

In this paper, we review the design aspects and implementation steps of the hardware and software infrastructure of the data repository storage and its main aspects, covering technical details, major design decisions and the solutions chosen. We will present the role and benefits of connecting to the ELKH Cloud system, how data security can be guaranteed by using different levels of redundancy and data replication, and how primary data collection processes at research institutes can be made even more convenient and efficient, adapting to possible unique, local work processes.

Kulcsszavak: adatrepozitórium, felhő infrastruktúra, FAIR-adattárolás, FAIR-adatkezelés, adatbiztonság

Keywords: data repository, cloud infrastructure FAIR-data storage, FAIR-data management, data security

BEVEZETÉS

Az ARP-adatrepozitóriumot egy központi adattárolóként terveztük megvalósítani, amely az adatokhoz, metaadatokhoz való hozzáférést és kereshetőséget egyetlen közös webes felületen teszi elérhetővé. Amennyiben szükséges, a felhasználók egyetlen feltöltési programozói végponton (API) keresztül tudják feltölteni a metaadatokkal ellátott nyers adatokat. Cikkünkben arról számolunk be, hogy hogyan lehet megtervezni és megvalósítani egy olyan kutatási infrastruktúrát, amely lehetővé teszi nagy mennyiségű kutatási adat tárolását, archiválását, gyors feldolgozását, kereshetőségét, és biztosítja a felhasználás és újrafelhasználás lehetőségét a magyar és nemzetközi kutatási projektek számára.

TERVEZÉSI SZEMPONTOK, MEGVALÓSÍTÁSI DÖNTÉSEK

A központosított elérhetőség a felhasználók számára könnyebbé teszi a munkát, mert látszólag egyetlen kiszolgáló géppel tartják a kapcsolatot, a háttérben azonban egy rendkívül komplex, elosztott infrastruktúra szolgálja ki a felhasználói kéréseket.

A komplexitást a megfogalmazott követelmények teljesítése indokolja. Az ARP-adatrepozitórium a felhasználói igényeken túl magas fokú elérhetőséget, rendelkezésre állást kell hogy biztosítson (más kontextusban ezt a szolgáltatási szint megállapodás [SLA] részének tekintik), és természetesen az adatok biztonsága rendkívül fontos szempont. Az adatbiztonságon ezúttal nem a jogosulatlan hozzáférés megakadályozását értjük, vagy egyéb titkosítási megoldások alkalmazását, hanem az adatok sérülésének vagy elvesztésének megelőzését.

Mivel az ARP-adatrepozitórium nagyszámú felhasználó egyidejű keresési és adat le- és feltöltési kérését kell, hogy kiszolgálja, a megvalósított rendszernek alkalmasnak kell lennie ezek konkurens kiszolgálására (megfelelően méretezett CPU- és memóriakapacitással), és megfelelően nagy méretű hálózati sávszélességre van szüksége, azért hogy az esetlegesen párhuzamosan futó, nagy méretű adatok mozgatása során ne hátráltassák egymást túlzott mértékben.

Ezeken túlmenően, az ARP-adatrepozitórium megvalósítása során felmerülő talán legfontosabb szempont és kihívás, hogy hogyan lehet hatékonyan és biztonságosan, hosszú távon tárolni rendkívül nagy mennyiségű adatot (mind terabyte, mind petabyte nagyságrendben), hogyan lehet kialakítani egy karbantartható infrastruktúrát (esetleges hardvermeghibásodások, szoftverfrissítések esetén is) anélkül, hogy adatot veszítenénk, vagy a rendszer elérhetetlenné válna a javítások idején.

A nyers kutatási adatok archiválásán, megőrzésén, hivatkozhatóságán és letölthetőségén túl fontos követelmény volt az ARP-adatrepozitóriumban való kereshetőség biztosítása is, amely újabb kutatásokhoz szolgáltathat alapadatokat.

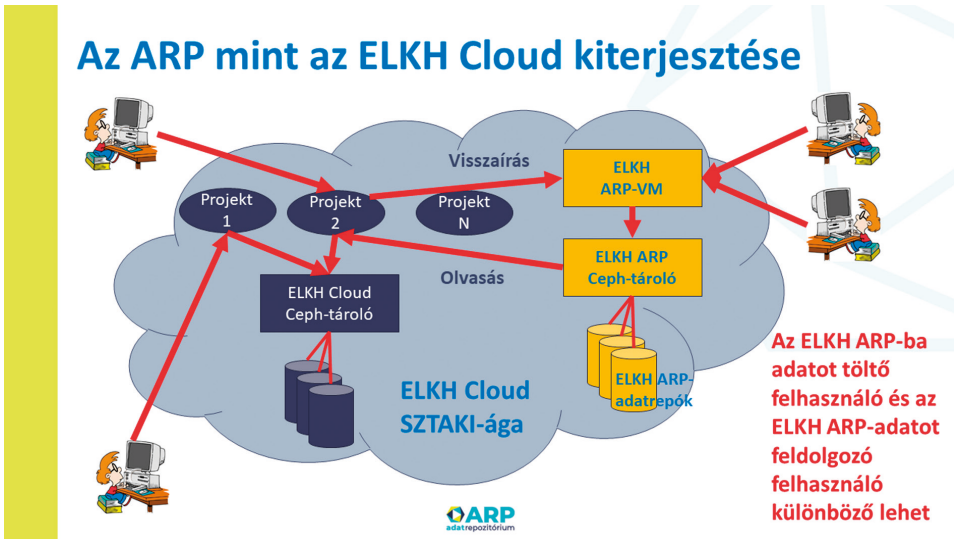
Szükséges szemantikai információ hiányában és az adatok óriási tömegét figyelembe véve ez szinte lehetetlen feladat lenne. Ezért lett kialakítva a nyers adatokhoz társított metaadatrendszer, amely a nyers adatokról tartalmi, értelmezési és egyéb leíró információkat tárol, amelyek segítségével egyszerű, de akár összetett szempontok szerint már kereshetővé válnak az adatok. A nyers adatokat tehát „metaadatolva” tároljuk az adatrepozitóriumban, a metaadatok a nyers adatokhoz vannak társítva, de a megvalósítás során nem feltétlenül kerülnek ugyanarra az adattárolóra, adattárhelyre, mert a metaadatoknak a hatékony kereshetőséget kell biztosítaniuk tetszőleges szempontrendszer szerint (belső metaadatmező-tartalmakra), míg a nyers adatoknak csak elérhetőséget kell biztosítaniuk (és nem szükséges a belső struktúrához, tartalomhoz hozzáférni).

A fenti szempontok szerinti rendszer kialakítása láthatóan nem egyszerű feladat, különösen, ha egy megfelelő háttér-infrastruktúra még nem áll rendelkezésre. Az ELKH Cloud azonban megfelelő alapként szolgált az adatrepozitórium megvalósításához. Az ELKH-felhő önmagában biztosítja a megfelelő rendelkezésre állás, adatbiztonság (redundancia), rugalmas hardverigények (CPU¹, memória), nagy külső és belső sávszélesség biztosítása, felhasználók és projektek kezelése, monitorozás és karbantartás kritériumait. Ezért az ARP-adatrepozitóriumot az ELKH Cloud kiterjesztéseként kívántuk megvalósítani. Az adatrepozitórium adatai számára az ELKH Cloudban már meglévő Ceph-tároló mellett egy új, önálló adattároló rész, ún. ARP Ceph-tároló lesz kialakítva, megfelelően nagyszámú fizikai lemezegység hozzáadásával és a Ceph-rendszer (Weil et al., 2006) telepítésével, konfigurálásával (*1. ábra*). Az ARP Ceph-tároló fizikailag az ELKH-felhőhöz lesz csatolva, így ez nem válik el helyileg (hálózati szempontból), hanem elérhető, kvázi lokálisan az ELKH számítási felhőből. Ennek a tervezési döntésnek az a nagy előnye, hogy amennyiben az adatfeldolgozó algoritmusokat az ELKH-felhőben futtatjuk, a nyers adatokat már nem kell egy megelőző lépésben ide átmozgatni, hanem a számítások azonnal a helyileg elérhető nyers adatokon elvégezhetőek. Ez a „Moving Compute to Data”, azaz a mozgassuk a számítást az adathoz elvnek megfelelő feldolgozási modell alkalmazását teszi lehetővé. Ez a paradigma rendkívül fontos, amennyiben extrém nagy mennyiségű adat feldolgozása szükséges, mert ekkor előfordulhat, hogy pusztán az adatok mozgatásának költsége (és ideje) nagyságrendekkel haladja meg a tényleges, a valójában értékes adatokat szolgáltatató feldolgozási költséget.

Az ELKH-felhő felhasználói felülete (OpenStack [URL1] alapon) lehetővé teszi, hogy az egyes kutatócsoportok elkülönülten (külön projekteken, rugalmas autorizációs feltételekkel) dolgozzanak a saját érdeklődésükre számot tartó adatfeldolgozási projektjeiken, mégis mindegyik önálló projekt hozzáfér az ARP-adat-

¹ CPU – central processing unit, központi feldolgozóegység, processzor

repozitórium tetszőleges adattartalmához. Fontos megjegyezni, hogy az adatokat feldolgozó személy vagy szervezet nem feltétlenül kell hogy azonos legyen az eredeti adatokat az adatrepozitóriumba feltöltő, publikáló személlyel vagy szervezettel. A számításokhoz szükséges helyi tárolókat és diszkeket (virtuálisigéplemezek, közties adatok és eredmények átmeneti tárolására) az ELKH-felhő saját tárolórendszere biztosítja.



1. ábra. Adat betöltése, tárolása és feldolgozása az ELKH ARP-adatrepozitóriumhoz kapcsolódóan (mindegyik ábra a szerzők szerkesztése)

ADAT- ÉS METAADAT-BIZTONSÁG, REDUNDANCIA ÉS GEOLOKÁCIÓ

Az adathordozó hardverek fizikai meghibásodásából eredő adatvesztés elleni védelem egyik legfontosabb eszköze az adatok replikálása több, különálló adathordozón, azaz a megfelelő redundancia kialakítása. Tovább javítja az adatbiztonságot, ha a replikák nemcsak egyszerűen különböző adathordozó egységeken vannak tárolva, hanem azok helyileg (földrajzilag, geolokációs, elektromos ellátás és internet hálózati szempontból) is elkülönülnek.

Az adatrepozitórium az első szintet tekintve az adatbiztonságot háromszoros redundancia (RAID² 3 [Patterson et al., 1987]) biztosításával segíti elő.

² RAID – Redundant Array of Inexpensive Disks vagy Redundant Array of Independent Disks, redundáns tömb olcsó lemezekből vagy redundáns tömb független lemezekből

Az ELKH Cloudot valójában két felhő federációja alkotja. A felhő egyik része a SZTAKI-ban van (SZTAKI Cloud), a másik része a Wigner Fizikai Kutatóközpontban üzemel (Wigner Cloud), jóllehet a felhasználók egyetlen egységes felületen érik el mindkettőt, számukra transzparens módon (URL2 /webes eléréssel), a nyílt forráskódú OpenStack-keretrendszerbe integrálva.

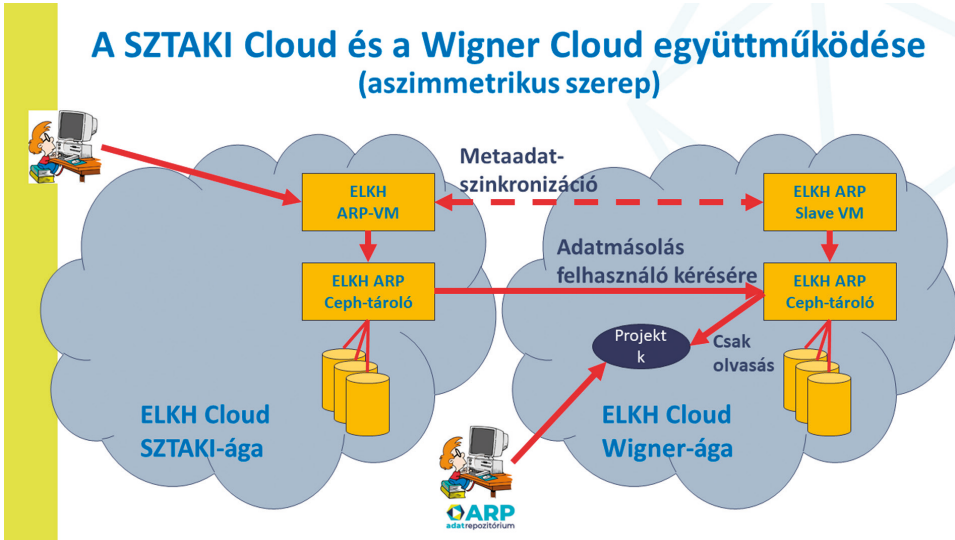
Az ARP-adatrepozitórium és a nyers adatok elsődleges tárolási helye a SZTAKI-felhőben lesz, de a Wigner-felhőben is telepítésre kerül és üzemel egy, a SZTAKI-felhőbeli adatrepozitóriummal azonos felépítésű és funkciókkal bíró komponens, ún. Slave VM³, amely mögött szintén rendelkezésre áll nagy mennyiségű adattároló kapacitás. A két rendszer, összekapcsolásuk révén, képes lesz egymással kommunikálni, adatokat cserélni, megfelelő szinkronizációs műveleteket egymás között elvégezni, automatizált módon.

Az összes nyers adat automatikus és folyamatos replikálása a SZTAKI-felhőből a Wigner-felhőbe azonban szükségtelenül nagy költséget és hálózati leterheltséget jelentett volna (a helyi redundancia önmagában kellő adatbiztonságot ad). A metaadatok mérete azonban lényegesen kisebb, a metaadatok replikációja, tükrözése a Wigner-felhőben automatikusan megtörténik, ezzel növelve a metaadatok geolokációs redundanciáját. A metaadatok (és egyéb rendszer-konfigurációs adatok) biztonságát növeli az automatikus napi mentés különálló adathordozóra.

Előfordulhat, hogy egy adott adatfeldolgozást egy felhasználó vagy intézet a Wigner-felhőben kívánja elvégezni, mert vagy kapacitási, vagy egyéb speciális hardverigények miatt itt van lehetőség, itt indokolt ezeket elvégezni, de a nyers adatok helye fizikailag továbbra is a SZTAKI-felhő. Azért, hogy elkerüljük azt, hogy a felhasználónak magának kelljen az adatokat átmozgatni a SZTAKI-felhőből a Wigner-felhőbe (vagy alacsony sávszélességgel tudja csak elérni a távoli adatokat), az adatrepozitórium igény esetén lehetőséget biztosít teljes nyers adatrepozitóriumok átmozgatására. Ezt az adattranszfert a rendszer automatikusan elvégzi, de csak felhasználói (projekt) kérésre. A nyers adatok tükrözése után a feldolgozás már a Wigner-felhőben is rendkívül hatékony módon történhet meg, helyi adathozzáféréssel. Ugyanakkor, mivel az elsődleges adattárolás és adatbevitel helye a SZTAKI-felhő, ezért a Wigner-felhőben futó feldolgozások csak olvasni tudják a Wigner-felhőben tárolt adatrepozitóriumok adatait, ahogy a 2. ábra mutatja. Ha a feldolgozás eredményeképpen előállított új adatokat is be kell írni az ARP-be, akkor ezt csak a SZTAKI-felhőben működő ARP-VM segítségével lehet elvégezni.

A Wigner-felhőben található adatok számára további lehetőség az adatbiztonság fokozására az adatok szalagos egységre történő kiírása. Ezt egy SL3000-es egység segítségével lehet megtenni, szintén felhasználói kérés esetén.

³ VM – virtual machine, virtuális gép; Slave VM, virtuális szolgagép



2. ábra. Adat elérése és feldolgozása a Wigner-felhőben

LOKÁLIS, INTÉZETI ADATKEZELÉS TÁMOGATÁSA

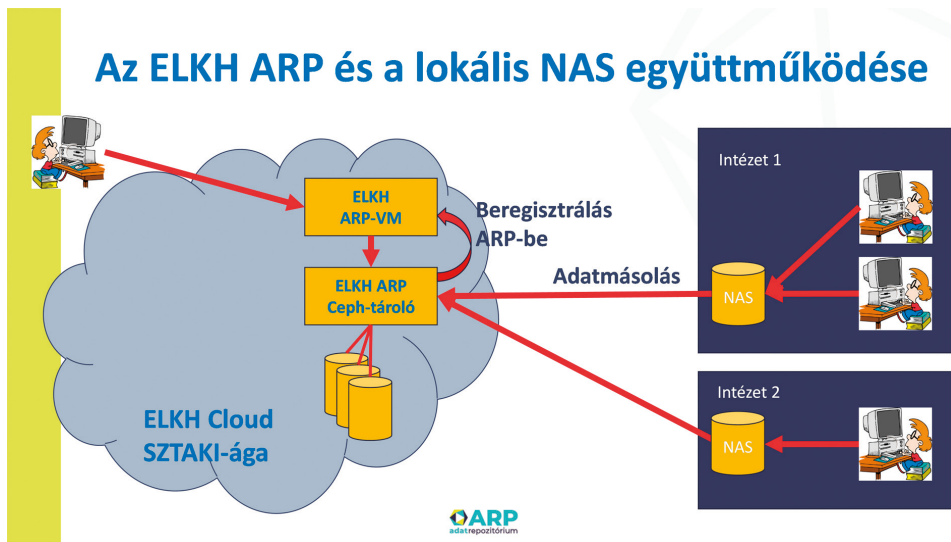
Az intézeti adatkezelés támogatására, igény esetén, az ELKH ARP-adatrepozitórium egy lokálisan, az adott intézetbe kihelyezésre kerülő és ott üzemelő hálózati adattárat (Network Attached Storage, NAS) tud biztosítani. Ennek nagy előnye, hogy beruházás, kapacitásbővítés és komolyabb rendszertelepítés nélkül azonnal rendelkezésre fog állni az intézet számára egy nagy méretű tárterület, ahova az adatokat elhelyezhetik, átmásolhatják azokat, illetve közvetlenül erre történhet a primer adatgyűjtés. Ezzel az eszközzel az intézeti adatkezelés, annak munkafolyamata is korszerűsödhet, akár egyszerűsödhet ezáltal. Az adattár átmenetileg (időszakonként vagy a projekt végéig) tárolja majd a kutatási adatokat, amíg azok be nem kerülnek végleges helyükre, a központi ARP-adatrepozitóriumba.

A NAS-ról a központi ARP-adatrepozitóriumba való adattranszfer többféle módon történhet. Amennyiben rendelkezésre áll nagy sávszélességű internetkapcsolat a központi repozitórium felé, az adatok feltöltése történhet a szokásos módon, a mindenki számára elérhető webes vagy API⁴-szolgáltatások valamilyen, azzal a különbséggel, hogy a forrásadatok helye ezúttal a NAS-on található. Lehetőség van továbbá kötegelte feltöltésre is a NAS és az adatrepozitórium között egy megfelelő szkript lefuttatásával, amely az összes adatot és metaadatot egyetlen parancs kiadásával feltölti. Amennyiben a metaadatok még nem készül-

⁴ API – Application Programming Interface, alkalmazásprogramozási felület

tek el, de már nagy mennyiségű primer adat áll rendelkezésre, az ELKH-felhőn indítható egy nagy tárolási kapacitású virtuális gép (egy, az erre a célra előkészített lemezkép kiválasztásával), ahova az adatokat átmeneti jelleggel már át lehet transzferálni (ezt ún. kétlépcsős adatfeltöltésnek hívjuk). Amint minden adat és a metaadatok is elkészültek, és feltöltődtek a virtuális gépre, innen már nagy sávszélességű kapcsolaton keresztül, felhasználói jelzésre, automatizáltan kerülhetnek át az adatok a központi ARP-adatrepozitóriumba. Nem utolsósorban, lehetőség van az eszköz fizikai átszállítására is a SZTAKI-ba, ahol azután a repozitóriumba való adatmásolás helyileg történhet, amely bár nagyságrendjében nem, működésében hasonló az Amazon Snowball (URL3) szolgáltatásához.

Ezek a lokális NAS-eszközök 60 TB-os tárterületet biztosítanak (RAID 6 redundanciával), előre telepített, Windows Server operációs rendszerrel, de összekapcsolhatók más, nem Windows-rendszerekkel is. Elsőként a Régészeti Intézetben lévő NAS-adattáron teszteltük a kétlépcsős adatfeltöltést. Az ELKH ARP végső célja, hogy a jelenlegi pilot megoldás általános megoldásként minél több intézet számára legyen bejáratos út a központi adatrepozitórium felé.



3. ábra. Adat feltöltése, adatmásolás a lokális intézeti NAS-tárolókból

KONKLÚZIÓ, JÖVŐBELI TERVEK

Az adatorientált kutatás előtérbe kerülése, a FAIR-kultúra gyakorlatainak elterjedése motiválta azt, hogy hazánkban is megteremtsük, és minél szélesebb körben elérhetővé tegyük a kutatók számára az ehhez szükséges feltételeket. Számos

szervezet felismerte ennek fontosságát, és támogatta is az ilyen jellegű törekvéseket nemzetközi szinten (Európai Bizottság, *European Open Science Cloud* – EOSC, Research Data Alliance – RDA) és hazai szinten is (Hungarian Research Data Alliance – HRDA). Magyarországon az ELKH támogatással indult el az ELKH Adatrepozitórium Platform projekt, a SZTAKI, a Társadalomtudományi Kutatóközpont (TK) és a Wigner kutatóintézetek részvételével, amelynek egyik fontos célkitűzése egy megfelelő, nagy mennyiségű kutatási adat tárolását, archiválását, gyors feldolgozását, kereshetőségét biztosító infrastruktúra kialakítása.

Ebben a cikkben ennek az infrastruktúrának a koncepcionális és tervezési megfontolásait, döntéseit mutattuk be, a konkrét megvalósítás néhány fontosabb műszaki részletével és megoldásával együtt. A megvalósított infrastruktúrának számos magas szintű követelményt, elvárást kell teljesítenie, mint például a magas rendelkezésre állás, adatbiztonság és redundancia, geolokáció, az adatfeldolgozás minél hatékonyabb támogatása, az adattranszfer és az intézeti, helyi adatkezelés támogatása. Épp ezért volt nagyon fontos stratégiai döntés, hogy az adatrepozitórium műszaki megvalósítását ne nulláról kezdjük, hanem a nagyon sikeres és stabil működésű ELKH Cloud továbbfejlesztéseként oldjuk meg. Ez garancia arra, hogy az adatrepozitórium is eleget fog tenni a magas műszaki elvárásoknak.

IRODALOM

- Patterson, D. A. – Gibson, G. A. – Katz, R. H. (1987): *A Case for Redundant Arrays of Inexpensive Disks (RAID)*. Technical Report No. UCB/CSD-87-391. Berkeley: EECS Department, University of California, <https://www2.eecs.berkeley.edu/Pubs/TechRpts/1987/CSD-87-391.pdf>
- Weil, S. – Brandt, S. – Miller, E. et al. (2006): Ceph: A Scalable, High-Performance Distributed File System. *OSDI '06: Proceedings of the 7th Symposium on Operating Systems Design and Implementation*, Nov., 307–320. <https://www.ssrc.usc.edu/media/pubs/6ebbf2736ae06c66f-1293b5e431082410f41f83f.pdf>

URL1: *OpenStack*. <https://www.openstack.org/>

URL2: <https://science-cloud.hu/>

URL3: *AWS Snowball*. <https://aws.amazon.com/snowball>