

cesnet  
.....

METACENTRUM

Miroslav Ruda

CESNET

---

Seminář gridového počítání, duben 2023

.....

- Krátké představení MetaCentra
- Novinky za poslední rok
- Rozvojové plány = rychlé osvětlení výběru přednášek na dnešek

Národní distribuované výpočetní prostředí, koordinace CESNET

- zdroje CESNET i CERIT-SC (50.000 jader CPU)
- výpočetní zdroje umístěné na CESNETu, univerzitách, AV ČR
  - původní motivace sdílení zdrojů (HW) stále platí
  - poskytování dočasně volných zdrojů pro vzdálené uživatele
  - využití vzdálených zdrojů v případě naléhavého/velkého zatížení
  - a použití jiných zdrojů během výpadku
  - zdroje pro spuštění projektu, kontrola vhodnosti HW
  - myšlenka funguje i pro cloud, drahé komerční licence
- komunitní přístup, centrální správa a AAI
- grid, cloud a map-reduce výpočetní modely
- virtualizační platforma pro vysoce dostupné služby

- distribuovaný model užitečný pro různé případy použití
  - integrace zdrojů vlastněných jinými VI (ELIXIR)
  - zpracování velkých dat - není potřeba vzdáleného přenosu
  - architektura kompatibilní s datovými úložišti v EOSC CZ
- NGI v evropské e-infrastruktuře EGI, mandated org. v EOSC-A
- cílená podpora velkých projektů (VI, ESFRI)
- vývoj nových služeb/nástrojů podle potřeb uživatelů/projektů
  - OnDemand, Jupyter, Galaxy, Kubernetes
  - podpora zpracování citlivých dat
- služby volně dostupné pro akademické využití a spolupracující partnery ze zahraničí
  - "placení" publikacemi s poděkováním, určují prioritu uživatelů

- grid, centrálně spravované HTC, HPC clustery
  - dávkové, dlouhé (dny/týdny) úlohy
  - HTC i paralelní výpočty (PBSPro)
  - včetně interaktivních úloh, GUI
    - Galaxy, Jupyter, OnDemand
  - trvalé úložiště (GPFS+NFS) a místní/sdílené úložiště scratch
  - výpočty také v kontejnerech (Singularity)
    - HPC přístup, podpora pro non-root obrazy Dockerů
    - Software NVIDIA GPU Cloud
  - distribuované clustery členů a partnerů e-INFRA CZ
    - vývoj a výzkum v oblasti plánování
    - počátky vývoje Perunu (AAI)
  - podmnožina zdrojů dostupných také v infrastruktuře EGI

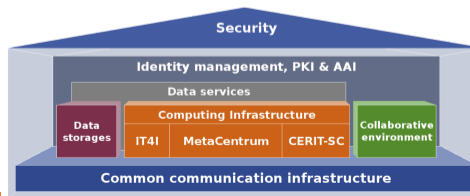
- **MetaCentrum cloud – virtuální stroj místo úloh**
  - obrazy poskytované MetaCentrem, EGI, projekty, uživateli
  - cloudové výpočty a služby spojené s výpočty (OpenStack)
    - ale také školení, výuka, bezpečnostní polygon KYPO
  - Terraform nebo EGI Infrastructure Manager pro virtuální clustery/K8s
  - centrální instalace v Brně, společné úsilí CESNET a CERIT-SC
    - od roku 2022 vývoj nové distribuce OpenStack
    - vzniká druhá instalace v IT4I
  - přístupné také prostřednictvím EGI FedCloud, EOSC Marketplace
- **MapReduce – Hadoop/Spark, jiný model zpracování dat**
  - vyhrazený cluster nahrazen virtuálními obrazy v cloudu

- Containerová platforma – containery místo úloh
  - Kubernetes pro mikroslužby, správa virtuálních prostředí
  - spravovaná služba, non-root kontejnery, GUI Rancher
  - silná podpora interaktivních požadavků a požadavků workflows
    - SaaS přístup pro Matlab, RStudio, NextFlow
  - vývoj v oblasti converged computing
    - integrace dávkového systému a Kubernetes
- GPU NVIDIA dostupné na všech platformách (až do A100)
- Sensitive cloud – práce na případech použití souvisejících se zpracováním citlivých dat
  - v současné době Kubernetes, později také OpenStack
- služba integrace zdrojů dalších organizací do MetaCentra
  - centrální správa, prioritní přístup vlastníků





- MetaCentrum je součást infrastruktury CESNETu od roku 1998
  - původně projekt MU, ZČU, UK, VUTBR
- oddělené VI CESNET a VI CERIT-SC od 2012
  - MetaCentrum stále integruje zdroje CERIT-SC
  - společný vývoj prostředí, včetně cloudových služeb
  - integrace zdrojů dalších organizací, včetně VI ELIXIR CZ
- od roku 2020 jediná společná e-infrastruktura e-INFRA CZ
  - CESNET, CERIT-SC (MU), IT4Innovations (VŠB)
  - sbližování prostředí MetaCentra a IT4Innovations



## ■ první integrační kroky v rámci e-INFRA CZ

- od minulého roku společné AAI - jednotný účet v e-INFRA CZ
- vzniká společná dokumentace <https://docs.e-infra.cz>
  - přednáška A. Melounová o <https://docs.metacentrum.cz>
- prototyp druhé cloudové instalace v IT4I - přednáška A. Rošince

## ■ nové modules pro instalaci softwarových balíků

- nástroj Spack pro automatickou kompilaci na více platformem (CPU architektura, operační systém)
  - příprava na možnost optimalizovaných balíků i pro IT4I
  - přebírání receptů z center v USA a EU
- instalované moduly dostupné přes CVMFS i pro další centra
  - AFS zůstává pro licencovaný software
- přehlednější ovládání, podpora balíků pro Python, Perl, R

- nový hardware v tomto roce
  - CESNET - HD cluster, 6144 CPU jader, probíhá akceptace
    - 96 uzlů, každý 2xAMD EPYC 7543, 512 GB RAM, 7TB disk
  - CERIT-SC - DGX-2, 8xH100 NVIDIA
  - v rámci e-INFRA CZ jsou dostupné i "komplementární systémy"
    - cluster 8xArm A64FX, dva servery s AMD FPGA (Xilinx)
- dotazník se spokojeností služeb MetaCentra, díky za odezvu
  - 2300 oslovených, 376 odpovědí, 550 částečných
  - podklad pro další rozhodování při plánování HW a SW
  - spokojenost a vnímání služeb, komunikace, školení, podpory
  - postupně reagujeme na konkrétní dotazy a připomínky
  - tento rok pod hlavičkou MetaCentra a Datových úložišť
    - příští rok snad pro celé e-INFRA CZ

- projekty EOSC CZ a NRP
  - e-INFRA CZ koordinuje vznik projektů na vytvoření národní datové infrastruktury a repozitářové platformy
    - viz prezentace ze včerejšího semináře a [prezentace L. Matysky](#)
    - součástí návrhu je i blízkost úložných a výpočetních zdrojů pro potřeby zpracování dat
- rozvoj všech základních platforem v MetaCentru
  - grid - Singularity containery a AI [přednáška A. Křenek](#) a [hands-on sekce](#)
  - cloud - nová instalace OpenStacku v IT4I [přednáška A. Rošinec](#)
  - containery - vývoj v oblasti Kubernetes [přednáška L. Hejtmánek](#)

- **nárůst využití MetaCentra o AI/ML výpočty**
  - grid 122xT4, 88xA40, 20xA100, 28x 2080
  - cloud 32xT4, 8xA40
  - kubernetes 6xA10, 24xA40, 12xA100
  - clustery partnerů z VI ELIXIR, ZČU a UK
  - uživatelská [přednáška o AI metodách v MetaCentru](#)
- **Matlab - nejpoužívanější komerční aplikace v MetaCentru**
  - speciální podpora v gridu, cloudovém prostředí Kubernetes
  - ve spolupráci s EGI i v prostředí Jupyter Notebooků
  - řada univerzit vlastní licence pro svoje studenty a zaměstnance
    - chceme dovolit využití "domácích"licencí uživatelů (BYOL model)
  - další informace [přednáška Shubo Chakrabarti](#)

- workflow systémy a interaktivních rozhraní v projektech
  - projekt EuroScienceGateway – podpora usegalaxy.eu a námi provozovaná usegalaxy.cz, přednáška A. Křenek
  - projekty EGI a služba EGI Notebooks (PaNOSC, EOSC Beyond, EOSC Procurement), přednáška V. Nykl
- interaktivní výpočty a nástroje s vlastním GUI
  - podpora OnDemand rozhraní pro PBS (a výhledově cloudové služby) - přednáška R. Leontovč
  - podpora v Kubernetes, Rancher GUI - přednáška L. Hejtmánek
  - posilujeme podporu Jupyter Notebooku v gridu, cloudu i Kubernetes přednáška V. Nykl
  - podporujeme Galaxy rozhraní přednáška A. Křenek

# Děkuji za pozornost

`support@metacentrum.cz`

`https://www.metacentrum.cz`

`https://docs.metacentrum.cz/`

`https://docs.e-infra.cz/`

