



Centrum CERIT-SC

Tomáš Rebok

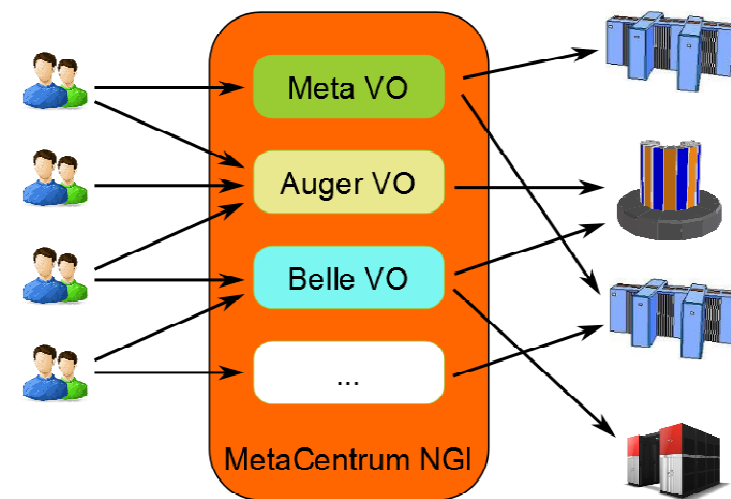
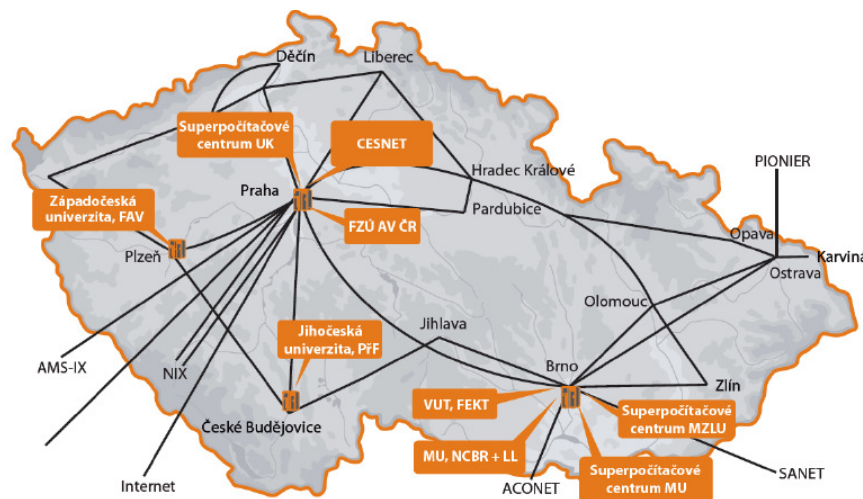
CERIT-SC, Masarykova Univerzita

MetaCentrum, CESNET z.s.p.o.

(rebok@ics.muni.cz)

MetaCentrum NGI @ CESNET

- koordinátor **národní gridové infrastruktury**
 - integruje velká/střední HW centra (**clustery, výkonné servery a úložiště**) několika univerzit/organizací v rámci ČR
 - → prostředí pro (spolu)práci v oblasti výpočtů a práce s daty
 - uživatelé sdružováni do tzv. **virtuálních organizací**
 - např. MetaVO



MetaCentrum VO (Meta VO)

- přístupné zaměstnancům a studentům VŠ/univerzit, AV ČR, výzkumným ústavům, atp.

- nabízí:

<http://metavo.metacentrum.cz>

- výpočetní zdroje
- úložné kapacity
- aplikační programy

- po **registraci** k dispozici **zcela zdarma**

- „placení“ formou publikací s poděkováním
 - → prioritizace uživatelů při plném vytížení zdrojů

- součást budované e-infrastruktury CESNETu

- datová úložiště, kolaborativní prostředí, ...



Meta VO – dostupný hardware

- výpočetní zdroje: cca **5700 jader** (x86_64)
 - klasické HD uzly (2x4-8 jader) i SMP stroje (32-80 jader)
 - paměť až 512 GB per uzel
 - Infiniband pro nízkolatenční komunikaci (MPI)
- **400 TB** pro semipermanentní data
 - úložiště 3x v Brně a 1x v Plzni, přístupné na všech clusterech
 - výhledově napojení na **PB** úložiště permanentních dat
- dostupnost specializovaného vybavení
 - např. NVIDIA CUDA karty v Plzni, 35TB scratch pro pracovní data

Meta VO – dostupný hardware

- výpočetní zdroje: cca **5700 jader** (x86_64)

- klasické
- paměť
- Infiniband

- **400 TB**

- úložiště
- výhled

- dostupné

- např. M

Co se od dubna změnilo?

- nové **clustery** -- počet jader **navýšen** o cca **1700**
- probíhá **výměna clusterů na JČU** (další jádra)
- finišuje **nákup stroje s 1 TB RAM**
- navyšujeme počty **uzlů s GPU kartami**
- příprava nákupu **2 nových ~ 100 TB polí** (Praha, Budějovice) pro semipermanentní data
- průběžná realizace **napojení na PB úložiště dat** (zatím zprovozněno jen zálohování)

- ...

Meta VO – trocha čísel ☺

- *HW výbava:*
 - ~ 2-8 GB na jádro (většina jader cca 6-8 GB na jádro)
 - viz <http://metavo.metacentrum.cz/pbsmon2/hardware>
- *sdílený scratch 35 TB*
 - k němu připojeno 14 uzlů → 900 CPU, 4GB na jádro
- *leden – září 2012: propočítáno **cca 20mil. corehours***
 - = cca 2,2 mil. corehours za měsíc
- *se stávajícím HW (navýšeno) lze teoreticky propočítat až 50mil. corehours za rok*
 - reálně dosažitelné vytížení je cca 70-80 %, tj. **cca 40 mil. c/h / rok**
 - počet jader budeme významně navyšovat (CERIT-SC)
- *příklad aktivity uživatelů (celkem ~ 550 uživatelů):*
 - max. **1,25mil. corehours / uživatel za** (posledního) **půl roku**
 - max. **350tis. corehours / uživatel za měsíc**



Meta VO – software

- ~ **160 různých aplikací**
 - viz <http://meta.cesnet.cz/wiki/Kategorie:Aplikace>
- průběžně udržované **vývojové prostředí**
 - GNU, Intel, PGI, ladící a optimalizační nástroje (TotalView, Alinea), ...
- generický **matematický software**
 - Matlab, Maple, gridMathematica, ...
- komerční i volný **software pro aplikační chemii**
 - Gaussian 09, Amber, Gamess, ...
- **materiálové simulace**
 - Wien2k, Fluent CFD, ...
- **strukturní biologie, bioinformatika**
 - řada volně dostupných balíčků
- **hledáme náměty na další sdílitelný/generický software**
 - i komerční

Meta VO – software

- ~ **160 různých aplikací**
 - viz <http://meta.cesnet.cz/wiki/Kategorie:Aplikace>

- průběžně udržované **vývoiové prostředí**

- G
- gen
 - M
- kom
 - c

- **mat**
 - v

- **stru**
 - ř

- i
- **hlec**
 - i

- i
- **hlec**
 - i

- i
- **hlec**
 - i

Co se od dubna změnilo?

- **Matlab (8.0)**, navýšení licencí **Matlabu** o 100 ks

- ladící nástroje **TotalView 8.10**, **Allinea DDT 3.2**

- **Ansys CFD 14.0** (Fluent + CFX), **Ansys HPC**

- **Maple 16**

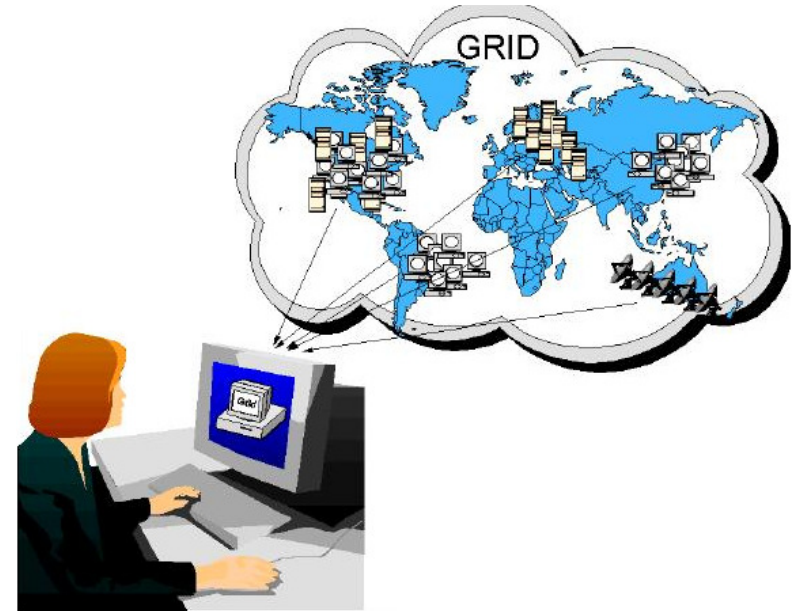
- **PGI CDK 12.4**

- navýšení počtu licencí **Intel CDK 12**

- **SciLab, CMAQ, Moses, Mosaik, Gromacs, QEspresso, ...**

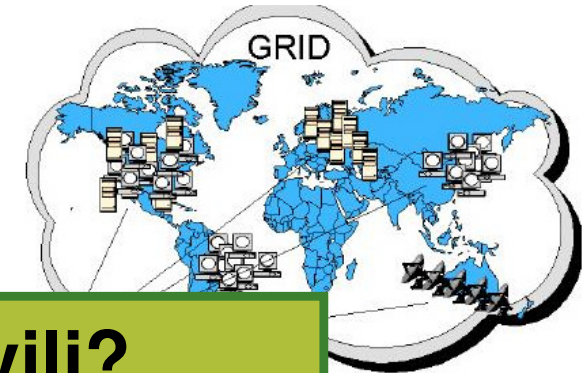
Meta VO – výpočetní prostředí

- *dávkové úlohy*
 - popisný skript úlohy
 - oznámení startu a ukončení úlohy
- *interaktivní úlohy*
 - **textový** i **grafický** režim
- *cloudové rozhraní*
 - základní kompatibilita s Amazon EC2
 - uživatelé **nespouští úlohy**, ale **virtuální stroje**
 - opět zaměřeno na vědecké výpočty
 - možnost vyladit si obraz a přenést ho do MetaCentra (Windows, Linux)
 - podpora pro aplikace, kterým gridový přístup nevyhovuje



Meta VO – výpočetní prostředí

- *dávkové úlohy*
 - popisný skript úlohy
 - oznámení startu a ukončení úlohy



■ *inte*

□ *t*

■ *clo*

□ *z*

□ *u*

Co nového jsme mj. připravili?

- školící hands-on semináře
 - výjezdní semináře (Brno, 2x Budějovice, 2x Olomouc)
 - zaměřeno na *nové a středně pokročilé uživatele*
 - úzká skupina uživatelů se společným zájmem
 - praktická část semináře přizpůsobena potřebám skupiny
- v první polovině příštího roku vícero výjezdů (Ostrava?)



Centrum CERIT-SC

- výzkumné centrum budované na ÚVT MU
 - transformace **Superpočítačového centra Brno (SCB)** při Masarykově univerzitě do nové podoby
 - zaměřeno na aplikaci špičkové ICT ve vědě
- prioritní projekt *Cestovní mapy velkých infrastruktur ČR*
 - *spolu s VI Cesnet a IT4Innovations*
- **významný člen/partner** národního gridu (∈ MetaVO)
 - I. poskytovatel **HW zdrojů**
 - SW výbava totožná s MetaVO
 - II. služby nad rámec „běžného“ HW centra –

zázemi pro kolaborativní výzkum



Cíle Centra

Studium a posun možností špičkové eInfrastruktury úzkou výzkumnou spoluprací mezi informatiky a uživateli takovéto infrastruktury. Podpora experimentů s novými formami a konfiguracemi eInfrastruktury.



Cíle Centra

*Studium a posun možností špičkové eInfrastruktury úzkou výzkumnou spoluprací mezi informatiky a uživateli takovéto infrastruktury. **Podpora experimentů s novými formami a konfiguracemi eInfrastruktury.***

■ infrastruktura

- vysoce flexibilní, experimentům příznivé výpočetní a úložné prostředí
 - instalované technologie slouží **primárně vývoji a experimentům**
 - až ve druhé rovině pro klasické výpočty a ukládání dat
 - hrubá síla přístupná integrací do MetaCentra a mezinárodních infrastruktur (EGI)
- **základní principy:**
 - analogie *best effort* známého z počítačových sítí
 - **minimální administrativa** pro uživatele (**žádné žádosti**)
 - **podpora špičkové vědy** - prioritizace přístupu na základě dosažených výsledků



Cíle Centra

*Studium a posun možností špičkové eInfrastruktury úzkou výzkumnou spoluprací mezi informatiky a uživateli takovéto infrastruktury. **Podpora experimentů s novými formami a konfiguracemi eInfrastruktury.***

■ infrastruktura

□ HW výbava:

- SMP uzly (1600 jader)
- HD uzly (580 jader,
cíl Q1/2013 >2500 jader)
- úložné kapacity (~ 3,2 PB,
cíl Q1/2013 >3,5 PB)





Cíle Centra

Studium a posun možností špičkové eInfrastruktury úzkou výzkumnou spoluprací mezi informatiky a uživateli takovéto infrastruktury. Podpora experimentů s novými formami a konfiguracemi eInfrastruktury.

■ výzkum a vývoj

- **vlastní** – zaměřený na principy a technologie eInfrastruktury a její optimalizaci
 - plánování v distribuovaném prostředí
 - řízení virtualizovaného prostředí a cloudů
 - bezpečnost
 - optimalizace výpočtů pro paralelní/distribuované prostředí
 - atp.



Cíle Centra

Studium a posun možností špičkové eInfrastruktury úzkou výzkumnou spoluprací mezi informatiky a uživateli takovéto infrastruktury. Podpora experimentů s novými formami a konfiguracemi eInfrastruktury.

■ výzkum a vývoj

- **kolaborativní s uživateli/partnery** – algoritmy, modely, prostředí a nástroje pro efektivní využití eInfrastruktury
 - **analýza problémů partnerů**, uživatelů eInfrastruktury
 - společný „projekt“, **skutečná vědecká spolupráce**, ne jen nabízení výkonu
 - společné týmy, zapojení nejen odborných pracovníků, ale zejména studentů z obou stran
 - potenciál transformace do skutečných externě financovaných projektů



Cíle Centra

Studium a posun možností špičkové infrastruktury úzkou výzkumnou spoluprací mezi informatiky a uživateli takovéto infrastruktury. Podpora experimentů s novými formami a konfiguracemi infrastruktury.

■ výzkum a vývoj

□ odborné zázemí Centra:

- organizačně součást Ústavu výpočetní techniky MU
- dlouholetá tradice **spolupráce s Fakultou informatiky MU**
 - personální i technická (společná laboratoř SITOLA, společný VZ)
- dlouholetá tradice **spolupráce se sdružením CESNET**
 - SCB (nyní CERIT-SC) je zakladatel MetaCentra



Příklady spolupráce s partnery I.

■ Rekonstrukce stromu z jeho laserového skenu

- **partner:** *CzechGlobe* (prof. Marek, doc. Zemek, dr. Hanuš, dr. Kaplan)
- **cíl projektu:** návrh algoritmu pro rekonstrukci stromu (smrků)
 - z mraku nasnímaných 3D bodů
 - strom nasnímán laserovým snímačem LIDAR
 - výstupem jsou souřadnice XYZ + intenzita odrazu
 - *očekávaný výstup:* struktura popisující strom (Blender)
- **hlavní problémy:** překryvy (mezery v datech)





Příklady spolupráce s partnery I.

- **Rekonstrukce stromu z jeho laserového skenu**
 - **stávající řešení:** poloautomatické metody zpracování
 - překryvy je nutno ručně doplnit
 - **námi navržené a implementované řešení:**
 - *plně automatická* metoda rekonstrukce kmene a hlavních větví
 - následné postupné dorůstání zelené biomasy L-systémem
 - **předpokládané ukončení: I/2013**





Příklady spolupráce s partnery II.

■ Použití neuronových sítí pro doplňování chybějících dat eddy-kovariančních měření

□ **partner:** *CzechGlobe* (prof. Marek, dr. Pavelka)

□ **cíl projektu:** nalezení nové, plně automatické metody pro doplňování chybějících měření

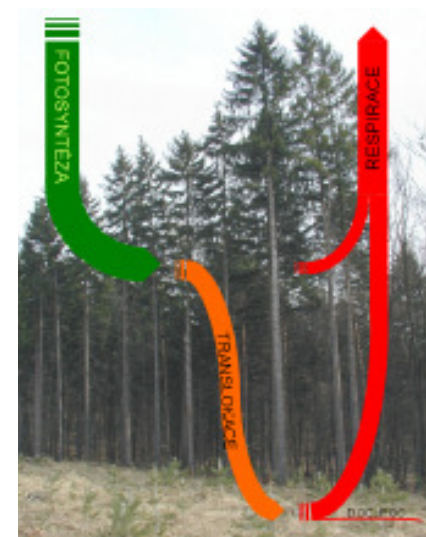
■ formou učení na historických datech

□ *doprovodné charakteristiky* – teplota, tlak, vlhkost, ...

□ **hlavní problémy:**

■ nutnost brát v úvahu i historická data

■ les se vyvíjí (roste)





Příklady spolupráce s partnery II.

■ Použití neuronových sítí pro doplňování chybějících dat eddy-kovariančních měření

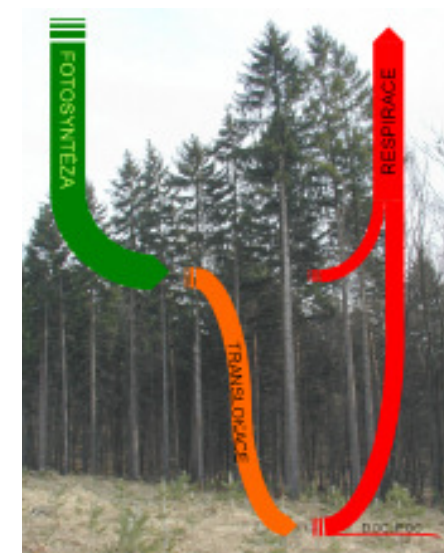
□ stávající řešení:

- jednoduché empirické modely (okamžitý stav mikroklimatu)
- hledání období s podobnými hodnotami mikrometeorologických parametrů
- statistické metody

□ naše řešení:

- již realizovaná neuronová síť
 - *slibné výsledky*

□ předpokládané ukončení: I/2013

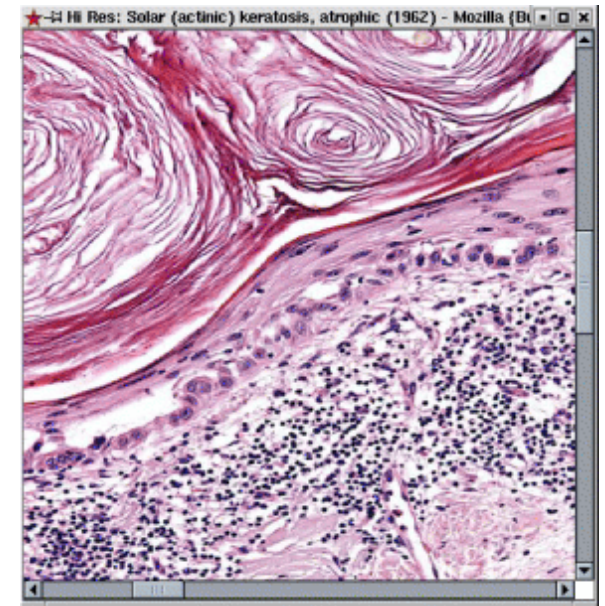




Příklady spolupráce s partnery III.

■ Virtuální mikroskop, patologické atlasy

- **partner:** *LF MU* (doc. Feit)
- **cíl projektu:** implementace virtuálního mikroskopu pro dermatologický atlas (webová aplikace)
 - zobrazuje skeny tkání
 - rozlišení až 170000x140000 pixelů
 - složeno z dlaždic (až 30000 ks)
 - umožňuje „doostřovat“ jako skutečný mikroskop
- **hlavní problémy:**
 - optimalizace zpracování snímků, autentizace





Příklady spolupráce s partnery III.

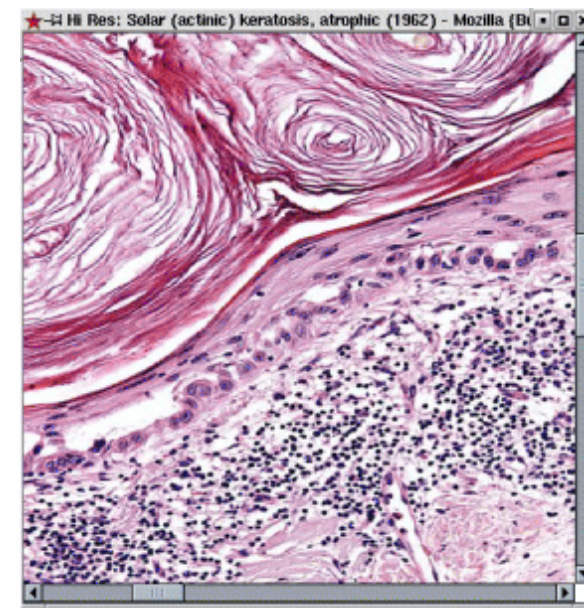
■ Virtuální mikroskop, patologické atlasy

□ náš přínos:

- doba zpracování snímků zkrácena z cca 2 dnů na cca 8 hodin
- nasazení formátu JPEG2000
- GPU implementace

□ společný projekt:

- *Výuka patologie pomocí hypertextových výukových materiálů a obrazových databází pro magisterské, pregraduální, doktorské i postgraduální studium (OPVK CZ.1.07/2.2.00/28.0045)*





Příklady spolupráce s partnery IV.

- **ELIXIR (ESFRI projekt) – bio-informační infrastruktura**
 - **partner:** *ÚOCHB AV ČR* (dr. Vondrášek), *BIOMED AV ČR* (dr. Vohradský)
 - **cíl projektu:** vybudování českého uzlu evropské bio-informační infrastruktury
 - data z oblastí přírodovědy a medicíny, případně i z příslušně orientovaných průmyslových odvětví
 - potřeba s daty organizovaně zacházet
 - sdružování, správa a přístup k datům
 - nutná celoevropská a celosvětová koordinace
 - CERIT-SC je zakládající člen konsorcia ELIXIR_CZ



Příklady spolupráce s partnery V.

- **Biobanka klinických vzorků (BBMRI_CZ)**
 - *partner: Masarykův onkologický ústav, Recamo*
- **Modely šíření epileptického záchvatu a dalších dějů v mozku**
 - *partner: LF MU, ÚPT AV, CEITEC*
- **Fotometrický archív astronomických snímků**
- **Extrakce fotometrických údajů o objektech z astronomických snímků**
- **Automatické fitování kontinua echelletovských spekter**
 - *3x partner: Ústav teoretické fyziky a astrofyziky PŘF MU*
- **Bioinformatická analýza dat z hmotnostního spektrometru**
 - *partner: Ústav experimentální biologie PŘF MU*
- **Identifikace oblastí zasažených geometrickými distorzemi v leteckých skenech krajiny**
- **Synchronizace časových značek v leteckých snímcích krajiny**
 - *2x partner: CzechGlobe*

...



Shrnutí

■ Centrum CERIT-SC

- nové jméno, dlouhá historie (SCB *1994)
- transformace z poskytovatele zdrojů na **výzkumného partnera**
 - snaha o inteligentní využití zdrojů (ne jen hrubou sílu)
- **silná výzkumná složka**
 - kapacita ne samoúčel, ale jako nezbytná podmínka pro ověřování nových myšlenek
- Centrum postaveno na **komunikaci a intenzivní spolupráci s uživatelskými skupinami**



Děkuji Vám za pozornost!

rebok@ics.muni.cz



Pozvánka: Školící hands-on seminář MetaCentra (14.+15.11., Olomouc)

www.cerit-sc.cz

www.cesnet.cz

www.metacentrum.cz