

Gridy v České republice

Luděk Matyska
Masarykova univerzita v Brně
CESNET, z.s.p.o.
Ludek.Matyska@muni.cz

Co jsou Gridy

- **Gridy** jsou rozsáhlé distribuované systémy, tvořené výpočetními, datovými a informačními zdroji, propojenými vysokorychlostní sítí
- **Základní koncept:** sdílení zdrojů

Gridy a sítě

- Požadavky na propojovací síť
 - Vysoká propustnost
 - End to end parametry
- Akademická síť jako prototyp
- Dedikované optické sítě
 - Splňují požadavek kapacity a end2end
 - Neškálují jako IP sítě
 - V podstatě se jedná o spojované sítě

METACentrum

- Distribuované výpočetní prostředí v ČR
 - Založeno v roce 1996
 - Ještě jako Metapočítač
 - Od roku 1998 součást výzkumných záměrů sdružení CESNET
- Propojuje zdroje
 - MU Brno
 - UK Praha
 - ZČU Plzeň
 - CESNET

METACentrum – vlastnosti

- Poskytuje výpočetní výkon
 - Výkonné SMP počítače SGI
 - Clustery IA-32 (přes 300 CPU)
 - Experimentální
 - IBM Power4+, AMD Opteron, IA-64 Itanium 2
- Poskytuje datové sklady
 - Přímě připojené k výpočetním zdrojům
 - Distribuované prostředí DiDaS

METACentrum – přístup

- Zdroje METACentra dostupné výzkumné veřejnosti ČR
 - <http://meta.cesnet.cz>
 - Best effort (v závislosti na zatížení)
- Cca 100 aktivních uživatelů
 - Registrovaných významně více

METACentrum – služby (1/2)

- METACentrum poskytuje
 - Spravované zdroje
 - I formou outsourcingu, např. pro NCBR MU
 - Datovou kapacitu
 - Řádově 40 TB on-line
 - Úložnou kapacitu
 - Do konce roku bude k dispozici 2*200 TB na dvou páskových robotech (Plzeň a Brno)
 - Žádný firewall, vše přímo přístupné z Internetu

METACentrum – služby (2/2)

- Zdroje propojeny dávkovým systémem **PBSPro**
 - Je možná i interaktivní práce
 - Fronty
 - Krátké/rychloobrátkové
 - Dlouhé
 - Extrémně dlouhé
 - Prioritní

METACentrum – programové vybavení

- Operační systém(y)
- Překladače a vývojová prostředí
 - Open i komerční překladače
 - pgi, Intel
 - Open i komerční ladicí prostředí
 - Vampir, Vampirtrace, ...

METACentrum – aplikace

- Aplikační programové vybavení
 - Matlab a většina toolboxů
 - Maple (pro některé univerzity)
 - Gaussian 03
 - Fluent
 - ... (viz <http://meta.cesnet.cz>)
- Nezbytná spolupráce s uživatelskou komunitou

METACentrum – datové kapacity

- Lokální disky u výpočetních systémů
 - Lokální domovské adresáře
 - Scratch prostor
- Přímou připojená disková pole
 - Domovské adresáře
- Disková kapacita přístupná přes AFS
 - Globální jmenný prostor
- Distribuovaná disková kapacita DiDaS

DiDaS

- Distribuované Datové Sklady
- Internet Backplane Protocol
 - Best effort služba pro data
- Sedm míst připojených k páteři
 - Celková kapacita 20 TB
 - Distribuováno přes celou republiku
 - Vhodné pro šíření dat

Uživatelé

- Primárně Computational Science
 - Výpočetní chemie
 - Fyzika materiálů
- Fluidní mechanika
 - Spíše omezené využití
- Obecné výpočty

Formy spolupráce (1/2)

- „Jednoduché“ nabízení výpočetní a/nebo úložné kapacity
 - Omezená kapacita
- Outsourcing
 - Správa clusterů
 - Integrace do METACentra

Formy spolupráce (2/2)

- **Expertní spolupráce**
 - Konzultace, případně společné zapojení do projektů (národních i mezinárodních)
 - Např. projekt MediGRID (IS AV ČR)
 - Zprostředkování mezinárodních kontaktů
 - Napojení do mezinárodních Gridových aktivit

Mezinárodní spolupráce

- Od roku 2000 zapojeni do EU infrastrukturních projektů
 - DataGRID
 - EGEE
- Zapojeni i do výzkumných projektů EU
 - GridLab
 - CoreGRID

EGEE (1/4)

- Budování pan-evropské Gridové infrastruktury
 - Primární uživatelé částicovní fyzici (HEP)
 - Koordinátorem CERN
 - Postavena na IA-32 a posléze IA-64 clusterech
 - Primárně dávkový systém
 - Hlavní problém – manipulace s daty

EGEE (2/4)

- Celkem 70 partnerů (včetně USA a Ruska) s příspěvkem 35 MEuro na 2 roky
 - CESNET zastupuje ČR
- Federativní (regionální) organizace
 - ČR součástí Středoevropské federace
 - Project Management Board tvořen zástupci federací
 - Středoevropskou teď zastupuje zástupce CESNETu (autor přednášky)

EGEE (3/4)

- Nefinancuje vlastní hw vybavení
 - To chápáno jako příspěvek partnerů/států
- Primárně mezinárodní koordinace
 - Regional Operating Center
 - U nás HQ v Krakově, CESNET působí jako alternativní HQ
- Vývoj (re-engineering) middleware
 - Middleware – programové vybavení, propojující a zpřístupňující jednotlivé uzly Gridu
 - CESNET jediným partnerem mimo UK, Itálii a CERN

EGEE (4/4)

- Uživatelé
 - Primárně komunita HEP
 - Bioinformatika
 - Nedaří se podchytit menší komunity
- Virtuální organizace
 - Aplikačně orientovány
 - Přes 30, avšak většina neaktivní
 - VO CE (Central Europe)
 - Nabízena uživatelům „bez“ VO

EGEE II

- EGEE končí 31. 3. 2006
- Podán návrh na pokračování
 - Na další 2 roky, do 7. rámcového programu
- Větší důraz na další aplikační skupiny
 - **Sdílení zdrojů**
- CESNET opět partnerem
 - Tlak na „jedna země jeden partner“

Výzkum

- Primárně sledování Gridu
 - Monitoring
 - Logging and Bookkeeping (EGEE)
 - Job Provenance (dlouhodobé úložiště informací o proběhlých úlohách)
- Bezpečnost v distribuovaném prostředí
- Plánování na Gridech
 - Komplexní plánovací algoritmy
- Kontakty na mezinárodní Gridovou komunitu

Příklad mezinárodního experimentu

- Interactive Remote Visualization across the Louisiana Optical Network and the National Lambda Rail
 - Tři místa: Brno, Baton Rouge, San Diego
 - Distribuovaná vizualizace kolize černých děr
 - Vzdálené ovládání vizualizace ze tří míst
 - Paralelní videokonference v High Definition kvalitě



7. 11. 2005

Seminář CESNET

24(#)

Shrnutí

- METACentrum (součást VZ sdružení CESNET) buduje národní Gridovou infrastrukturu
 - Součinnost s budováním sítě CESNET2
- Zapojeno do rozsáhlých mezinárodních aktivit (včetně spoluúčasti na rozhodování)
- Vlastní výzkum v oblasti Gridů
- Zájem o spolupráci
 - Zapojení do infrastruktury
 - Propojení dalších zdrojů

Děkuji za pozornost