

EGEE-II/EGEE-III

Aleš Křenek

2. dubna 2008

Co je EGEE

Enabling Grids for e-science in Europe

Enabling Grids for e-science

Enabling Grids for e-science

Enabling Grids for e-science

„EGEE is providing a production quality Grid infrastructure spanning more than 30 countries with over 150 sites to a myriad of applications from various scientific domains, including Earth Sciences, High Energy Physics, Bioinformatics and Astrophysics.” (Bob Jones, ředitel projektu)

Historie, současnost, blízká budoucnost

- DataGrid, 2001–2003, ~20 MEur, 21 partnerů
- EGEE, 2004–2005, ~ 30 MEur, ~ 70 partnerů
- EGEE-II, 2006–2007(8), ~ 30 MEur, ~ 90 partnerů
- EGEE-III, 2008–2010, ~ 30 MEur, ~ 90 partnerů

Historie, současnost, blízká budoucnost

- DataGrid, 2001–2003, ~20 MEur, 21 partnerů
 - formulace požadavků, základní vývoj
- EGEE, 2004–2005, ~ 30 MEur, ~ 70 partnerů
- EGEE-II, 2006–2007(8), ~ 30 MEur, ~ 90 partnerů
- EGEE-III, 2008–2010, ~ 30 MEur, ~ 90 partnerů

Historie, současnost, blízká budoucnost

- DataGrid, 2001–2003, ~20 MEur, 21 partnerů
 - formulace požadavků, základní vývoj
- EGEE, 2004–2005, ~ 30 MEur, ~ 70 partnerů
 - nasazení prototypu
- EGEE-II, 2006–2007(8), ~ 30 MEur, ~ 90 partnerů
- EGEE-III, 2008–2010, ~ 30 MEur, ~ 90 partnerů

Historie, současnost, blízká budoucnost

- DataGrid, 2001–2003, ~20 MEur, 21 partnerů
 - formulace požadavků, základní vývoj
- EGEE, 2004–2005, ~ 30 MEur, ~ 70 partnerů
 - nasazení prototypu
- EGEE-II, 2006–2007(8), ~ 30 MEur, ~ 90 partnerů
 - produkční provoz
- EGEE-III, 2008–2010, ~ 30 MEur, ~ 90 partnerů

Historie, současnost, blízká budoucnost

- DataGrid, 2001–2003, ~20 MEur, 21 partnerů
 - formulace požadavků, základní vývoj
- EGEE, 2004–2005, ~ 30 MEur, ~ 70 partnerů
 - nasazení prototypu
- EGEE-II, 2006–2007(8), ~ 30 MEur, ~ 90 partnerů
 - produkční provoz
- EGEE-III, 2008–2010, ~ 30 MEur, ~ 90 partnerů
 - spuštění LHC

Hlavní hráči

- Uživatelé
 - částicová fyzika, věda o zemi, chemie, lékařské aplikace, ...
 - aktivity projektu NA2–5

Hlavní hráči

- Uživatelé
 - částicová fyzika, věda o zemi, chemie, lékařské aplikace, ...
 - aktivity projektu NA2–5
- Provoz
 - koordinace správy zdrojů (nejsou vlastněny EGEE), monitorování
 - podpora uživatelů (GGUS)
 - správa virtuálních organizací
 - aktivita SA1

Hlavní hráči

- Uživatelé
 - částicová fyzika, věda o zemi, chemie, lékařské aplikace, ...
 - aktivity projektu NA2–5
- Provoz
 - koordinace správy zdrojů (nejsou vlastněny EGEE), monitorování
 - podpora uživatelů (GGUS)
 - správa virtuálních organizací
 - aktivita SA1
- Middleware
 - vývoj a údržba gLite
 - aktivity JRA1 a SA3

- bezpečnost: VOMS, LCAS, ...
- výpočetní zdroje: LCG-CE, CREAM
- datová úložiště: DPM
- správa úloh: WMS, L&B
- datové služby: LFC, FTS, GFAL

Infrastruktura

- CPU a diskový prostor

DataGrid	1841 CPU	
EGEE	14000 CPU	
EGEE-II	41000 CPU	5 PB
- 30 (stabilně), 100 (špička) tisíc úloh za den

Lidová tvořivost

- sklony aplikačních skupin řešit problémy po svém
 - AliEn, Ganga, Dirac, Taverna, Moteur, Charon, . . .

Lidová tvořivost

- sklony aplikačních skupin řešit problémy po svém
 - AliEn, Ganga, Dirac, Taverna, Moteur, Charon, . . .
- zpravidla efektivnější a rychlejší řešení specifických problémů
 - mnoho vynalezených trakařů
 - jak nejlépe tento vývojářský potenciál využít?

Lidová tvořivost

- sklony aplikačních skupin řešit problémy po svém
 - AliEn, Ganga, Dirac, Taverna, Moteur, Charon, . . .
- zpravidla efektivnější a rychlejší řešení specifických problémů
 - mnoho vynalezených trakařů
 - jak nejlépe tento vývojářský potenciál využít?
- analogie v sítích
 - proč už nemáme ATM
 - co je “IP” v Gridech?

Cíle EGEE-III

- Oficiální formulace:
 - To expand, optimize and simplify the use of Europe's largest production Grid by continuous operation of the infrastructure, support for more user communities, and addition of further computational and data resources;
 - To prepare the migration of the existing Grid from a project-based model to a sustainable federated infrastructure based on National Grid Initiatives.

Cíle EGEE-III

- Oficiální formulace:
 - To expand, optimize and simplify the use of Europe's largest production Grid by continuous operation of the infrastructure, support for more user communities, and addition of further computational and data resources;
 - To prepare the migration of the existing Grid from a project-based model to a sustainable federated infrastructure based on National Grid Initiatives.
- nárůst infrastruktury
 - 125 MSI2k (2–3×), 33 TB (6×)
 - financováno partnery, předpokládaný rozpočet 50 MEur
- podpora „strategických“ aplikačních oblastí
 - HEP, life sciences, earth science, chemie
- důraz na produkci, útlum vývoje
 - jen JRA1 (proti JRA1–4 v EGEE)
 - redukovaný rozpočet (38%) – jen údržba

Zapojení CESNETu

- NA*
 - web, tiskové zprávy, ...
 - školení uživatelů
 - podpora aplikací
- SA1 – provoz
 - cluster Skurut
 - ▶ CESNET
 - ▶ původně $32 \times 2 \times$ P-III
 - ▶ nyní $10 \times 2 \times 4$ Xeon
 - ▶ 13 TB disků
 - cluster Goliáš
 - ▶ $200 \times$ P-III
 - ▶ 240 jader Intel/AMD
 - ▶ 32 TB disku
 - slíbeno do EGEE-III: 350 kSI2k, 30 TB
 - virtuální organizace VOCE a Auger

Zapojení CESNETu (2)

- JRA1 a SA3
 - Logging and bookkeeping (L&B)
 - ▶ sledování úloh za běhu
 - Job Provenance
 - ▶ archiv dat o úlohách
 - bezpečnost
 - ▶ obnovování X509 proxy
 - ▶ účast v koordinačních skupinách