



cesnet
.....

METACENTRUM - VĚDECKÉ VÝPOČTY A VIRTUALIZACE

Miroslav Ruda

CESNET

CESNET Day, září 2024



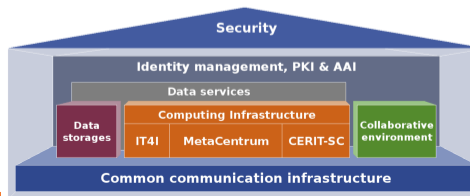
- MetaCentrum - infrastruktura pro vědecké výpočty
- Virtualizace výpočetních služeb v MetaCentru a CESNETu

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
 - 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, IaaS cloud a PaaS cloud
- 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

- **MetaCentrum je součást infrastruktury CESNETu od roku 1998**
 - původně projekt MU, ZČU, UK, VUT
- **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)
 - sblížování prostředí MetaCentra a IT4Innovations



Spolupráce s projekty jako dlouhodobá motivace vývoje služeb

- **HEP , astrofyzika - LHC, Auger, CTA, Belle**
 - původní motivace gridu, nadále aktivní v EGI
- **Life Science - ELIXIR (OpenScreen, CCT, Czech Bioimaging)**
 - spolupracující VI, zdroje začleněné do MetaCentra
 - spolupráce při provozu služeb ELIXIRu
 - life-science je největší konzument zdrojů
- **ELI, BBMRI, LINDAT/CLARIN, ICOS**
 - zejména na úrovni EGI
- **výzkumná centra CzechGlobe, CEITEC, Recetox (EIRENE)**
 - dlouhodobí uživatelé, vazba přes CERIT-SC
- **ESA – CollGS, Data Relay Hub**
 - řada skupin a projektů využívajících Copernicus data

Grid, centrálně spravované HTC a HPC clustery

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

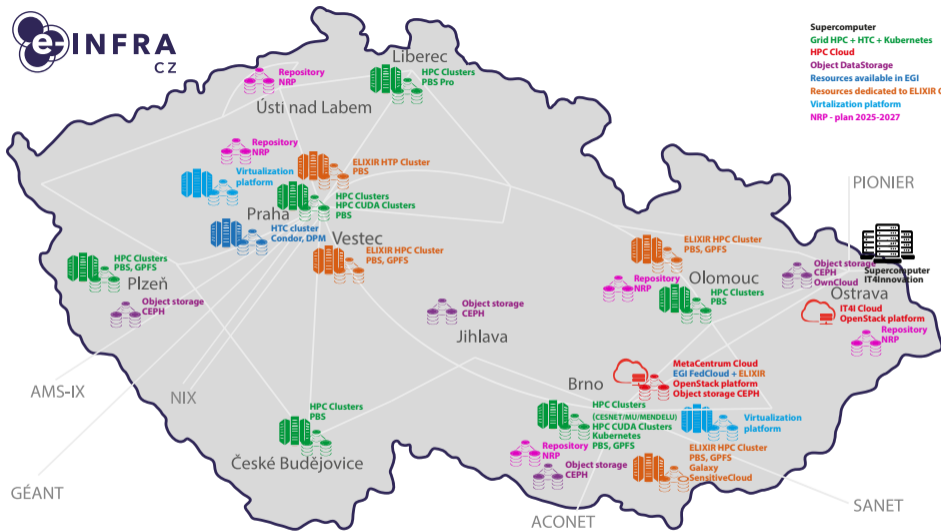
MetaCentrum cloud – virtuální stroj místo úloh

- obrazy poskytované MetaCentrem, EGI, projekty, uživateli
- cloudové výpočty a služby spojené s výpočty (OpenStack)
 - ale také školení, výuka, bezpečnostní polygon KYPO
- Terraform nebo EGI IM pro virtuální clustery/K8s
- centrální instalace v Brně, společné úsilí CESNET a CERIT-SC
 - od roku 2022 vývoj nové distribuce OpenStack - Bescar Cloud
 - od roku 2023 druhá instalace v IT4I
- přístupné i prostřednictvím EGI FedCloud, EOSC Marketplace
- v tomto prostředí provozujeme i služby EGI Jupyter a Binder
 - včetně EOSC EU uzlu a projektů ENVRI-Hub a EOSC Beyond

Containerový cloud, PaaS řešení

- **containers místo úloh/virtuálních strojů**
 - metoda balení softwarových balíčků, zapouzdření služeb
 - reprodukovatelnost, provenance
 - jak v HPC (Singularity) tak v cloudu (Docker, Kubernetes)
- **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

- Sensitive cloud – práce na případech použití souvisejících se zpracováním citlivých dat
 - koordinuje CERIT-SC
 - v současné době Kubernetes, později také OpenStack
 - od roku 2024 s certifikací ISO 27.000
- služba integrace zdrojů dalších organizací do MetaCentra
 - centrální správa, prioritní přístup vlastníků
 - dlouhodobě pro HPC clusterly
 - od roku 2024 i pro OpenStack cloud



50.000 CPU jader, 20 PB, 460 GPU, 3.000 uživatelů, 300+ WoS

- HD uzly - 32-128 CPU jader (x86_64), 256-1024 GB RAM
- SMP servery - 2-3 TB RAM
 - specializované servery s 6/10 TB RAM
- NVIDIA GP-GPU - 160 uzlů, 460 karet
 - NVIDIA DGX2 s 8xH100 GPU cards, cluster s 28xH100
 - NVIDIA T4, 1080 Ti, 2080 Ti, A100, A40, H100
- CPU jádra: grid - 32.500+6.500, IaaS 6.500, Kubernetes 2.300
- CPU roky: grid - 22.000+13.000, IaaS 10.000, Kubernetes 2.000
- clustery vlastněné rostoucím počtem partnerů
 - CESNET 23.000 a CERIT-SC 8.000
 - VI ELIXIR 5.500
 - FZU 5.400 for LHC/EGI
 - ZČU, MU, UK, TUL, AV ČR

	2012	2014	2016	2018	2019	2020	2021	2022	2023
Počet uživatelů MC	613	1112	1611	2020	2185	2225	2606	2710	3055
Noví uživatelé (Meta)	312	605	742	713	762	774	792	767	850
Počet úloh [miliony] Meta/EGI	1,1/ n/a	3,9/ n/a	3,6/ 6	5/ 6,7	8,6/ 6,8	13,1/ 10	12,1/ 9,3	11,1/ 14,2	11,7/ 5,4
CPU čas [CPU let] Meta/EGI	2500/ n/a	6403/ n/a	9475/ 5963	11357/ 4074	13129/ 4531	16630/ 9160	22647/ 9581	27547/ 9218	31858/ 14770
Počet CPU jader vč. EGI	6028	14164	17234	21344	26602	29874	34084	44088	47748
Počet GPU					255	322	434	455	462

Projekty OP JAK , e-infrastructura e-INFRA CZ a VI ELIXIR CZ

- v součtu za 2024-2026 přibližně 200 mil. Kč na clustery a semi-permanentní úložiště
 - modernizace infrastruktury -> obnova výpočetních zdrojů
 - podpora moderních trendů (GPU, AI, big data)

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>
- druhá cloudová OpenStack instalace v IT4Innovations
 - nová distribuce Bescar Cloud

Interaktivní výpočty a nástroje s vlastním GUI

- podpora OnDemand rozhraní pro PBS, interaktivní úlohy
 - RStudio, Jupyter, Fluent, Matlab
- podpora v Kubernetes, Rancher GUI
- Galaxy workflows, zprovoznění usegalaxy.cz
- podpora Jupyter Notebooků v gridu, cloudu i Kubernetes
 - včetně podpory v EOSC EU node a EGI infrastruktuře

Nový hardware v letech 2023 a 2024

- CESNET 2023 - HD cluster, 96 serverů, 6144 CPU jader
- CERIT-SC 2023 - server DGX-2 8x NVIDIA H100
- CERIT-SC 2024 - GPU cluster
 - 18 uzlů, 28x NVIDIA H100, 1PB scratch
- CESNET 2024 - probíhající výběrové řízení
 - GPU servery (80+ NVIDIA L40s) a Ceph (6x120 TB)
- CESNET 2024 Elixir - probíhající výběrové řízení
 - 40 HD/SMP serverů, 20 Ceph uzlů

Děkuji za pozornost

`support@metacentrum.cz`

`https://www.metacentrum.cz`

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

`https://metavo.metacentrum.cz/cs/state/
stats/2023/index.html`

Vědecké výpočty

- Grid - Singularity containery
- OpenStack - virtualizace na úrovni celého stroje
- Kubernetes - Docker, orchestrace většího počtu containerů

Služby pro správu vědeckých výpočtů

- webové služby v OpenStacku, servery zajišťující sběr a vystavení dat
- složitější služby složené z microservices - Kubernetes, OS s K8s

Software-as-a-Service

- Galaxy, Jupyter Notebooky
- řada dalšího software v Kubernetes

Zajištění důvěrnosti a bezpečnosti

- Sensitive Cloud - zpracování citlivých (medicínských dat)
 - služba koordinovaná CERIT-SC, CESNET se postupně zapojuje
- striktně oddělení uživatelé, více zabezpečený přístup, včetně fyzických opatření, data nesmí opustit prostředí
- v případě, že stačí ukládat šifrovaná data, má CESNET i řešení v oddělení Datových úložišť (šifrované zálohy)

Zabezpečení vysoké dostupnosti

- virtualizační platforma CESNETu
- důraz není na vysoký výkon, ale dostupnost při výpadku libovolné HW komponenty i jednoho centra

- VMWare vSphere 8 pro provoz vysoce zabezpečených služeb
- důsledně redundantní konfigurace
 - dvě autonomní centra (Praha, Brno), v každém dvě linky do páteře CESNETu
 - v každém centru zdvojené switche, servery, pole, ...
 - živá migrace virtuálního stroje mezi centry
 - zálohy do druhé centra, diskové pole, datová úložiště CESNETu
- personální průnik s VMWare skupinou na VŠB
 - spolupráce s dalšími univerzitami
- původně pro zabezpečení služeb CESNETu
- postupně i služby našich projektů, univerzit (placená služba)

- novinky za poslední roky
 - možnost využití aplikačního firewallu Fortigate
 - možnost Anycast routování i pro IP rozsahy univerzit
 - podpora TPM čipů pro uložení klíčů a šifrování
 - Veeam - zálohování do S3 úložišť CESNETu
- stav k září 2024 - 740 virtuálních strojů, 90+ za Fortigate fw
- vzhledem k nejasnosti licencování VMWare testujeme náhrady
 - ProxMox, HA OpenStack
 - možnost semináře během konference CESNET v listopadu

Děkuji za pozornost

`https://virtualizace.cesnet.cz`

`cesnet-virtual@cesnet.cz`

