

Úvod do Gridu (Introduction to Grid)

Miroslav Dobrucký
Ústav informatiky
Slovenská akadémia vied
Bratislava

- **Čo je to gridové počítanie?**
- **Charakteristiky gridu**
- **Aplikácie**
- **Gridové projekty**

Vedci a výskumníci vykonávajú svoje aktivity v kooperácii v rámci celého sveta, komunikujú medzi sebou a zdieľajú spolu údaje.

Grid: zosieťované centrá na spracovanie údajov a "middleware" ako "lepidlo" na spojenie zdrojov.



Virtuálna organizácia

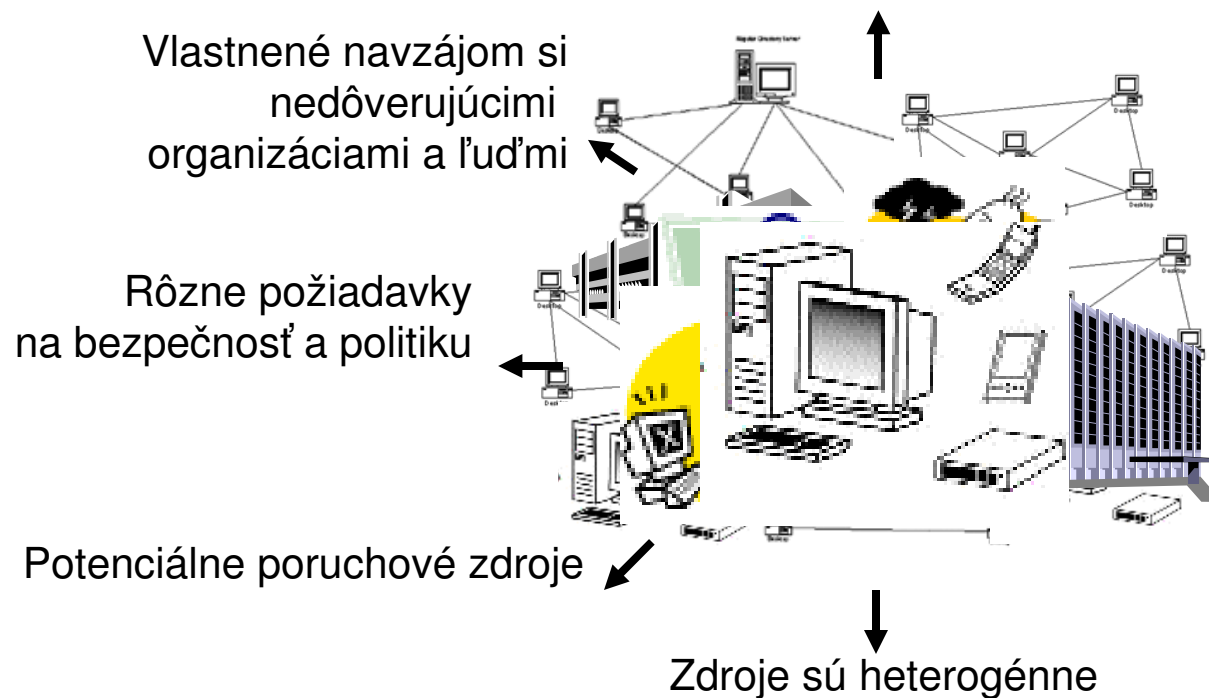
Vedecké nástroje a experimenty poskytujú obrovské množstvo údajov.

- **Súčasné distribuované aplikácie:**
 - Smerujú k *špecializovaným* systémom
 - Sú zamerané na *jeden* účel alebo skupinu používateľov

- **Gridy idú ďalej a berú do úvahy:**
 - Rôzne druhy zdrojov
 - Nie je vždy rovnaký hardvér, údaje a aplikácie
 - Rôzne druhy *interakcií*
 - Skupiny používateľov alebo aplikácie chcú interagovať s gridom rôznymi spôsobmi
 - *Dynamická* povaha
 - Zdroje a používatelia sú pridávaní/odoberaní často

- Čo je to gridové počítanie?
- **Charakteristiky gridu**
- **Aplikácie**
- **Gridové projekty**

Velké množství zdrojov



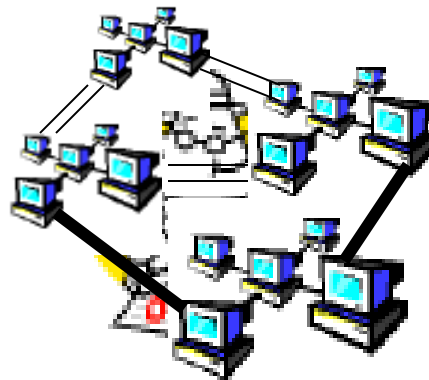
Štandardy

Veľké množstvo zdrojov

Vlastnené navzájom si nedôverujúcimi organizáciami a ľuďmi

Rôzne požiadavky na bezpečnosť a politiku

Potenciálne poruchové zdroje



Prepojené heterogénnymi, viac-úrovňovými sieťami

Rôzne politiky manažmentu zdrojov

Geograficky oddelené

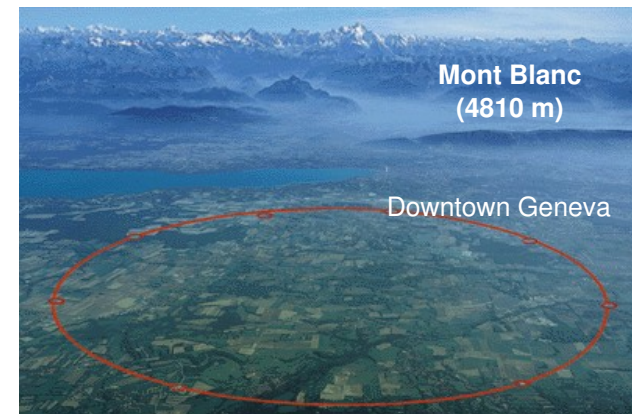
Zdroje sú heterogénne

- Čo je to gridové počítanie?
- Charakteristiky gridu
- **Aplikácie**
- **Gridové projekty**

- **Medické/zdravotnícke** (*snímkovanie, diagnostika a liečenie*)
- **Bioinformatické** (*skúmanie ľudského genómu aby sme porozumeli genetickým chorobám*)
- **Nanotechnológie** (*návrh nových materiálov na molekulárnej úrovni*)
- **Strojárstvo** (*optimalizácia návrhu, simulácie, analýza porúch a vzdialená manipulácia a riadenie*)
- **Prírodné zdroje a životné prostredie** (*predpoveď počasia, pozorovanie zeme, modelovanie a predikcia zložitých systémov*)



- **The Large Hadron Collider (LHC)**
 - doteraz postavený **najvýkonnejší nástroj** na vyšetovanie elementárnych častíc
- **Údajová výzva:**
 - **10 Petabajtov/rok** údajov !!!
 - 20 miliónov CD-čiek každým rokom!
- **Simulácia, rekonštrukcia, analýza:**
 - Údaje z LHC vyžadujú výpočtový výkon porovnateľný s **~100,000** najrýchlejšími **súčasnými PC** procesormi!



- Čo je to gridové počítanie?
- Charakteristiky gridu
- Aplikácie
- **Gridové projekty**

Mnohé snahy vyvíjať Grid — na celom svete

- NASA Information Power Grid
- DOE Science Grid
- NSF National Virtual Observatory
- NSF GriPhyN
- DOE Particle Physics Data Grid
- NSF TeraGrid
- DOE ASCI Grid
- DOE Earth Systems Grid
- DARPA CoABS Grid
- NEESGrid
- DOH BIRN
- NSF iVDGL

- DataGrid (CERN, ...)
- EuroGrid (Unicore)
- DataTag (CERN,...)
- Astrophysical Virtual Observatory
- GRIP (Globus/Unicore)
- GRIA (Industrial applications)
- GridLab (Cactus Toolkit)
- CrossGrid (Infrastructure Components)
- EGSO (Solar Physics)

- UK – OGSA-DAI, RealityGrid, GeoDise, Comb-e-Chem, DiscoveryNet, DAME, AstroGrid, GridPP, MyGrid, GOLD, eDiamond, Integrative Biology, ...
- Netherlands – VLAM, PolderGrid
- Germany – UNICORE, Grid proposal
- France – Grid funding approved
- Italy – INFN Grid
- Eire – Grid proposals
- Switzerland - Network/Grid proposal
- Hungary – DemoGrid, Grid proposal
- Norway, Sweden - NorduGrid

European DataGrid (EDG)

www.edg.org



LHC Computing GRID (LCG)

cern.ch/lcg



CrossGRID

www.crossgrid.org



DataTAG

www.datatag.org



GridLab

www.gridlab.org



EUROGRID

www.eurogrid.org

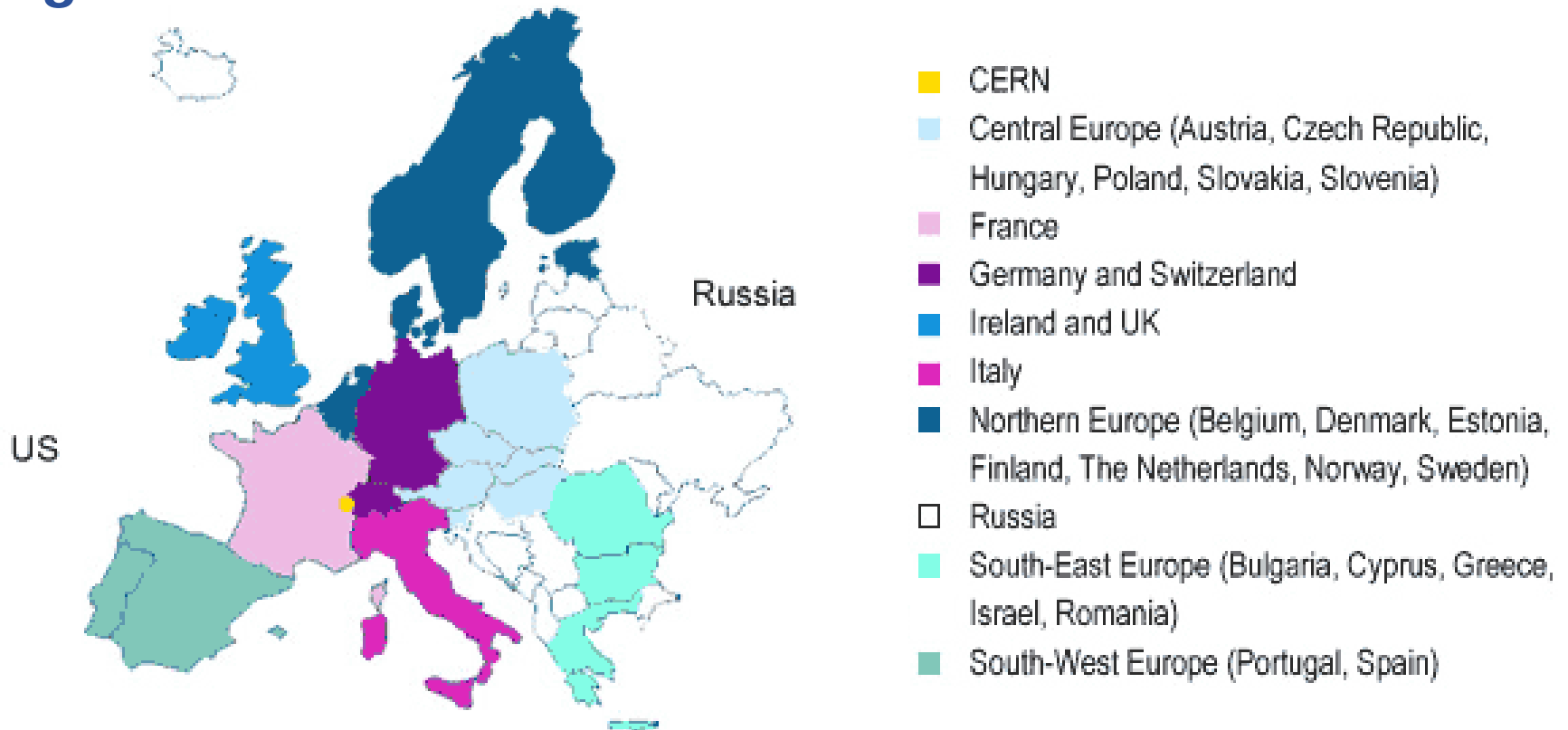


Európske národné projekty:

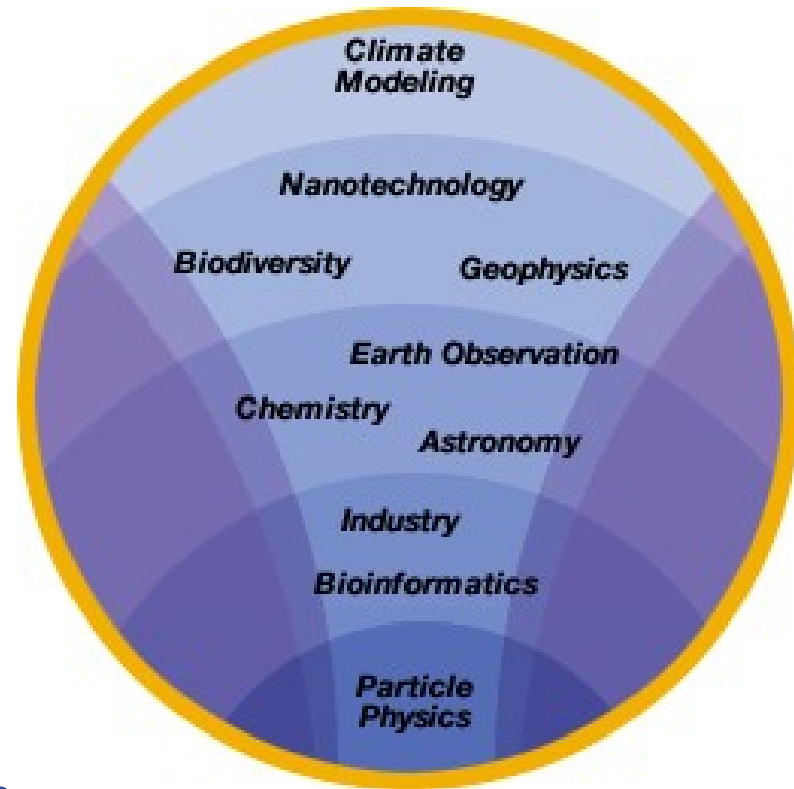
- INFN
- UK e-Science Programme
- NorduGrid



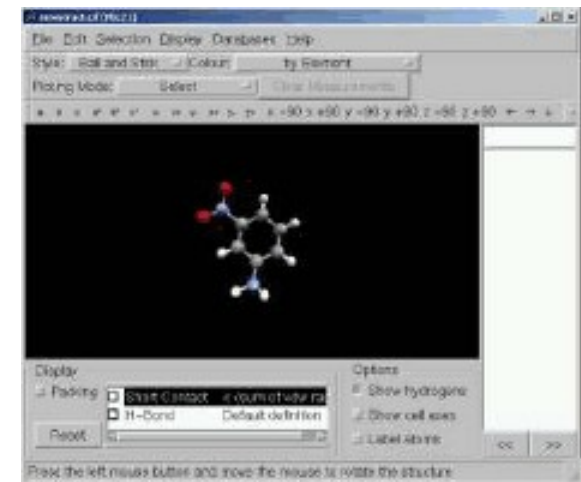
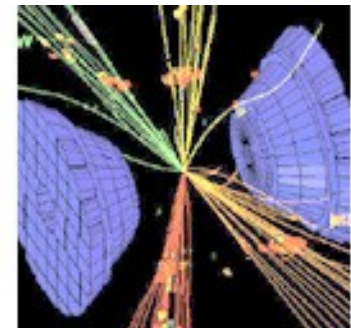
- Spoločné využívanie národných zdrojov za účelom dosiahnutia vyššieho úžitku pre Európu
- 70/91 inštitúcií z 27 krajín, združených v regionálnych gridoch



- Rozsah projektu EGEE: plná podpora pre akademické aplikácie (je taktiež otvorený pre priemyselný a socio-ekonomický svet)
- Hlavné kritérium úspechu EGEE projektu: koľko spokojných používateľov z koľkých rôznych oblastí ?
- Cieľ: 5000 používateľov (3000 po 2 rokoch) aspoň z 5-tich disciplín
- Dve vybrané pilotné aplikácie budú určovať implementáciu a zhodnotia výkonnosť a funkčnosť vyvíjanej infraštruktúry: Fyzika & Bio-informatika



- **Fyzika vysokých energií (High Energy Physics)** s výpočtovým gridom LHC (www.cern.ch/lcg) využíva gridovú infraštruktúru na ukladanie a analýzu petabajtov (10^{15} bajtov) reálnych a simulovaných údajov. LCG projekt je hlavným zdrojom požiadaviek a tvrdých termínov, ktoré sa nedajú konvenčne riešiť.
- **V biomedicíne** mnohé komunity čelia rovnako skľučujúcim výzvam zvládnuť záplavu bioinformatických a zdravotníckych údajov. Potrebujú mať prístup k veľkým a distribuovaným nehomogénnym údajom a majú dôležitú požiadavku na okamžité výkonné počítanie.



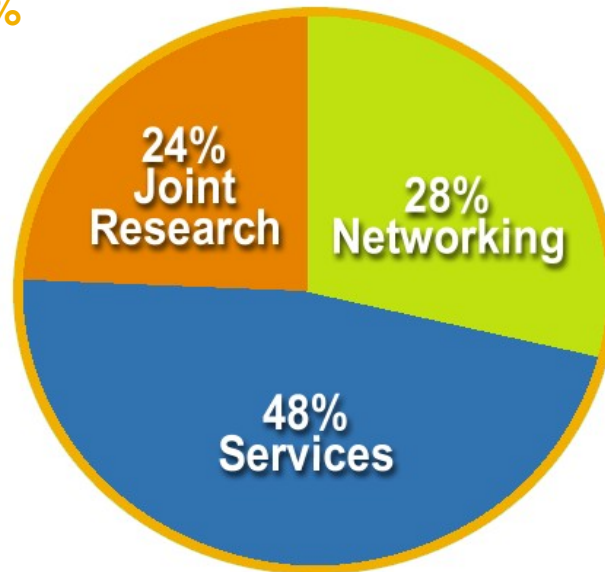
24% Združený výskum

JRA1: Príprava middleware a integrácia – 17%

JRA2: Zabezpečenie kvality – 1.5%

JRA3: Bezpečnosť – 3%

JRA4: Vývoj sieťových služieb - 2.5%



48% Služby

SA1: Prevádzka gridu, podpora a správa

SA2: Zabezpečovanie sieťovej infraštruktúry

28% Sieťové aktivity

NA1: Manažment

NA2: Disseminácia a propagácia

NA3: Trénovanie používateľov a vzdelávanie

NA4: Identifikácia aplikácií a ich podpora

NA5: Politika a medzinárodná spolupráca

Dôraz v EGEE je na prevádzke produkčného gridu a na podpore koncových používateľov

Ďakujem za pozornosť.

egee.ui{AT}sav.sk

Miroslav Dobrucký

Ústav informatiky

Slovenská akadémia vied

Bratislava