



Enabling Grids for
E-science in Europe

www.eu-egee.org

Úvodní seminář EGEE, 26. října 2004

Použití EGEE v dalších vědeckých disciplínách

Jan Kmuníček
podpora uživatelů a aplikací



EGEE je projekt financován Evropskou Unií (smlouva IST-2003-508833)

Obsah

Typy výpočetních problémů

Aplikační struktura v rámci EGEE

Aktuální stav – aplikační portfolio

Další potenciální vědecké oblasti využívající EGEE Grid

Jak přizpůsobit vlastní aplikaci prostředí EGEE Gridu

Typy výpočetních problémů

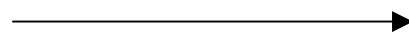
- Gridy jsou obecně vhodné pro **paralelní (současné) zpracování úloh**.
- **Paralelní výpočetní úlohy** takové úlohy, jejichž
 - zpracování na jediném CPU trvá neúnosně dlouho
 - celkové paměťové nároky jsou příliš velké
- Primární aplikační určení Gridů:
 - velké množství malých, navzájem nezávislých úloh
 - příklad: parametrické studie
 - zpracování rozsáhlých datových souborů „po částech“
 - příklad: digitální vyhlazení satelitních snímků

Typy výpočetních problémů

- **Granularita výpočetních problémů**
 - Výpočty s **malou granularitou**
 - každý podproblém vysoce závislý na výsledku dalších podproblémů
 - přenos do gridového prostředí je velmi složitý
 - nutno vymyslet algoritmy minimalizující množství a četnost dat, které si mezi sebou procesory vyměňují
 - příklad: výpočet počasí, který může být rozdělen do mnoha malých výpočtů počasí v malých objemech atmosféry
 - vhodné řešit na velkých paralelních superpočítačích nebo alespoň velmi **„pevně vázaných“ (tightly coupled)** clusterech s procesory propojenými extrémně rychlou sítí
- **„high-performance computing“**

Typy výpočetních problémů

- **Granularita výpočetních problémů**
- Výpočty s **velkou granularitou**
 - každý podproblém nezávislý na ostatních
 - přenos do gridového prostředí je relativně jednoduchý
 - příklad: Monte Carlo simulace, kde se obměňují parametry komplexního modelu reálného systému a výsledky se studují statistickými technikami – jedná se druh výpočetního experimentu
 - vhodné řešit pomocí sítě **"volně vázaných" (loosely coupled)** počítačů, jelikož prodlevy při získání výsledků z jednotlivých procesorů neovlivní práci ostatních procesorů



„high-throughput computing“

Typy výpočetních problémů

- **Reálné výpočetní problémy**
 - kombinace obou extrémních typů
- Gridové prostředí je **nejvhodnější** pro řešení takto **komplexních úloh**
 - poskytuje vhodné výpočetní prostředky pro podúlohy s nízkou granularitou
 - jejich výsledky lze kombinovat formou podúloh s vysokou granularitou
 - příklad: komplexní modelování podnebí Země
 - hledání informace, jak výpočty závisí na jednotlivých parametrech použitých v modelech – spuštění mnoha podobných výpočtů
 - každý z nich je paralelní výpočet s malou granularitou
 - velké množství v podstatě nezávislých výpočtů může být současně distribuováno na mnoho různých clusterů na Gridu

—————> průběh simulace v kratším čase

Aplikační struktura v rámci EGEE

- Formální členění aplikací v projektu EGEE
 - **Pilotní aplikace**
prvotní implementace a testování výkonu i funkčnosti vznikající gridové infrastruktury
 - LCG (LHC Computational Grid)
 - biomedicínské Gridy
 - **Interní aplikace**
národní projekty se zkušeností v oblasti gridových výpočtů
 - **Externí aplikace**
explicitní podpora vybraných uživatelských skupin tak, aby jejich aplikace byly přeneseny do gridového prostředí

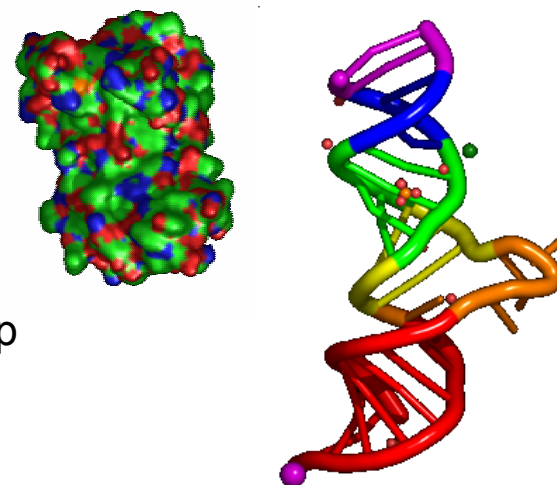
Aplikační struktura v rámci EGEE

- Typy aplikací v projektu EGEE
 - **HEP aplikace**
experimenty částicové fyziky pod záštitou CERNu
 - **Biomedicínské aplikace**
interní i externí projekty zabývající se zpracováním lékařských a medicínských dat
 - **Obecné aplikace**
všechny ostatní aplikační oblasti

Aktuální stav – aplikační portfolio

- **Biomedicínské pilotní aplikace**

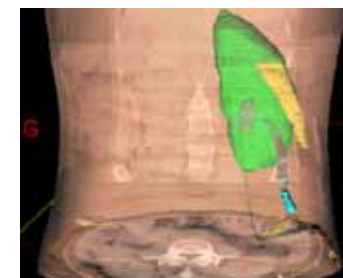
- Clinical Decision Support System (CDSS)
 - diagnostický nástroj umožňující extrahovat lékařsky relevantní informace s cílem zjednodušit klinickou praxi
- Grid Protein Sequence Analysis (GPS@)
 - molekulárně bioinformatický portál poskytující možnost analýzy proteinových sekvencí, jejich alignment, predikci sekundárních struktur a přístup do databáze 3D-struktur proteinů
- Geant4 Application for Tomographic Emission (GATE)
 - Monte-Carlo simulace pro radioterapii a nukleární medicínu umožňující modelování a plánování ošetřování karcinogenních nádorů před vlastním radioterapeutickým zákrokem



Aktuální stav – aplikační portfolio

- **Biomedicínské interní aplikace**

- 3D MRI Simulator (SiMRI3D)
 - paralelní simulátor obrazů magnetické rezonance
- xmipp_MLrefine
 - analýza 3D makromolekulárních struktur
- DICOM data analysis tool (gPTM3D)
 - nástroj pro interaktivní analýzu radiologických dat



- **Biomedicínské externí aplikace**

- MammoGrid
 - distribuovaná mamografická analýza pro diagnózu a terapii nádorů prsu

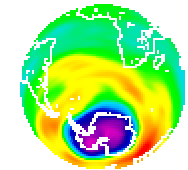
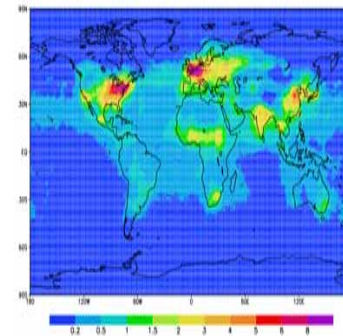
Aktuální stav – aplikační portfolio

- **Výběr nových aplikací**
 - zaslání žádosti o zapojení do výběrového řízení obecných aplikací
 - výběr a schválení výběrovou komisí EGAAP (EGEE Generic Application Advisory Panel)
 - získání oficiální podpory vybrané uživatelské komunity po omezenou dobu, během níž dojde k portování vybrané aplikace do prostředí EGEE Gridu
 - výsledkem prvního EGAAP zasedání 4 nové aplikační oblasti

Aktuální stav – aplikační portfolio

- **Vědecký výzkum Země**

- zpracování a validace měření globálních změn atmosférického ozonu pomocí satelitu GOME

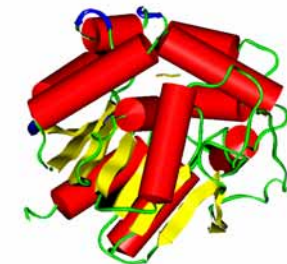
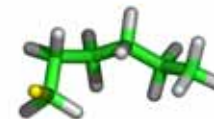


- **Geofyzika**

- 3D analýza zemského podloží s cílem hledání zdrojů ropy a zemního plynu

- **Výpočetní chemie**

- atomární a molekulární simulace použitím molekulárně-mechanických a kvantově-chemických metod



- **Fyzika astro-částic**

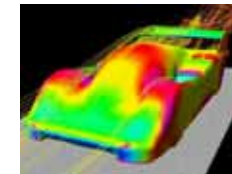
- Monte Carlo simulace dat z teleskopu MAGIC



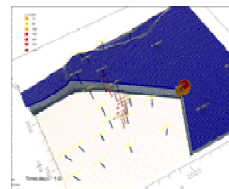
Další potenciální vědecké oblasti využívající EGEE Grid

- Spektrum aplikačních oblastí vhodných pro EGEE Grid

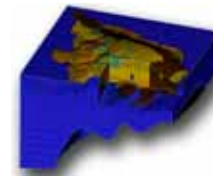
- Technické a materiálové simulace



- Hydrologie



- Seismologie



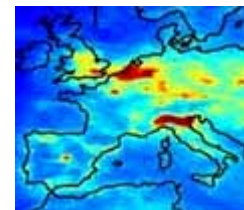
- Gridové vyhledávací nástroje



- Simulátory burzovních trhů

Další potenciální vědecké oblasti využívající EGEE Grid

- Spektrum aplikačních oblastí vhodných pro EGEE Grid
 - Digitální zpracování videa
 - Distribuované datové sklady
 - Real-time rendering & raytracing pro architekty, počítačové grafiky a umělce
 - Evropský úřad pro vesmír



Jak přizpůsobit vlastní aplikaci prostředí EGEE Gridu

- Zvolit typ aplikací, které jsou pro Vás klíčové a jež byste rádi používali v prostředí EGEE Gridu
- Zvážit vynaložené úsilí na portaci a výhody plynoucí z možnosti použití vybraných aplikací na EGEE Gridu
- Diskutovat detailní kroky nezbytné transformace Vašich aplikací



(kmunicek@ics.muni.cz)

- Sledovat up-to-date informace na gridovém portále EGEE

!!! egee.cesnet.cz !!!