Tier 2 Center in Prague

Jiří Chudoba

Institute of Physics AS CR

Tier 2 Centre in Prague

- Grid projects in the CR (local and international Collaboration)
- Network in the Czech Republic
- Main grid applications and user communities, resources
- Plans

Grid projects in CZ

- Grid-like project from 1990th
  - MetaCentre
- Grid projects
  - EDG – European DataGrid
  - GRIDLAB, COREGRID - CESNET
  - LCG – LHC computing Grid
  - EGEE – Enabling Grids for E-science in Europe
  - SAMGRID/JIM (working data handling system with grid functions and job monitoring) – developed at FNAL for D0 and CDF – starting in FZU (D0)

Project MetaCentre (1/2)

- MetaCentre – project of CESNET with more collaborators
  - CESNET z.s.p.o.
    - Association of legal bodies formed by Czech Universities and Academy of Sciences of the Czech Republic (AS CR), provider of Czech research network, research in advanced network technologies and applications
  - Distributed environment for extensive computing
    - Super/Cluster/Grid computing
  - Collaborates with international computing projects
    - CESNET collaborates with international Grid projects through MetaCentre
  - Main HW resources in
    - Supercomputing Centre at Masaryk University in Brno
    - Supercomputing Centre at Charles University in Prague
    - Supercomputing Centre at West Bohemian University in Pilsen

Project MetaCentre (2/2)

- High speed network connectivity
- Unified environment with
  - Single user login, same user interfaces and application interfaces
  - Shared file system OpenAFS with Kerberos
  - Interactive access and batch queue system PBS, support for parallel and distributed computing.
  - Special applications supported on specialized university locations – e.g. computational chemistry, molecular modeling, technical and material simulations

EDG in CZ

- EDG – European DataGrid Project
  - 2001 – 2003 (prolonged to end March 2004)
  - CZ participation
    - CESNET
    - Institute of Physics AS CR (FZU)
      - Research institute for solid state and particle physics
  - Contribution to work packages
    - WP1 – Workload Management System (scheduling and resource management)
      - CESNET – logging and book keeping service
    - WP6 – Testbed and Demonstrator
      - Grid farms at CESNET and FZU
      - Certification Authority created at CESNET for scientific computing
    - WP7 – Network monitoring
      - CESNET

LCG in CZ (1/2)

- LCG – LHC Computing Grid (joined 2002)
  - Computing environment for LHC experiments by deploying a worldwide computational grid service, integrating the capacity of scientific computing centres spread across Europe, America and Asia into a virtual computing organisation
  - CZ participation in GDB – Grid Deployment Board
    - forum within the LCG project where the computing management of the Experiments and the Regional Centres can discuss and take, or prepare, the decisions necessary for planning, deploying and operating the LCG Grid
    - FZU participates
      - Hosting Tier-2 Regional Computing Center
      - Currently installed MW is LCG-2 and AliEn

LCG-2 status map

2 centers in CZ - FZU and CESNET (EGEE)

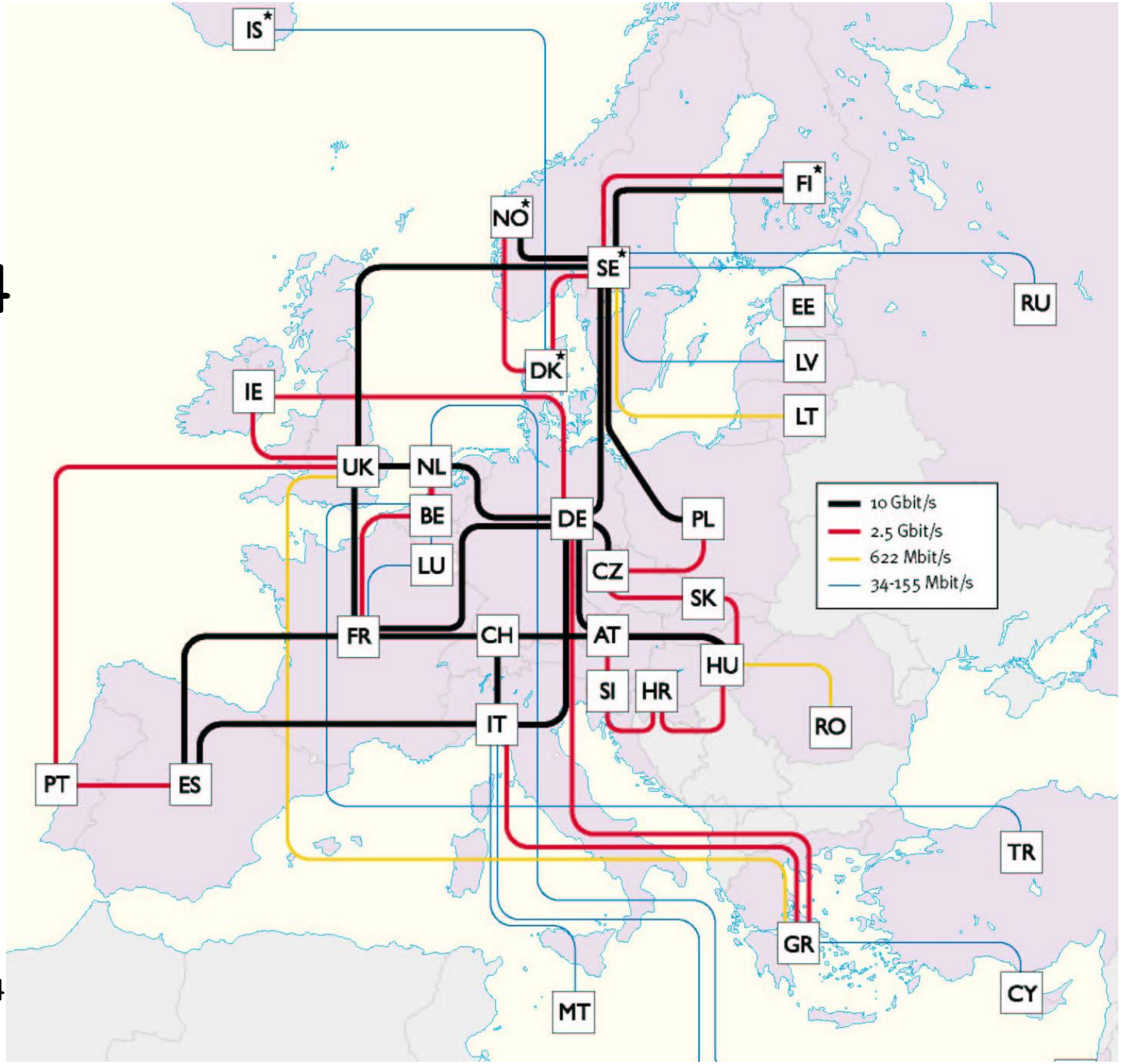
EGEE in CZ

- EGEE – Enabling Grids for E-science in Europe
  - For 2004-2005
  - integrate current national, regional and thematic Grid efforts to create a seamless European Grid infrastructure for the support of the European Research Area
  - CZ participation in EGEE – CESNET
    - SA1 - European Grid Support, Operation and Management
      - Operation of LCG-2 certified farms FZU and CESNET
      - Local support
    - NA3 - User Training and Induction
    - NA4 - Application Identification and Support
      - computational chemistry, technical and material simulations
    - JRA1 - Middleware Re-engineering and Integration

Network in the CZ

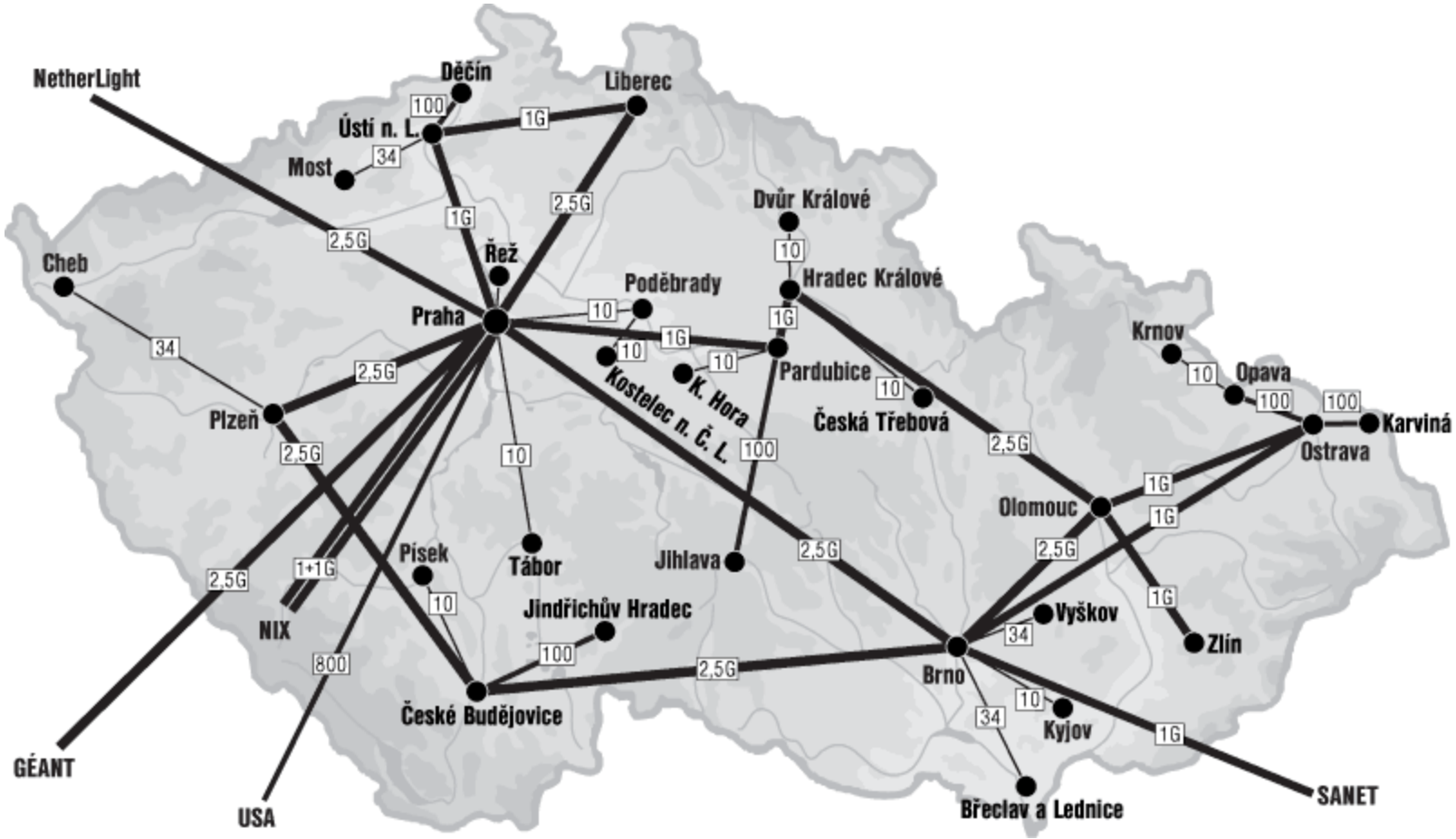
- External connections
  - 2.5 Gbps line to GÉANT, used for academic traffic
  - 800 Mbps line to Telia, used for commodity traffic
  - 10 Gbps line to NetherLight for experimental traffic
- Internal network
  - Two connected stars
  - Redundant connections
    - Effectively connections of network rings with small number of hops

GEANT Topology May 2004



Prague, 9.9.2004

CESNET Topology April 2004



Main grid applications in CZ

- HEP - High Energy Physics (i.e. Particle Physics)
  - Well established
- Computational chemistry, Technical and material simulations
  - Preparing environment and looking for users in the framework of EGEE

Particle Physics Applications (1/6)

- Particle Physics in the Czech Republic
  - Charles University in Prague
  - Czech Technical University in Prague
  - Institute of Physics of the Academy of Sciences of the Czech Republic
  - Nuclear Physics Institute of the Academy of Sciences of the Czech Republic
- Main Applications
  - Projects ATLAS, ALICE, D0, STAR, TOTEM
  - Groups of theoreticians
  - Approximate size of the community in 2004
    58 scientists, 22 dipl. engineers, 21 technicians and 43 students and PhD students

Particle Physics Applications (2/6)

- Computing/ grid resources
- From 2004 usage of the CESNET farm
  - Skurut
    - 16 nodes PIII, 2 processors, 700 MHz, U2, 1 GB per node
- Today Skurut
  - LCG2 certified farm
  - shared with other applications

Particle Physics Applications (3/6)

- Computing farm GOLIAS in FZU
  - 34x dual 1.13Ghz PIII
  - 1GB RAM per node
  - 1TB disc array
  - 10TB disc, 3 x 3 TB
  - Power 34 kSI2000
  - Plan for usage:
    - 50% LCG (ATLAS+ALICE), 50% D0

Particle Physics Applications (4/6)

- New farm from July 2004
  - New computer place in FZU with 150 kW electric power for computers
  - Air condition, UPS, Diesel,
  - Network connection
    - Standard 1 Gbps connection to research network CESNET
      - via metropolitan Prague research network PASNET
    - Direct optical connection 1 Gbps to CzechLight (CESNET- Amsterdam-Geneva)
      - With BGP fallback to GEANT
  - 2 x 9 racks
  - Shared with FZU – main user
  - + room for operator

Particle Physics Applications (5/6)

- New equipment (from July 2004), installation
  - 49x dual Intel Xeon 3.06 GHz with Hyper Threading – computing elements
  - 2x dual Intel Xeon 2.8 GHz with Hyper Threading – frontend
  - 3x dual AMD Opteron 1.6 GHz
    - 1 file server (64 bits)
    - 2 computing elements
  - Disc array 30 TB, ATA discs, RAID5
  - All nodes connected via 1 Gbps
  - 3x HP ProCurve Networking Switch 2848
- New power ~100 kSi2000

Particle Physics Applications (6/6)

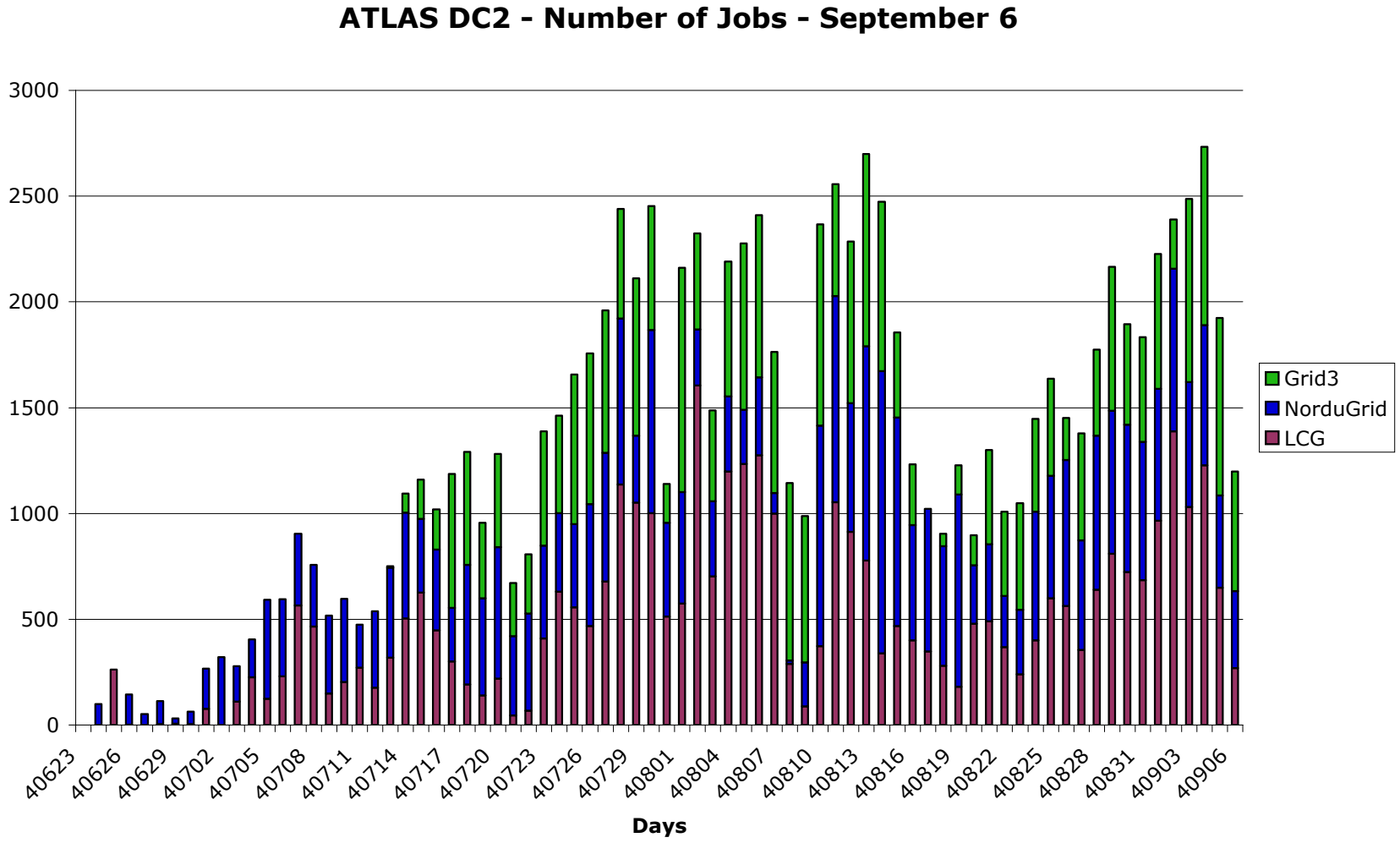
- **LCG2** (LCG-2_2_0) installed and certified
- Queuing system
  - PBSPro 5.4.0
  - Task submission either locally or via grid (LCG2)
  - For LCG2 we have prepared interface to PBSPro
    - Changes to middleware published LCG web
  - queues
    - General: short, normal, long
    - Special for experiments: D0, ATLAS, ALICE
    - For LCG2: lcgshort, lcglong, lcginfinite

- Disk array partitions
  - Success with creation one 8TB partition from smaller partitions
    - using LVM2 from linux kernel 2.6, can't be easily done with 2.4 kernel due to block device size limit
  - At least kernel 2.6.6 is needed, older versions have stability problems when NFS goes under heavy load on >1TB partitions

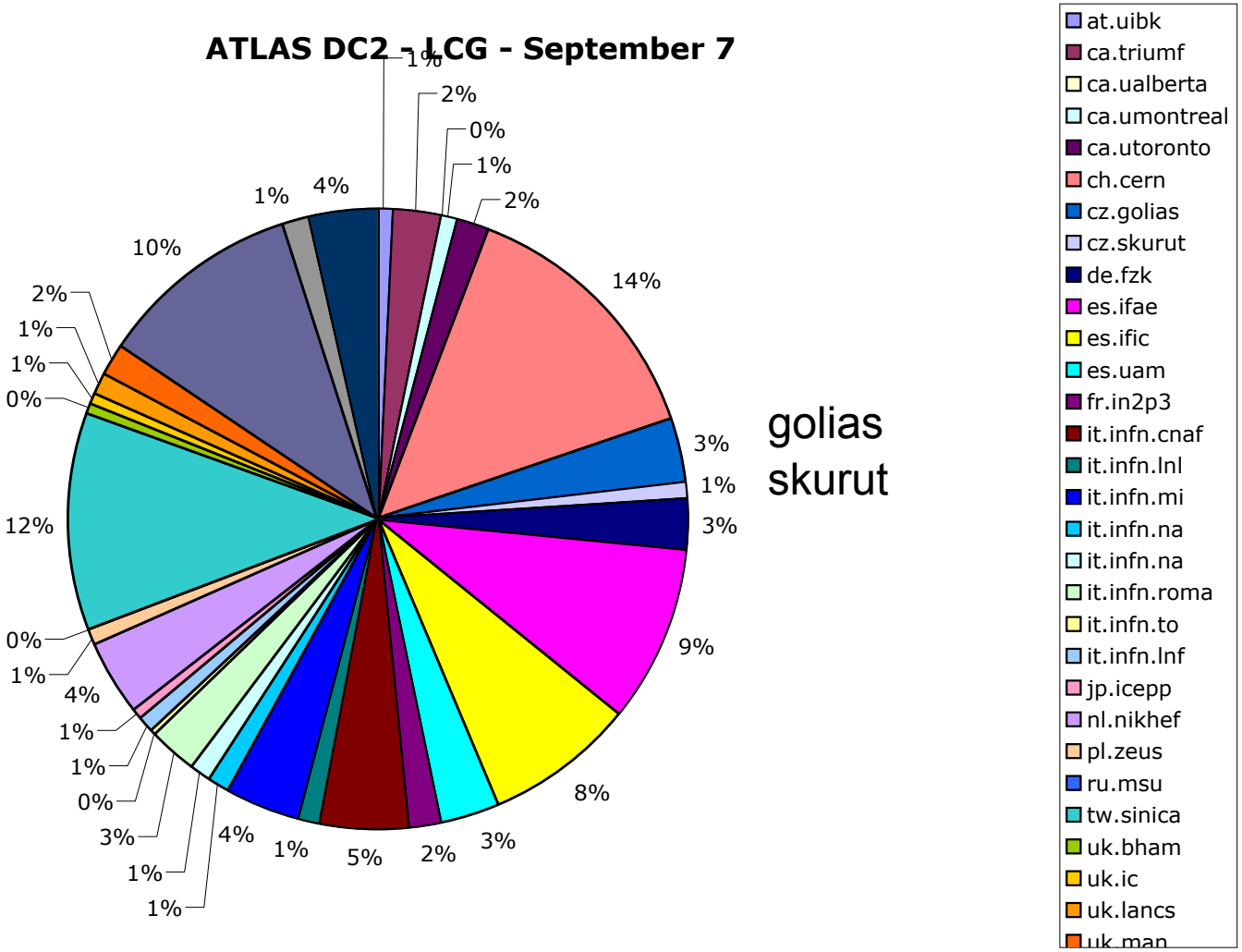
- On both farms Skurut and Golias, instlalled and certified
  - LCG-2_2_0 http://goc.grid-support.ac.uk/gppmonWorld/gppmon_maps/CERN_lxn1188.html
  - ATLAS SW http://mbranco.home.cern.ch/mbranco/cern/lcg2.html
  - ALIEN

Participation in Atlas DC

- effectively started in July, during Golias farm upgrade
- anyway we managed to participate from the beginning:
  - new nodes were gradually installed and added to the farm
  - disk space is big enough, 30TB disk array will be added later

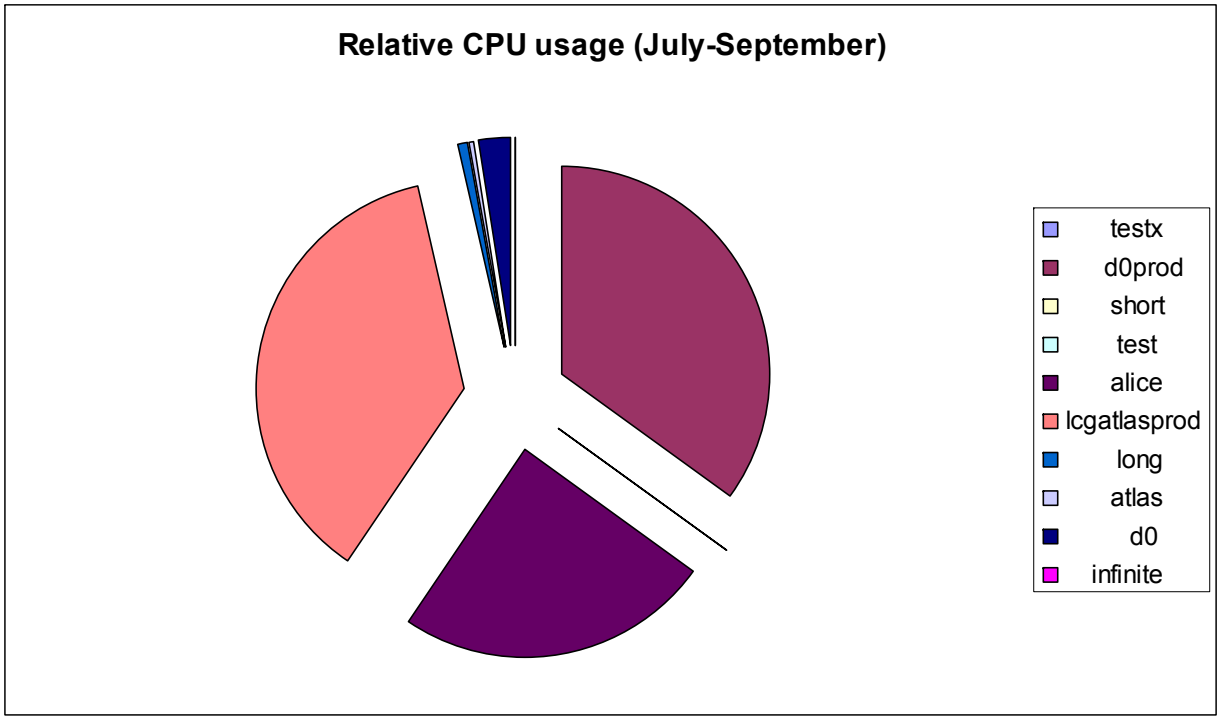

ATLAS DC2 - Number of Jobs - September 6
3000
2500
2000
1500
1000
500
0
40623
40626
40629
40702
40705
40708
40711
40714
40717
40720
40723
40726
40729
40801
40804
40807
40810
40813
40816
40819
40822
40825
40828
40831
40903
40906
Days
Grid3
NorduGrid
LCG

ATLAS DC2 - LCG part



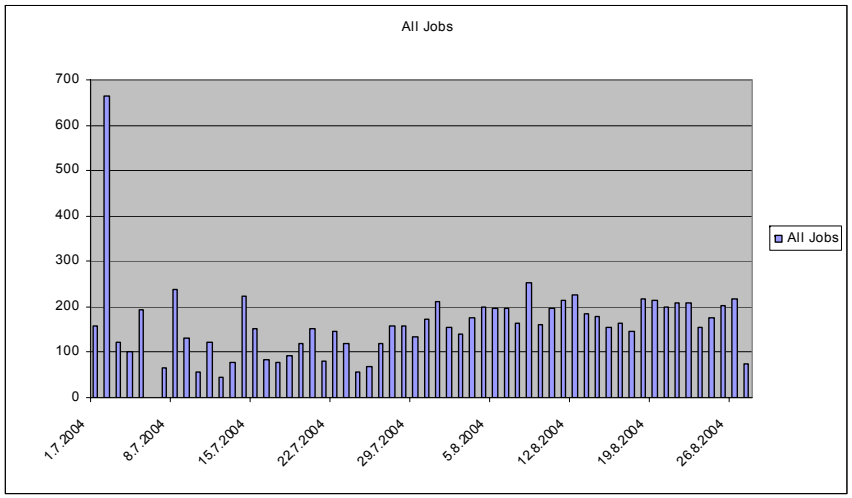
golias
skurut

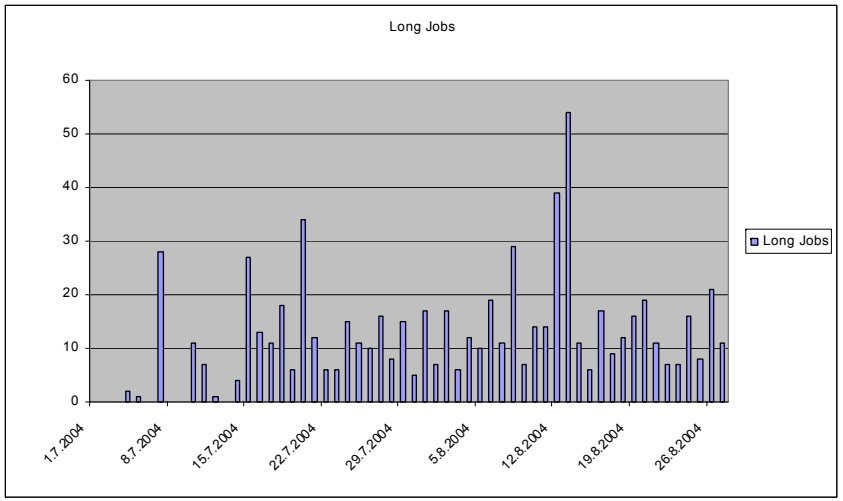
Jobs statistics - Golias

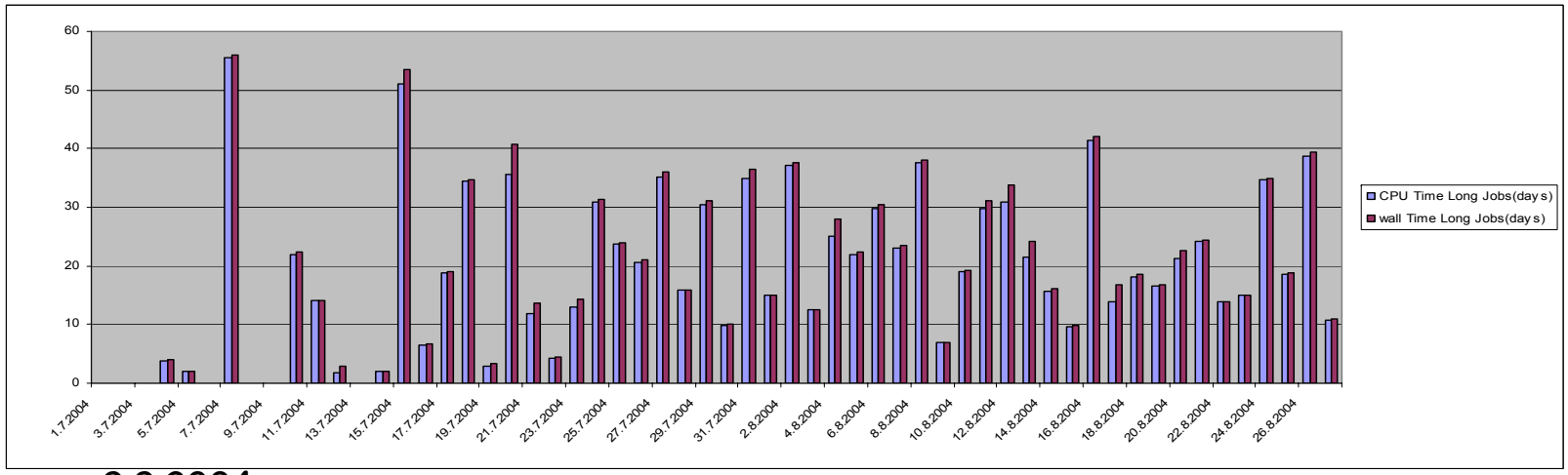


Golias farm usage, July-September

Skurut in ATLAS DC2

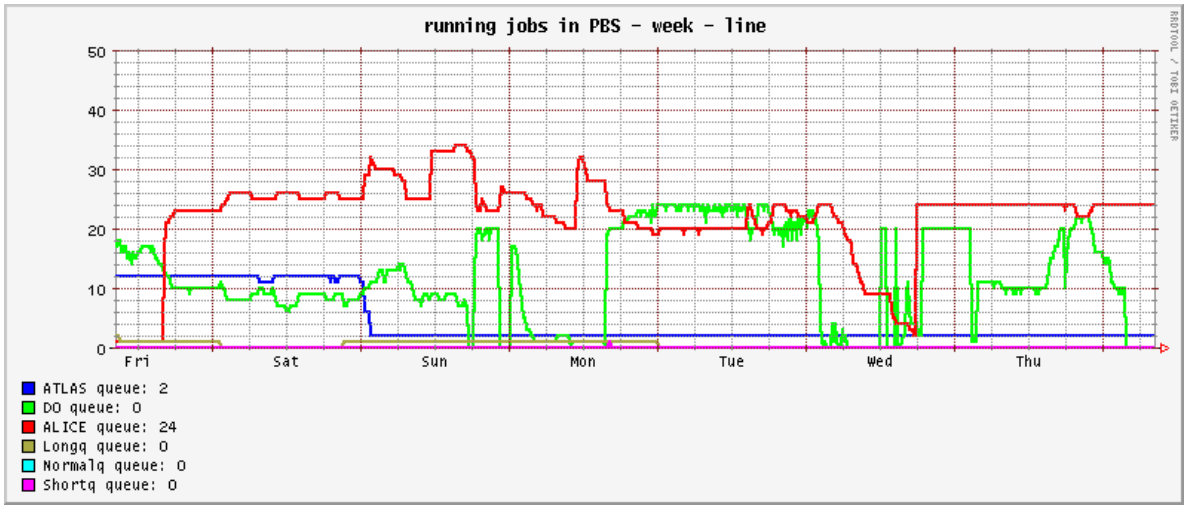






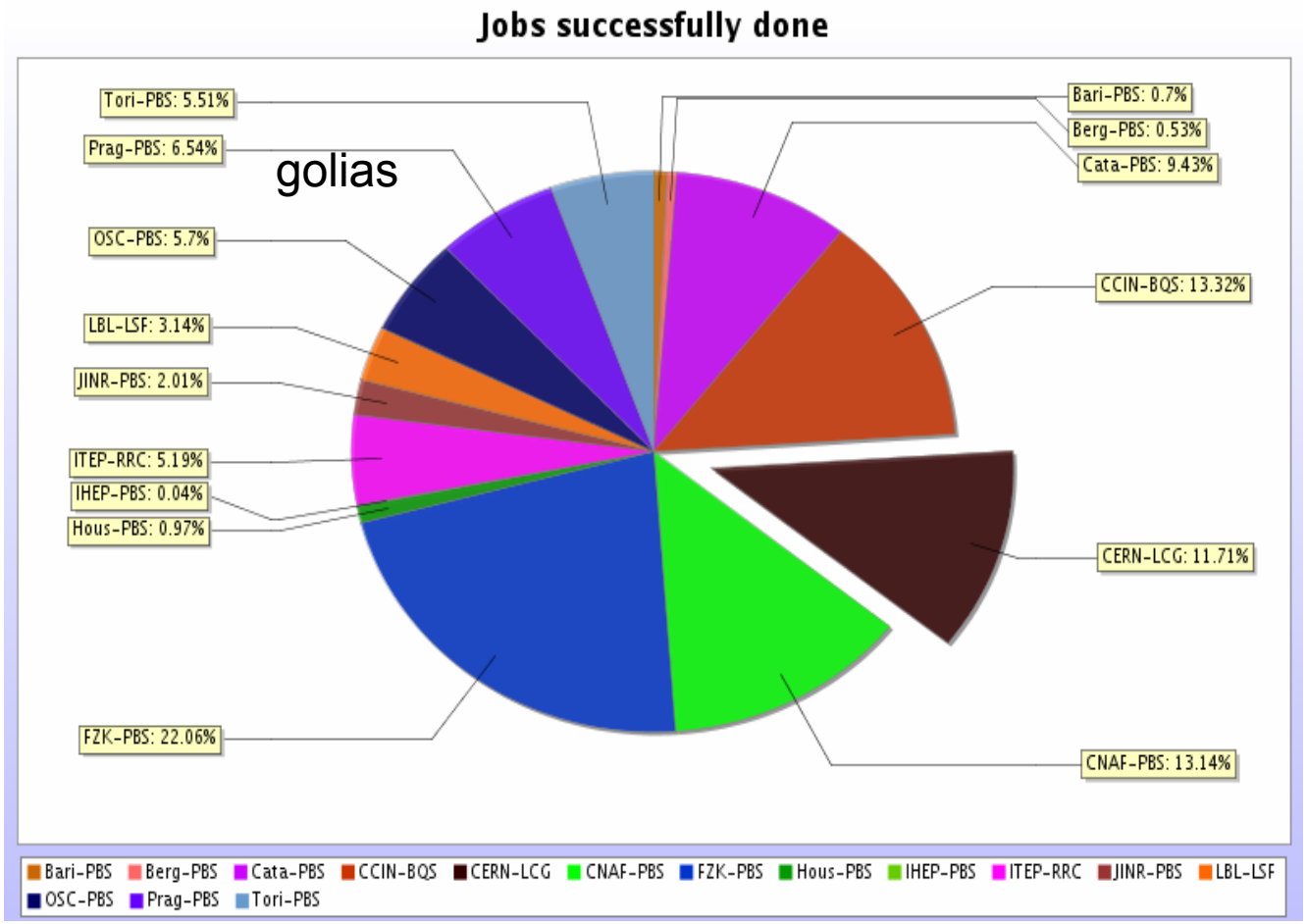
ALICE PDC Phase1
Prague – Golias (before upgrade)

- Data transfer: 1.8 TB over 50 days active running
- CPU work: 20.3 MSI-2K hours
- Number of CPUs: Max 32, avarage 16



Number of running jobs on Golias farm during ALICE DC
(1 week snapshot)

ALICE PDC – Phase2

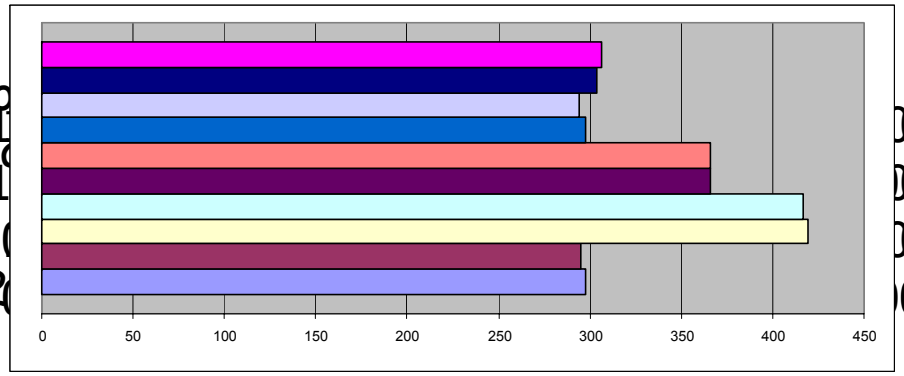


Technical problems to solve

- Fair-share usage in heterogeneous environment:
  - just reserved number of CPUs not good
  - smaller difference in the CPU power for jobs with big IO requirements
- OpenPBS vs PBSPro
- Optimal usage of the facility
- HyperThreading: on or off?

HyperThreading

	noHT	HT	HT with scheduling
	1	1.13	1
2 jobs, parallel	297 ±1 s	337 ± 48 s	296 ± 2 s
			1.73
4 jobs, parallel			514 ± 3 s
			2
2+2			592 s



CERN RH 7.3.3, kernel 2.4.20, AliRoot v4-01-05, 1000 tracks HIJINGParam, Real time

ftp://ftp.kernel.org/pub/linux/kernel/people/rml/cpu-affinity/ + http://freshmeat.net/projects/sched-utils/

Financial resources

- Financial resources for the Grid activities in the Czech Republic
  - No dedicated Government support for the Grid projects
  - CESNET supports Grid projects from its research budget and from EU grants (EGEE)
  - Institutions exceptionally get small grants
  - Occasional supports from the Institutes
    - Like the new computing room in FZU
  - Marginal support from application projects needing computing

Plans

- EGEE
  - Should help to attract other applications to use working grid infrastructure
- HEP
  - Existing environment serves at the level of TIER-2 regional centre
  - We have to find resources to be able to upgrade it ~ten times to have required computing resources in 2007

Conclusion

- The Grid infrastructure in the Czech Republic established
  - Profit from earlier experience and know-how of the MetaCentre project
  - Established EDG, LCG and EGEE projects and some smaller ones
  - Applications
    - HEP – well established local community actively using grid environment for LCG
    - Other applications – looked for inside EGEE project, local experience already with Computational chemistry, Technical and material simulations
  - No specialised resources for grid projects, partial financing from different research projects or Institutions ad hoc contributions