



CERIT-SC

9 měsíců po OP VaVpl a co dál

Aleš Křenek

Seminář MetaCentra 2. 12. 2014



EVROPSKÁ UNIE
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
INVESTICE DO VAŠÍ BUDOUCNOSTI



OP Výzkum a vývoj
pro inovace

- ▶ SCB založeno 1994, projekt FRVŠ

- ▶ SCB založeno 1994, projekt FRVŠ
- ▶ doba superpočítačová
 - ▶ 1994–1999
 - ▶ SGI Challenge, Origin, procesory MIPS

- ▶ SCB založeno 1994, projekt FRVŠ
- ▶ doba superpočítačová
 - ▶ 1994–1999
 - ▶ SGI Challenge, Origin, procesory MIPS
- ▶ doba clusterová
 - ▶ od 2000 dosud
 - ▶ IA32, (IA64), x86_64

- ▶ SCB založeno 1994, projekt FRVŠ
- ▶ doba superpočítačová
 - ▶ 1994–1999
 - ▶ SGI Challenge, Origin, procesory MIPS
- ▶ doba clusterová
 - ▶ od 2000 dosud
 - ▶ IA32, (IA64), x86_64
- ▶ Metacentrum a doba gridová
 - ▶ už původní projekt FRVŠ pokrýval více center
 - ▶ 1996 koncept MetaCentra
 - ▶ od 2001 EDG, EGEE, EGI

- ▶ SCB založeno 1994, projekt FRVŠ
- ▶ doba superpočítačová
 - ▶ 1994–1999
 - ▶ SGI Challenge, Origin, procesory MIPS
- ▶ doba clusterová
 - ▶ od 2000 dosud
 - ▶ IA32, (IA64), x86_64
- ▶ Metacentrum a doba gridová
 - ▶ už původní projekt FRVŠ pokrýval více center
 - ▶ 1996 koncept MetaCentra
 - ▶ od 2001 EDG, EGEE, EGI
- ▶ doba cloudová – CERIT-SC

Jak jsme je viděli v roce 2011

- ▶ **Infrastruktura**
 - ▶ poskytnout pro potřeby národního a mezinárodního výzkumu
 - ▶ vysoce flexibilní výpočetní prostředí
 - ▶ odpovídající úložné kapacity
 - ▶ budovat cloudové prostředí pro VaV
 - ▶ udržet pozici nejvýznamnějšího národního gridového uzlu
- ▶ **Výzkum a vývoj**
 - ▶ v ICT: technologie pro e-infrastrukturu (výpočty, data, bezpečnost) a její vývoj
 - ▶ s uživateli/partnery: algoritmy, modely, prostředí, nástroje a prostředí pro efektivní využití e-infrastruktury
 - ▶ obecně: synergický vývoj e-infrastruktury a způsobů jejího využití
- ▶ **Výuka**

- ▶ Vlastní výzkumná témata
 - ▶ plánování v distribuovaném prostředí
 - ▶ řízení virtualizovaného prostředí
 - ▶ bezpečnost
 - ▶ PKI (Public Key Infrastructure)
 - ▶ federace identit a její další vývoj
 - ▶ autorizace
 - ▶ důvěra a její správa
 - ▶ optimalizace výpočtů pro paralelní a distribuované prostředí
 - ▶ výpočetní a úložné prostředí (GPU, nové protokoly, ...)
- ▶ Budování a rozvoj doktorské školy
 - ▶ s využitím zázemí FI a ÚVT

- ▶ Spolupráce se sdružením CESNET (MetaCentrum, DÚ, bezpečnost, ...)
 - ▶ poskytování flexibilních zdrojů, důsledná virtualizace
 - ▶ spolupráce při vývoji metod a nástrojů pro technicky složitější prostředí s vyšší distribuovaností zdrojů
 - ▶ práce s daty

- ▶ 100 MKč investic, 5/2011 – 4/2014
- ▶ clustery
 - ▶ HD: 192 uzlů, 8–16 jader (E5-2620 – E5-2643), 96–128 GB RAM
 - ▶ SMP: 20 uzlů, 80 jader (E7-[24]860), 512–1250 GB RAM
 - ▶ UV2: 288 (+ 384) jader, 2x 6 TB RAM
- ▶ úložiště
 - ▶ scratch 60–250 GB/jádro
 - ▶ 260+370 TB standardní /home
 - ▶ 4 PB MAID HSM
- ▶ síťové propojení IB + 10GE

- ▶ uživatelé
 - ▶ všichni registrovaní uživatelé MetaCentra
 - ▶ uživatelé spolupracujících institucí bez nároku na členství v Meta
 - ▶ v omezené míře „anonymní“ uživatelé komunitních služeb (webové aplikace)

- ▶ uživatelé
 - ▶ všichni registrovaní uživatelé MetaCentra
 - ▶ uživatelé spolupracujících institucí bez nároku na členství v Meta
 - ▶ v omezené míře „anonymní“ uživatelé komunitních služeb (webové aplikace)
- ▶ disková pole /home
 - ▶ pracovní prostor pro clustery
 - ▶ dostupné standardně v metacentru (/storage/*-cerit)
 - ▶ export NFSv4, zvažujeme CIFS (postupná integrace s úložišti na MU)

- ▶ uživatelé
 - ▶ všichni registrovaní uživatelé MetaCentra
 - ▶ uživatelé spolupracujících institucí bez nároku na členství v Meta
 - ▶ v omezené míře „anonymní“ uživatelé komunitních služeb (webové aplikace)
- ▶ disková pole /home
 - ▶ pracovní prostor pro clustery
 - ▶ dostupné standardně v metacentru (/storage/*-cerit)
 - ▶ export NFSv4, zvažujeme CIFS (postupná integrace s úložišti na MU)
- ▶ MAID HSM
 - ▶ experimentálnější, než jsme čekali (i pro dodavatele)
 - ▶ archiv, resp. aplikace nad velkými repozitáři
 - ▶ near-real-time repliky jiných úložišť
 - ▶ další experimenty

- ▶ **clustery**
 - ▶ dávkové zpracování úloh
 - ▶ maximální možná shoda s MetaCentrem (debian, /software, instalované balíky, ...)
 - ▶ Torque s rozvrhovým plánovačem (viz přednáška Vaška Chlumského)
 - ▶ dedikované uzly
 - ▶ pouze speciální případy (100% využití jinak node-lock licencí)
 - ▶ OpenNebula – cloudová střelnice
 - ▶ projektové experimenty, uživatelské obrazy VM
 - ▶ rozšiřování ručně, na úkor dávkového prostředí
 - ▶ OpenNebula – robustnější cloudové prostředí
 - ▶ provoz uživatelských služeb po odladění
 - ▶ konzervativnější režim
 - ▶ robustní diskové úložiště

- ▶ Gordon E. Moore nám nepřeje
 - ▶ 1994: grond.ics.muni.cz 1.8 GFLOPS, nejvýkonnější v ČR, v TOP 500
 - ▶ 1996: první superpočítač 1 TFLOPS (Intel ASCI Red)
 - ▶ 2008: první TFLOPS GPU

- ▶ Gordon E. Moore nám nepřeje
 - ▶ 1994: grond.ics.muni.cz 1.8 GFLOPS, nejvýkonnější v ČR, v TOP 500
 - ▶ 1996: první superpočítač 1 TFLOPS (Intel ASCI Red)
 - ▶ 2008: první TFLOPS GPU
- ▶ cluster s dostatečným výkonem si postaví mnohé výzkumné skupiny samy
- ▶ hrubým výkonem neoslovíme, musíme dát víc

- ▶ Gordon E. Moore nám nepřeje
 - ▶ 1994: grond.ics.muni.cz 1.8 GFLOPS, nejvýkonnější v ČR, v TOP 500
 - ▶ 1996: první superpočítač 1 TFLOPS (Intel ASCI Red)
 - ▶ 2008: první TFLOPS GPU
- ▶ cluster s dostatečným výkonem si postaví mnohé výzkumné skupiny samy
- ▶ hrubým výkonem neoslovíme, musíme dát víc
- ▶ aktuální formulace poslání CERIT-SC (FormB_CERIT-SC.docx): „The most important for the RI CERIT-SC are the collaborative aspects of its research, supporting the leading edge research in non-informatics disciplines, taking care of the informatics-related part and allowing thus researchers and their teams to move more quickly with their own research.“

- ▶ > 20 uzavřených dohod o spolupráci
 - ▶ typicky využití zdrojů + konkrétní témata ke spolupráci
 - ▶ monitorovací indikátor projektu
- ▶ publikace
 - ▶ 2011–14: cca. 75 článků s autorstvím širšího týmu centra
 - ▶ více než 3/4 ve spolupráci s uživateli
- ▶ studenti
 - ▶ 9 zahájených doktorských studií na související téma
 - ▶ desítky Mgr. a Bc. prací

- ▶ rekonstrukce 3D struktury stromu z lidarových dat
- ▶ mnohonásobné urychlení korektoru sekvenačních chyb
- ▶ simulace chování sběru dat z inteligentních elektroměrů
- ▶ simulace dopadu silničních uzavírek
- ▶ funkční modely mozku
- ▶ ... *už jsme o nich mluvili jinde*

- ▶ Laboratoř proteomiky CEITEC
 - ▶ instalace a konfigurace laboratorního systému
 - ▶ napojení na systémy správy identit
 - ▶ postupně se objevují problémy na zpracování většího objemu dat
- ▶ Ústav geografie PŘF MU
 - ▶ městský klimatický model Muklimo
 - ▶ zprovoznění programu v prostředí CERIT-SC
 - ▶ možnosti velkého zrychlení paralelizací
- ▶ Mendelovo centrum proteomiky a genomiky rostlin
 - ▶ registrace sekvence mikroskopických snímků živých buněk
 - ▶ zpracování většího objemu dat (10 GB/snímek)
 - ▶ příprava automatizovaného workflow v prostředí CERIT-SC

- ▶ FP7 SDI4Apps: *Uptake of Open Geographic Information Through Innovative Services Based on Linked Data*
 - ▶ navrhnout vhodnou architekturu, postavit testbed
- ▶ FP7 Thalamoss: *THALAssaemia MOdular Stratification System for Personalized Therapy of beta-Thalassemia*
 - ▶ odpovědnost za správu a zpracování získaných dat
- ▶ TAČR: *Platforma pro poskytování specializovaných meteo-predikací pro oblast energetiky*
 - ▶ návrh infrastruktury, provoz testbedu
- ▶ GA MU: *Development of new interdisciplinary approaches to study function and regulation of transcriptome*
 - ▶ vývoj nových kombinovaných výpočetních metod ve strukturní biologii

Nikdy není tak zle, abychom si nemohli napsat další projekt.

Nikdy není tak zle, abychom si nemohli napsat další projekt.

- ▶ remake #4, 2016–2022, 57 MKč ročně (včetně spoluúčasti)
- ▶ stabilní úroveň reinvestic
 - ▶ G. E. Moore hraje s námi
 - ▶ dostatečně atraktivní technologie
 - ▶ stačí pro experimentální použití a přetok
 - ▶ mírný pokrok v technologiích
- ▶ plnohodnotná realizace principů spolupráce a flexibilní infrastruktury
- ▶ výrazný nárůst týmu – 12–15 FTE
 - ▶ stabilní jádro
 - ▶ posíleno dalšími projekty, studenty
- ▶ důraz na mezinárodní spolupráci
- ▶ probíhající hodnocení v prosinci 2014

- ▶ VI Elixir CZ
- ▶ H2020 OncoSteer
- ▶ GAČR: Rozsáhlá analýza sensitivity parametrů u specifických vysoce deformovatelných modelů měkkých tkání
- ▶ GAČR: Enhanced Sampling Simulations for Complex Systems
- ▶ GAČR: Rostliny – tradiční inspirace pro moderní medicínu

- ▶ H2020 Elixir-Excelerate
- ▶ H2020 Instruct VRE
- ▶ H2020 modely měkkých tkání
- ▶ ...

- ▶ OP VVV Zázemí pro ÚVT MU, CERIT-SC je jeden z hlavních směrů

- ▶ cíle z roku 2011 naplněny
 - ▶ centrum se podařilo vybudovat
 - ▶ technické vybavení v provozu
 - ▶ fungující tým
 - ▶ vysoká míra využití
- ▶ základní myšlenka – spolupráce s uživateli
 - ▶ ukazuje se jako správná
 - ▶ začínají se objevovat společné publikace
 - ▶ dokladem jsou zejména společné projekty
- ▶ mírný finanční optimismus
 - ▶ provozní náklady 2015 zajištěny
 - ▶ 2016–2022 spoléháme na kombinaci VI, OP VVV a dalších projektů