From the MTA Cloud research infrastructure to ELKH Cloud

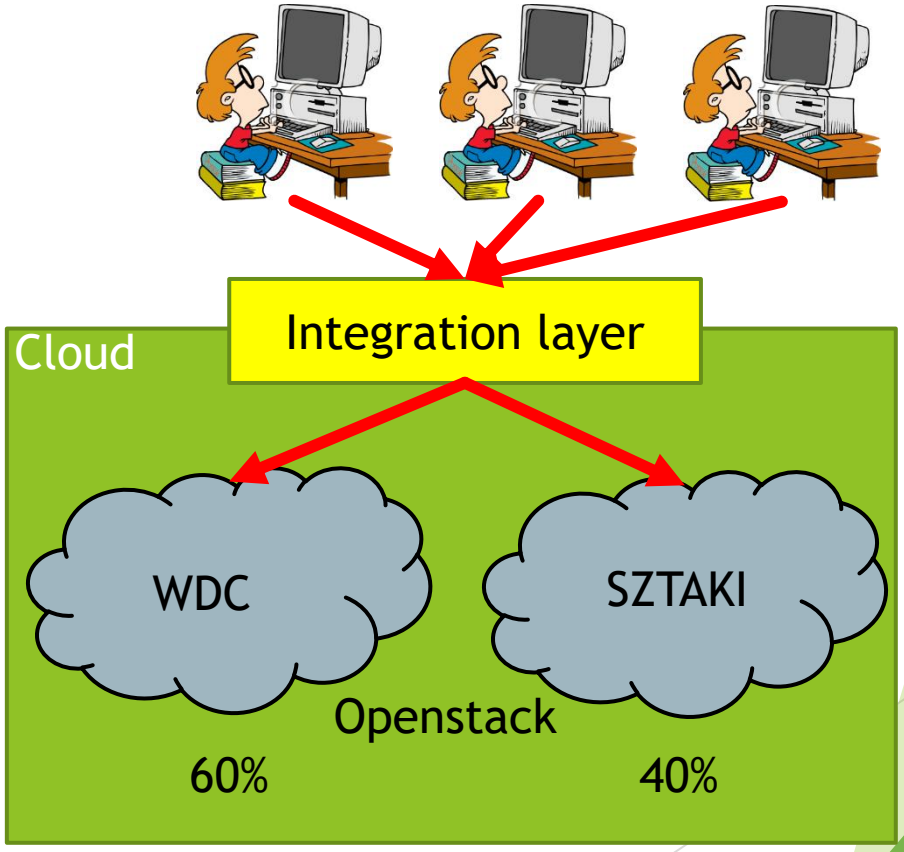
Prof. Dr. Péter Kacsuk
Technical director of ELKH Cloud
SZTAKI

The story of MTA/ELKH Cloud in a nutshell

- **1 Oct 2016: MTA Cloud full production operation started**
- 1 Sep 2019: Move from MTA (Academy) to ELKH (Eötvös Lóránd Research Network)
- 1 July 2020: The 1-year ELKH Cloud project has been started
- **1 Oct 2020: Official new name: ELKH Cloud**
- 1 July 2021: The new extended ELKH Cloud will be open for the users

Capacity of MTA Cloud and current capacity of ELKH Cloud

	Current capacity
vCPU (max)	1368
GPU core	12
vGPU (max)	12
RAM (TB)	3,25
SSD storage (TB)	0
HDD storage (TB)	527
Tensor GPU performance (PFLOPS)	~0
Floating point GPU performance (PFLOPS)	~0
Network bandwidth (Gbps)	10

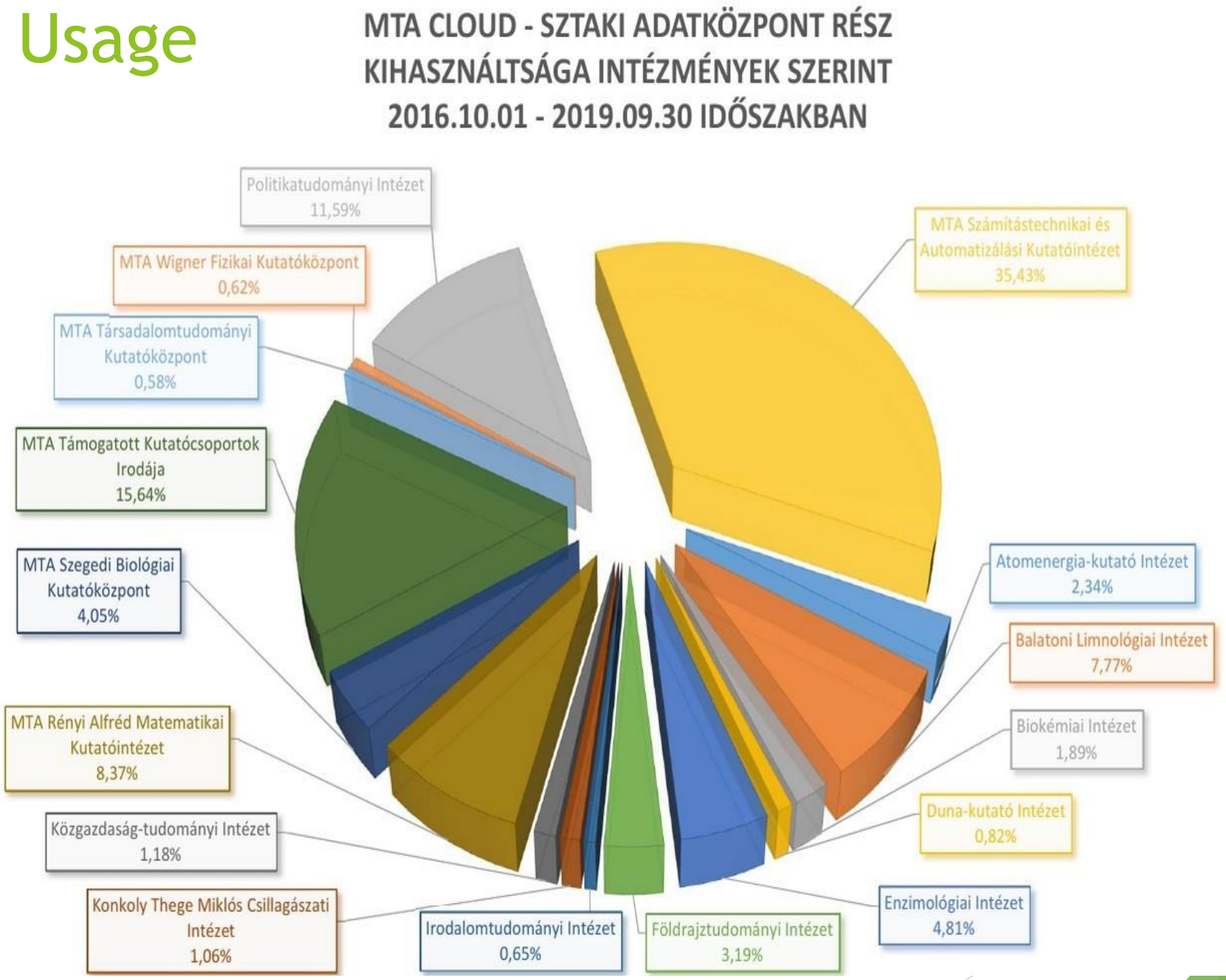


Integration level among the two sites

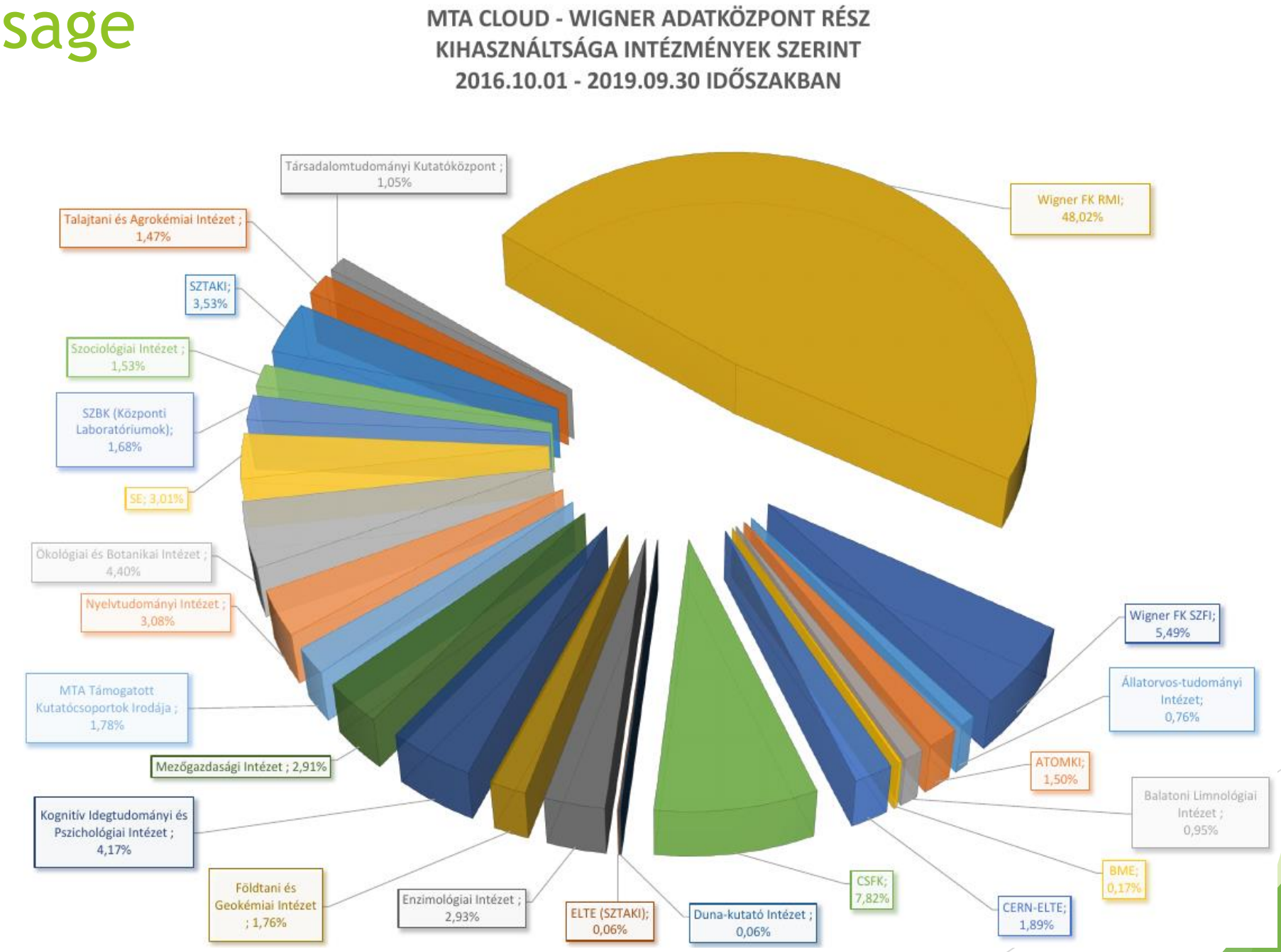
- Common cloud technology: Openstack
- LDAP based AAI
- Common cloud access policy for distributing capacity among users
- Common web page
- Common training
- Common user support

- Negotiations are ongoing for connecting new sites in the near future
- Hybrid cloud project with Microsoft Hungary to demonstrate the joint usage of ELKH Cloud and Azure Cloud

Usage

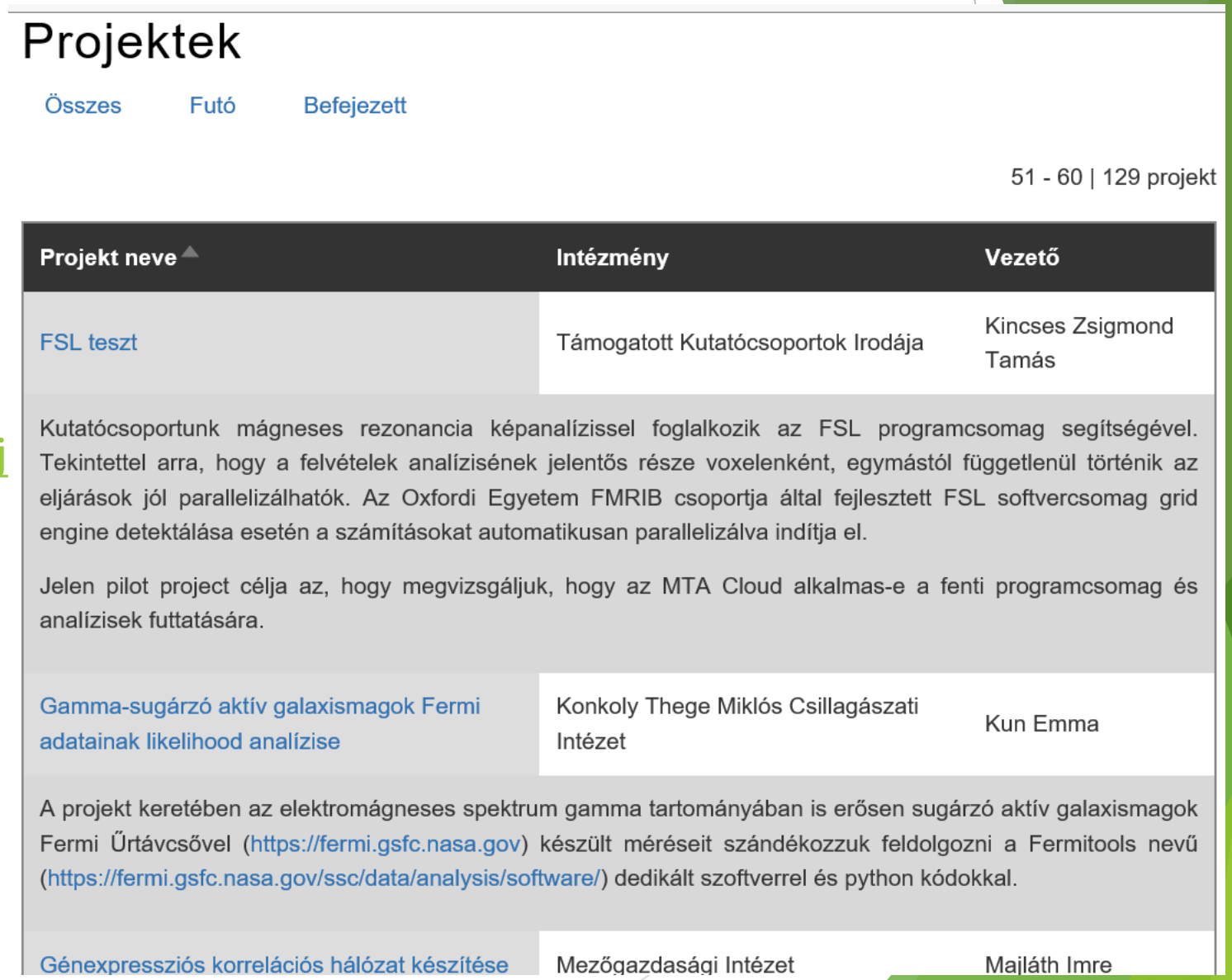

MTA CLOUD - SZTAKI ADATKÖZPONT RÉSZ KIHASZNÁLTSÁGA INTÉZMÉNYEK SZERINT 2016.10.01 - 2019.09.30 IDŐSZAKBAN
Politikatudományi Intézet 11,59%
MTA Számítástechnikai és Automatizálási Kutatóintézet 35,43%
MTA Wigner Fizikai Kutatóközpont 0,62%
MTA Társadalomtudományi Kutatóközpont 0,58%
MTA Támogatott Kutatócsoportok Irodája 15,64%
MTA Szegedi Biológiai Kutatóközpont 4,05%
Atomenergia-kutató Intézet 2,34%
Balatoni Limnológiai Intézet 7,77%
MTA Rényi Alfréd Matematikai Kutatóintézet 8,37%
Biokémiai Intézet 1,89%
Közgazdaság-tudományi Intézet 1,18%
Duna-kutató Intézet 0,82%
Konkoly Thege Miklós Csillagászati Intézet 1,06%
Enzimológiai Intézet 4,81%
Irodalomtudományi Intézet 0,65%
Földrajztudományi Intézet 3,19%

Usage


MTA CLOUD - WIGNER ADATKÖZPONT RÉSZ
KIHASZNÁLTSÁGA INTÉZMÉNYEK SZERINT
2016.10.01 - 2019.09.30 IDŐSZAKBAN
Társadalomtudományi Kutatóközpont ; 1,05%
Talajtani és Agrokémiai Intézet ; 1,47%
SZTAKI; 3,53%
Szociológiai Intézet ; 1,53%
SZBK (Központi Laboratóriumok); 1,68%
SE; 3,01%
Ökológiai és Botanikai Intézet ; 4,40%
Nyelvtudományi Intézet ; 3,08%
MTA Támogatott Kutatócsoportok Irodája ; 1,78%
Mezőgazdasági Intézet ; 2,91%
Kognitív Idegtudományi és Pszichológiai Intézet ; 4,17%
Földtani és Geokémiai Intézet ; 1,76%
Enzimológiai Intézet ; 2,93%
ELTE (SZTAKI); 0,06%
Duna-kutató Intézet ; 0,06%
CSFK; 7,82%
CERN-ELTE; 1,89%
BME; 0,17%
Balatoni Limnológiai Intézet ; 0,95%
ATOMKI; 1,50%
Állatorvos-tudományi Intézet; 0,76%
Wigner FK SZFI; 5,49%
Wigner FK RMI; 48,02%

Current usage of ELKH Cloud

- Number of started projects: 138
- Number of completed projects: 58
- Web page shows every relevant information (in Hungarian): https://cloud.mta.hu/projektek



Projektek

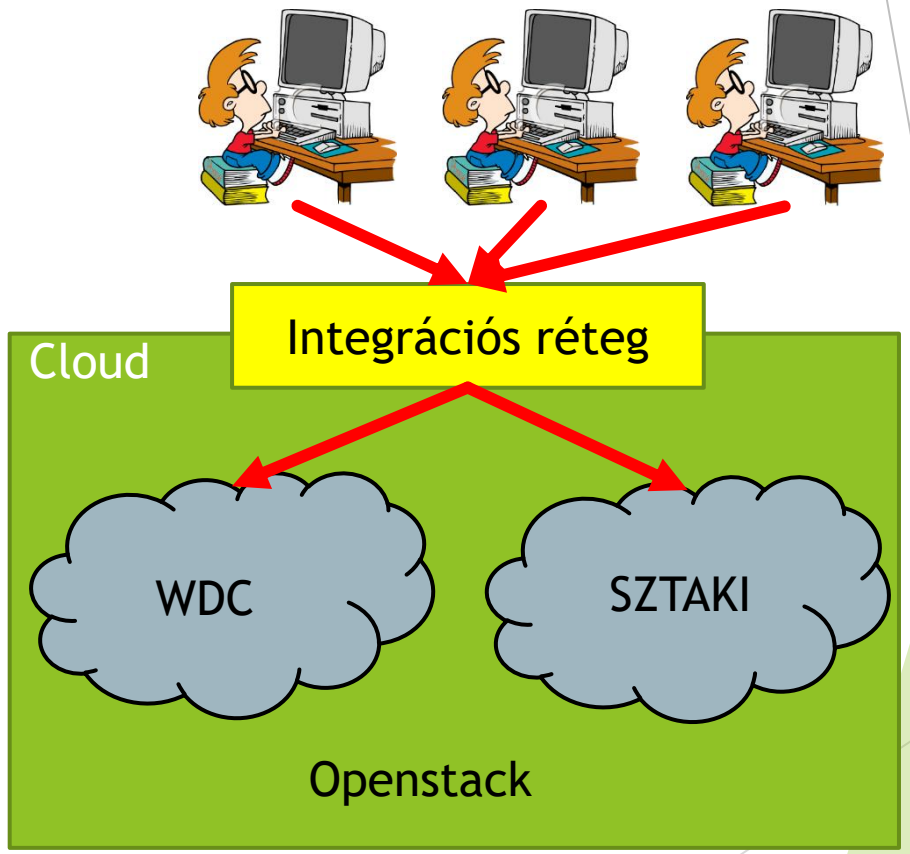
Összes Futó Befejezett

51 - 60 | 129 projekt

Projekt neve	Intézmény	Vezető
FSL teszt	Támogatott Kutatócsoportok Irodája	Kincses Zsigmond Tamás
Kutatócsoportunk mágneses rezonancia képanalízissel foglalkozik az FSL programcsomag segítségével. Tekintettel arra, hogy a felvételek analízisének jelentős része voxelenként, egymástól függetlenül történik az eljárások jól parallelizálhatók. Az Oxfordi Egyetem FMRIB csoportja által fejlesztett FSL softvercsomag grid engine detektálása esetén a számításokat automatikusan parallelizálva indítja el. Jelen pilot project célja az, hogy megvizsgáljuk, hogy az MTA Cloud alkalmas-e a fenti programcsomag és analízisek futtatására.		
Gamma-sugárzó aktív galaxismagok Fermi adatainak likelihood analízise	Konkoly Thege Miklós Csillagászati Intézet	Kun Emma
A projekt keretében az elektromágneses spektrum gamma tartományában is erősen sugárzó aktív galaxismagok Fermi Űrtávcsővel (https://fermi.gsfc.nasa.gov) készült méréseit szándékozzuk feldolgozni a Fermitools nevű (https://fermi.gsfc.nasa.gov/ssc/data/analysis/software/) dedikált szoftverrel és python kódokkal.		
Génexpressziós korrelációs hálózat készítése	Mezőgazdasági Intézet	Majláth Imre

Comparison of the capacity of MTA Cloud and extended ELKH Cloud

	MTA	ELKH
vCPU (max)	1368	4000
GPU core	12	76
vGPU (max)	12	2060
RAM (TB)	3,25	11
SSD storage (TB)	0	153
HDD storage (TB)	527	1500
Tensor GPU performance (PFLOPS)	~0	7.16
Floating point GPU performance (PFLOPS)	~0	0.89
Network bandwidth (Gbps)	10	100



The concept of ELKH Cloud

- It is not a **job submission cloud** receiving and executing various incoming jobs
  - E.g. PBS cluster, which accepts jobs from various users no matter from which project they come and what size of quota they have.
- **It is not a multi-user service provision cloud**:
  - E.g. a science gateway that can simultaneously be accessed by many users no matter from which project they come and what size of quota they have.

The concept of ELKH Cloud

- **ELKH Cloud is an e-infrastructure hosting framework:**
  - Not user-oriented, rather **project-oriented**
    - Primarily projects are registered
    - Users register for existing projects
  - A registered project receives a certain amount of capacity called **kvota.** Within the given quota the project can build any e-infrastructure.
    - E.g. A project can build a PBS cluster for the users registered for this project (other users cannot access this PBS cluster)
    - E.g. A project can build a multi-user science gateway for the users registered for this project (other users cannot access this gateway)

The concept of ELKH Cloud

- Building e-infrastructures in the cloud is not an easy job. Therefore, the ELKH Cloud projekt provides
  - **Reference architectures** for building tipical, frequently used e-infrastructures (e.g. Kubernetes cluster, Hadoop cluster, Spark cluster, Tensorflow with GPU, etc.)
  - Direct **user support** for building non-tipical e-infrastructures and creating a corresponding reference architecture for these e-infrastructures
  - A repository for storing these reference architectures. Any ELKH Cloud user can access and use these stored reference architectures.

Current reference architectures for creating services

1. Occopus cloud orchestrator
2. JupyterLab
3. Docker-Swarm cluster
4. Kubernetes cluster
5. Apache Hadoop cluster
6. Apache Spark cluster with RStudio
7. Apache Spark cluster with Python
8. TensorFlow with Jupyter Notebook
9. TensorFlow with Jupyter Notebook and GPU usage
10. DataAvenue for large volume data transfer
11. gUSE/WS-PGRADE Science gateway framework for HTC and workflow applications
12. Flowbster workflow system for large volume data processing
13. CQueue cluster for serving VM based HTC applications
14. MiCADO for building and running scalable cloud applications

Advantages of ELKH Cloud

- There is no long and complicated procurement procedure
- Users can build different size and type of e-infrastructures
  - either from scratch as they want
  - Or based on the available reference architectures
- Large storage capacity can be accessed and used
- **Very reliable IaaS cloud services**: average 2 days/year out of work
- Training events in every month (80 – 120 users/event):
  - odd months: for beginners
  - Even months: for advanced users in specific subjects (AI, Big data, etc.)

Comparison of MTA Cloud and ELKH Cloud

MTA Cloud

Capacity:

- Small capacity cloud

Users:

- MTA researchers

International access:

- Very limited

ELKH Cloud

Capacity:

- Mid-size capacity cloud

Users :

- ELKH researchers
- University researchers
- Commercial researchers

International access : Participation in large European programmes and projects:

- EOSC (EGI-ACE)
- RDA
- ESFRI (SLICES, SLICES-SC)

Plans for integration of ELKH Cloud with EOSC Compute Platform

- Steps of integration
  - Step 1: Completing the capacity extension project of ELKH Cloud (deadline: 31 Oct 2021)
  - Step 2: EGI Check-in (deadline: 31 Dec 2021)
  - Step 3: AppDB (deadline: 28 Feb 2022)
  - Step 4: DataHub (deadline: 30 April 2022)
  - Step 5: Usage accounting (deadline: 30 June 2022)
  - Step 6: investigating the integration options for CVMFS, RUCIO, Data Transfer (deadline: 30 Sep 2022)

Further information>

- Web: https://science-cloud.hu/
- FAQ: https://science-cloud.hu/gyakran-ismetelt-kerdesek
- E-mail: info@science-cloud.hu

Thank you

Questions?