



Enabling Grids for
E-science in Europe

www.eu-egee.org

Úvodní seminář EGEE, 26. října 2004

Celoevropský Grid EGEE a Další Projekty EU

Luděk Matyska



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

Přehled Prezentace



- Grid
- Projekt EGEE
 - Cíl, struktura a přístupy
- Další projekty EU
- Co můžeme očekávat

- Primární motivace
 - Spojování nedostatečně výkonných počítačů
 - Konec 80. Let minulého století
- Omezené možnosti
 - Kapacita sítě
 - Počet připojených (super)počítačů
- Výhody
 - Definovaná skupina zdrojů i uživatelů
- Parallel virtual machine (PVM)

- Exponenciální nárůst počtu připojených počítačů
- Web
 - První krok ke *sdílení*
 - Logické propojení počítačů
- Stále přetrvává zájem o maximální výkon
- Message passing interface (MPI)
- I-way (1995)

Změna Pohledu

- Od high performance k high throughput
 - Condor
- Využití možností daných velkým počtem počítačů propojených stále rychlejší sítí
 - Sít' (a počítače) jako kvalitativně *nový* počítač
- Roste zájem o *koordinaci* zdrojů

- První generace – *výpočetní grid*
 - Stále důraz na výpočty
 - Grid chápán jako množina výpočetních uzlů propojených počítačovou sítí společně řešící konkrétní problém či množinu problémů
- 2. Generace – *třívrstevný grid*
 - Kromě výpočetního i *informační a znalostní*
- Gridy postupně chápány jako nástroj rozvoje vědy (nikoliv pouze informatiky)

- Důraz na *spolupráci*
- *Prostředí pro spolupráci*
 - Gridy propojují pracoviště a lidi, uzly již zdaleka nejsou jen počítače
- Parafráze britského programu e-science:
„Nemůžete dělat kvalitní vědu, pokud ji děláte sami“

- Definice, kdo s kým bude spolupracovat
 - Potřeba flexibilních a jednoduchých nástrojů
 - Podpora bezpečnosti
 - Autentizace – kdo jsou uživatelé/zdroje
 - Autorizace – kdo smí co použít
 - Podpora spolehlivosti
 - Ve velké organizaci vždy něco nefunguje
 - Peer to peer sítě
- Grid umožňuje tvorbu a využití virtuálních organizací

- Zajištění interoperability
- Inspirace z webu
 - Webové služby
- Rozšíření webových služeb
 - Podpora *zpracování* (procesy)
 - Robustnost (soft stav)
 - Virtuální organizace

- Výzkumná výzva pro informatiku
 - Jak grid spravovat
 - Včetně sledování „stavu“
 - Jak grid programovat
 - Škálovatelné distribuované algoritmy
 - Jak delegovat
 - Včetně tvorby virtuálních organizací a jejich podpory
- Praktická výzva
 - Produkční gridy

- Dnešní název:
Enabling grids for e-science
- Hlavní cíl:
Vybudovat produkční gridovou infrastrukturu na celoevropské úrovni
- Komplementární cíle:
 - Zprovoznit aplikace
 - Stabilizovat gridové programové vybavení (middleware)
- Zahájen 1. Dubna 2004

Projekt EGEE – Základní Údaje



- Hlavní řešitelská organizace: CERN
 - Hlavní řešitel: fabrizio gagliardi
- 70 partnerů explicitně zahrnutých do projektu:
 - Pokryty všechny státy EU
 - USA, rusko
 - Zájem o další rozšiřování (asie, jižní amerika)
- Pouze dvouletý projekt, rozpočet 35+ meuro
- ČR reprezentuje CESNET, z.S.P.O.

Projekt EGEE – Organizace



- Rozdělení do 7 „federací“
 - Zpravidla skupina „sousedících“ států
 - Ovšem holandsko+severské země
 - CERN samostatná federace
 - ČR/CESNET součástí středoevropské federace (CE federation)
- Hierarchické řízení
 - Resp. Alespoň snaha o ně

Projekt EGEE – Aktivita



- Rozděleny do tří skupin:
 - Joint research activities, JRA
 - Vývoj middleware, bezpečnost, sítě
 - Rovněž postupy k zajištění kvality
 - Specific service activities, SA
 - Garantují vlastní cíl projektu: zajištění gridu jako služby
 - Networking activities, NA
 - Řízení, školení, diseminace, podpora aplikací, mezinárodní spolupráce, ...

Projekt EGEE – Souvislosti



- Navazuje na projekt datagrid
 - Indukován potřebami částicové fyziky
- Primárně high throughput model
- Důraz na manipulaci s velkými objemy dat
- Intenzivní mezinárodní spolupráce prostřednictvím hlavní aplikační skupiny – fyziků zabývajících se fyzikou částic

Projekt EGEE – Infrastruktura

- Postavit produkční evropský grid
 - Měl by sloužit jako inspirace dalším
 - Stabilita je primární cíl
- Není omezen pouze pro částicovou fyziku
 - Ale ta je primární tažnou silou
- Hledání dalších aplikací/silných mezinárodních aplikačních skupin
 - Příležitost pro ty, kteří mají rozsáhlé mezinárodní kontakty a potřebují počítat a/nebo pracovat s velkými objemy dat

Projekt EGEE – Provoz



- Omc, cern
 - Centrální dohled
- CIC, zatím 4
 - Zajištění služeb (např. VO management)
 - Žádná ve střední evropě
- Roc
 - Regionální operační centrum, u nás HQ v polsku (krakow)
- Rc
 - Vlastní výpočetní zdroje (není financováno z EGEE)

Projekt EGEE – ROC

- Distribuovaná struktura ve střední Evropě
- Každý partner se nějak podílí
- Odpovědnost
 - Instalace sw
 - Dohled na RC
 - Včetně monitoring infrastruktury
 - Bezpečnost
 - Lokalizace, ...
- V současné době se dokončuje

- Centrální server pro správu požadavků
 - Karlsruhe
 - Jediné rozhraní pro všechny uživatele nezávisle na jejich poloze
- Koordinace s lokálními systémy správy požadavků (např. RT systém)
 - Součást aktivity ROC

- EGEE hlavní infrastrukturní gridový projekt
 - Má uloženo spolupracovat s ostatními
- Podpora EU jde i do rozvoje komplexních systémů, využívajících grid jako nástroj:
Grids for complex problem solving
- Expertní zpráva:
Next generation grid
 - Podpora i infromatického výzkumu (přesněji jeho celoevropské integrace)

Gridy V 5. Rámčovém Programu

- Infrastruktura
 - Datagrid
 - Damien
 - Datatag
 - Eurogrid
 - Mamogrid
 - Lege-wg
 - Biogrid
 - Cog
 - Gemss

Gridy V 5. Rámčovém Programu

- Podpora řešení složitých problémů
 - Asp-bp
 - Crossgrid
 - Flowgrid
 - Grace
 - Grasp
 - Gria
 - Gridlab
 - Gridstart
 - Grip
 - Openmolgrid
 - Selene
 - Websi

Gridy V 6. Rámčovém Programu



- 1. Září spuštěno 12 nových gridových projektů v rámci podpory gridů pro řešení složitých problémů
 - Celkem 52 meuro

Seznam Nových Projektů

- Coregrid – síť excellence
- Akogrimo
- Datamininggrid
- Gridcoord
- Hpc4u
- Intelligrid
- K-WF grid
- Nextgrid
- Ontogrid
- Provenance
- Simdat
- Unigrids

Situace V Členských Zemích

- Zpravidla národní iniciativa/program
 - ČR v podstatě výjimka – nemá žádný oficiální program
- Nejrozsáhlejší UK: e-science
 - Depositář middleware
- V současné době se rozbíhá d-grid (SRN)
- Maďarsko
 - Národní grid z 20 instalací na vysokých školách
 - 100 CPU každé místo
 - Ve dne pro výuku, v noci propojeny do gridu
 - Ovládány systémem condor

- Dlouhá historie
 - Metacentrum od roku 1996
 - Virtuální SP2
- Primárně infrastrukturní orientace
- Zapojení do řady EU projektů
- Hlavní výzkumný zájem

Grid infrastructure monitoring and management

Gridy Mimo Evropu

- Gridy jako pojem pochází z USA
- Posun od superpočítačových projektů k prostředí pro spolupráci
 - Významná role fyziky částic
 - Netriviální komerční využití
 - Ke gridům se hlásí řada velkých firem (sun, IBM)
- Mimořádný rozvoj v souvislosti s cyberinfrastructure
 - Zajištění/robustnost a bezpečnost

- Pokračující posun k podpoře spolupráce
 - Využití gridů ve výuce
- Stálé hledání *killer* aplikace
 - Může se jí stát zpracování videa?
- V infrastrukturní oblasti důraz na robustnost a kvalitu
- Ve výzkumu propojování s peer to peer sítěmi

- Gridové prostředí bez podpory státu
- EGEE jako jedna z variant gridové infrastruktury
- Zvyšování robustnosti
- Propojení se sousedními zeměmi
 - Vo ce
 - Sdílení zdrojů
- Zájem na podpoře konkrétních aplikací

- Gridy jsou nástroje pro zajištění špičkové vědy
- Vytváří zázemí *výzkumného (a výukového) prostoru*
- Nezbytné zejména pro mezinárodní spolupráci
 - Návyky získané v ČR se zúročí při zapojení do mezinárodních aktivit
- Zvyšují potenciál společného výzkumu

Děkuji za pozornost