On the Way from Distributed Computing to e-Infrastructure for Research

M. Kuba, D. Antoš, J. Kmuniček, I. Křenková
CESNET and Masaryk University Brno
Outline

• MetaCentrum transformation to NGI
• what are EGI, NGI, VO
• MetaCentrum NGI, MetaCentrum VO
• expected new resources in NGI
• NGI and structural biology
MetaCentrum transforms to NGI

- MetaCentrum was originally established as an alliance of supercomputing centers of MU, UK and ZČU – hence meta-centrum
- Over time other centers joined, became the Czech academic distributed computing infrastructure coordinated by CESNET
- 300 machines, 1560 CPU, 600TB storage
- Now is transforming to National Grid Initiative (NGI) under European Grid Initiative (EGI)
- The era of the European grid begins :-(
• three EGEE (Enabling Grids for E-sciencE) projects *created* the world’s largest grid infrastructure (267 sites in 55 countries, 150000 CPUs, 69PB, 14k users)

• EGI (European Grid Initiative) will maintain *sustainable* infrastructure

• each country will have NGI (National Grid Initiative), the Czech NGI will be CESNET’s MetaCentrum
Virtual Organizations

- the grid is organized in VO (Virtual Organizations)
- a VO has users (typically groups from various real organizations) and access to resources (computing and storage)
- resources belong to resource owners
MetaCentrum roles

• three better separated roles
  – MetaCentrum NGI – coordinates national VOs and resource owners
  – MetaCentrum VO – free catch-all organization for everybody in the Czech academia
  – MetaCentrum as CESNET resource owner

• other VOs
  – NCBR, LL, AUGER, BELLE, VOCE, ...

• other resource owners
  – NCBR, LL, CERIT, MU, ZČU, UK, VUT, JU, ...
Planned new resources in NGI

- **CERIT** Center of Education, Research and Innovation in ICT in Brno
  - Supercomputer Center with about 10000 CPU and 15-20 petabytes of storage

- **CESNET** storage infrastructure
  - large storage services in Internet points of presence (8-12 PB)

- integrated resources from large infrastructures
NGI and structural biology

• NGI provides
  – data storage for backup and archiving
  – common infrastructure for cooperation with foreign partners
  – collaboration environment for geographically dispersed teams and individuals (next slide)
  – computational resources (second next slide)
Computational resources for

- long term (bio)molecular simulations
- protein engineering of enzymatic structures
- QM/MM studies of catalytic reactions
- QM calculations of interaction energies and NMR parameters
- free energy calculations of intermolecular interactions and conformational changes of biomolecules
- detailed exploration of binding sites and processes in biomolecules
- prediction of 3D structure of macromolecules (proteins and nucleic acids)
- biostatistics analysis and multidimensional parametric studies
- analysis of genome and its evolution
- bioinformatics methods of virtual screening of biologically active compounds
- quantitative structure activity / function relationships studies
- ...

MetaCentrum VO

- free membership in MetaCentrum VO for everybody in academia
- demand for resources is prioritized by scientific results
- new resources are welcome
- we offer maintenance of computing resources for IT non-experts
- international cooperation through user groups within EGI
Top users’ time

January – November 2009

CPU years

benedikt | ettich | briza | lukecz | seidj1am | judit | pepap | melicher | zeleny | jsebera

0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0
Summary

- MetaCentrum evolves from SC alliance to NGI in the upcoming EGI
- the “free access” part will become MetaCentrum VO
- the “restricted” parts will be other VOs
- collaboration tools and storage will be in the NGI, available to all VOs
Thank you