



Enabling Grids for E-science

Od klastrov ku Gridom

Stavebné bloky EGEE infraštruktúry

Marek Ciglan, Ján Astaloš
Ústav informatiky,
Slovenská akadémia vied



www.eu-egee.org





- **Motivácia a úvod**
- **Dávkové systémy**
 - Condor
- **Gridová štruktúra**
- **Životný cyklus gridovej úlohy**
- **Možné konfigurácie výpočtových klastrov pre EGEE infraštruktúru**
- **Middleware**
- **EGEE Infraštruktúra**
- **Pripojenie nového výpočtového centra do EGEE infraštruktúry**



- **Nové oblasti (medicína, životné prostredie, materiálový výskum, kryptografia, ...)**
- **Stále dokonalejšia reprezentácia reality**
- **Rastúce nároky na výpočtovú kapacitu**

- **Rôzne požiadavky na výpočtový systém**
 - paralelné simulácie (High Performance Computing)
 - parametrické štúdie (High Throughput Computing)
 - dátovo intenzívne simulácie
 - interaktívne



- **Faktory vplývajúce na vývoj:**
 - rast výkonu počítačov triedy PC
 - zvyšovanie priepustnosti sieťových prepojení
 - zvyšovanie kapacity pamätí
 - znižovanie cien
- **Trend: PC klaster - “poor man's supercomputer”**



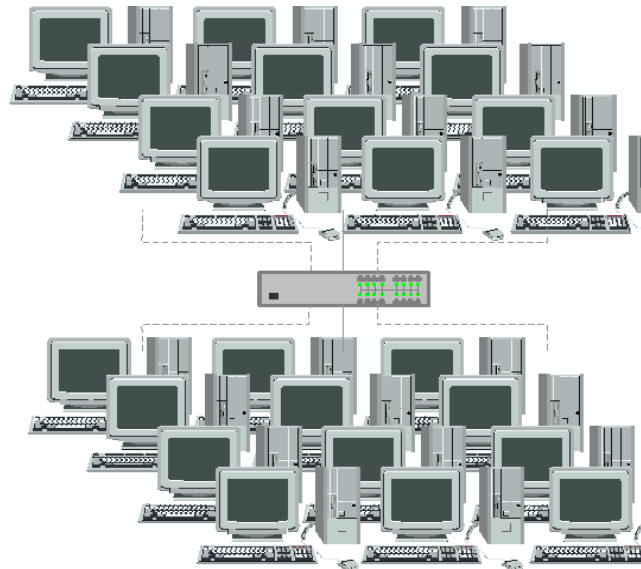
- **Faktory vplývajúce na vývoj:**
 - stále viac nevyužitého výkonu
 - globalizácia

- **Trendy:**
 - dobrovoľné poskytovanie voľnej výpočtovej kapacity
 - združovanie geograficky distribuovaných prostriedkov

Vysoko výkonné HPC klastre



Sieť pracovných Staníc (Condor)



Distribuované výpočtové Gridy (EGEE)



Prepojenie výpočtových centier



- **Torque+Maui** (<http://www.supercluster.org>) - Torque je open source verzia populárneho systému PBS (Portable Batch System). Maui scheduler je sofistikovaný plánovač ktorým je možné nahradiť štandardný plánovač v Torque. Tento systém je vhodný najmä pre klastre v nepretržitej prevádzke.
- **PBS Pro** (<http://www.pbspro.com/>) - Komerčná verzia systému PBS s rozšírenou funkcionalitou (v oblasti plánovania, poskytuje lepšiu škálovateľnosť, má podporu pre Gridy).
- **LSF - Load Sharing Facility** (<http://www.platform.com>) - komerčný systém ktorý používajú firmy ako AMD, Peugeot Citroen a ďalšie. Je rozšírený v oblasti vysokovýkonných paralelných výpočtových systémov (HPC).
- **Condor** (<http://www.cs.wisc.edu/condor/>) - je vyvíjaný takmer 20 rokov. Pôvodne bol určený najmä na využitie voľnej výpočtovej kapacity pracovných staníc (je známy ako „hunter of idle workstations“). Je voľne dostupný (aj keď nie je open source)



- **špecializovaný dávkový systém na spracovanie výpočtovo náročných úloh**
- **dokáže využiť voľnú výpočtovú kapacitu pracovných staníc**
- **jednoduché používanie a administrácia**
- **heterogénne prostriedky**
 - Linux, IRIX, Solaris, HP-UX, OSF/1, Windows NT
- **typy úloh**
 - parametrické simulácie (rádovo stovky úloh)
 - závislosti medzi úlohami (DAG)
 - paralelné úlohy v PVM
- **Globus+Condor -> správa prostriedkov Gridu**



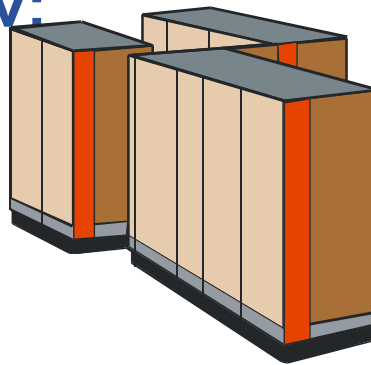
- **Checkpointing (periodické odkladanie stavu úloh)**
 - migrácia úloh pri odpojení pracovnej stanice
 - spoľahlivé ukončenie dlhotrvajúcich úloh
- **Vzdialené systémové volania**
 - úlohy pracujú s dátami na centrálnom serveri
- **ClassAds (classified advertisements)**
 - správa prostriedkov na princípe inzercie
 - škálovateľná na tisíce úloh a výpočtových prostriedkov
- **Systém priorít**
 - pridelený čas závisí od priority používateľa a času ktorý spotreboval



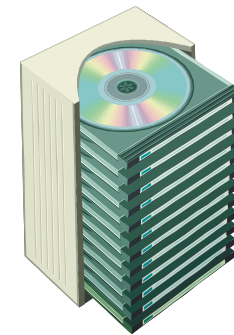
- **Stavebné prvky Gridu**
 - Výpočtové prostriedky, úložné systémy, siete
- **Typy gridových uzlov:**



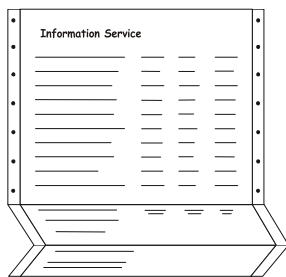
User Interface (UI)



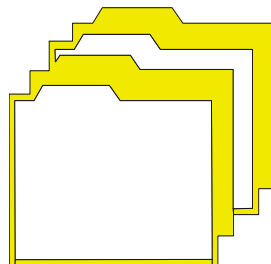
Computing Element (CE)



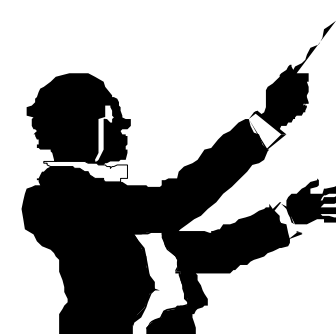
Storage Element (SE)



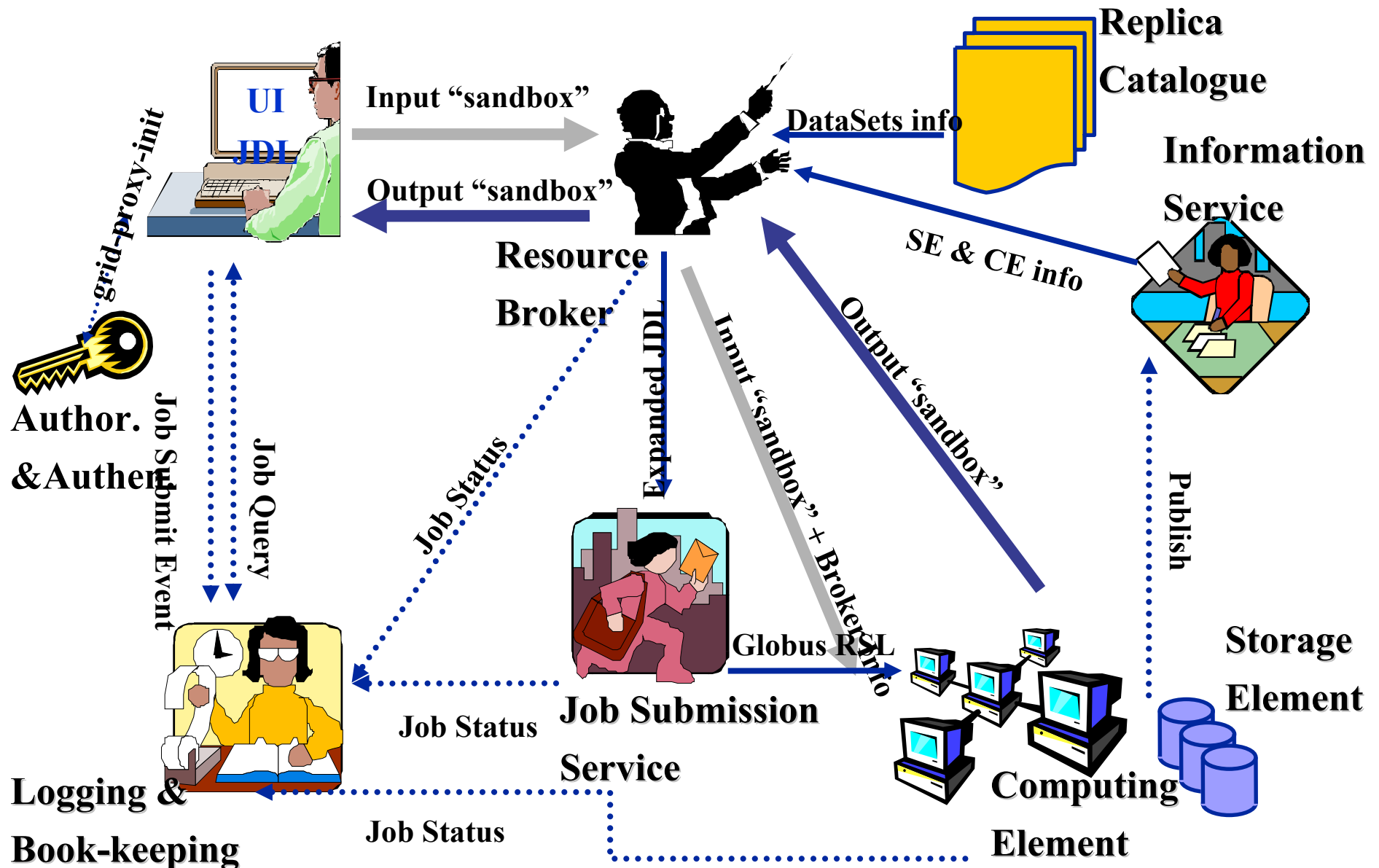
Information Service (IS)



Replica Catalog (RC, RLS)



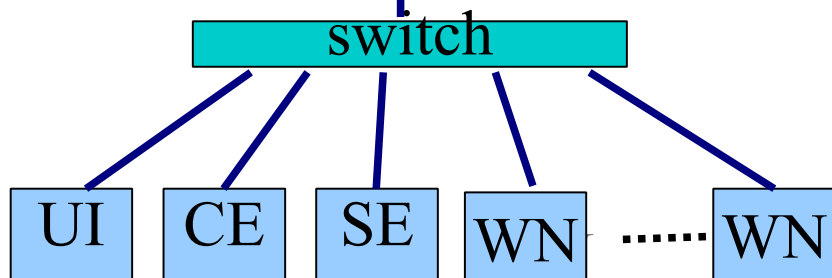
Resource Broker (RB)



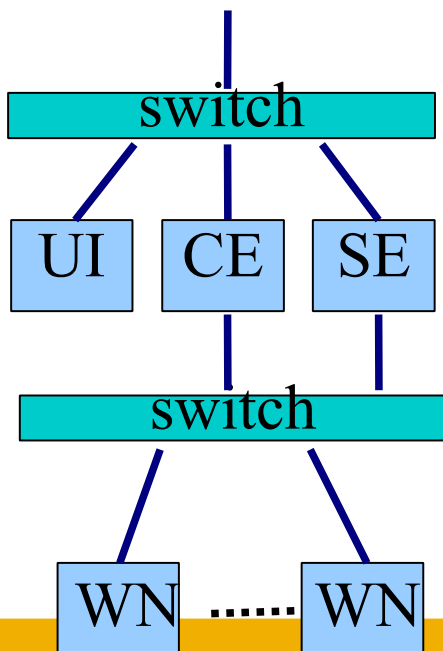


- **Výpočtový klaster sa skladá z uzlov:**
 - **Computing Element (CE)** – inak nazývaný gatekeeper, centrálny uzol klastra
 - **Storage Element (SE)** – rozhranie na úložný priestor (disk, páskový robot, ...). Môže byť integrovaný s CE.
 - **Worker Node (WN)** – výpočtový uzol
 - **User Interface (UI)** – uzol z ktorého používatelia prístupujú do Gridu. Na ňom majú uložené gridové certifikáty. Nemusí byť súčasťou výpočtového uzla ak používatelia majú prístup na UI niektorej z virtuálnych organizácií.

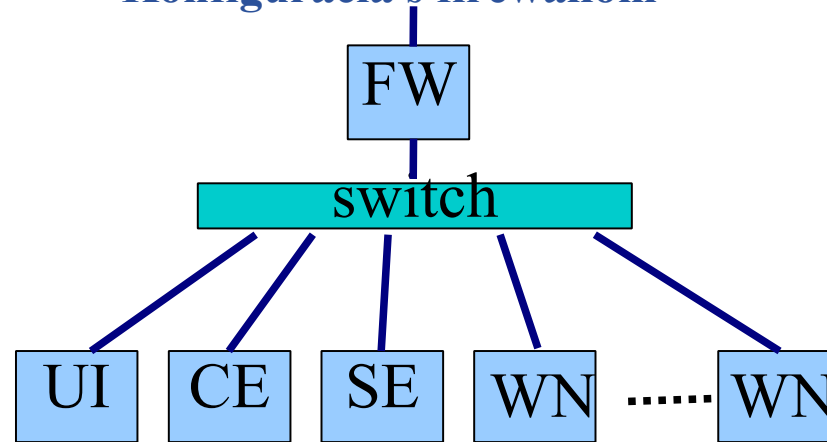
Štandardná konfigurácia:
výpočtové uzly s verejnými IP adresami



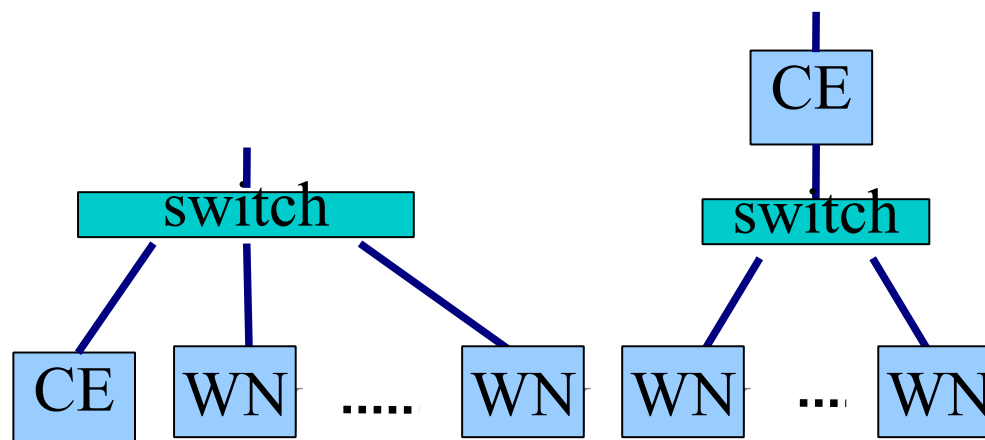
Výpočtové uzly s privátnymi IP adresami



Konfigurácia s firewallom



Minimálna konfigurácia





- **Výpočtové uzly na privátnej sieti**
 - + môžu byť bezdiskové (nižšia cena a spotreba energie)
 - + možnosť filtrovania komunikácie – zvýšená bezpečnosť
 - zložitejšia inštalácia a administrácia
- **Firewall**
 - + možnosť filtrovať komunikáciu (vrátane komunikácie smerom von) – vyššia bezpečnosť
 - zložitejšia administrácia



- **EGEE middleware poskytuje generické Grid služby:**
 - Informačné služby
 - Zadávanie úloh (Job submission)
 - Dátový management
 - Bezpečnosť
 - Logovanie
 - Monitorovanie
- **EGEE podporuje zdieľanie výpočtových cyklov a dátových úložísk viacerými Virtuálnymi Organizáciami**



Aplikácie

Lokálne Aplikácie

Lokálne Databázy

MIDDLEWARE

Gridová Aplicačná Vrstva

Manažment
úloh

Dátový
Manažment

Metadata
Manažment

Kolektívne Služby

Plánovač
(Scheduler)

Manažér
Replík

Info & Monitorovacie
Služby

Podporné Gridové Služby

SQL
Databázové
Služby

Computing
Element
Služby

Storage
Element
Služby

Katalóg
Replík

Autorizácia
,Autentifikácia,
Výkazy

Indexová
Služba

Štruktúrálna
vrstva

Štruktúrálné Služby

Manažment
Zdrojov

Manažment
Konfigurácií

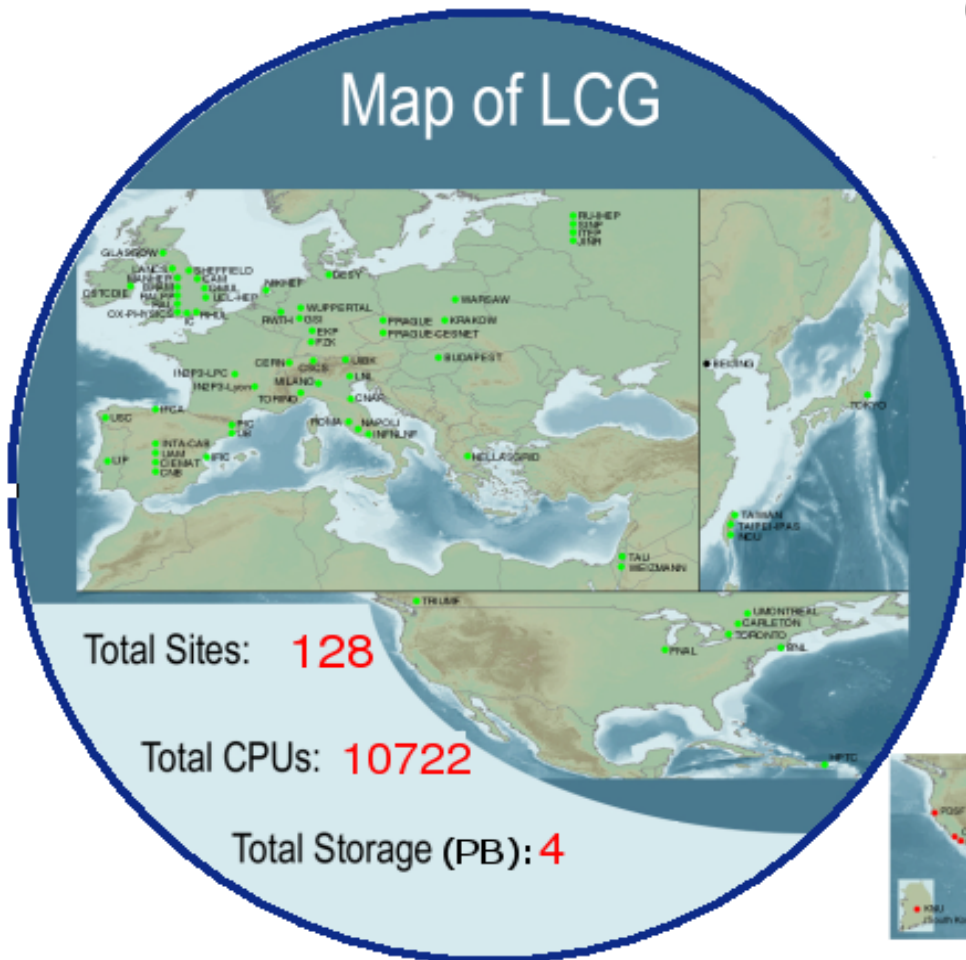
Monitorovanie

Manažment a
Inštalácia uzlov

Manažment
úložísk



- **Globus Toolkit**
 - <http://www.globus.org>
- **LCG**
 - <http://lcg.web.cern.ch/LCG>
- **gLite**
 - <http://glite.web.cern.ch/glite>



Collaborating with LCG

NorduGrid



Grid3





- **Základné infraštruktúrne služby:**
 - Spravovanie základných gridových služieb
- **Monitorovanie a kontrola Gridu:**
 - Proaktívne monitorovanie operačného stavu a výkonnosti
 - Zahajovanie opravných akcií
- **Zavádzanie Middlewaru:**
 - Validácia a zavádzanie nových verzií middlewaru
 - Nasatvenie operačných procedúr pre nové zdroje
- **Podpora užívateľov a poskytovateľov zdrojov:**
 - Koordinácia riešenia problémov užívateľov a zdrojových centier
 - Filtrácia a združovanie problémov, poskytovanie riešení
- **Grid manažment:**
 - Koordinácia Regional Operations Centres (ROC) a Core Infrastructure Centres (CIC)
 - Riadenie vzťahov medzi poskytovateľmi zdrojov na úrovni servisných dohôd
- **Medzinárodná kolaborácia:**
 - Posilňovanie kolaborácie s obdobnými organizáciami v USA a Ázisko-Pacifickom regióne
 - Zabezpečenie interoperability gridových infraštruktúr a služieb pre viacdoménové VO
 - Účasť v štandardizačných telesách v širšej gridovej komunite



1 Operations Management Centre – OMC

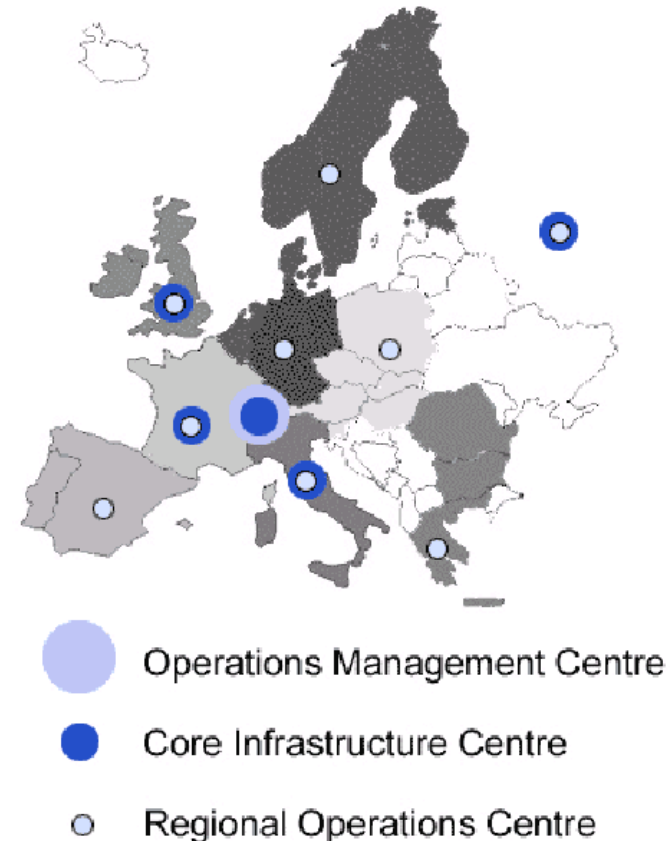
- Koordinácia CIC a ROC
- Tím pre operačný manažment – vyriešené problémy, výkonnostné ciele a pod.
- Operations Advisory Group - poradná skupina pre riadenie postupov

5 Core Infrastructure Centres – CIC

- Každodenné riadenie operácií – zavádzanie operačných postupov definovaných OMC
- Monitorovanie stavu, riadenie korekcií, 24x7 operačná grid infraštruktúra
- Vykazovanie používanosti zdrojov, koordinácia riešenia bezpečnostných incidentov, zabezpečenie postupov pre znovuzprístupnenie zdrojov

~11 Regional Operations Centres – ROC

- Priamo poskytujú podporu pre užívateľov a výpočtové centrá
- Podpora pre nové výpočtové centrá vstupujúce do EGEE infraštruktúry





- **Registračná procedúra**
 - Získanie osobných certifikátov pre administrátorov
 - Zaslanie registračných informácií ROCu
 - Zaradenie kandidátskeho zdrojového centra do infraštruktúry
- **Certifikácia zdrojového centra**



Vďaka za pozornosť.