



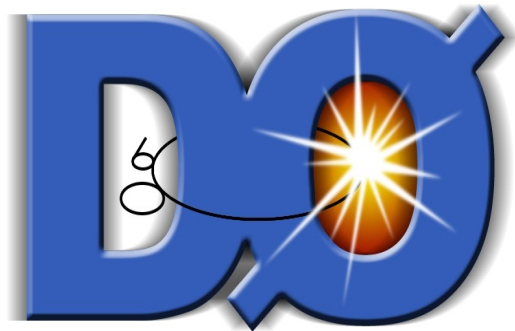
Prague Tier 2 monitoring progress

T. Kouba

Institute of Physics, ASCR
15th September 2011, Varna, Bulgaria

Overview - users

- Provides computing power and storage for
 - samgrid (D0 experiment)
 - WLCG (ATLAS, ALICE)
 - EGI (Pierre Auger observatory, Cherenkov Telescope array)
 - Local users (solid state physics)



Overview - Hardware and network

- The site is equipped with:
 - UPS, Air cooling, Water cooling
 - Neo8000 Tape library (100 TB)
 - 2PB on disk servers (DPM or NFSv3)
 - ~3500 cores on WNs, ~25000 HEPSpec
 - 10Gb backbone, 10Gb+9x1Gb uplinks

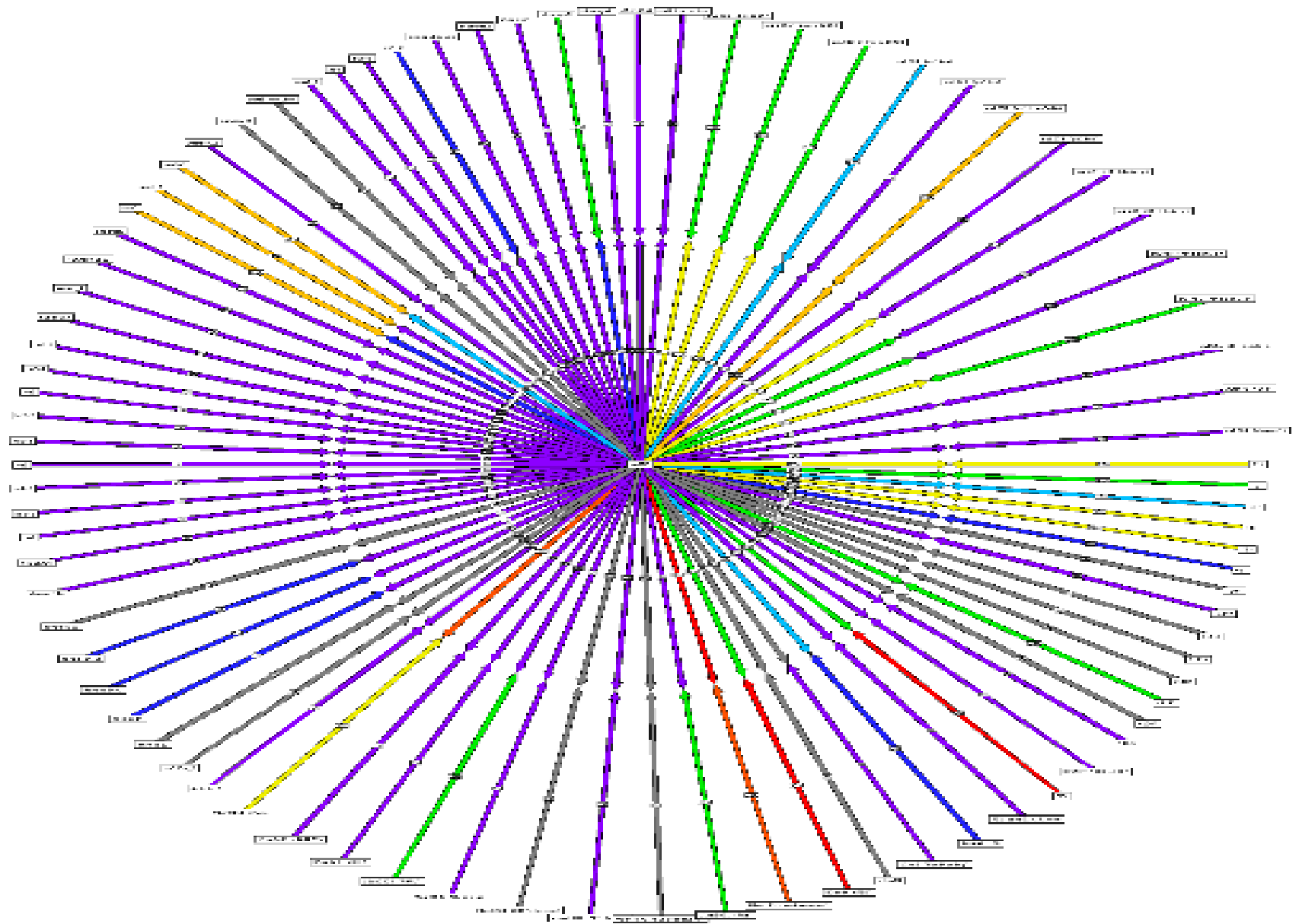
Overview - Services

- Torque + MAUI server
- UI, MonBox, Icg-CE, creamCE
- SE: DPM head with 15 DPM disk nodes
 - Everything configured by CFengine
- Alice VOBOX
- Two SAM stations
- Auger dashboard
 - Managed by VO experts

Local network

- 10Gb backbone
- Visualization by Weathermap
- MRTG – graphs of switches load
- Transfer graphs by ipac-ng, munin plugin
 - iptables based
- Netflow
 - ipt-netflow kernel module used on storage nodes to generate netflow data
 - Transfers graphed by flowtracker, accounted by site_stat

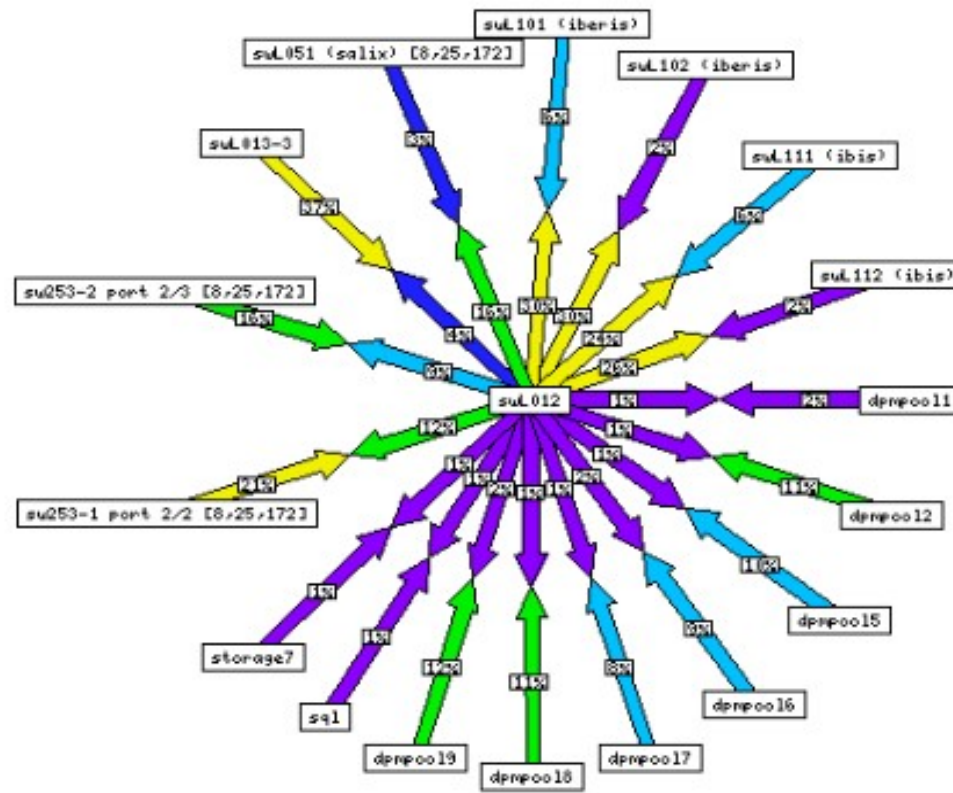
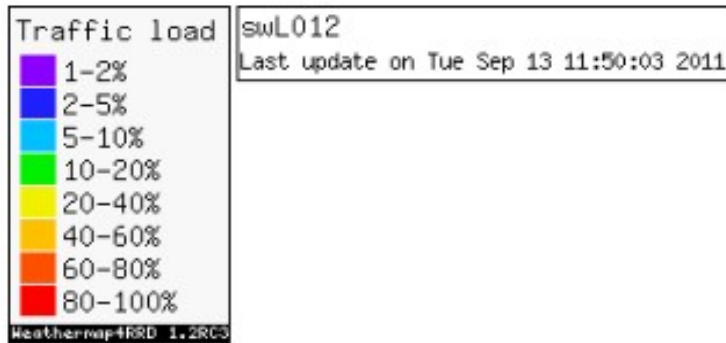
Local network – weather map, main router



2011-09-15

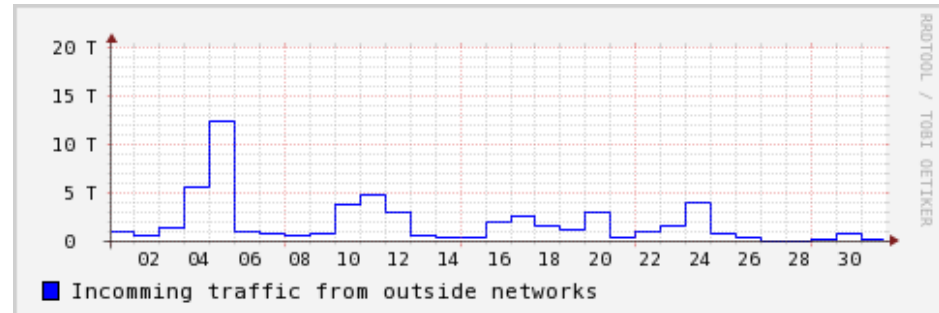
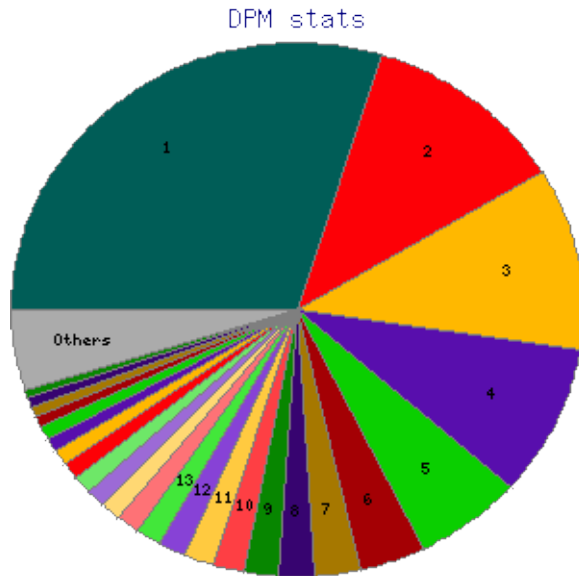
Prague Tier 2 monitoring progress

Local network – weathermap, 10Gb switch



Local network - site_stat

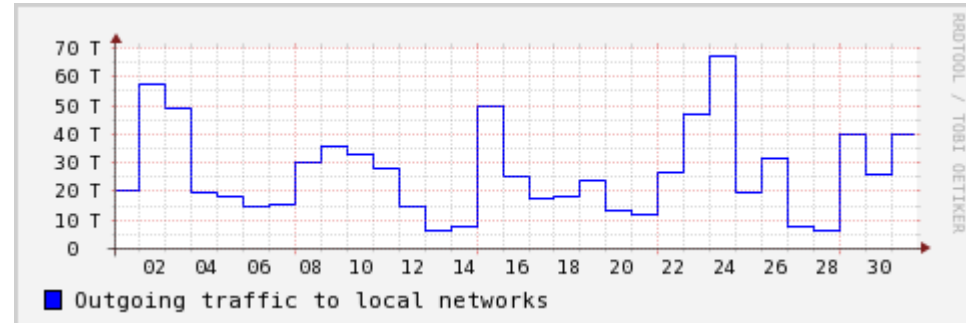
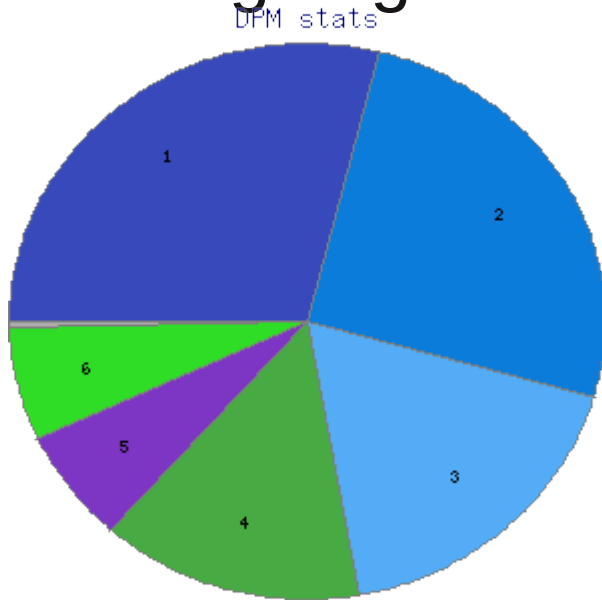
- Incoming traffic from outside networks (Aug 2011)



1.	192.108.45.0 - 192.108.47.255	16.85 TB	29.70%	Forschungszentrum Karlsruhe (FZK)
2.	129.187/16	6.63 TB	11.69%	Leibniz-Rechenzentrum (LRZ)
3.	134.158/16	6.25 TB	11.02%	Institut National de Physique Nucleaire et de Physique des Particules
...				
Total.		56.74TB		

Local network - site_stat

- Outgoing traffic to local subnets (Aug 2011)

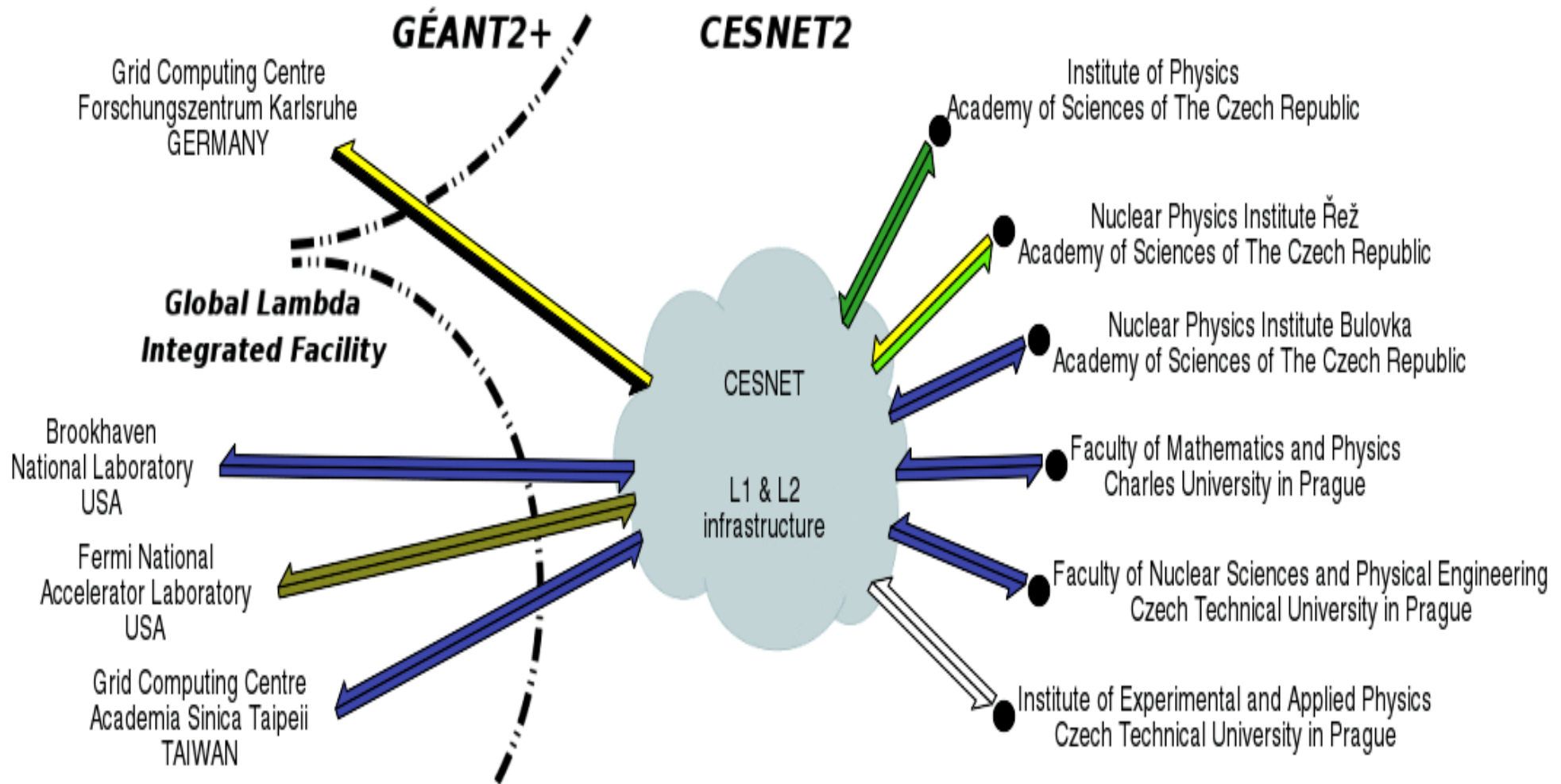


1.	172.16.2/24	231.10 TB	28.83%	subcluster iberis
2.	172.16.6/24	209.22 TB	25.55%	subcluster ibis
3.	172.16.7/24	146.24 TB	17.86%	subcluster ib
...				
Total.		818.93TB		

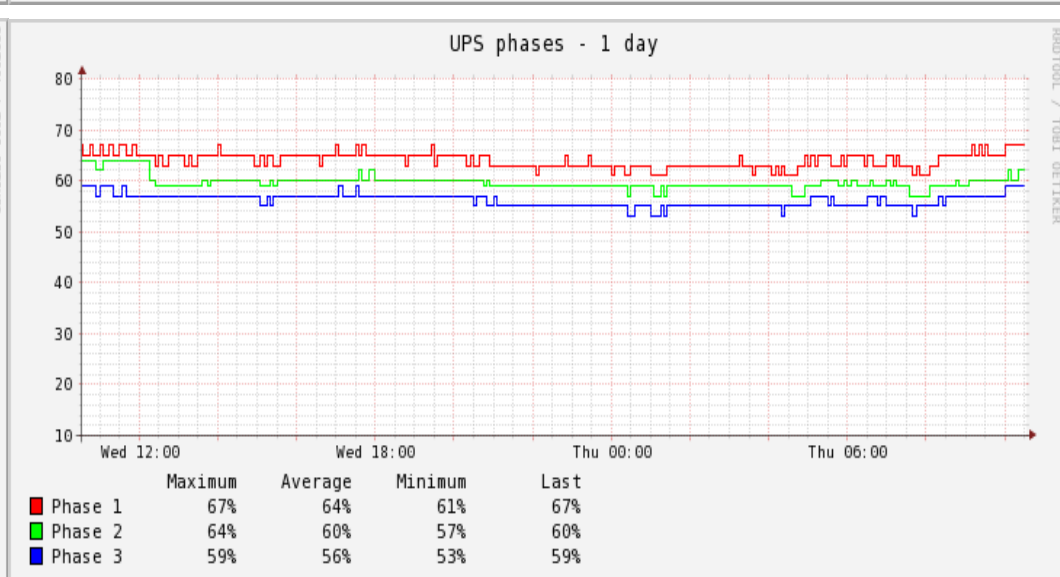
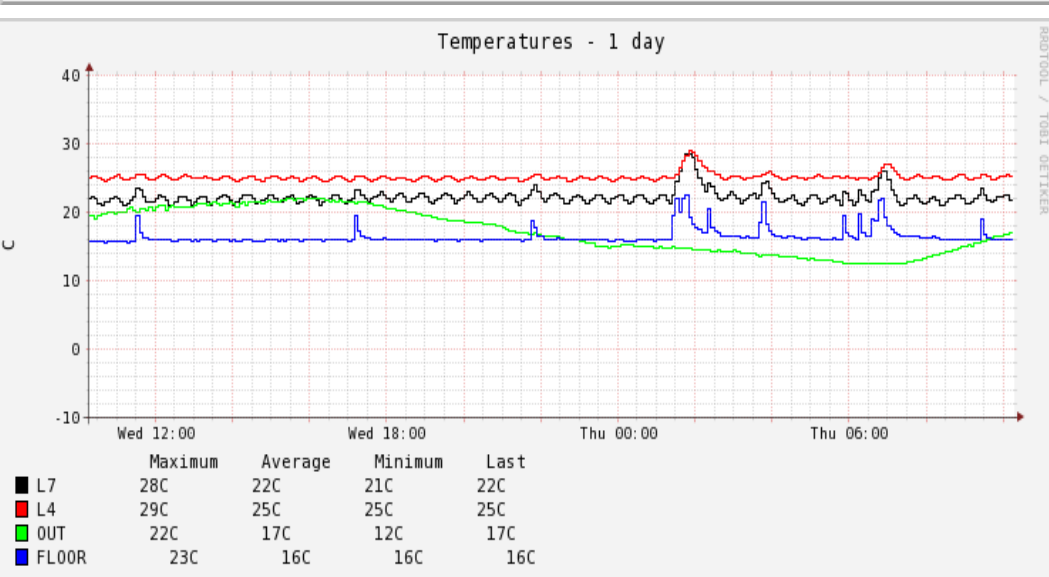
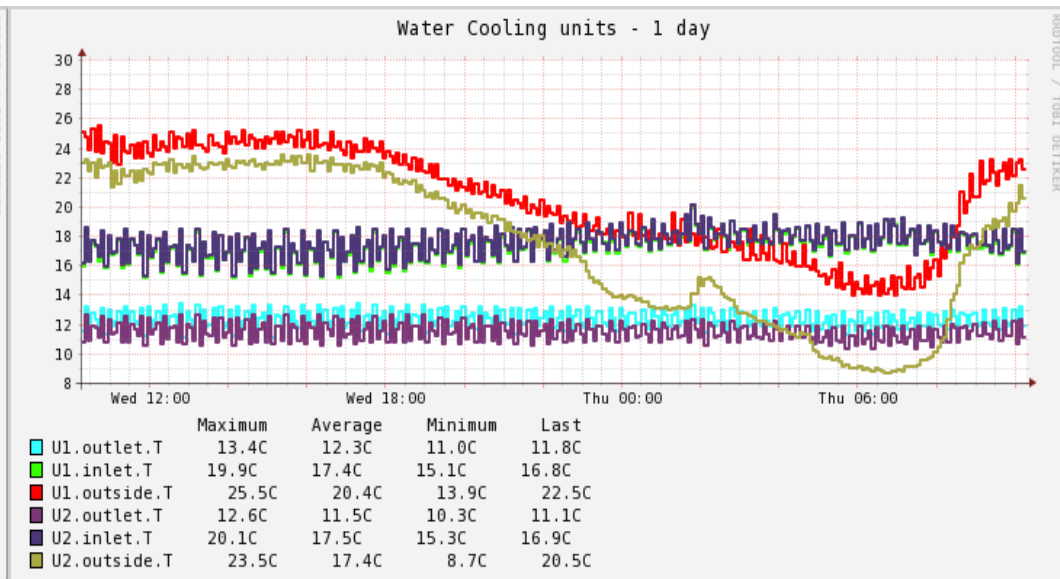
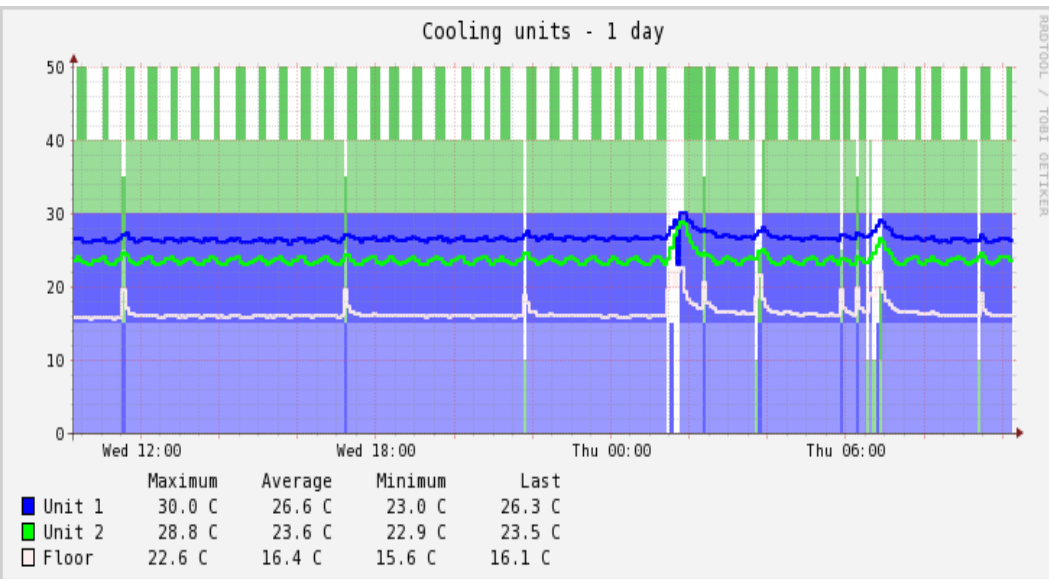
External network connectivity

- 10Gb connection to the Internet
- Dedicated λ (wavelength) with 1Gb bandwidth
 - 1Gb connection to FZK (ATLAS DE cloud)
- Dedicated λ with 10Gb bandwidth
 - 1Gb connection to BNL
 - 1Gb to FNAL
 - 1Gb to ASGC
 - Several 1Gb connections to Czech HEP institutes
- All the connections are provided by Cesnet (Czech NREN)
- Dedicated connections are monitored by G3 system reporter (product of Cesnet)

Prague Tier 2 – prague1cg2, network



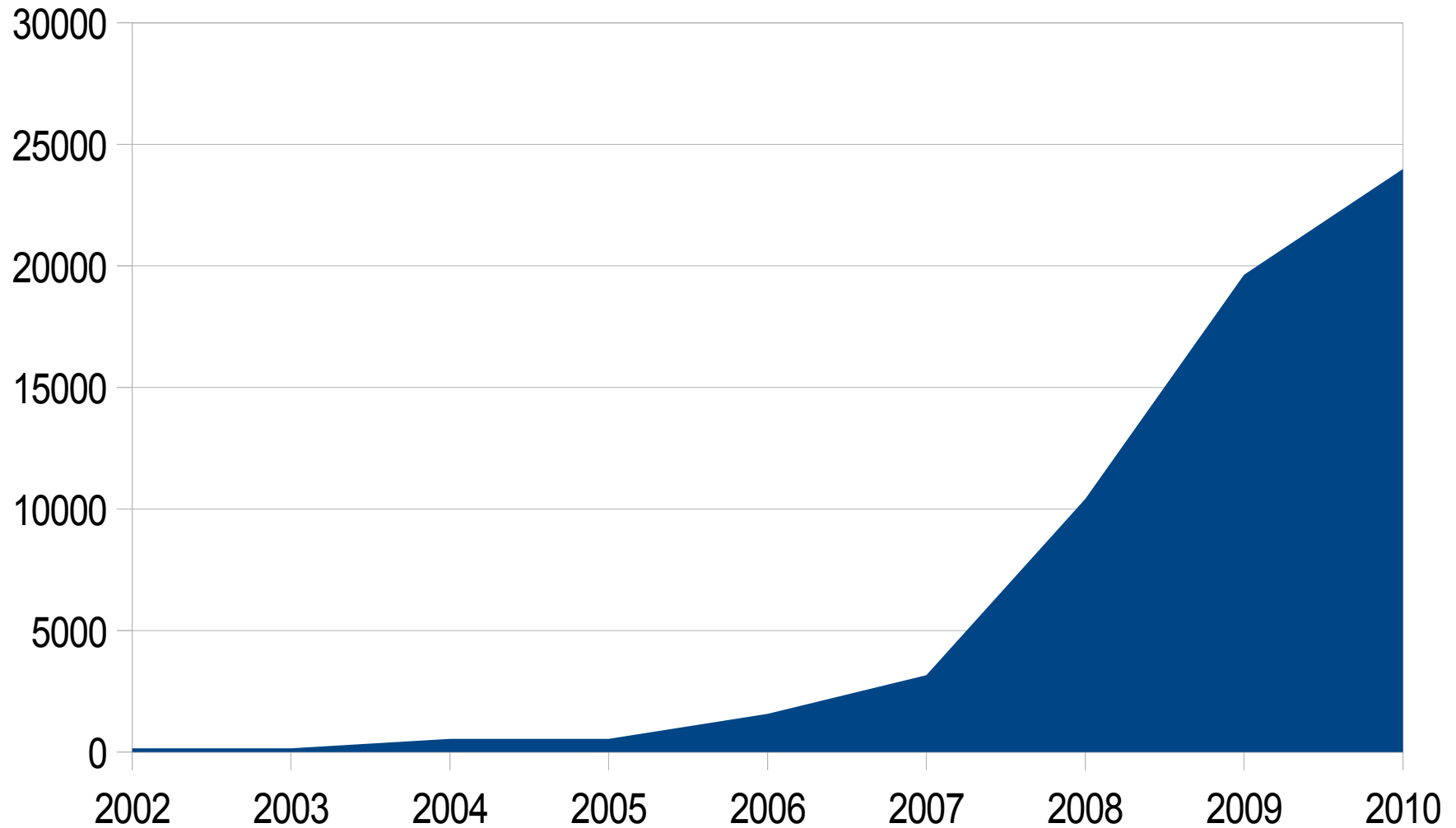
RRD graphs for UPS, Cooling, temperature



Worker nodes

- Several subclusters with homogenous HW
 - goliath (various HP blade servers) 66 nodes, 2374 HS
 - salix (SGI Altix xe310) 40 nodes, 2698 HS
 - iberis (IBM dx340) 76 nodes, 5302 HS
 - saltix (SGI Altix xe340) 18 nodes, 1710 HS
 - ibis (IBM dx360) 65 nodes, 7664 HS
 - ib (IBM dx360m3) 26 nodes, 5321 HS
- 291 nodes, 25069 HS in total


Prague Tier2 HEPSpec growth



Health monitoring - Nagios

- Nagios with basic sensors
 - load, memory consumption, running processes, ldap bindable, ssh listening, torque mom listening
- Many in house developed sensors
 - pakiti, cfagent work complete, yum finished, ghost jobs on torque server, certificate lifetime
- nsca for reporting syslog status
 - e.g. disk failure → system log → central syslog server → nsca on syslog server pushes the status into nagios
- livestatus from check_mk suite
- multiview interface from check_mk suite

Multiview web interface

multisite
check  v1.1.9i3

Tactical Overview ×

Hosts	Problems	Unloaded
467	0	0
Services	Problems	Unloaded
6358	20	20

Quicksearch ×

Views ×

- Hosts**
 - All hosts
 - All hosts (Mini)
 - All hosts (tiled)
 - Host search
- Hostgroups**
 - Hostgroups
 - Hostgroups (Grid)
 - Hostgroups (Summary)
- Services**
 - All services
 - Recently changed services
 - Serv. by host groups
 - Service search
- Servicegroups**
 - Servicegroups (Grid)
 - Servicegroups (Summary)
 - Services by group
- Problems**
 - Alert Statistics
 - Host problems
 - Pending Services
 - Service problems
 - Unchecked services
- Addons**
 - Search PNP graphs
- Other**
 - Comments
 - Downtimes
 - Host- and Service events
 - Search Global Logfile

EDIT

FZU

Name	Alias	Up	Dw	Un	Pd	O	W	C	U	P
afs_client	AFS client on Golias farm	18	0	0	0	270	0	0	0	18
alice_vobox	VOBOX service for alice	2	0	0	0	27	1	1	0	1
backup_server	Backup head node	1	0	0	0	16	0	0	0	0
bind	Bind server	2	0	0	0	25	0	0	0	2
bonding	Nodes with bonding configured	5	0	0	0	72	0	1	0	5
creamce	CREAM in Golias farm	2	0	0	0	30	0	0	0	1
dhcpd	DHCPD service	2	0	0	0	25	0	0	0	2
dpm_head	DPM head nodes	2	0	0	0	42	2	0	0	2
dpm_pool	DPM pool (disk) nodes	14	2	0	0	235	1	39	0	11
edge_cesnet_cz	Hosts in domain farm.particle.cz	31	0	0	0	400	4	15	0	25
farm_particle_cz	Hosts in domain farm.particle.cz	337	6	0	0	5067	6	127	1	317
fc_switch	FC switches	6	0	0	0	20	0	0	0	8
frontier_squid	Squid configured as frontier atlas service	1	0	0	0	13	0	0	0	1
ganglia_aggregators	Nodes with running ganglia aggregation daemons	17	1	0	0	261	1	17	0	16
gentoo	Hosts with running gentoo	3	0	0	0	32	0	0	0	3
glite_wms	gLite WMS	2	0	0	0	31	0	0	0	1
golias-monitor-eva	Golias monitoring eva box	1	0	0	0	1	0	0	0	1
golias-monitor-satabeast	Golias monitoring satabeast box	3	0	0	0	6	0	0	0	0
golias-monitor-ultamus	Golias monitoring ultamus box	5	1	0	0	10	0	1	0	1
golias-monitor-vtrakM610p	Golias monitoring vtrak box	1	0	0	0	2	0	0	0	0
golias-storage-box	Golias monitoring storage HW box	10	1	0	0	19	0	1	0	2
golias_all	All nodes on Golias farm	328	6	0	0	4977	7	120	1	307
golias_ce	Computing elements in Golias farm	1	0	0	0	16	0	0	0	0
golias_core	Core servers on Golias farm	41	3	0	0	619	5	58	1	33
golias_dom0	Golias Xen hosts with VMs	3	0	0	0	41	0	1	0	3
golias_ui	User interface nodes for golias farm	1	0	0	0	14	0	0	0	0
golias_wn	All worker nodes on golias farm	284	3	0	0	4317	1	61	0	273
ha	Xen nodes with shared underlying storage	3	0	0	0	41	0	1	0	3
has_cert_golias	Nodes requiring server certificate on Golias farm	26	2	0	0	424	3	41	0	19
has_cert_skurut	Nodes requiring server certificate on skurut farm	15	0	0	0	215	4	2	0	10
hw_3ware_raid	Hardware with 3ware raid controller	2	0	0	0	25	0	1	0	2
hw_areca_controller	Areca controllers	24	0	0	0	70	0	1	0	1
hw_bl20p	HP blade servers	6	0	0	0	96	0	0	0	6
hw_bl35p_275	HP blade servers	9	0	0	0	135	0	0	0	9
hw_bl35p_280	HP blade servers	19	0	0	0	286	0	0	0	18

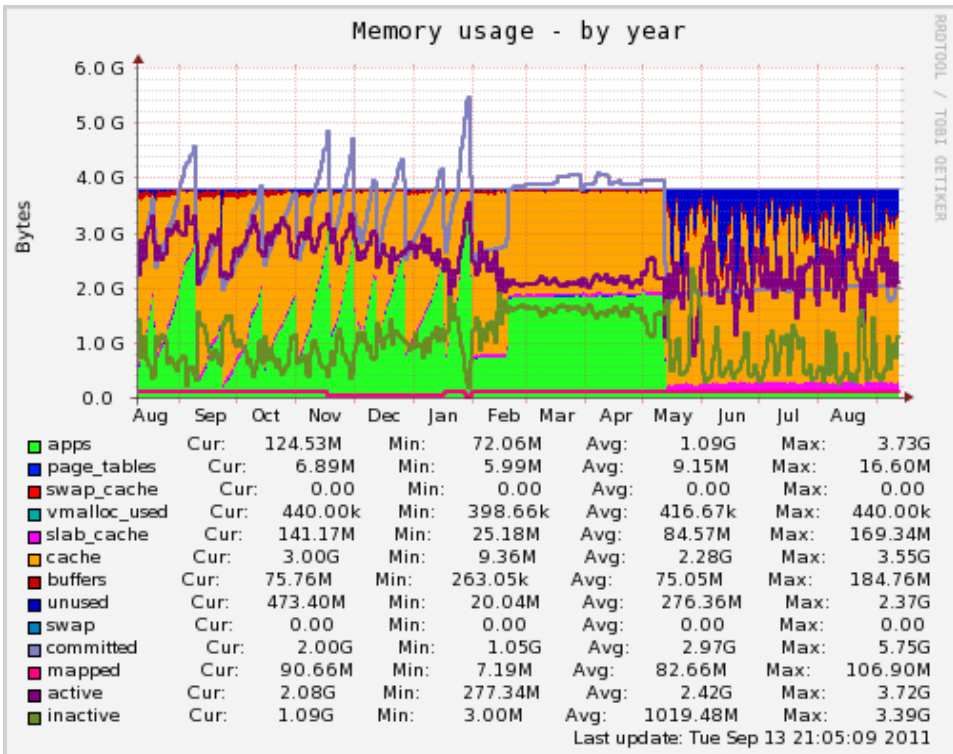
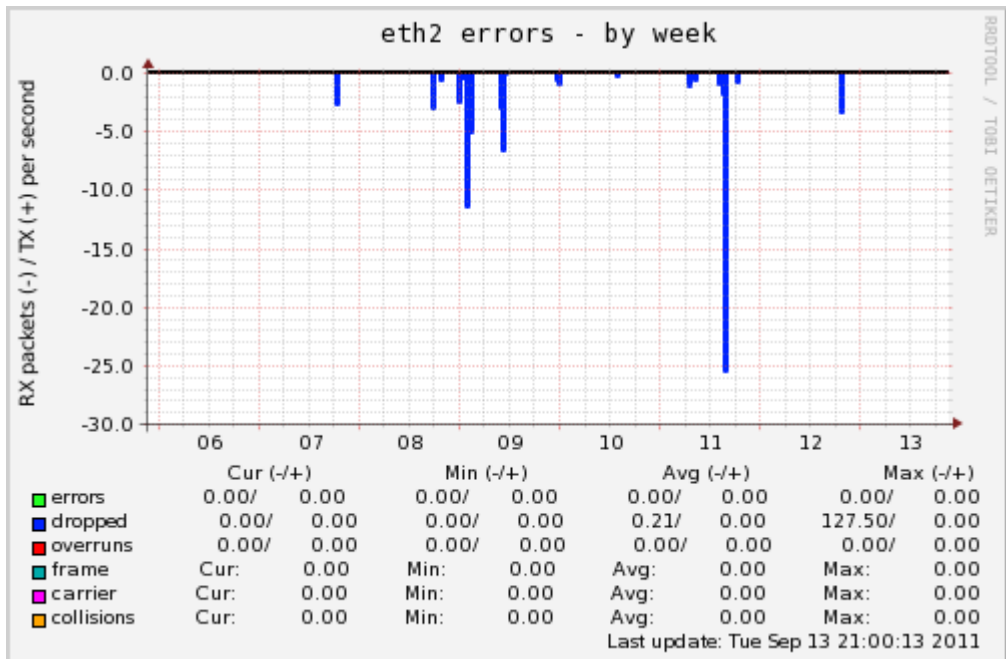
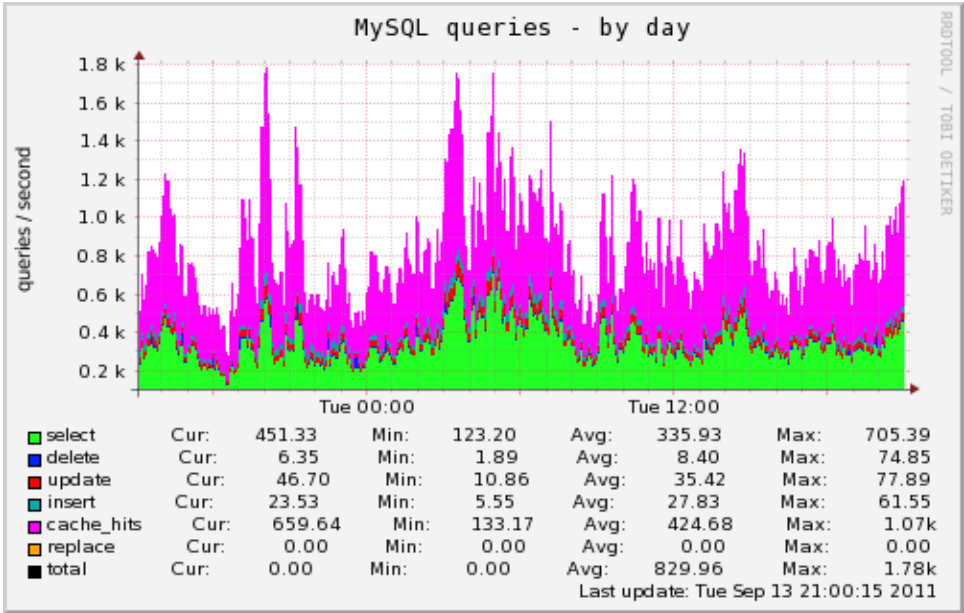
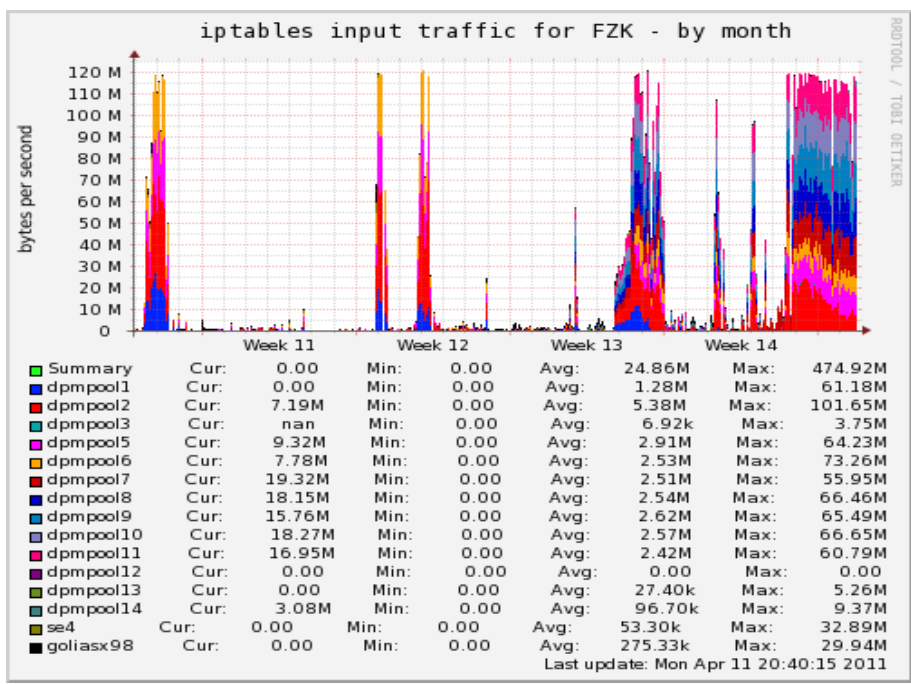
Multiview web interface

- Advantages
 - Easy massive downtime definition
 - Supports Nagios and Icinga
 - Custom views + Bookmarks
 - Integration with pnp4nagios
 - Better searching than default Nagios web
 - Active development
- Disadvantages
 - Not prepackaged in most linux distributions

Trends and graphs

- Nagios uses pnp4nagios
 - not flexible enough
- Munin
 - uses RRD internally
 - very simple to setup
 - server launched by cron, asking for data from clients
 - lightweight
 - sensors automatically configure graphing
 - we use it to monitor trends
 - data transfers, errors on interfaces, mysql performance, disk and memory consumption



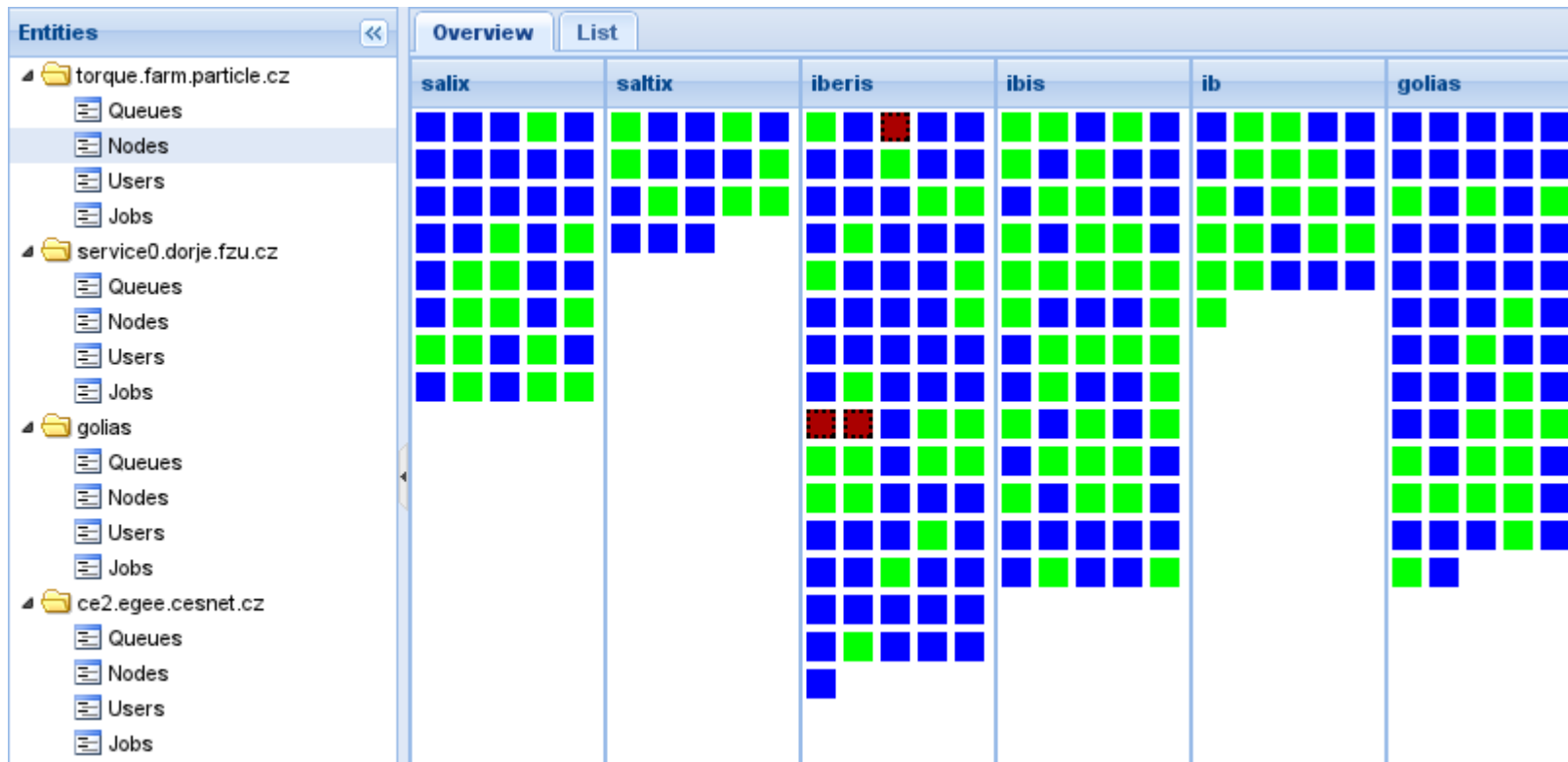


Torque status and accounting

- In-house developed “goove” (goliath overview)
- Daemon parsing accounting logs, putting data into mysql database
- Django python framework is used to query the database
- Ext 4 JS framework is used for presenting the data
- We plan to parse DPM logs in a similar way

Goove screenshots

Entities		List	Graphs						
		Name	Total jobs	Queued	Waiting	Running	Max. running	Started	Enabled
torque.farm.particle.cz	Queues	d0prod	2705	1653	0	1051	0	true	true
	Nodes	gridalice	1014	139	0	875	0	true	true
	Users	gridatlas	1232	425	5	802	0	true	true
	Jobs	gridauger	110	30	0	80	0	true	true
service0.dorje.fzu.cz	Queues	d0merge	8	0	0	8	0	true	true
	Nodes	solid	4	0	0	4	0	true	true
	Users	iatlas	3	0	0	3	0	true	true
	Jobs								



благодаря ви за вниманието
Q/A ?