

Grid jako superpočítač

Jiří Chudoba
Fyzikální ústav AV ČR, v. v. i.

Potřeba výkonných počítačů

- Vědecké aplikace
- Podnikové aplikace
- Internetové aplikace



Microsoft datová centra

Google datová centra

600 mil USD na datové centrum
36 center celkem v 2008
až 100 MW příkonu



HPC

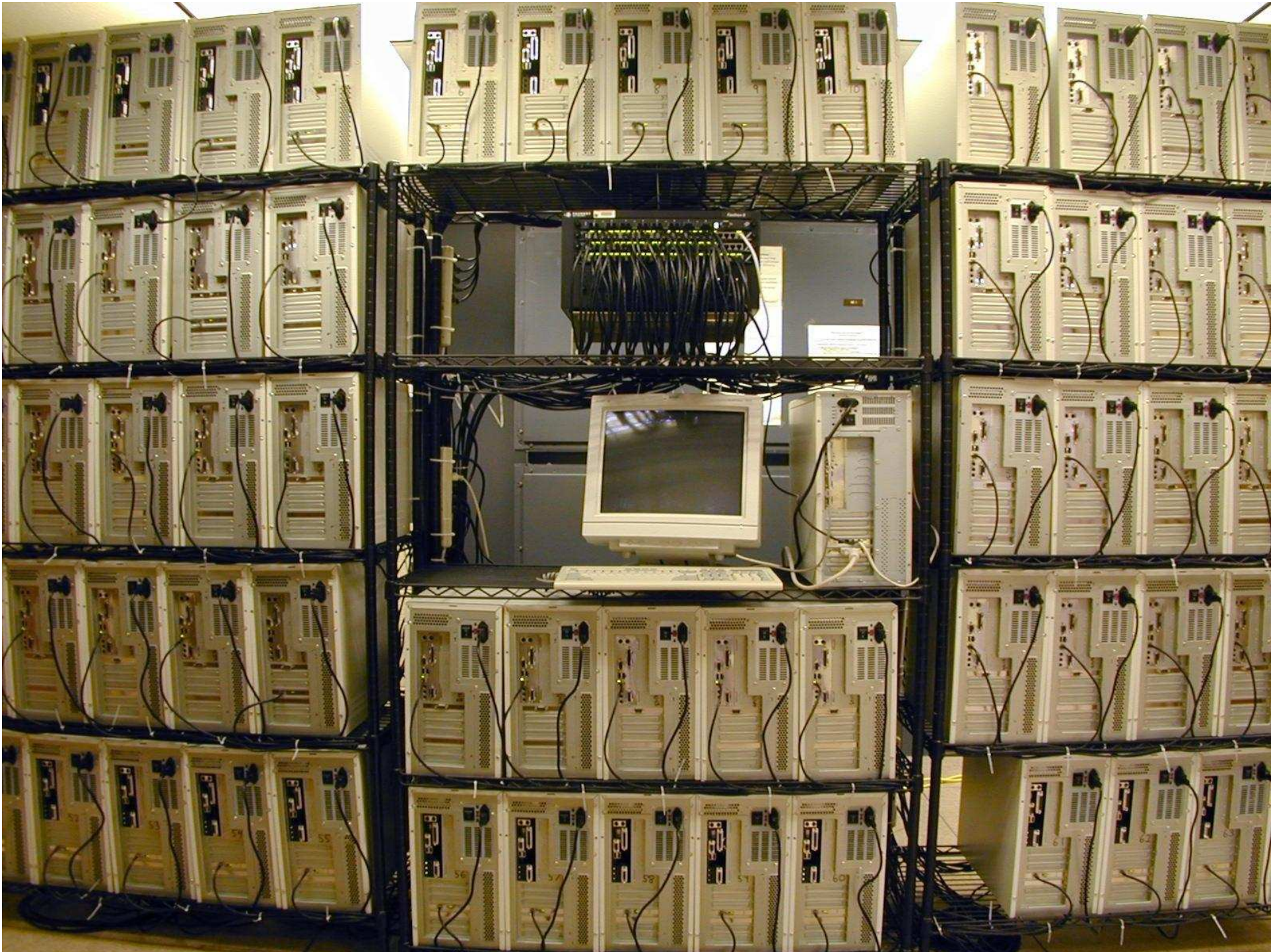
High Performance Computing

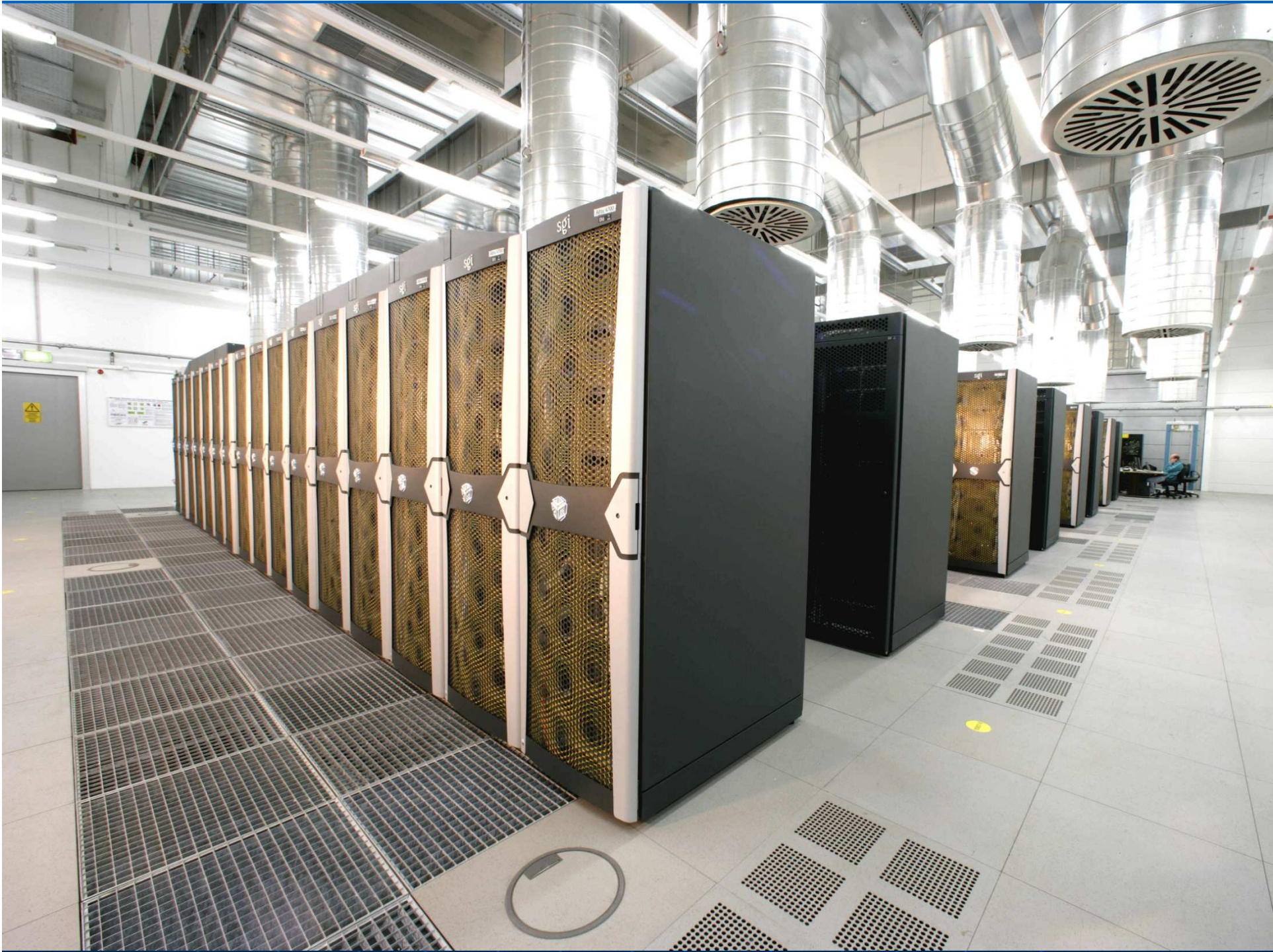
- Superpočítače vs klastry
 - DEISA (Distributed European Infrastructure for Supercomputing Applications)
 - EGI (European Grid Initiative)
- Řešení závisí na aplikacích
 - požadavky na paměť
 - komunikace mezi procesory

Klastry a superpočítače

- Paralelizace na úrovni
 - bitů (32, 64 bitů)
 - instrukcí
 - dat
 - úloh
- Vektorové procesory
 - SIMD single instruction, multiple data
 - Cray – nejvýkonnější počítače své doby
 - nyní návrat formou GPU
- SMP (Symmetric multiprocessing) vs MPP (massively parallel processing)









Top500

- Seznam nejvýkonnějších počítačů
- Měření výkonu ve Flops (floating point operations per second) pomocí LINPACK
 - teoretický výkon R_{peak} : počet jader * 4/frekvence
 - měřený výkon R_{max}

Top500 - Listopad 2009

1. Jaguar - Cray XT5-HE Opteron Six Core 2.6 GHz (Oak Ridge National Laboratory)
224162 jader, $R_{\max} = 1759$ TFlops, $R_{\text{peak}} = 2331$ TFlops, 7 MW
2. Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz, Voltaire Infiniband
122400 jader, $R_{\max} = 1042$ TFlops, $R_{\text{peak}} = 1376$ TFlops, 2.3 MW
410. iDataPlex, Xeon QC 2.5 GHz, GigEthernet / 2009 IBM
5376 jader, $R_{\max} = 23.19$ TFlops, $R_{\text{peak}} = 53.76$ TFlops, 173 kW
478. SGI Altix ICE 8200 Enhanced LX, Xeon X5560 quad core 2.8 GHz / 2009 SGI
2048 jader, $R_{\max} = 21.16$ TFlops, $R_{\text{peak}} = 22.94$ TFlops, 84 kW

Virtuální počítače a gridy

- BOINC - 5.128 PFlops (24.4. 2010)
- Folding@Home - 5 PFlops (17.3. 2009)
- WLCG – 100000 jader, asi 1 PFlops

Výpočetní středisko FZÚ

- Serverovna otevřena roku 2004
- 65 m², 18 skříní
- 200 kVA UPS, 380 kVA diesel, 120 kW vzduchové chlazení



Výpočetní středisko FZÚ - 2009



- 3000 jader (27 TFlops), 6 TB RAM, 300 TB disková kapacita, 150 kW spotřeba

IBM iDataPlex

1. systém (2008)

- 84 serverů = 168 procesorů Xeon E5440 2.83 GHz = 672 jader
- 1344 GB RAM
- $R_{\max} = 7.6 \text{ TFlops}$

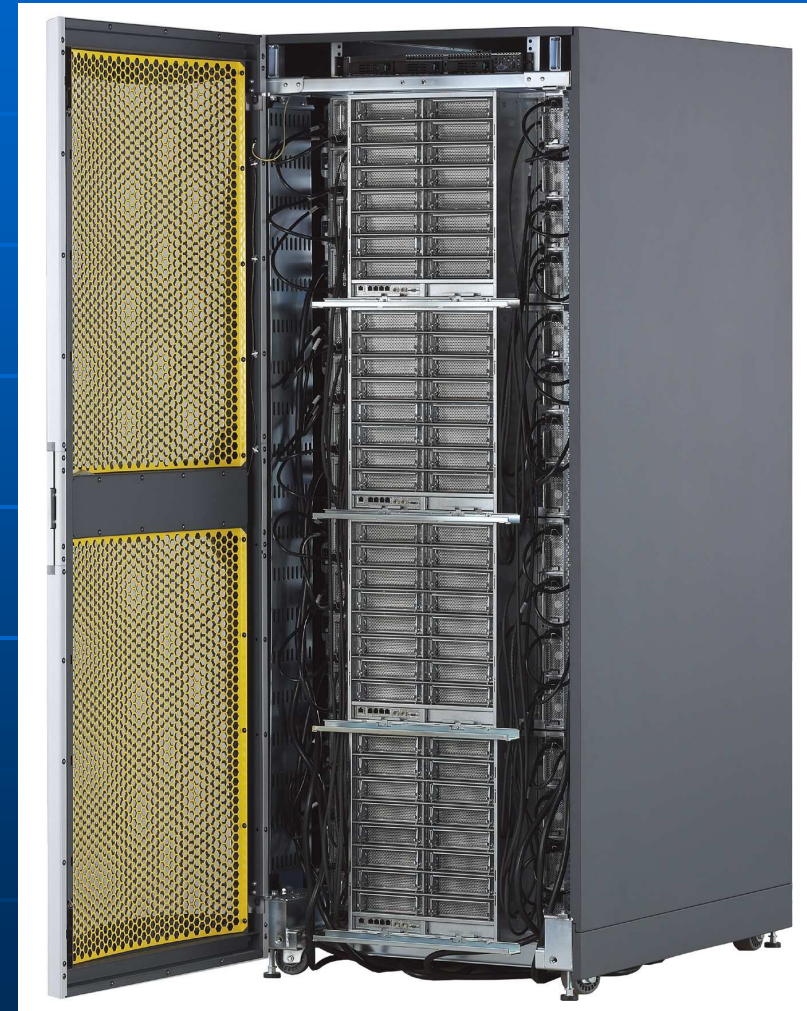
2. systém (2009)

- 65 serverů = 130 procesorů Xeon E5520 2.26 GHz = 520 jader (*2 s HT)
- 2080 GB RAM
- $R_{\max} = 4.7 \text{ TFlops}$



Dorje - Altix ICE 8200 ve FZÚ

- 64 výpočetních serverů = 128 procesorů Intel E5420 2.5GHz = **512 jader**
- RAM: 24 nebo 8 GB na server, **1 TB** celkem
- $R_{\text{peak}} = \mathbf{4.2 \text{ TFlops}}$
- $R_{\text{max}} = \mathbf{5.1 \text{ TFlops}}$
- Admin node, rack leader node, login node
- Diskové pole *SGI Infinite Storage 220*, 6 TB
- vodní chlazení
- SLES 10, SGI ProPack 5
- Torque/Maui, OpenMPI, MKL knihovny, Fortran 90/95, C/C++



VS FZÚ

Počet strojů	Typ	Procesor (počet jader)	Výkon skupiny strojů (GFLOPS)	RAM skupiny strojů (GB)
26	HP BL35p	2x Opteron 280 2,4GHz (4)	998	208
10	HP BL35p	2x Opteron 275 2,2GHz (4)	352	80
6	HP BL20p	2x Xeon 5160 3GHz (4)	288	60
8	HP BL460c	2x Xeon 5160 3GHz (4)	384	64
10	HP BL460c	2x Xeon 5420 2,5GHz (8)	800	160
12	HP BL465c	2x Opteron 2220 2,8GHz (4)	538	96
52	HP DL140	2x Xeon 3,06GHz (2)	1273	104
15	HP DL140	2x Xeon 3,06GHz (2)	367	60
40	SGI Altix XE 310	2x Xeon E5420 2,5 GHz (8)	3200	640
84	IBM DX340	2x Xeon E5440 2,83 GHz (8)	7607	1344
18	SGI Altix XE 340	2x Xeon E5520 2.26Ghz (8)	1300	288
65	IBM DX360	2x Xeon E5520 2.26Ghz (8)	4700	2080
64	Altix ICE 8200	2x Xeon E5420 2.50 GHz (8)	5120	1024

410

26927

6208

Další (mě) známé superpočítače v ČR

Současná konfigurace superpočítače Amálka (ÚFA AV ČR):

Typ	Počet procesorů	Počet jader	Výkon GFlops	RAM (GB)
Intel Pentium III Xeon 700	12	12	8	48
Intel® Xeon® Processor 2.80 GHz	140	140	392	70
Intel® Xeon® Processor E5140	84	168	1 568	84
Intel® Xeon® Processor E5345	56	224	2 090	112
Intel® Xeon® Processor L5520	64	256	2 324	256
CELKEM	356	800	6 383	570

ČHMÚ

nový výkonný počítač specializovaný na předpověď počasí
NEC SX-9 se špičkovým výkonem **1 600 GFlops**, 1 TB RAM,
diskové pole NEC 100 TB.

LUNA (FZÚ AV ČR)

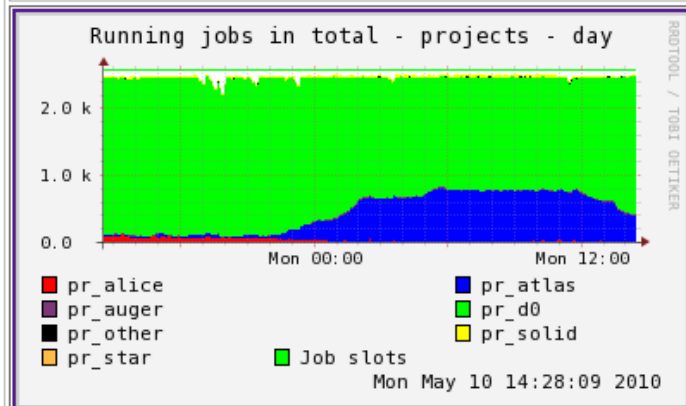
2x SUN Fire X4600, 64 jader, **512 GB RAM**,
4x SUN Fire v20z, 1x SUN Fire v40z
710 GFlops

Využití

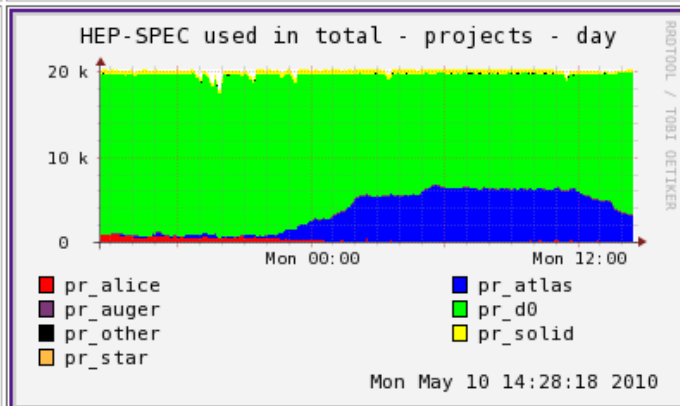
- HEP projekty:
 - D0 experiment, FERMILAB (USA)
 - LHC experimenty ATLAS, ALICE; CERN
 - STAR experiment, BNL (USA)
- Fyzika pevných látek
- Astročásticová fyzika
 - Pierre Auger Observatory, Argentina

[Service Availability](#)[Siteview](#)[GridMap](#)[Projects](#)[Queues](#)[Nodes](#)

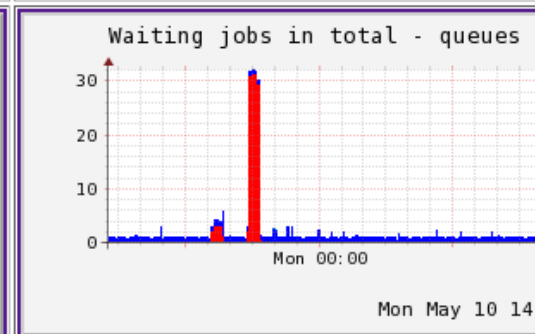
Projects Running



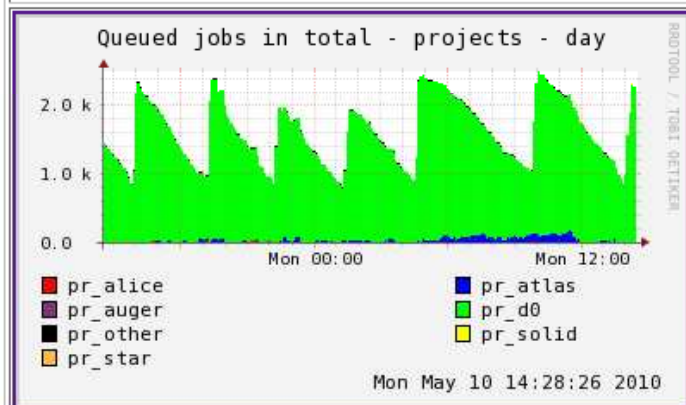
Projects HEP-SPEC



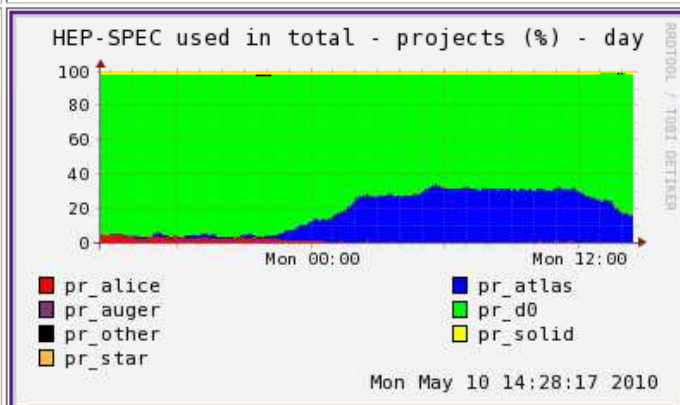
Waiting Jobs



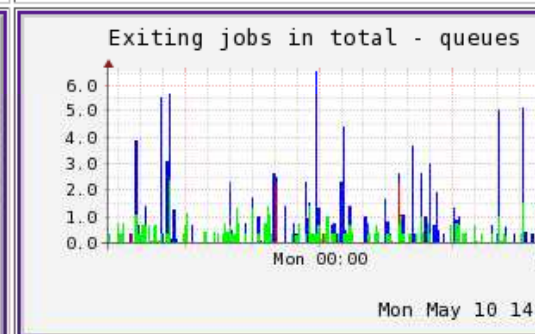
Projects Queued



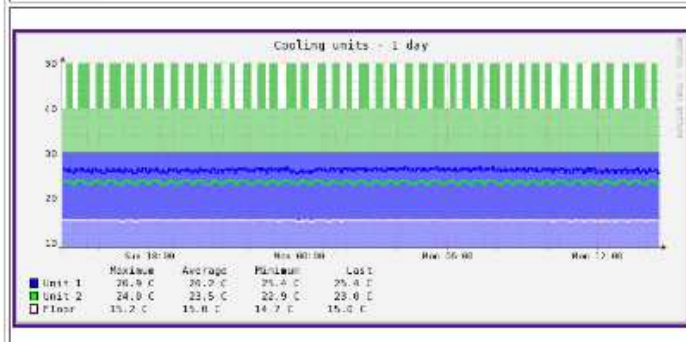
Projects HEP-SPEC Percentage



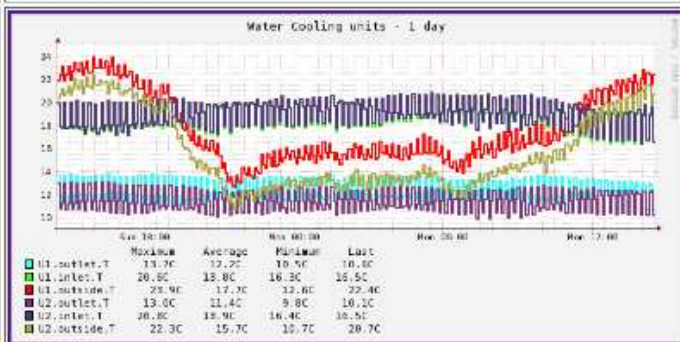
Exiting Jobs



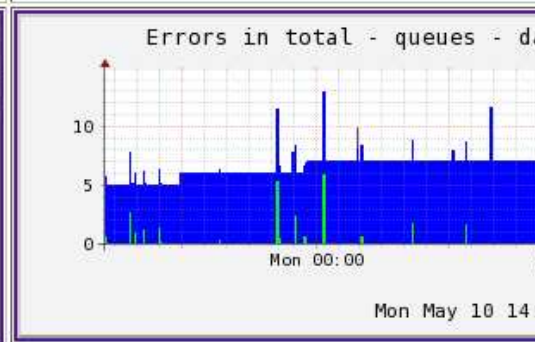
Cooling



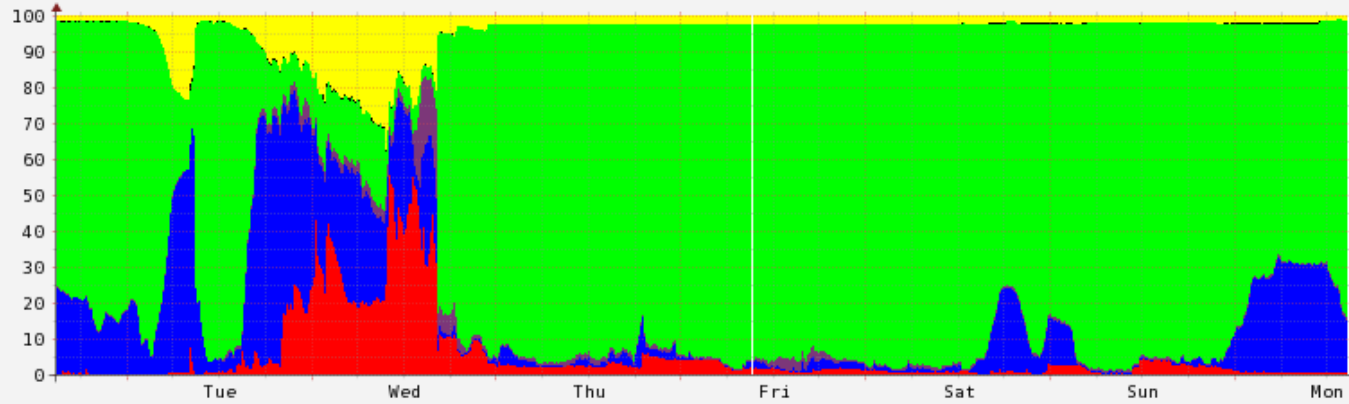
Aqua Cooling



Errors



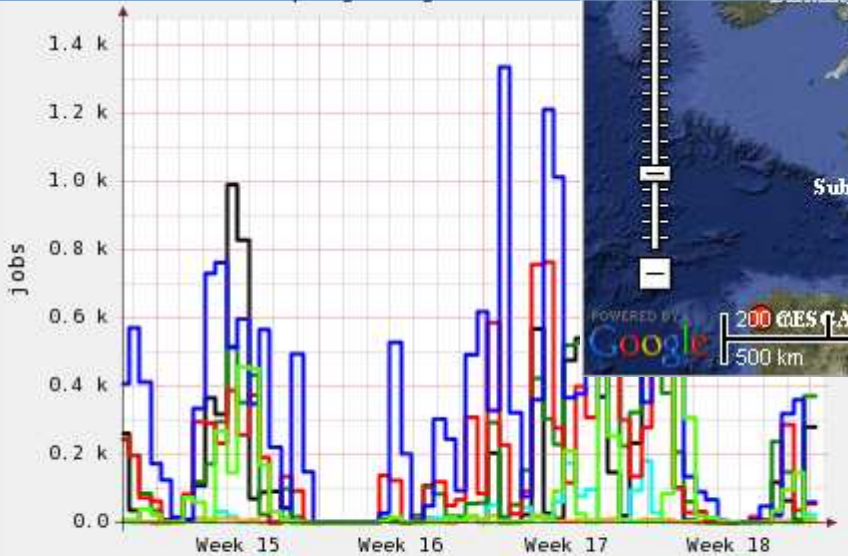
HEP-SPEC used in torque - projects (%) - week



ROOTOOL / TORI OETIKER

pr_alice	0
pr_atlas	15
pr_auger	1
pr_d0	83
pr_other	0
pr_solid	1
pr_star	0

prague1cg2 - month



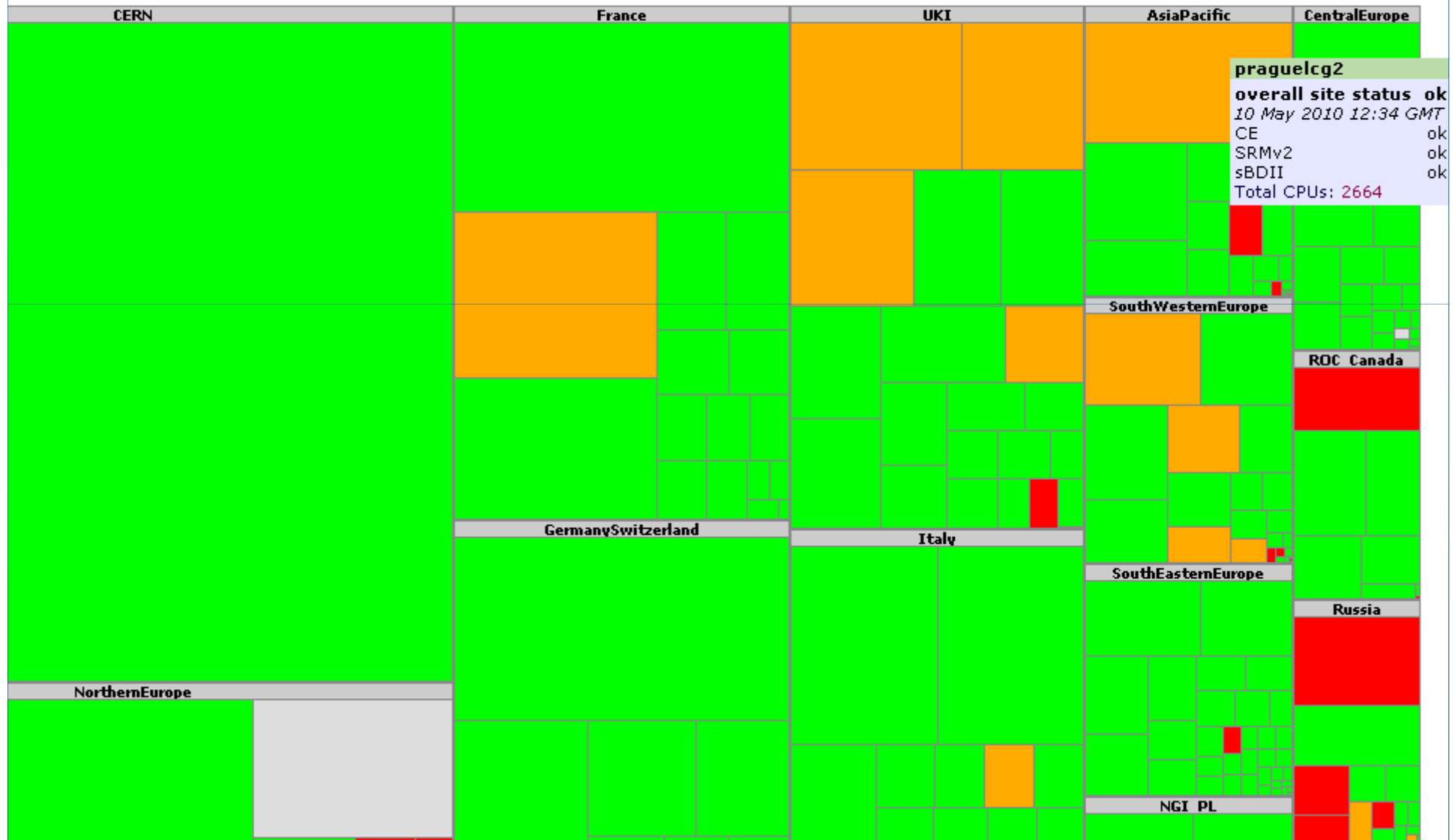
activated	assigned	holding	running
transferring	finished (12hrs)	failed (12hrs)	

Generated by TRIUMF-LCG2 (times in UTC)



Imagery ©2010 TerraMetrics - Terms of Use

Visualizing the "State" of the Grid



GridMap – Visualizing the "State" of the Grid

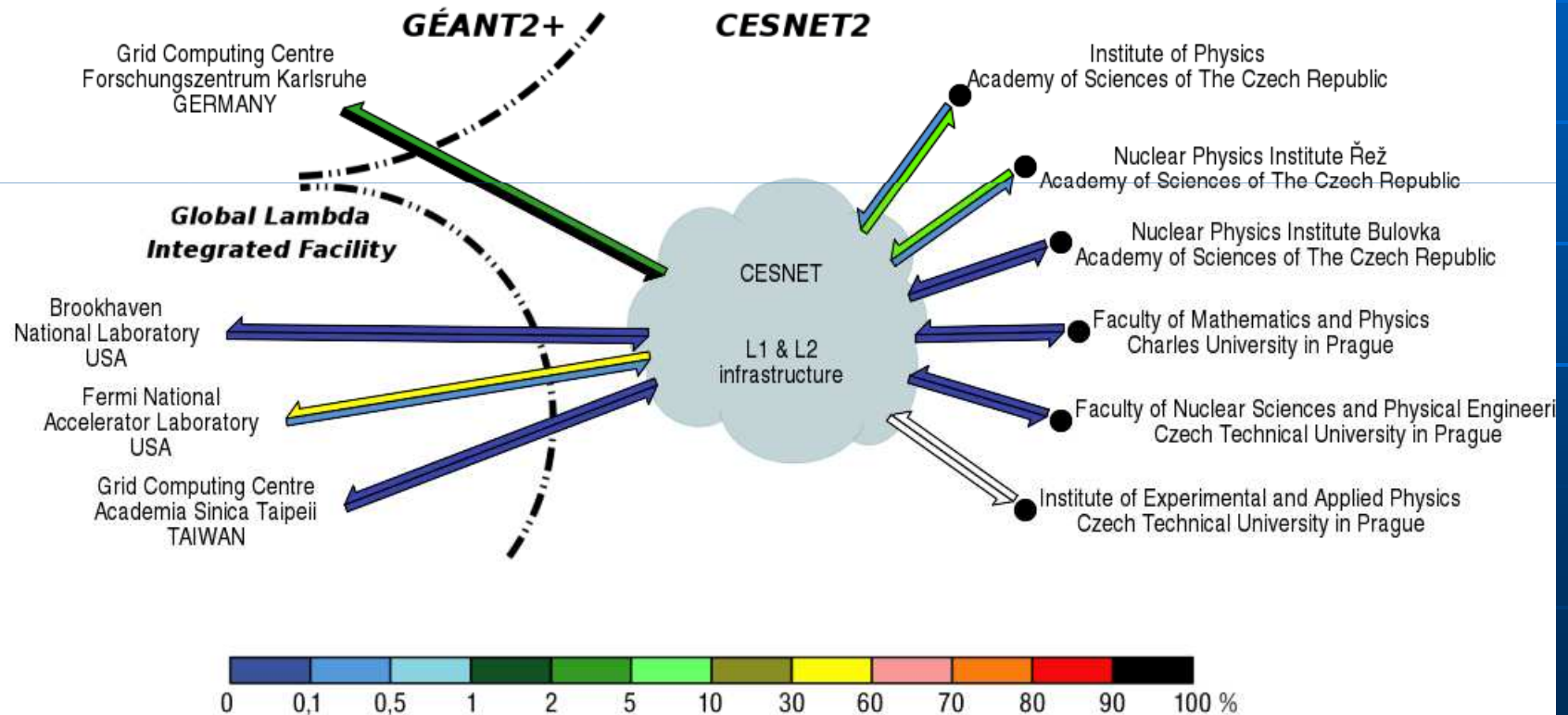


Sítové připojení

CESNET2: experimental facility for High Energy Physics

Average utilization for period: 2010/05/10 15:46:28 - 2010/05/10 16:46:28 MEST (Europe/Prague), GMT+2H

The map shows the actual utilization of the experimental facility for High Energy Physics in the CESNET2 network. The presented links are coloured by average load. Click on the line to obtain detailed statistics.

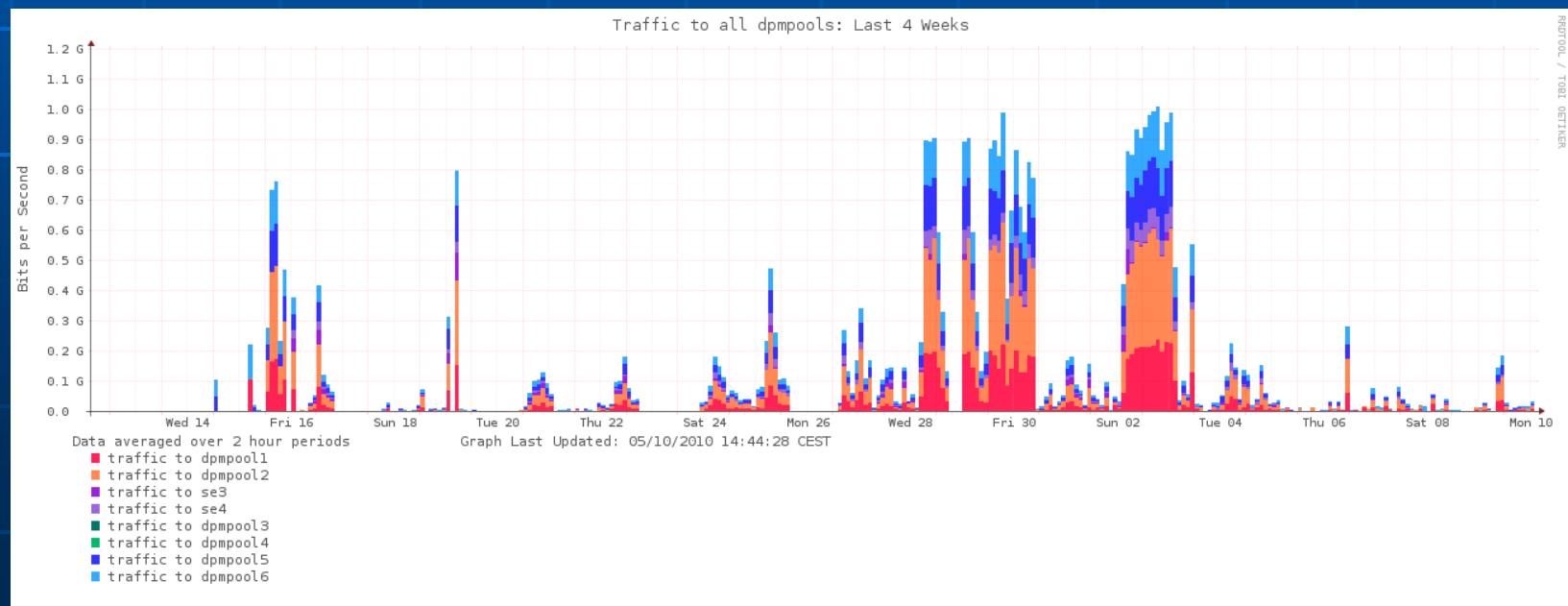


Other Information

Sítové připojení

The following table shows values of other relevant parameters measured at active network devices during the period: **2010/05/10 15:54:47 - 2010/05/10 16:54:47 MEST (Europe/Prague), GMT+2H**.

Link	Utilization (avr)	Capacity
Grid Computing Centre Forschungszentrum Karlsruhe GERMANY->CESNET	98.685 %	1000.000 Mbps
CESNET->Fermi National Accelerator Laboratory USA	35.146 %	1000.000 Mbps
CESNET->Institute of Physics Academy of Sciences of The Czech Republic	8.974 %	11.000 Gbps
Nuclear Physics Institute Řež Academy of Sciences of The Czech Republic->CESNET	7.534 %	1000.000 Mbps
CESNET->Grid Computing Centre Forschungszentrum Karlsruhe GERMANY	4.557 %	1000.000 Mbps
Fermi National Accelerator Laboratory USA->CESNET	0.414 %	1000.000 Mbps
Institute of Physics Academy of Sciences of The Czech Republic->CESNET	0.394 %	11.000 Gbps



WLCG

- LCG = LHC Computing Grid
 - úzce propojeno s evropskými projekty DataGrid, EGEE
- Propojit distribuované zdroje do zdánlivě jednoho klastru
- Virtuální organizace
 - uživatelský interface, certifikát
- Memorandum of Understanding



CERN



US-BNL



Amsterdam/NIKHEF-SARA



rtolucci, CERN



Bologna/CNAF



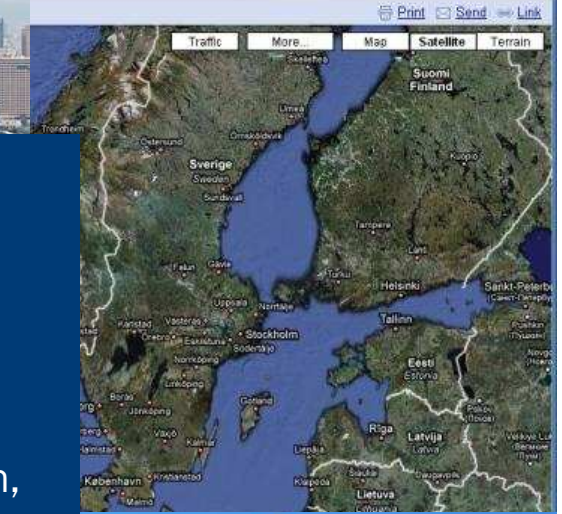
Ca-
TRIUMF

WLCG

Tier 0; 11 Tier 1; 61 Tier 2 federací (121 Tier 2 středisek)

Dodnes podepsalo MoU 49 institucí ze 34 zemí:

Australia, Austria, Belgium, Brazil, Canada, China, **Czech Rep**, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, India, Israel, Japan, Rep. Korea, Netherlands, Norway, Pakistan, Poland, Portugal, Romania, Russia, Slovenia, Spain, Sweden, Switzerland, Taipei, Turkey, UK, Ukraine, USA.



US-FNAL



De-FZK



Barcelona/PIC



Lyon/CCIN2P3



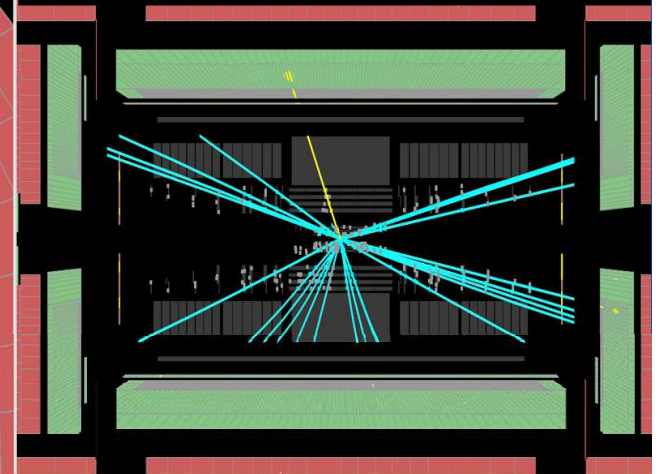
UK-RAL

Gridy fungují!

 **ATLAS**
EXPERIMENT

Run Number: 152409, Event Number: 596680

Date: 2010-04-05 06:54:50 CEST



**W \rightarrow ev candidate in
7 TeV collisions**

$$p_T(e^+) = 34 \text{ GeV}$$

$$\eta(e^+) = -0.42$$

$$E_T^{\text{miss}} = 26 \text{ GeV}$$

$$M_T = 57 \text{ GeV}$$

