



CERIT-SC

Luděk Matyska, David Antoš, Aleš Křenek

7. 11. 2011



EVROPSKÁ UNIE
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
INVESTICE DO VAŠÍ BUDOUCNOSTI



OP Výzkum a vývoj
pro inovace

- ▶ Historie a vznik centra CERIT-SC (CERIT Scientific Cloud)
- ▶ Centrum CERIT jako celek
- ▶ Cíle CERIT-SC
- ▶ Možnosti využití centra
 - ▶ hardwarové zdroje
 - ▶ režim využívání zdrojů centra
 - ▶ spolupráce s uživateli ve výzkumu a vývoji
- ▶ Projektové zajištění
 - ▶ OP VaVpl
 - ▶ národní infrastruktura pro výzkum a vývoj

- ▶ CERIT-SC vzniká transformací Superpočítačového centra Brno (SCB)
 - ▶ CERIT-SC oficiálně ustaven 1. 5. 2011
- ▶ SCB – součást Ústavu výpočetní techniky Masarykovy univerzity
- ▶ SCB vzniklo v roce 1994
 - ▶ jedno z největších superpočítačových center v ČR
 - ▶ spolupráce s UK, ZČU, VUT a VŠB-TUO
 - ▶ zakladatel MetaCentra
- ▶ Dlouhodobé propojení s Fakultou informatiky MU
 - ▶ personální i věcné (Laboratoř SITOLA)
 - ▶ výuka i výzkum – společný výzkumný záměr „Vysoce paralelní a distribuované výpočetní systémy“
- ▶ Dlouhodobá spolupráce se sdružením CESNET

- ▶ CERIT-SC je součástí širšího projektu CERIT
 - ▶ Centrum vzdělávání, výzkumu a inovací pro ICT
- ▶ Další součásti centra CERIT:
 - ▶ CERIT Science Park
 - ▶ podpora průmyslového výzkumu, technologického vývoje a inovací
 - ▶ spolupráce s průmyslovými partnery
 - ▶ CERIT Research & Education
 - ▶ posílení infrastruktury studia a výzkumu na Fakultě informatiky MU
 - ▶ primární investice do rozšíření budovy fakulty
- ▶ CERIT-SC je prioritním projektem *Cestovní mapy ČR velkých infrastruktur pro výzkum, vývoj a inovace*
 - ▶ Společně se sdružením CESNET a Centrem IT4Innovations
 - ▶ <http://www.cerit-sc.cz/>

- ▶ **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

- ▶ 2 základní typy výpočetních zdrojů
 - ▶ SMP – Symmetric MultiProcessing, ≥ 64 jader ve sdílené paměti
 - ▶ HD – High Density \approx cluster výkonných PC
- ▶ Q4/2011: SMP cluster (500 640 jader), úložiště (250 TB)
 - ▶ 8 uzlů po 80 jádrech a 256 GB RAM
- ▶ Q2/2012: HD cluster (500 jader), SMP cluster (500 jader)
- ▶ Q3/2012: hierarický úložný systém pro dlouhodobá data (4 PB)
- ▶ Q1/2013: HD cluster (2000 jader), úložiště pro pracovní data (350 TB)
- ▶ Software průběžně
 - ▶ vývojové nástroje, Matlab, Maple, ...
 - ▶ aplikační software podle okolností

počty jader a kapacity v současném poměru cena/výkon (reálně očekáváme více)

- ▶ Základní principy:
 - ▶ uživatelský režim vycházející z MetaCentra
 - ▶ analogie best effort známá ze sítí
 - ▶ prioritizace postavená na dosažených kvalitních výsledcích
 - ▶ minimální administrativa pro uživatele
- ▶ Výpočetní zdroje
 - ▶ grid a cloud rozhraní
 - ▶ široká škála režimů využití včetně bezprostřední interaktivní práce
- ▶ Datové zdroje
 - ▶ přímé využití koncovými uživateli
 - ▶ integrovány s výpočetním prostředím a jeho virtualizací
 - ▶ až po permanentní archivaci a sdílení dat
- ▶ Zapojení do národní a mezinárodní infrastruktury
 - ▶ uzel EGI
 - ▶ budoucí začlenění do cloud infrastruktury

- ▶ Stávající uživatelé MetaCentra dostanou přístup na zdroje CERIT-SC automaticky
 - ▶ operační systém Debian 6.0, software v modulech, ...
 - ▶ gridové rozhraní v souladu s politikami MetaCentra
- ▶ Systém správy úloh
 - ▶ oddělený Torque server s vlastními politikami
 - ▶ homogennější prostředí, méně vlastností strojů, pouze jedna fronta, sdílený /home
 - ▶ později automatické předávání úloh
- ▶ Cloudové rozhraní (OpenNebula)
- ▶ Vlastní diskový prostor /home (250 TB)
- ▶ Spuštění provozu bude oznámeno novinkou MetaCentra
- ▶ Formální slavnostní zahájení provozu začátkem prosince 2011

- ▶ Principy:
 - ▶ spolupráce, nikoliv pouhá nabídka výkonu
 - ▶ analýza problémů uživatelů spojených s využitím infrastruktury
 - ▶ společné projekty na jejich překonání
- ▶ Vytváření společných týmů
 - ▶ ze strany CERIT-SC řešení infromatických problémů v aplikačních oblastech
 - ▶ ze strany partnera řešení jeho výzkumných potřeb a požadavků
 - ▶ včetně intenzivní spolupráce na úrovni studentů
 - ▶ student z FI/CERIT-SC vyslán pracovat v partnerském týmu
 - ▶ student z partnerského týmu „na stáži“ v CERIT-SC
 - ▶ společné publikace
 - ▶ primárně v aplikační oblasti
 - ▶ podklad pro více ICT orientované výsledky
- ▶ Společná výuka
 - ▶ výchova odborníků na „obou stranách“

- ▶ Analýza dat z hmotnostního spektrometru
 - ▶ Ústav experimentální biologie PŘF MU
 - ▶ analyzovaný vzorek – směs peptidů
 - ▶ vyhledání odpovídajících položek v databázích podle hmotnostního spektra
 - ▶ vyloučení falešných pozitiv, stanovení důvěryhodnosti apod.
- ▶ Analýza obrazů jednovrstvých buněčných kultur
 - ▶ Ústav fyzikální biologie JČU
 - ▶ mikroskopické snímky živých buněčných kultur, vývoj v čase
 - ▶ segmentace obrazů, modely chování kultur
 - ▶ urychlení algoritmů na GPU/hybridních architekturách
 - ▶ prototypová aplikace pro další rozvoj obecných metod automatické optimalizace

- ▶ Interaktivní modelování molekul
 - ▶ NCBR, HCILab FI MU, Loslab
 - ▶ haptické modelování interakce biomolekul
 - ▶ zkoumání tunelů v proteinech
 - ▶ vysoké nároky na výpočetní výkon, modely v reálném čase
 - ▶ modelování molekul v prostředí vzdálené spolupráce
- ▶ Dermatopatologické atlasy
 - ▶ Ústav patologie LF MU
 - ▶ dlouhodobá spolupráce na konkrétním tématu
 - ▶ rozsáhlá obrazová data, extrémně velké množství souborů
 - ▶ workflow zpracování pořizovaných obrazů do podoby interaktivního atlasu
 - ▶ zabezpečení přístupu v mezinárodním kontextu – federované rozhraní

- ▶ Zpracování signálů z lékařských přístrojů
 - ▶ Ústav přístrojové techniky AV ČR, LF MU
 - ▶ signály zejména EEG, větší množství signálů (128), vysoká vzorkovací frekvence (kHz)
 - ▶ hledání korelací – kvadratický nárůst velikosti
 - ▶ zaběhané metody přestávají stačit
 - ▶ optimalizace implementace, vývoj nových metod
- ▶ Zpracování 3D scanu stromu
 - ▶ Centrum výzkumu globální změny AV ČR
 - ▶ nasnímaná 3D množina bodů včetně odrazivosti
 - ▶ rekonstrukce struktury stromu (kmen, větve, listy, ...)
 - ▶ existuje poloautomatická metoda, pracujeme na zdokonalení

- ▶ Triviální: začít využívat zdroje přes rozhraní MetaCentra
 - ▶ budeme sledovat efektivitu využití této infrastruktury
 - ▶ kontaktovat partnery s větším, neobvyklým nebo neefektivním využitím
- ▶ Bezprostřední: obrátit se na nás s konkrétními plány/potřebami
 - ▶ uvedli jsme příklady započaté spolupráce pro inspiraci
- ▶ „Nabídka“:
 - ▶ intenzivní spolupráce při odstraňování překážek efektivního využití e-infrastruktury
 - ▶ ochota přizpůsobit infrastrukturu potřebám konkrétní skupiny
 - ▶ významně vyšší flexibilita než striktně produkční prostředí
 - ▶ široké IT zázemí, přenositelnost výsledků (přímo i prostřednictvím MetaCentra)
- ▶ „Cena“:
 - ▶ zajímavé problémy, středně až dlouhodobá spolupráce
 - ▶ konkrétní výsledky (publikace, software, ...)
- ▶ Výuka

- ▶ Pro nastartování transformace SCB do centra CERIT-SC
- ▶ Projekt „CERIT-SC“ 3. osy OP VaVpl
- ▶ Realizace od května 2011 do října 2013
- ▶ Rozpočet cca. 130 mil. Kč (pouze vybavení a lidé)
- ▶ Projekt má pomoci
 - ▶ pořídit nadkritické množství standardních IT technologií
 - ▶ urychlit dříve popsané procesy
- ▶ Vlastní Centrum CERIT-SC personálně přesahuje stávající projekt

- ▶ Vznik centra CERIT-SC
 - ▶ transformace z SCB
 - ▶ první fáze s pomocí projektu VaVpl
- ▶ Důraz na inteligentní využití zdrojů, nikoli jen hrubou sílu
 - ▶ interaktivita, okamžitá dostupnost, různorodost, ...
- ▶ Silná výzkumná složka
 - ▶ kapacita je potřebná pro ověřování nových myšlenek
- ▶ Komunikace a intenzivní rozvoj spolupráce s uživatelskými skupinami
- ▶ CERIT-SC
 - ▶ není izolovaný projekt
 - ▶ ale organická součást národní e-infrastruktury
 - ▶ a národního a mezinárodního výzkumného prostoru



EVROPSKÁ UNIE
EVROPSKÝ FOND PRO REGIONÁLNÍ ROZVOJ
INVESTICE DO VAŠÍ BUDOUCNOSTI



OP Výzkum a vývoj
pro inovace

Projekt CERIT Scientific Cloud, reg. č. CZ. 1.05/3.2.00/08.0144,
je podporován v rámci 3. prioritní osy Operačního programu
Výzkum a vývoj pro inovace.