ROMANIAN SITES REPORT
RO-02-NIPNE, RO-14-ITIM, RO-16-UAIC

Gabriel Stoicea

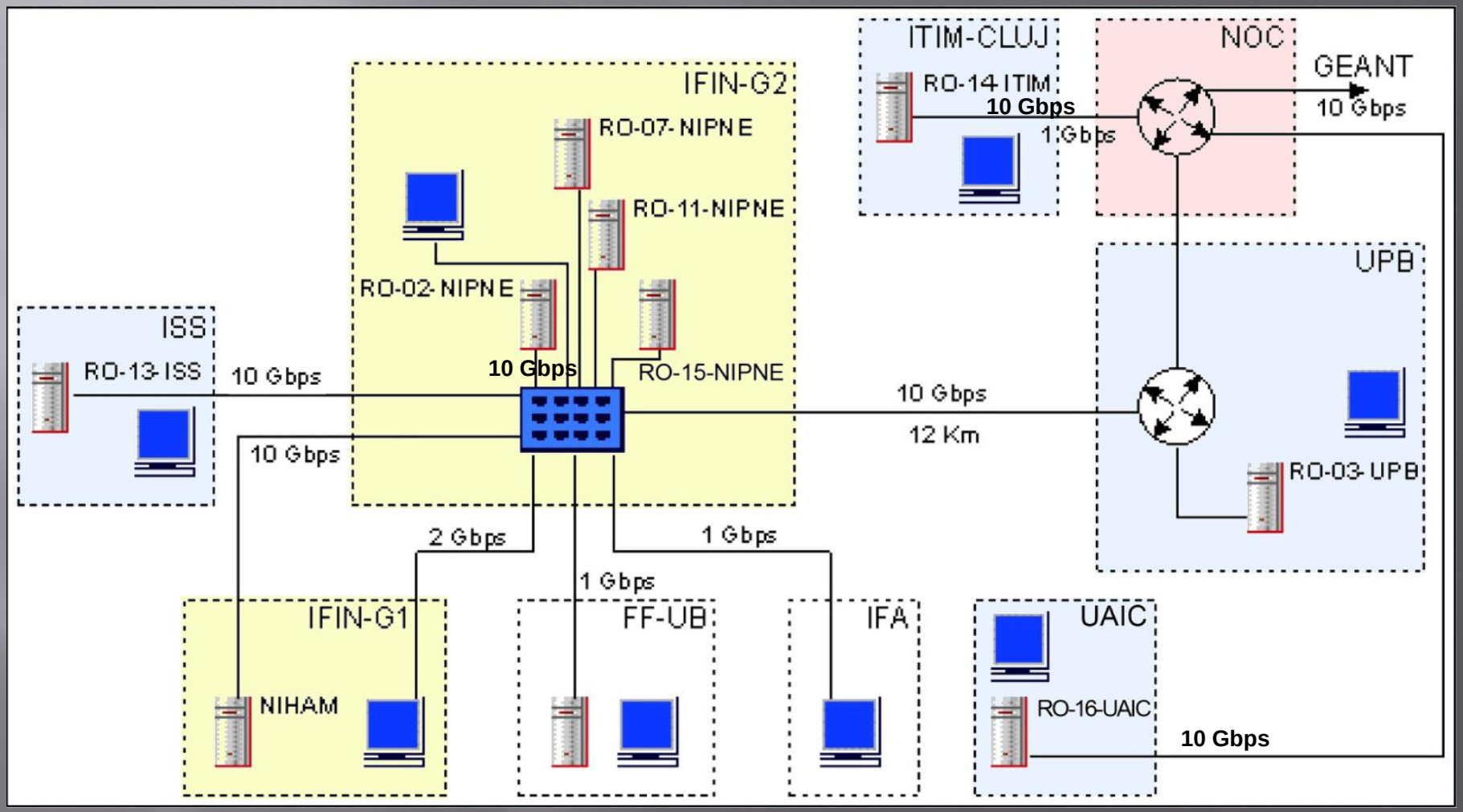
Particle Physics Department
National Institute of Physics and Nuclear Engineering "Horia Hulubei"
IFIN-HH

FR-cloud Regional Centers Meeting, LPNHE-Paris, 18.06.2012

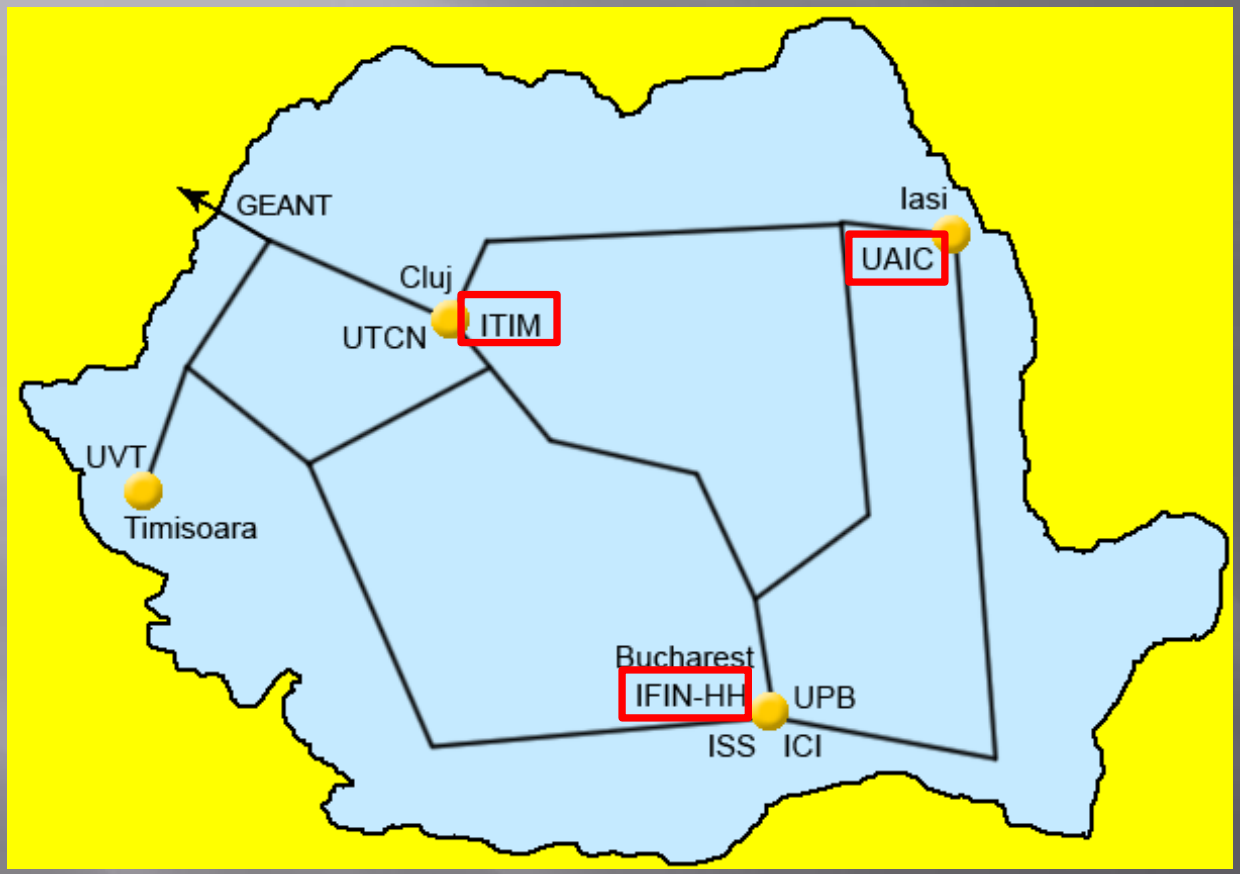
Outline

- RO-LCG Grid network infrastructure
- ATLAS Grid Infrastructure in Romania
- RO-14-ITIM and RO-16-UAIC Grid sites
- RO-02-NIPNE Grid site & Local Analysis Cluster – BAAF (Bucharest ATLAS Analysis Facility)

RO-LCG Grid Network Infrastructure



ATLAS Grid Infrastructure in Romania



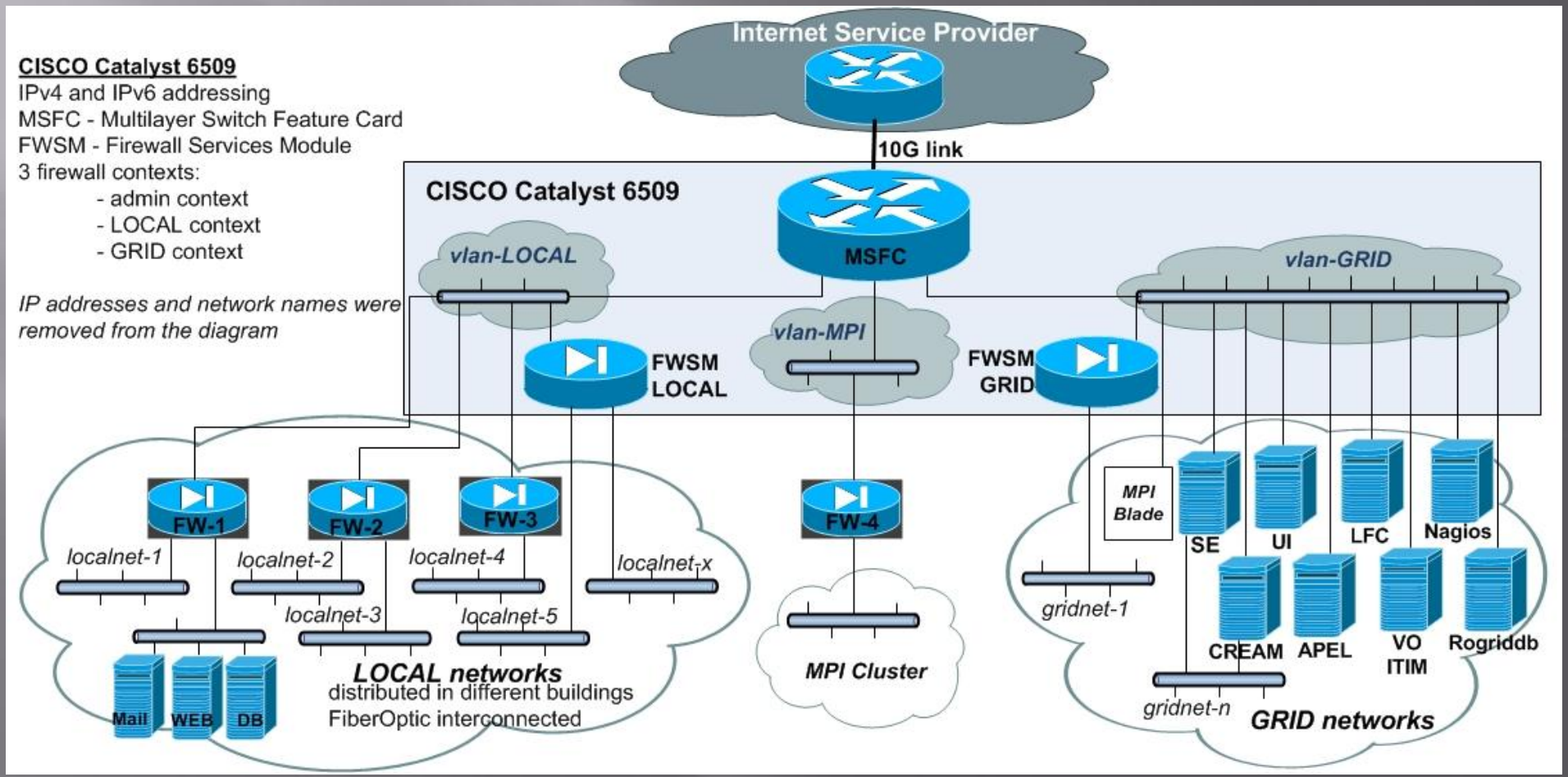
- 4 Grid Sites: **RO-02-NIPNE**, RO-07-NIPNE, RO-14-ITIM and RO-16-UAIC
- 2 sites dedicated to ATLAS: **RO-02-NIPNE** and RO-14-ITIM
- 2 sites running Production and Analysis jobs: **RO-02-NIPNE** and RO-07-NIPNE
- 2 sites running only Production jobs: RO-14-ITIM and RO-16-UAIC

RO-14-ITIM Grid site

2009 – certified & production

NATIONAL INSTITUTE FOR RESEARCH AND DEVELOPMENT OF ISOTOPIC AND MOLECULAR TECHNOLOGIES – Cluj Napoca

Network Configuration and Logical Schema of the Datacenter



RO-14-ITIM Grid site

Site manager:
Dr. Eng. Felix Fărcaş
felix@itim-cj.ro

- Named **RO-14-ITIM**
- Processing power: **core** 500, online **250**
- Storage capacity: 100 TB, online 5**0 TB**
- Technology **1U** + **Blade** system (IBM & HP)
- Virtual Organization (**ATLAS, ops, voitim**)
- Operations system Scientific Linux 64 bit
- Middleware we use is **gLite 3.2 for 64 bit**

Blade system and MSA storage

RO-16-UAIC Grid site

2008 – certified & production

"ALEXANDRU IOAN CUZA" UNIVERSITY OF IAŞI
Digital Communications Department

- 3 Gigabit Ethernet switches with 48 Gigabit ports and two Ten Gigabit Ethernet ports

- 50 servers used for WN with 8 cores (2.66GHz),16GB RAM, 2 Gigabit interfaces, 160 GB disk storage per computer
- 1 server used for NFS, DHCP, DNS with 8 cores (2.66GHz), 24 GB RAM, 4* 160 GB disks storage

- 1 storage with 8 TB in Raid 6 used for atlas software
- 1 storage server used for SE with 16 cores (2.4GHz), 24GB RAM(1333MHz), 6 Gigabit interfaces and 80TB in Raid 6

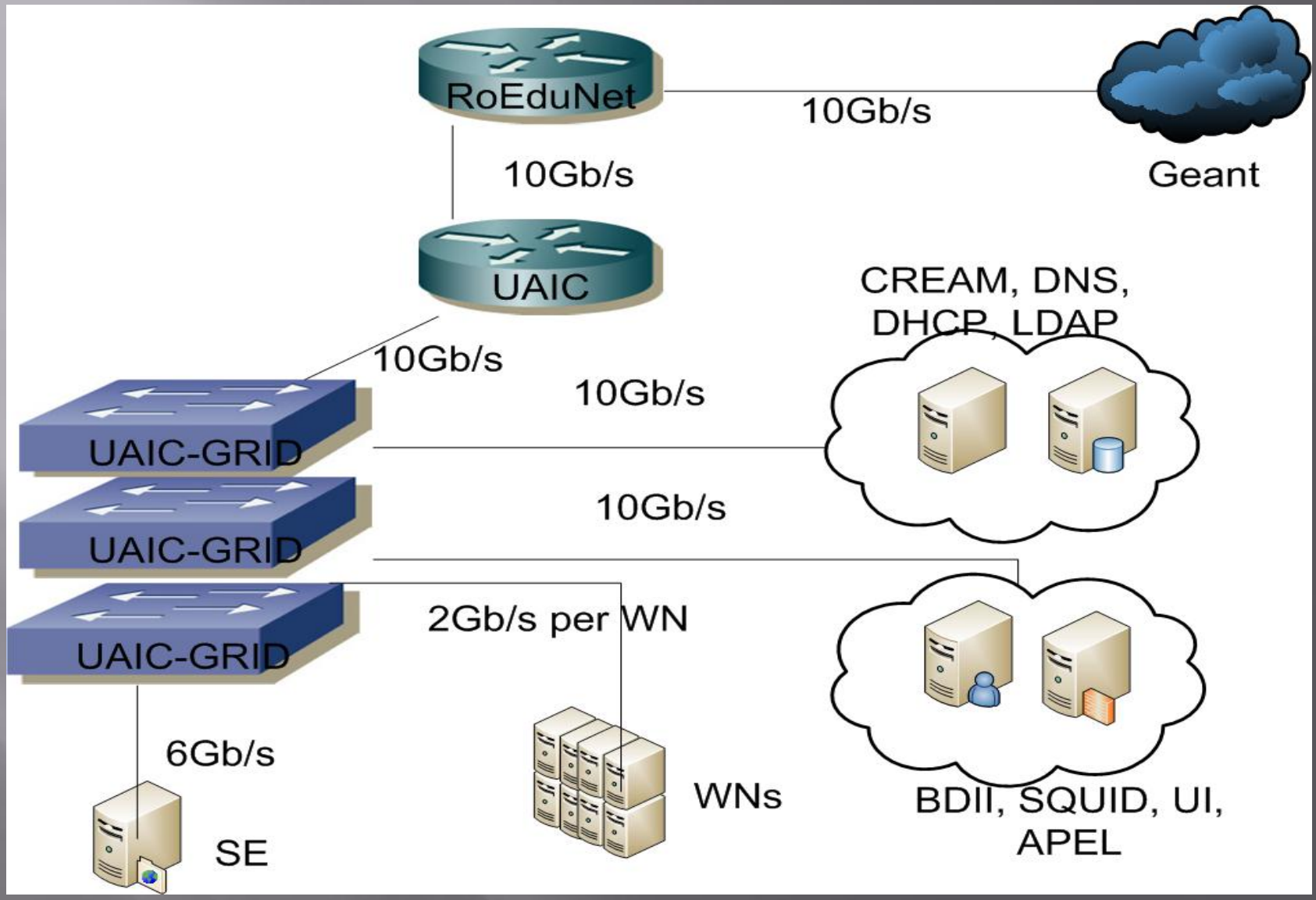
- 2 servers will be used for back-up virtualization system in CREAM, BDII, UI, SQUID, with 12 cores (2.66GHz), 32GB RAM, 2 Ten Gigabit Ethernet

- ✓ Virtual Organization : **atlas, ops, dteam**
- ✓ Operations system: **Scientific Linux 5.7 64 bit**
- ✓ Middleware: **gLite 3.2 for SE**

EMI for WN, APEL, UI, BDII, CREAM

System manager: Pînzaru Ciprian

RO-16-UAIC Network Infrastructure

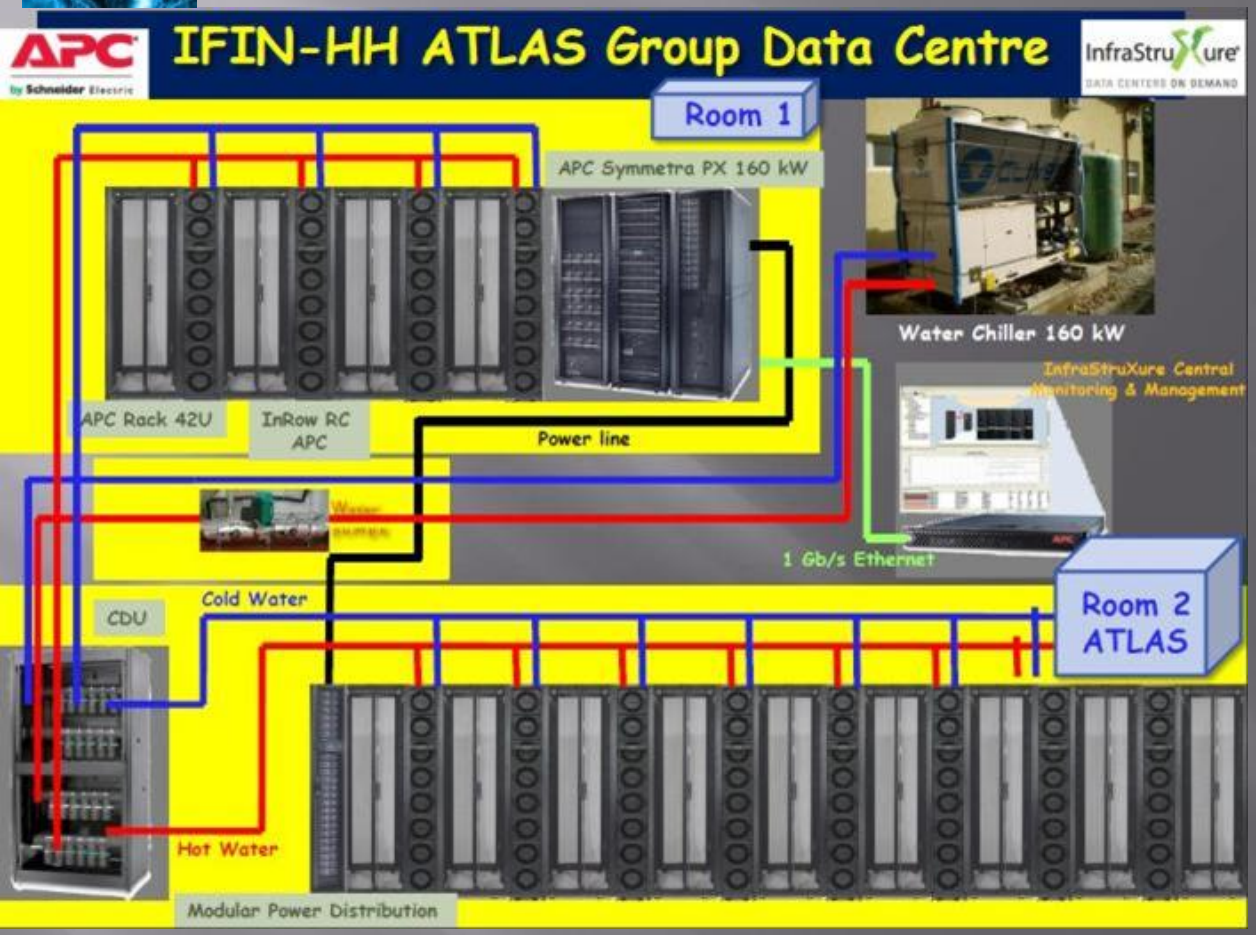


RO-16-UAIC Power Infrastructure

- One UPS have power 54 KW and the other two by 18 KW.
- One generator with a output power by 100KW.

RO-02-NIPNE Grid Site

2005 – certified & production



Uninterruptible Power Supply (UPS): **APC Symmetra PX 160KW** 400 V w/ Integrated Modular Distribution; high-efficiency 3-phase UPS with integrated modular distribution that can be right-sized to data center power requirements. With hot-scalable power, distribution and run-time this UPS scales with data center up to 160KW/160kVA.

144 KW and 1h21min run-time

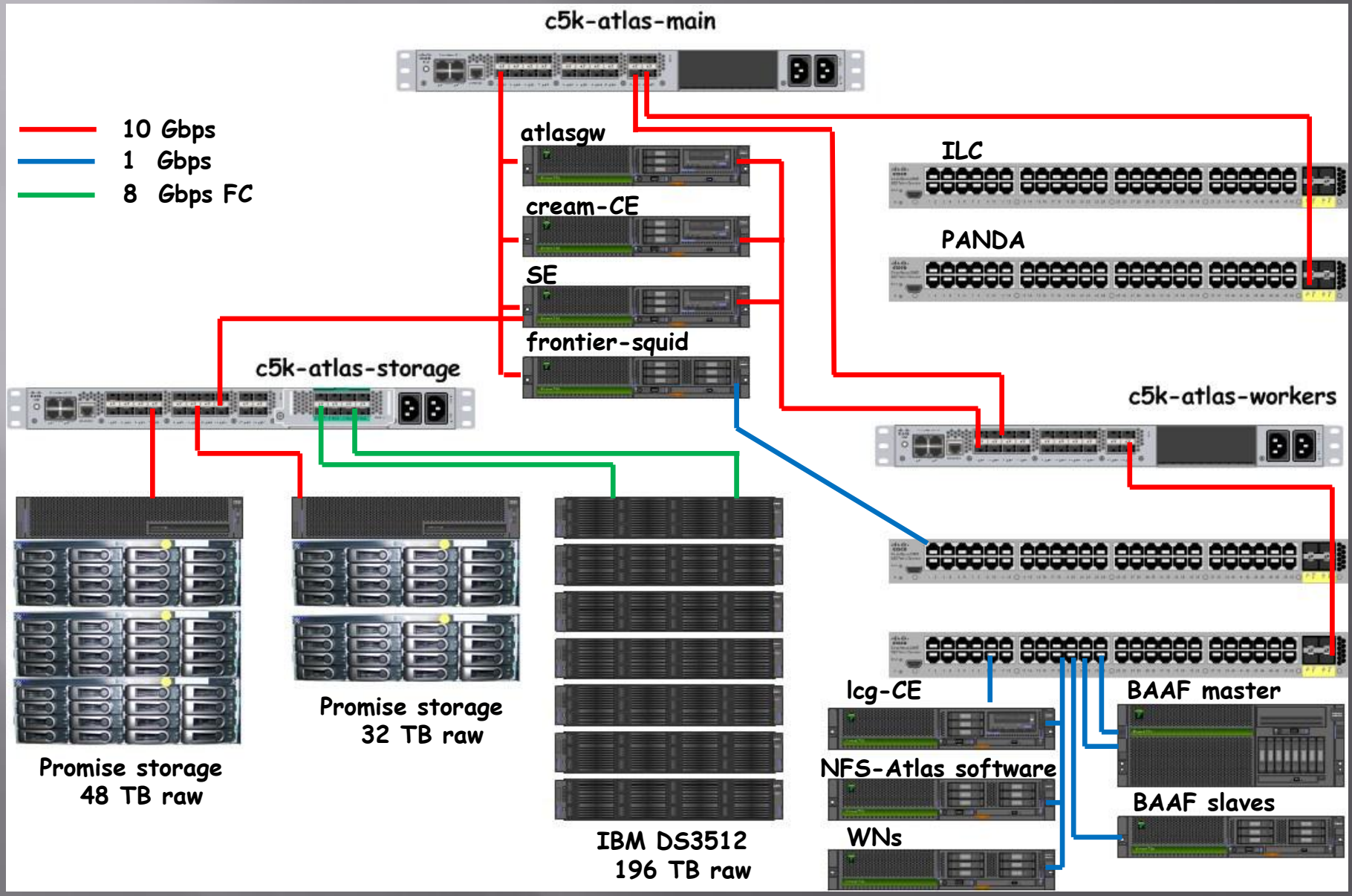
Institute Diesel Power Generator 1 MW

Cooling System APC – based on chilled water topology – hot-aisle/cold-aisle design (see next slide):

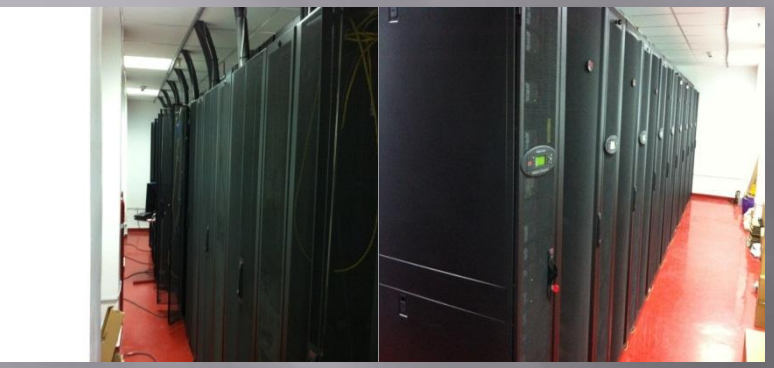
- ✓ Water Chiller -160 kW- with expansion tank – around two hours autonomy in case of power-cut.
- ✓ CDU (Cooling Distribution Unit) - Flexible chilled water and glycol distribution system for InRow cooling units - 12 Circuit, Bottom/Top Mains, Top Distribution Piping.
- ✓ 12 In Row RC Chilled Water units, 200-240V 50/60 Hz, IEC 309-16.

InfraStruXure Central Basic – management and monitoring system; It supports capacity management up to 20 racks and change management up to 1,000 IT assets - RHE Linux 5.0 – 1 GB RAM/160 GB HDD data storage

RO-02-NIPNE Grid Site – Network Infrastructure



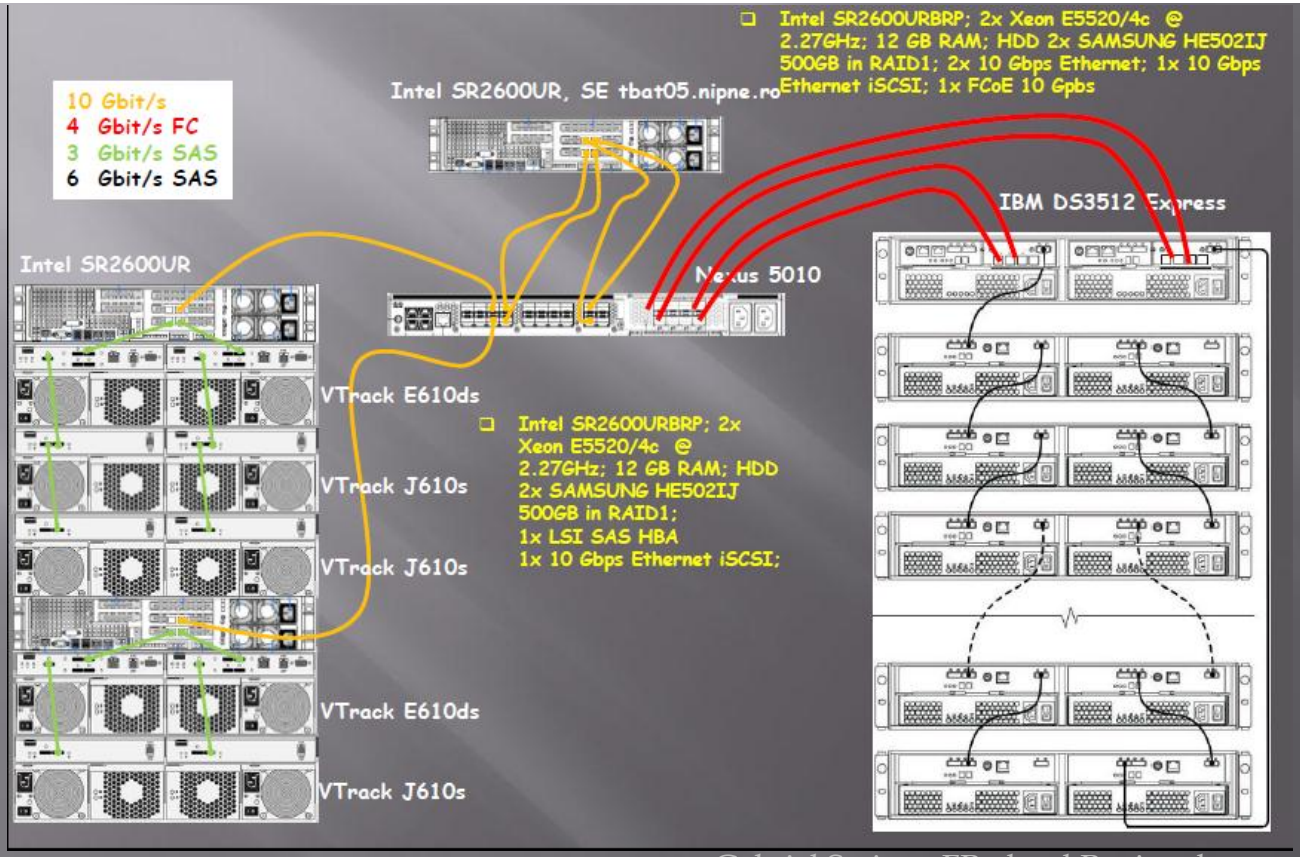
RO-02-NIPNE Grid Site – Grid Services and Computing Hardware

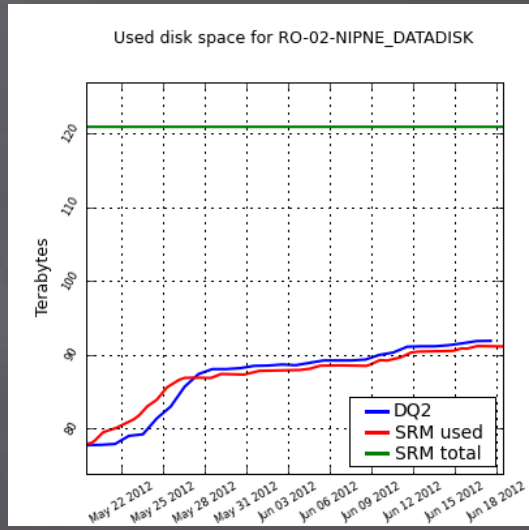


- Part of RO-LCG / Romanian Tier2 Federation
- Grid Workers:
  - 20 x Intel SR1500AL; 2xXeon 5030/2c 2.66GHz, RAM 4x2GB, HDD Seagate 80GB
  - 32 x Intel SR1500AL; 2xXeon E5335/4c 2.00 GHz, RAM 8x2GB, HDD Seagate ST3250620NS 250GB.
- Two CEs:
  - lcg-CE and site-BDII: Intel SR1500AL; 2 x Xeon E5420/4c @ 2.50GHz; 16 GB RAM; HDD Seagate ST3500320NS 500GB; 2 x 1 Gbps Ethernet
  - cream-CE: Intel SR2600URBRP; 2x Xeon E5520/4c @ 2.27GHz; 12 GB RAM; HDD 2x SAMSUNG HE502IJ 500GB in RAID1; 2x 10 Gbps Ethernet
- Atlasgw /Gateway machine: Intel SR2600URBRP; 2x Xeon E5520/4c @ 2.27GHz; 12 GB RAM; HDD 2x SAMSUNG HE502IJ 500GB in RAID1; 2x 10 Gbps Ethernet
- Storage system 250 TB raw capacity
- SE (DPM/SRM) with 200 TB on-line
- NFS server for ATLAS software: Intel SR2500AL; 1 x Xeon E5420/4c @ 2.50GHz; 8 GB RAM; HDD Seagate ST3250620NS 250GB; ATLAS soft space 1TB RAID1 2x Segate ST31000340NS; 1 Gbps Ethernet on private WNs network
- Frontier-squid-cache server serving RO-02, RO-14 and backup for RO-16: Intel SR2500AL; 1 x Xeon E5420/4c @ 2.50GHz; 8 GB RAM; HDD Seagate ST3500320NS 500GB; 1 Gbps Ethernet on private WNs network & 10 Gbps Ethernet on public network

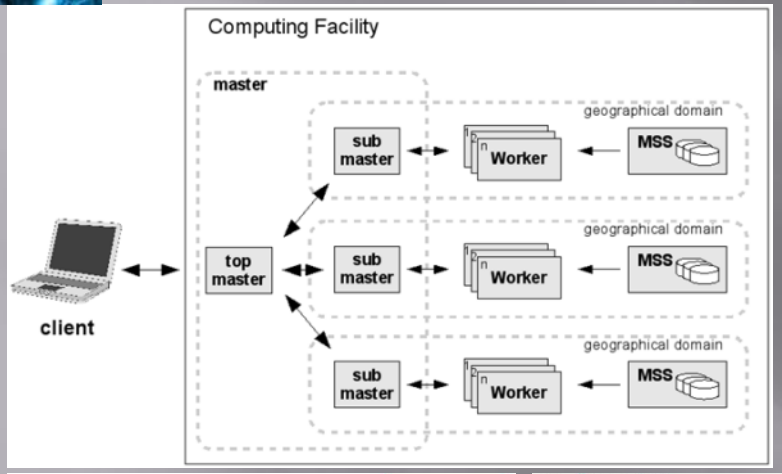
RO-02-NIPNE ATLAS Storage

Federation	Site	DATADISK		DATATAPE		GROUPDISK		HOTDISK		MCTAPE		PRODDISK		SCRATCHDISK	
RO-LCG	RO-02-NIPNE	120.9	492.6 (60%)	-	- (-%)	3.3	3.3 (0%)	2.2	8.8 (1%)	-	- (-%)	6.6	31.9 (3%)	5.5	16.5 (2%)
	RO-07-NIPNE	329.9		-		-		2.2		-		17.6		5.5	
	RO-14-ITIM	15.4		-		-		1.1		-		3.3		-	
	RO-16-UAIC	26.4		-		-		3.3		-		4.4		5.5	

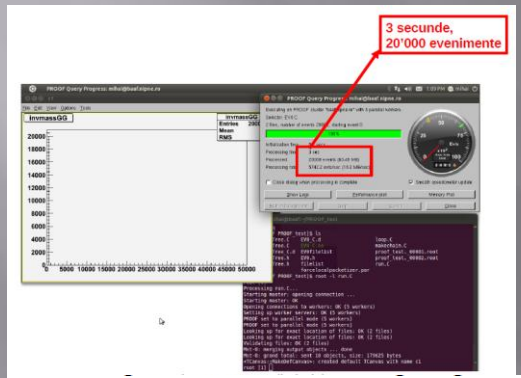


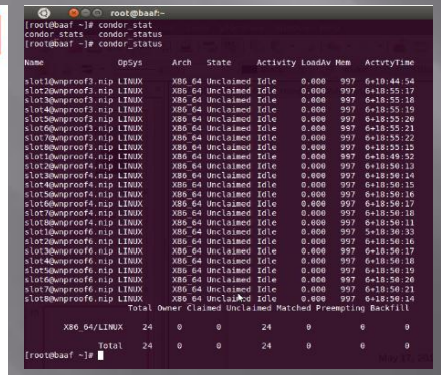


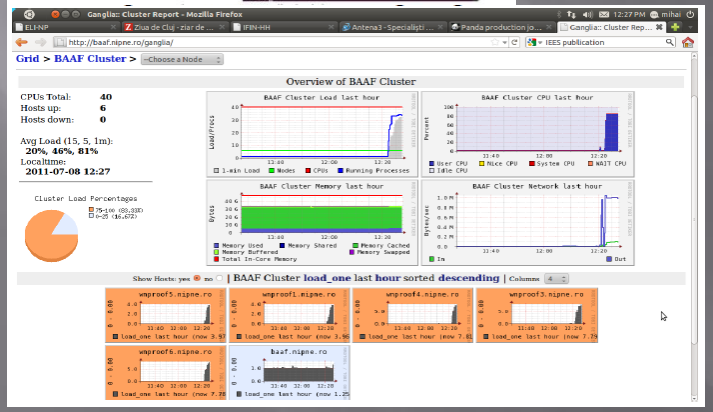
BAAF (Bucharest ATLAS Analysis Facility)



- consist of a non-homogeneous PROOF cluster
- computing cluster driven by Condor
- dedicated to members of Bucharest ATLAS group







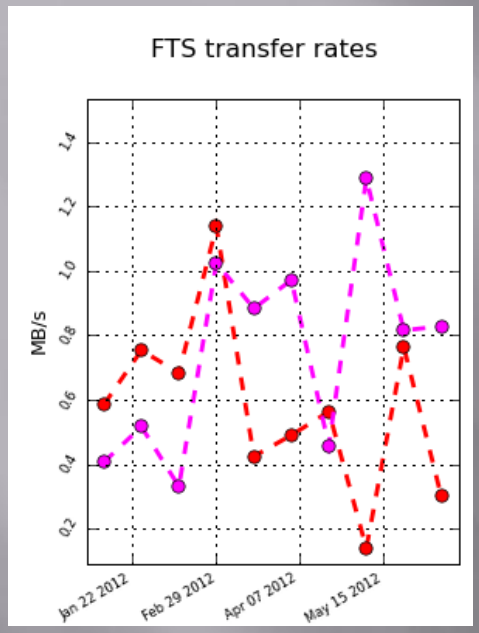
- Hardware overview:
- 1 Masternode – with 8 cores and 7 TB storage
  - configured as glite-UI with dq2-tools
  - CVMFS for ATLAS software
  - configured as masternode of the Proof cluster (for analysis)
  - manager of the Condor cluster; submitting and scheduling jobs (for simulations)
  - configured to store data via xrootd
- 6 Slavenodes
  - 3 with 8 cores / Dell PowerEdge 1950
  - 2 with 4 cores

Total of 40 cores

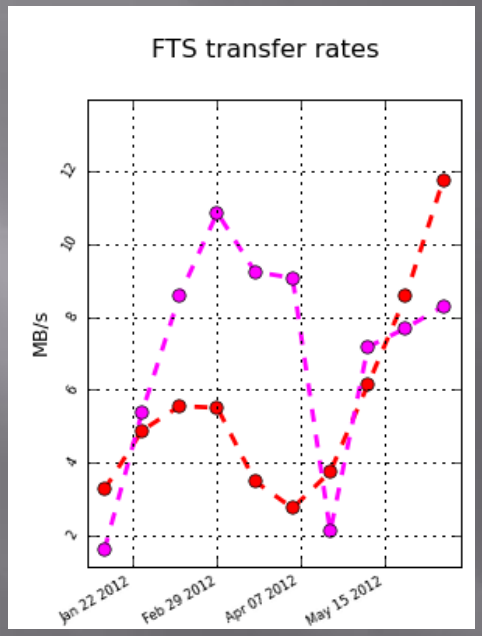
- Network – 1Gbps in a private network; only the masternode can be accessed from Internet
- Monitoring the cluster with Ganglia

Results for the Grid sites

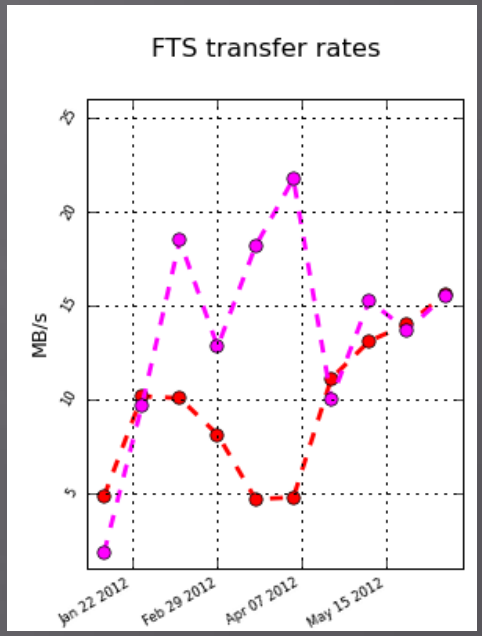
Small files (0 to 100 MB)

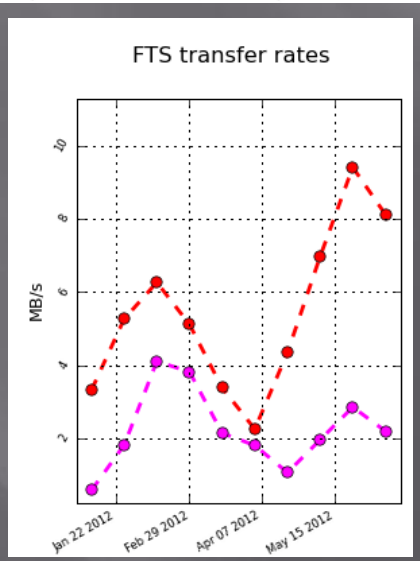


Medium files (100 MB to 1 GB)



Large files (1 GB to infinity)



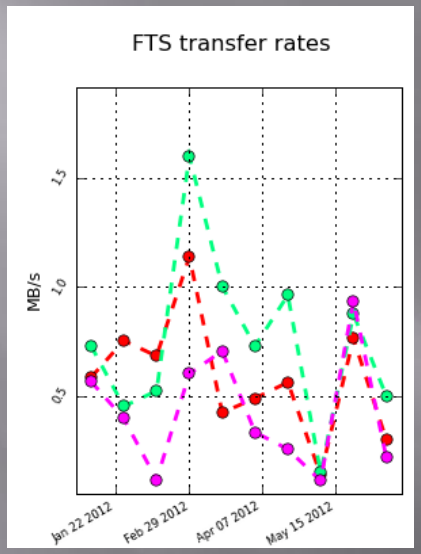


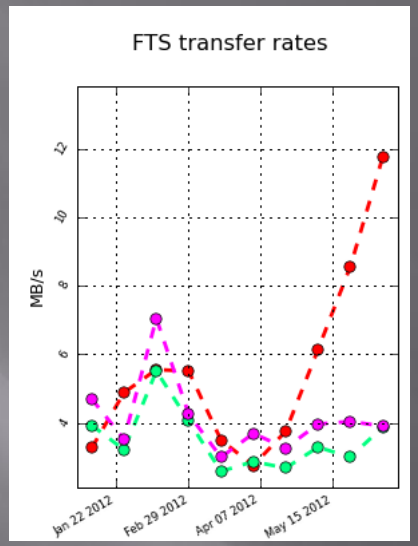
All files (0 to infinity)

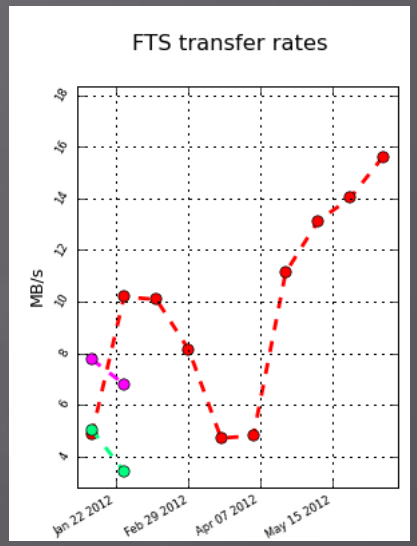
Results for the Grid sites

Small files (0 to 100 MB) | Medium files (100 MB to 1 GB) | Large files (1 GB to infinity)

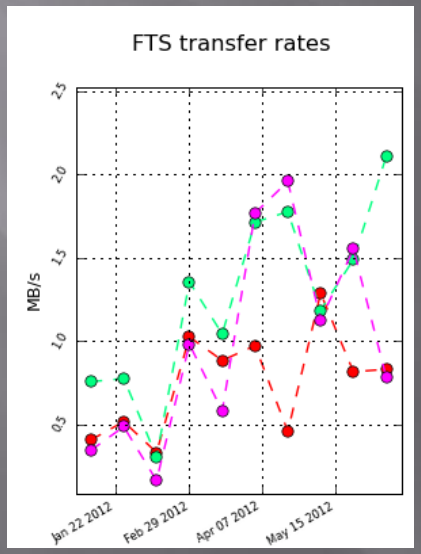
IN2P3-CC - RO-02-NIPNE
IN2P3-CC - RO-14-ITIM
IN2P3-CC - RO-16-UAIC

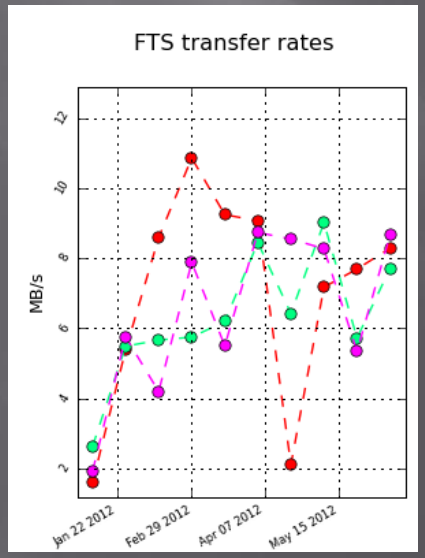


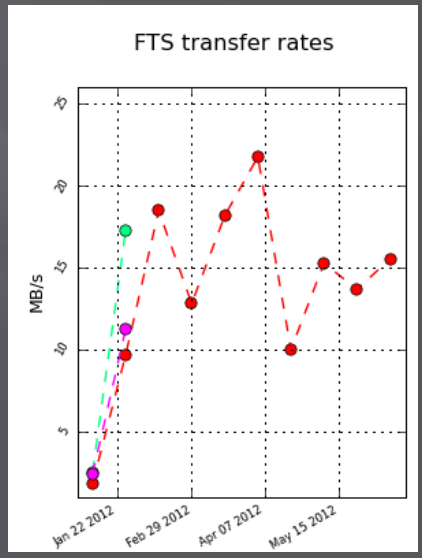




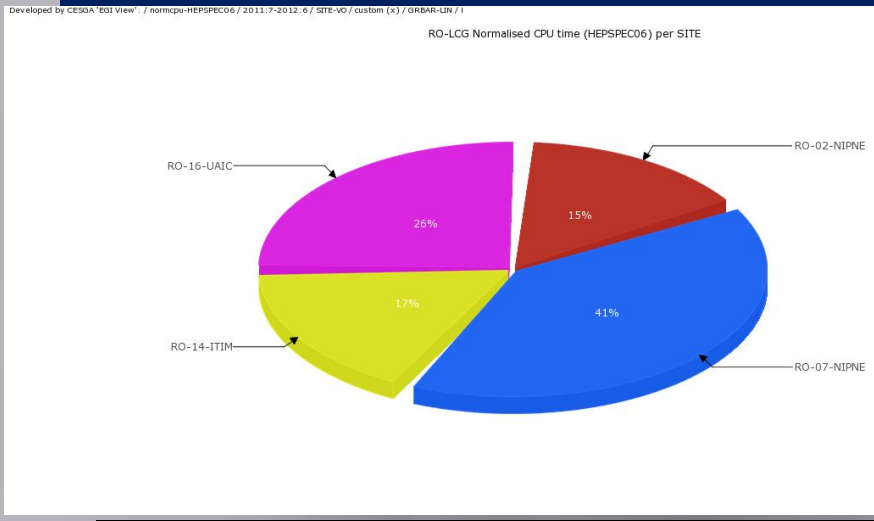
RO-02-NIPNE - IN2P3-CC
RO-14-ITIM - IN2P3-CC
RO-16-UAIC - IN2P3-CC

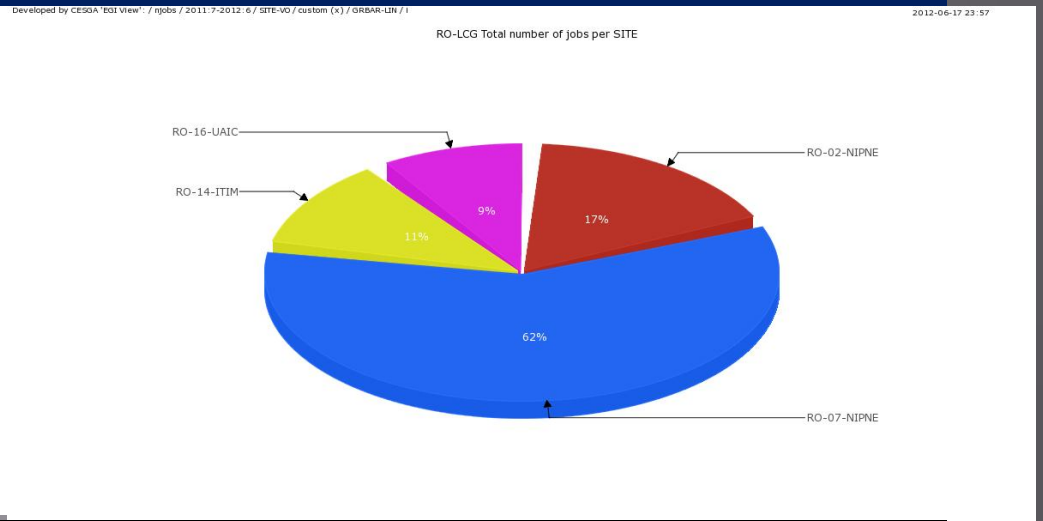






Results for the Grid sites

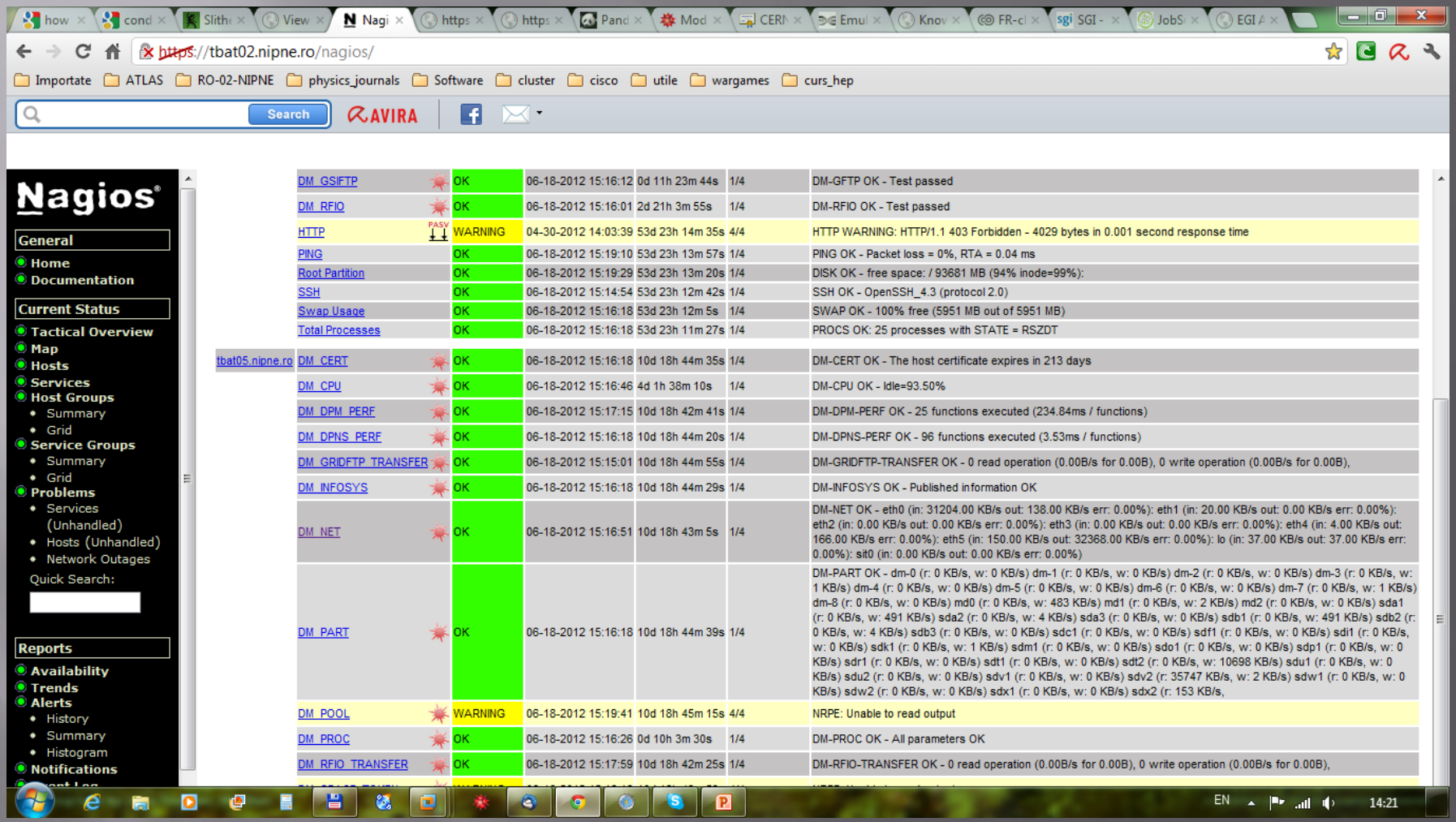




Total number of jobs run by SITE and VO

SITE	atlas		Total	%
RO-02-NIPNE	1,408,053		**1,408,053**	**17.48%**
RO-07-NIPNE	4,967,999		**4,967,999**	**61.69%**
RO-14-ITIM	925,808		**925,808**	**11.50%**
RO-16-UAIC	751,590		**751,590**	**9.33%**
Total	**8,053,450**		**8,053,450**	
Percentage	**100.00%**			

Monitoring



- Local monitoring using Nagios – first service is DPM
- Grid services monitored with Nagios at RO-LCG and NGI-RO level

Monitoring problems

- RO-02-NIPNE :

  - reconfiguring data centre network basic infrastructure; more segmented lowering the network traffic load on switches.
  - start tuning and optimization for the storage system.

  Principal errors for file transfers:

  - Failed to complete PutDone request [0cf466a8-846a-4eb0-8aa1-e851429e2f77] on remote SRM [httpg://tbat05.nipne.ro:8446/srm/managerv2]: [SRM_FAILURE] Failed for all SURLs.The PutDone request has been successfully aborted

  - No status updates received since more than [360] seconds. Probably the process serving the transfer is stuck

RO-07-NIPNE
status and current development

Mihai Ciubancan

IT Department
National Institute of Physics and Nuclear Engineering “Horia Hulubei”
IFIN-HH

CONTENT

RO-07-NIPNE general info
Networking
Running jobs
Monitoring

RO-07-NIPNE: GENERAL

➔Running ATLAS, LHCB, and also ALICE soon.

➔More then 1400 cores:

- ➔1100 available for ATLAS ,LHCB and GRIDIFIN(local VO)
- ➔300 available for ALICE

➔About 420 TB of storage, ~220TB used

- 330 TB DATADISK, 17,5 TB PRODDISK, 5,5 TB SCRATCHDISK, 2,2 TB HOTDISK, ~65 TB for other VO's

RO-07-NIPNE: GENERAL

- CE:
  - CREAM1 – dedicated to ATLAS, LHCB and GRIDIFIN.
  - CREAM2 – dedicated to ALICE
- SE: glite-SE_dpm_mysql, 15 DPM_DISKs
- WN for ATLAS, LHCB and GRIDIFIN:
  - 46x8 cores – 500GB HDD, 62,5GB/core, 2 GB/core
  - 31x16 cores – 1TB HDD, 62,5GB/core, 2GB/core
  - 7x32 cores – 1TB HDD, 31,25GB/core, 2GB/core
  - SLC5-x86_64

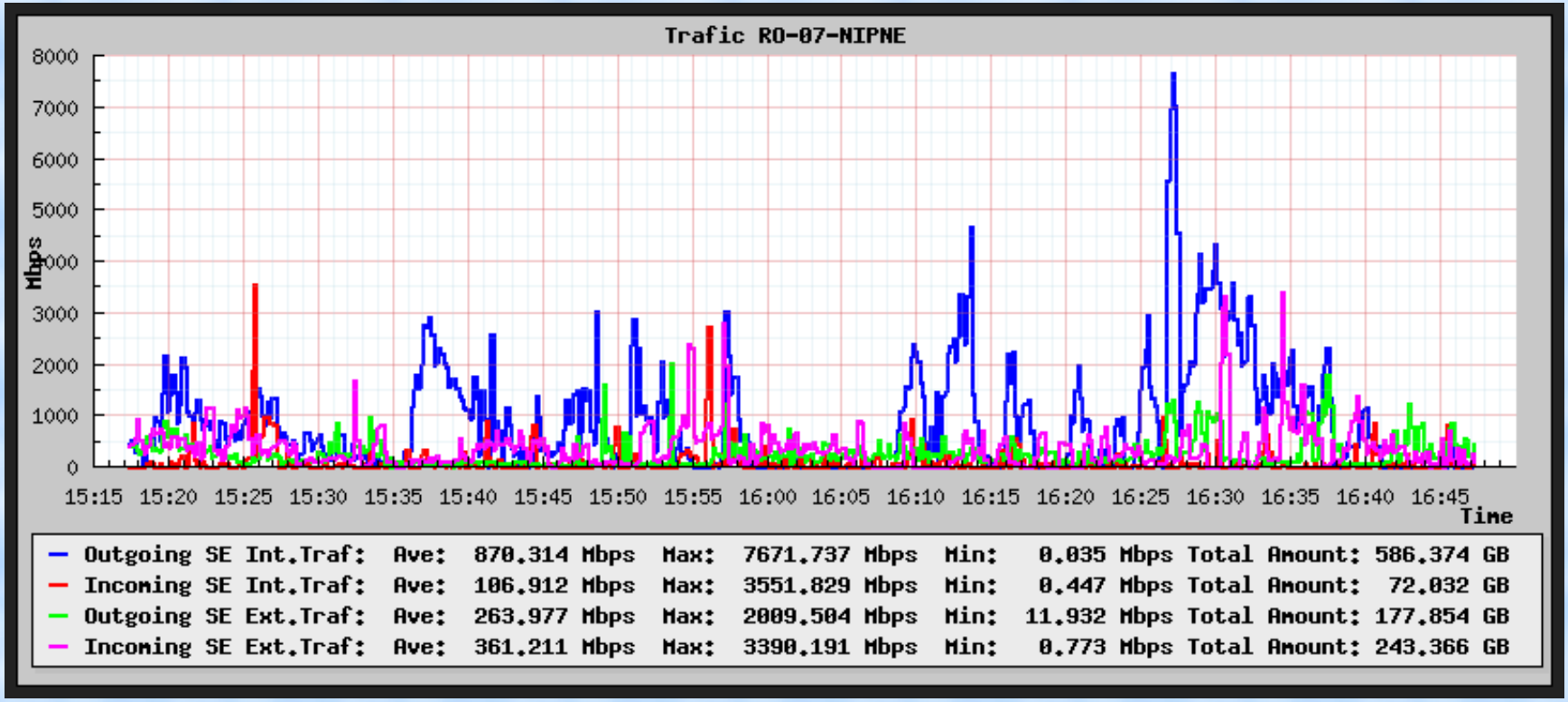
RO-07-NIPNE: GENERAL

- BDII-top: setup a local BDII-top which improved the successful jobs rate
- SQUID server for CVMFS on the WN's gateway
- CVMFS deployed on all worker nodes for ATLAS
- Still having NFS server for ATLAS, ALICE, LHCB and local VO
- VO-BOX installed for ALICE
- WN for ALICE:
  - 28x8 cores – 500GB HDD, 62,5GB/core, 2GB/core
- Deployed a LFC, VOMS, MyProxy and WMS for local VO's

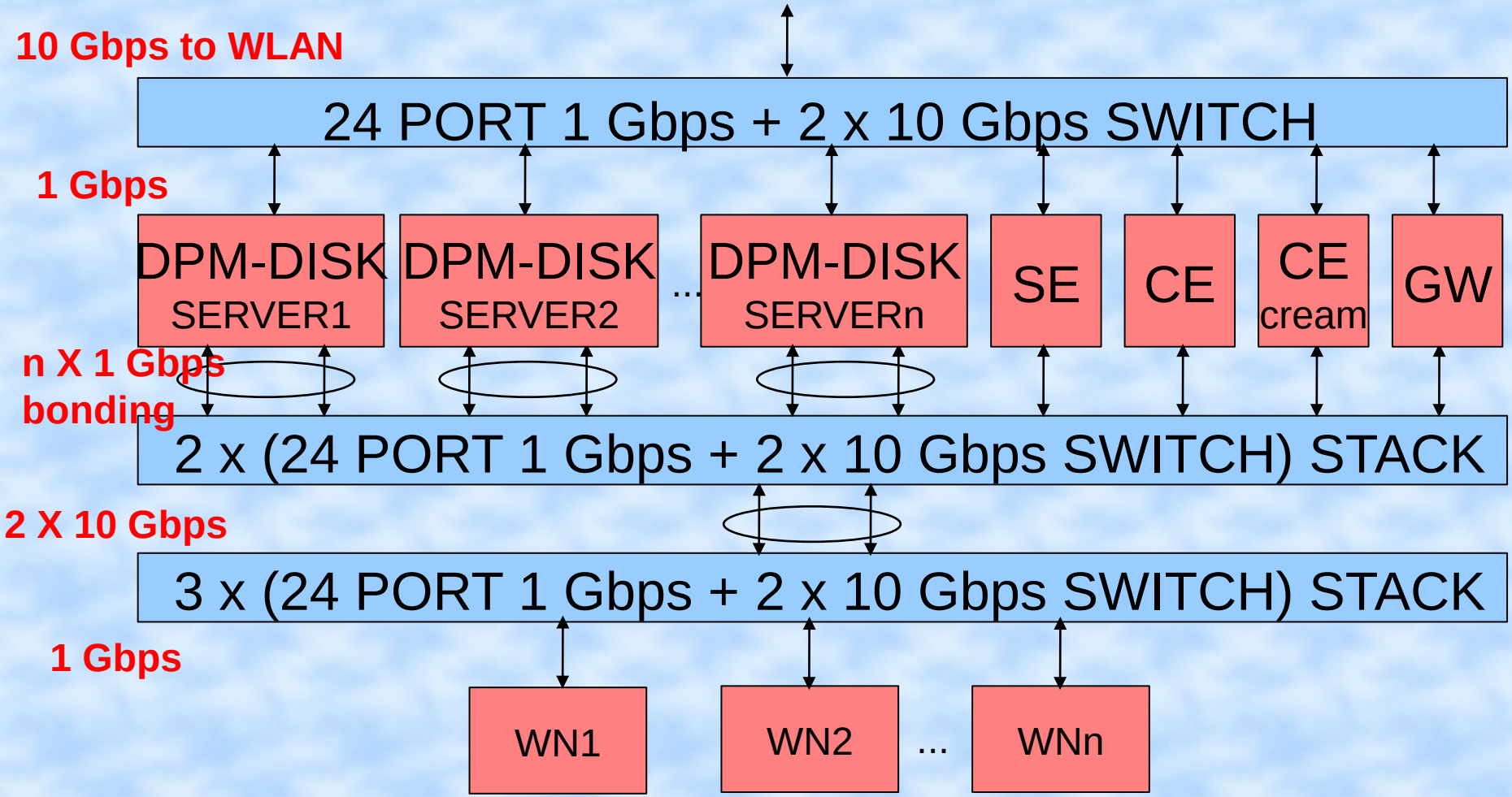
NETWORKING

- WAN link (site) 10 Gbps with the provider (ROEDUNET, (10 Gbps in GEANT).
- CREAM1: 1 Gbps
- CREAM2: 1Gbps
- SE:
  - internal link (WN's) : minimum 100Mbps/1TB of storage
- WN: 1 Gbps for 8 cores.
- GW: 1 Gbps
- Monitoring the traffic with Cacti and local developed tool

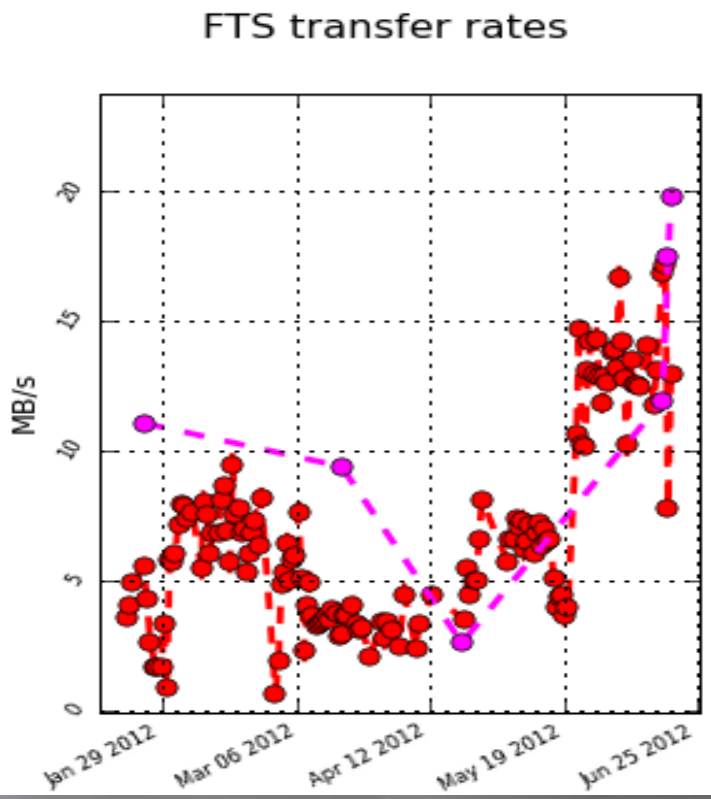
NETWORKING

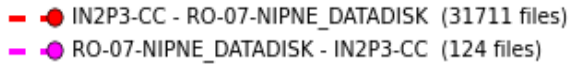

Trafic RO-07-NIPNE
8000
7000
6000
5000
4000
3000
2000
1000
0
Mbps
15:15 15:20 15:25 15:30 15:35 15:40 15:45 15:50 15:55 16:00 16:05 16:10 16:15 16:20 16:25 16:30 16:35 16:40 16:45
Time
Outgoing SE Int.Traf: Ave: 870.314 Mbps Max: 7671.737 Mbps Min: 0.035 Mbps Total Amount: 586.374 GB
Incoming SE Int.Traf: Ave: 106.912 Mbps Max: 3551.829 Mbps Min: 0.447 Mbps Total Amount: 72.032 GB
Outgoing SE Ext.Traf: Ave: 263.977 Mbps Max: 2009.504 Mbps Min: 11.932 Mbps Total Amount: 177.854 GB
Incoming SE Ext.Traf: Ave: 361.211 Mbps Max: 3390.191 Mbps Min: 0.773 Mbps Total Amount: 243.366 GB

RO-07-NIPNE NETWORK LAYOUT



NETWORKING

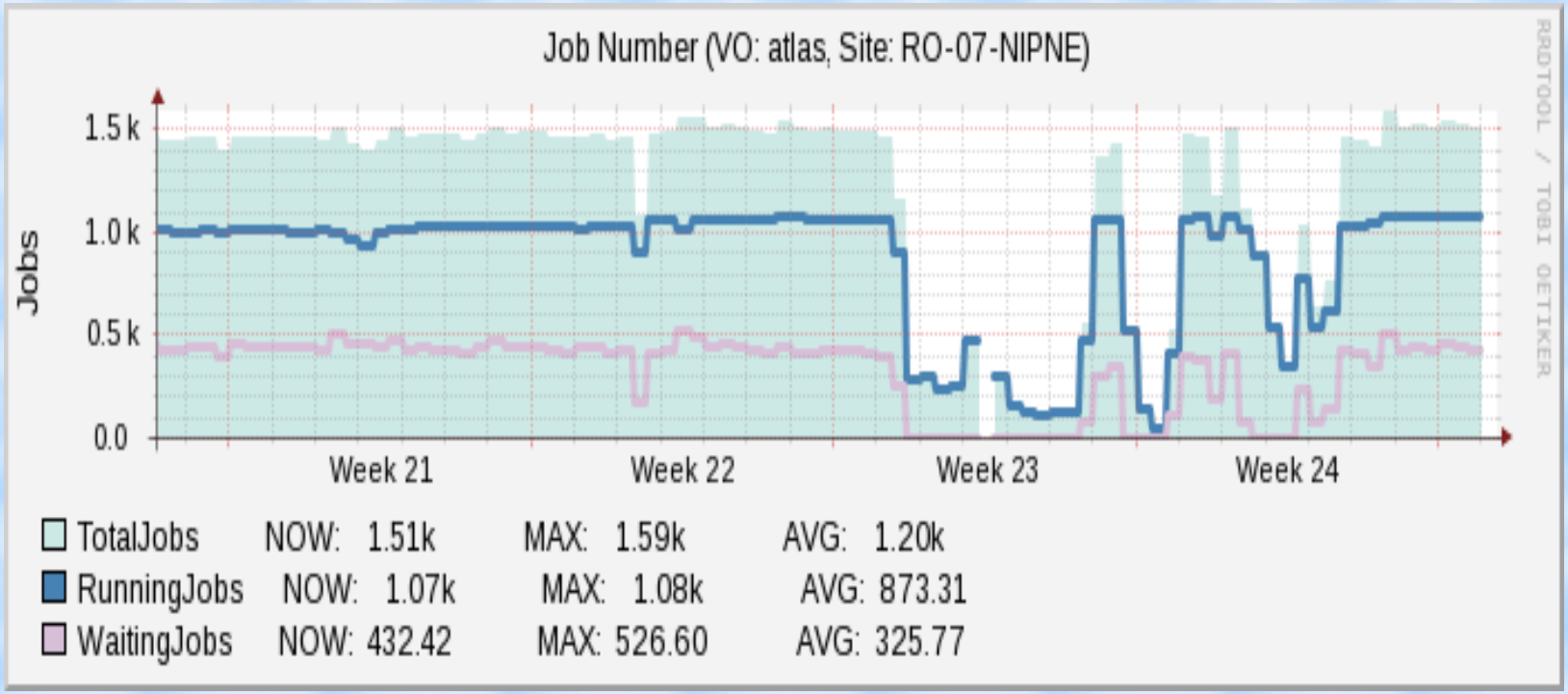

FTS transfer rates
MB/s
0
5
10
15
20
Jan 29 2012
Mar 06 2012
Apr 12 2012
May 19 2012
Jun 25 2012


IN2P3-CC - RO-07-NIPNE_DATADISK (31711 files)
RO-07-NIPNE_DATADISK - IN2P3-CC (124 files)

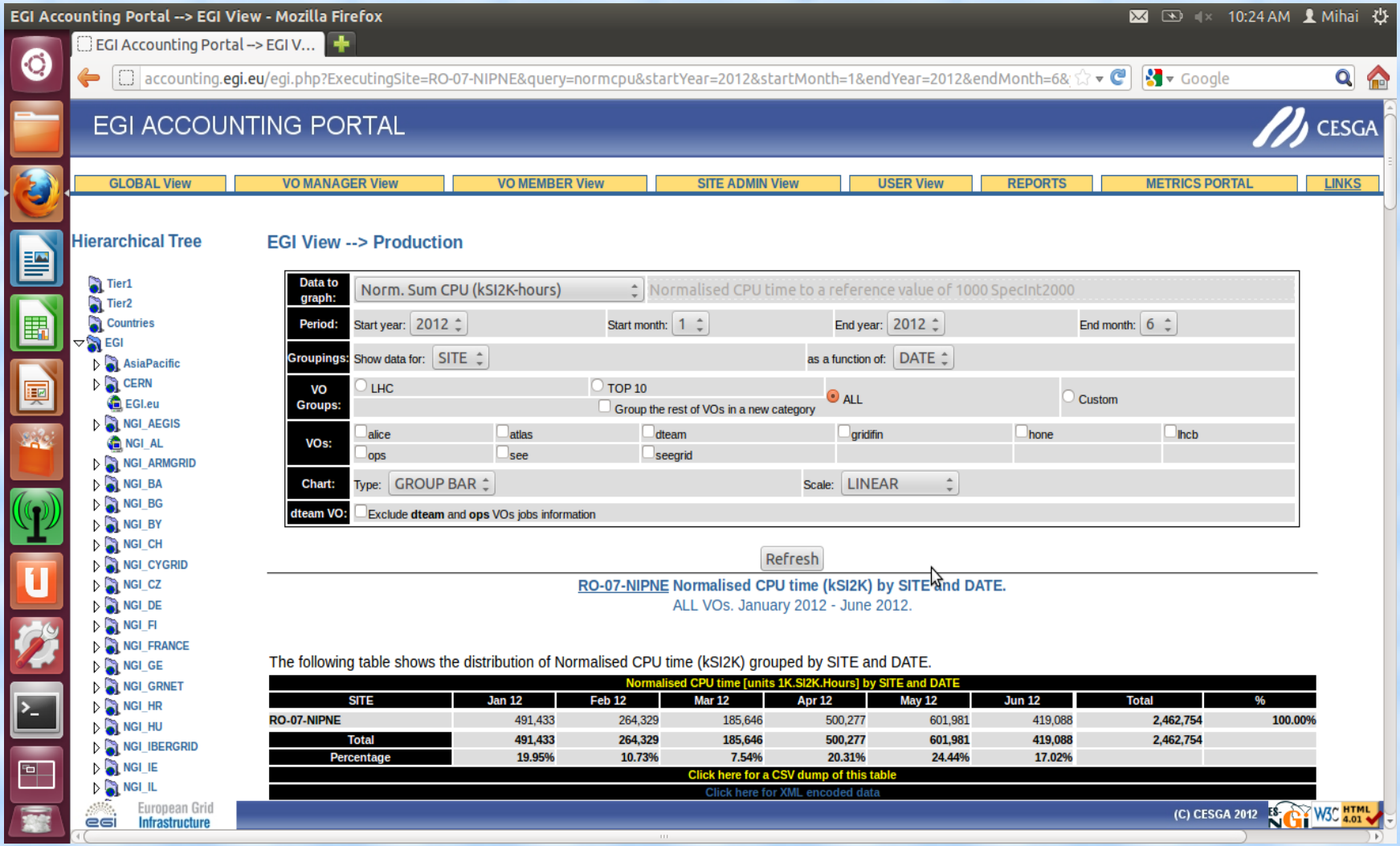
RUNNING JOBS

- Running mostly production jobs - ATLAS and LHCB.
- Running analysis jobs concurrently - maximum 150 simultaneous jobs.
- Maximum ATLAS load: 1090 jobs (May-June).
- Maximum LHCB load 430 jobs (January).

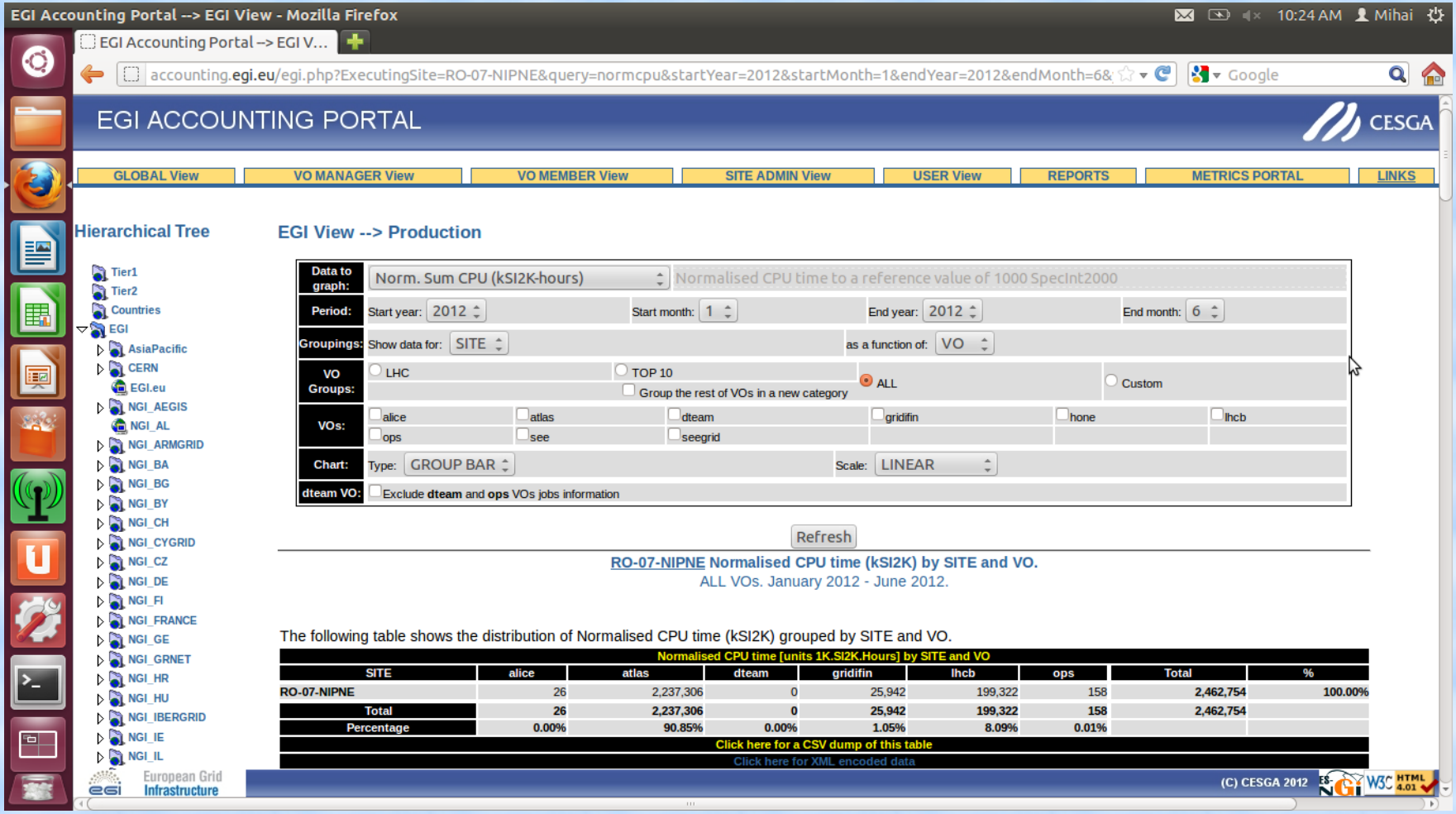
RUNNING JOBS



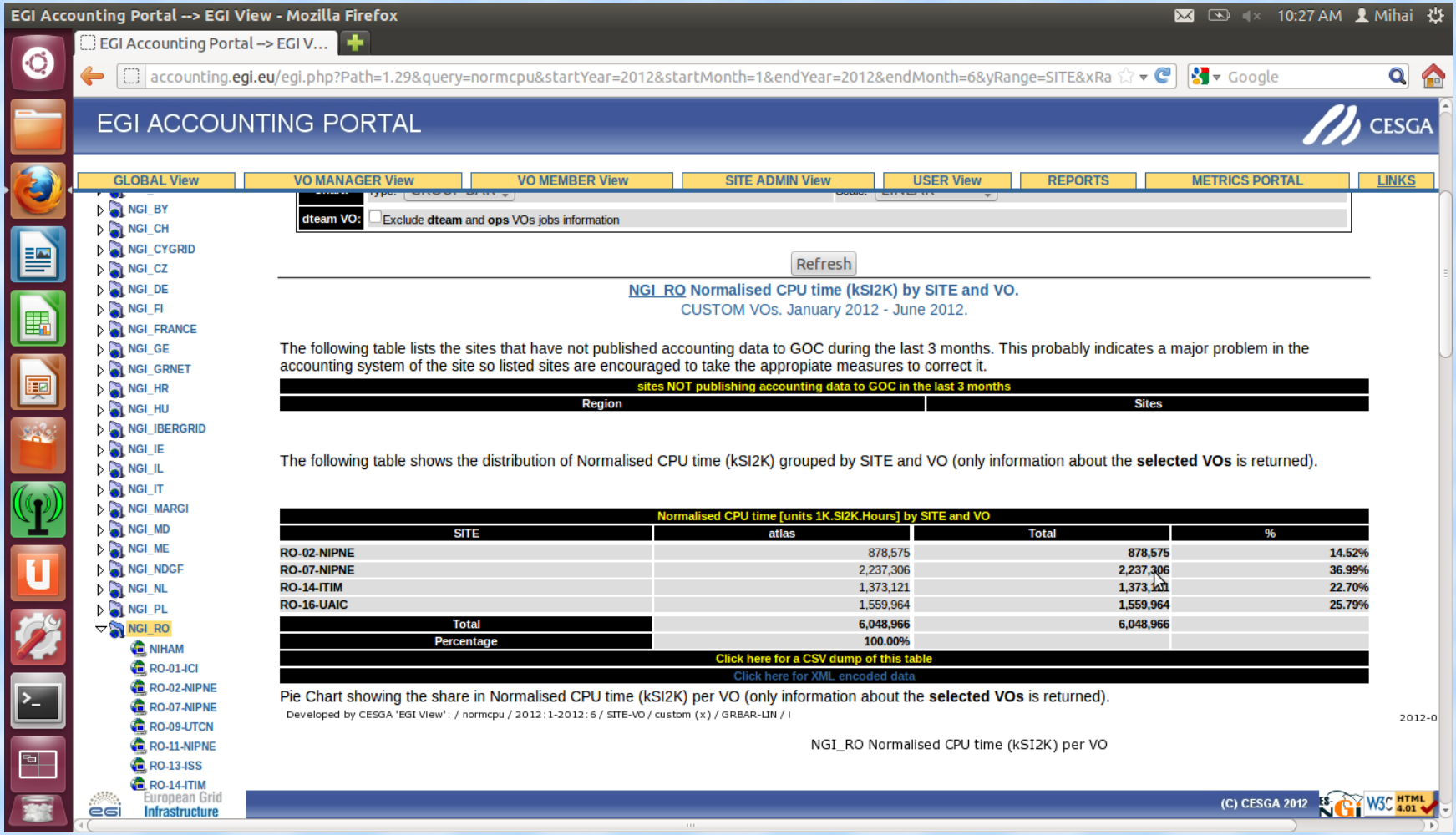
RUNNING JOBS



RUNNING JOBS



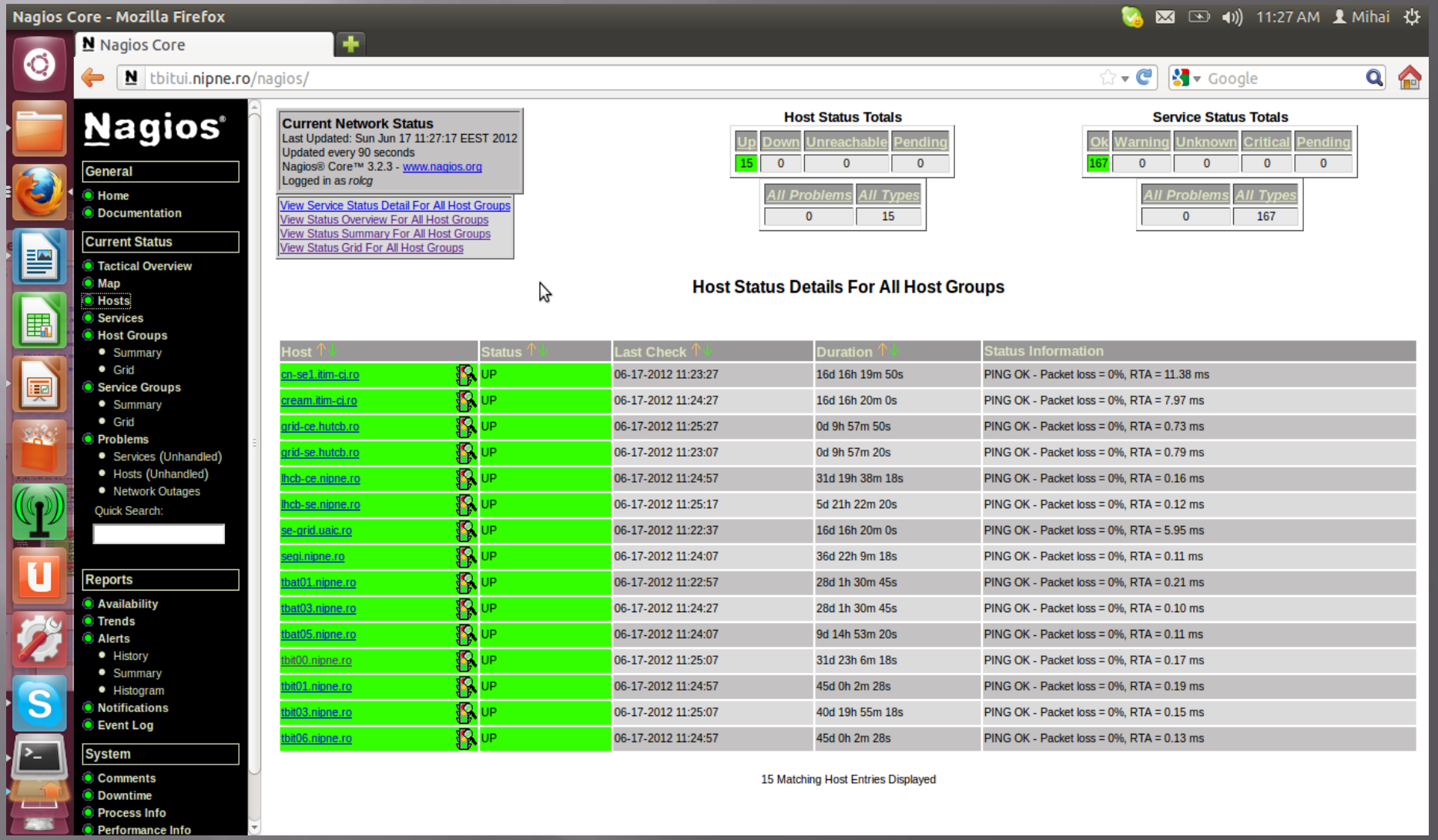
RUNNING JOBS



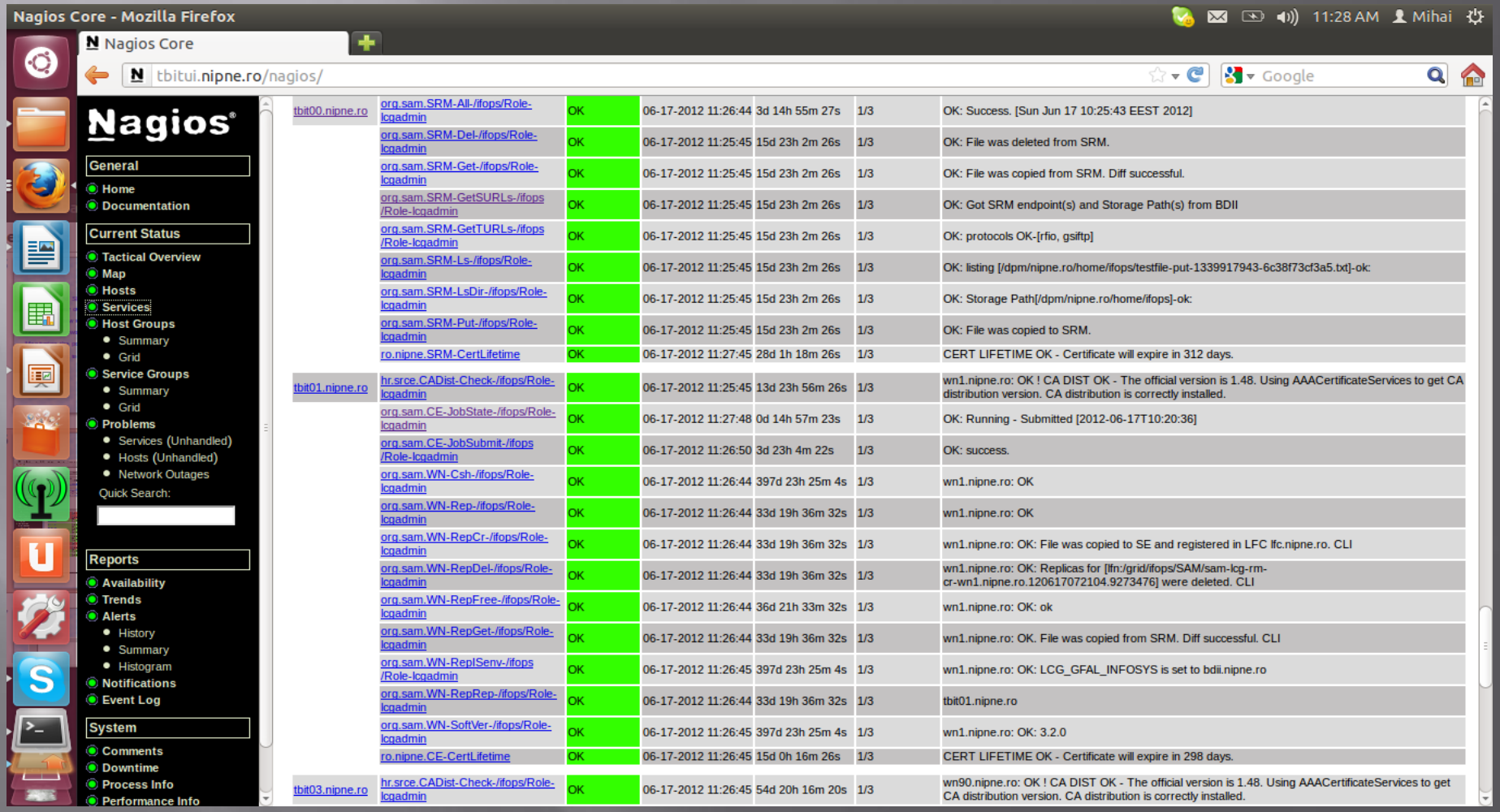
RO-07-NIPNE: MONITORING

- Monitoring tools:
- For sanity check with nagios done with local VO ifops for all the RO-LCG sites
- For services status of RO-07-NIPNE with nagios
- For number of jobs and ksi2k running for RO-07-NIPNE
- For Cooling and UPS(power consumption)
- E-mail alerts regarding network and temperature inside data centers

RO-07-NIPNE: MONITORING

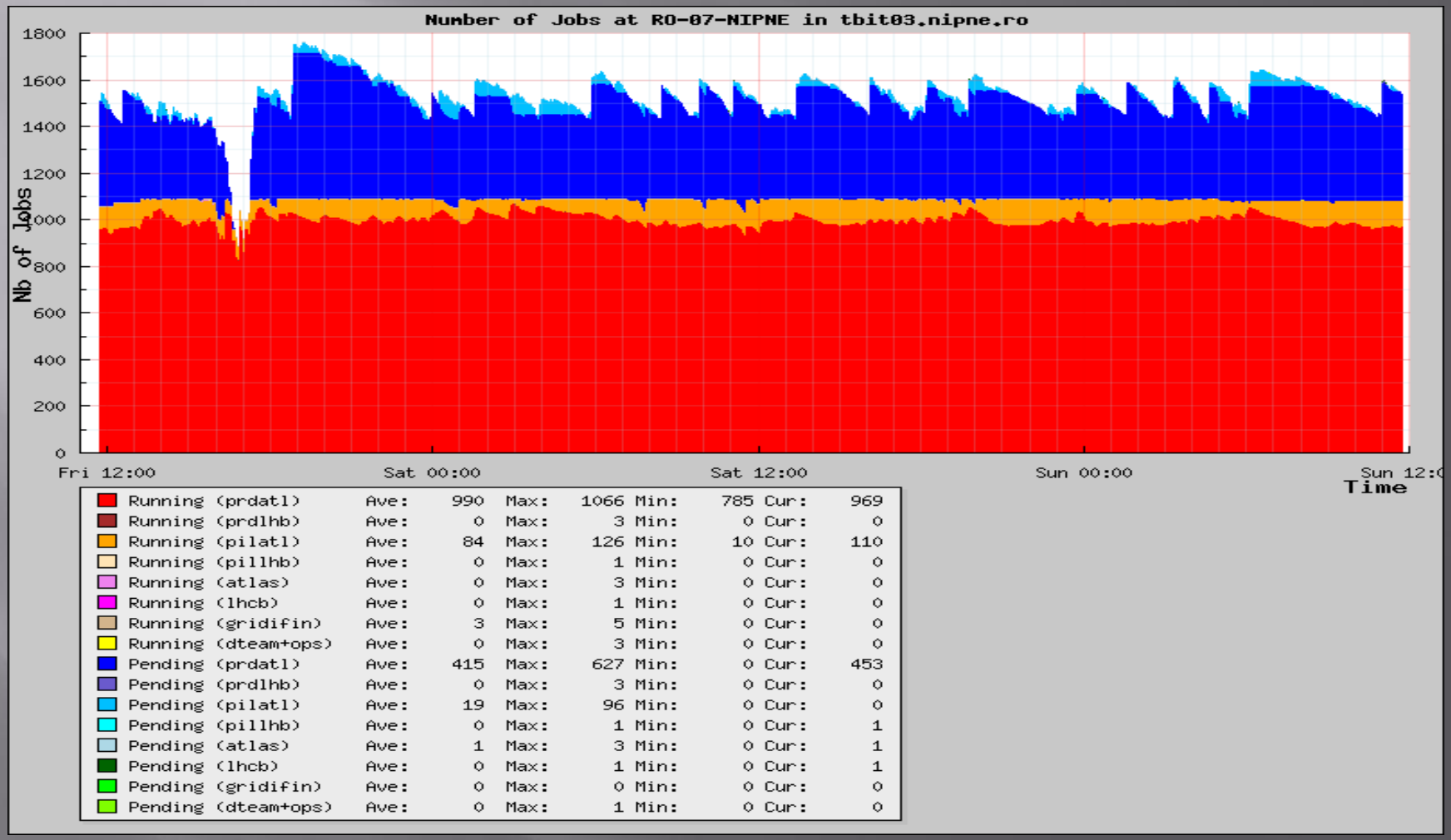


RO-07-NIPNE: MONITORING



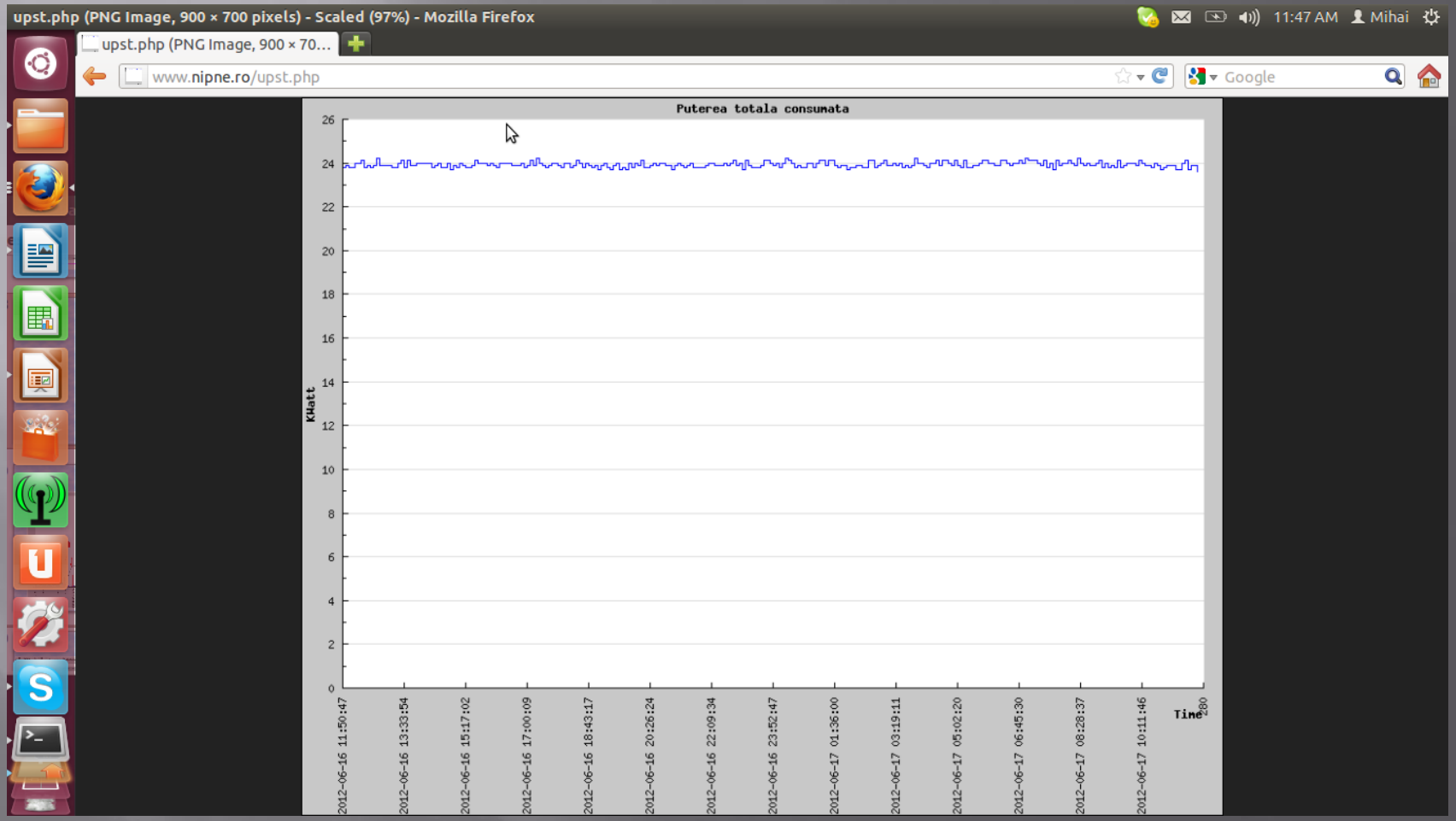
tbit00.nipne.ro	org.sam.SRM-All-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	3d 14h 55m 27s	1/3	OK: Success. [Sun Jun 17 10:25:43 EEST 2012]
	org.sam.SRM-Del-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: File was deleted from SRM.
	org.sam.SRM-Get-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: File was copied from SRM. Diff successful.
	org.sam.SRM-GetSURLs-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: Got SRM endpoint(s) and Storage Path(s) from BDII
	org.sam.SRM-GetTURLs-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: protocols OK-[rfio, gsiftp]
	org.sam.SRM-Ls-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: listing [/dpm/nipne.ro/home/ifops/testfile-put-1339917943-6c38f73cf3a5.txt]-ok:
	org.sam.SRM-LsDir-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: Storage Path[/dpm/nipne.ro/home/ifops]-ok:
	org.sam.SRM-Put-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	15d 23h 2m 26s	1/3	OK: File was copied to SRM.
	ro.nipne.SRM-CertLifetime	OK	06-17-2012 11:27:45	28d 1h 18m 26s	1/3	CERT LIFETIME OK - Certificate will expire in 312 days.
tbit01.nipne.ro	hr.srce.CADist-Check-/ifops/Role-lcgadmin	OK	06-17-2012 11:25:45	13d 23h 56m 26s	1/3	wn1.nipne.ro: OK ! CA DIST OK - The official version is 1.48. Using AAACertificateServices to get CA distribution version. CA distribution is correctly installed.
	org.sam.CE-JobState-/ifops/Role-lcgadmin	OK	06-17-2012 11:27:48	0d 14h 57m 23s	1/3	OK: Running - Submitted [2012-06-17T10:20:36]
	org.sam.CE-JobSubmit-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:50	3d 23h 4m 22s	1/3	OK: success.
	org.sam.WN-Csh-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	397d 23h 25m 4s	1/3	wn1.nipne.ro: OK
	org.sam.WN-Rep-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	33d 19h 36m 32s	1/3	wn1.nipne.ro: OK
	org.sam.WN-RepCr-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	33d 19h 36m 32s	1/3	wn1.nipne.ro: OK: File was copied to SE and registered in LFC lfc.nipne.ro. CLI
	org.sam.WN-RepDel-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	33d 19h 36m 32s	1/3	wn1.nipne.ro: OK: Replicas for [lfn:/grid/ifops/SAM/sam-lcg-rm-cr-wn1.nipne.ro.120617072104.9273476] were deleted. CLI
	org.sam.WN-RepFree-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	36d 21h 33m 32s	1/3	wn1.nipne.ro: OK: ok
	org.sam.WN-RepGet-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	33d 19h 36m 32s	1/3	wn1.nipne.ro: OK. File was copied from SRM. Diff successful. CLI
	org.sam.WN-ReplSenv-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:45	397d 23h 25m 4s	1/3	wn1.nipne.ro: OK: LCG_GFAL_INFOSYS is set to bdii.nipne.ro
	org.sam.WN-RepRep-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:44	33d 19h 36m 32s	1/3	tbit01.nipne.ro
	org.sam.WN-SoftVer-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:45	397d 23h 25m 4s	1/3	wn1.nipne.ro: OK: 3.2.0
	ro.nipne.CE-CertLifetime	OK	06-17-2012 11:26:45	15d 0h 16m 26s	1/3	CERT LIFETIME OK - Certificate will expire in 298 days.
tbit03.nipne.ro	hr.srce.CADist-Check-/ifops/Role-lcgadmin	OK	06-17-2012 11:26:45	54d 20h 16m 20s	1/3	wn90.nipne.ro: OK ! CA DIST OK - The official version is 1.48. Using AAACertificateServices to get CA distribution version. CA distribution is correctly installed.

RO-07-NIPNE: MONITORING



RO-07-NIPNE: MONITORING

RO-07-NIPNE: MONITORING



RO-07-NIPNE: MONITORING

Conclusions:

- February, March and April lots of network problem – independently of our Instiut
- End of April, beginning of May – power network problem in the data center
- Facing a new type of problem – if a DPM-disk dies, dies also the dpm deamon on the DPM-MySQL server
- More stable from beginning of May
- Never ran more than 150 analy jobs simultaneously , why?
- Power consumption in our data centers(there is also a Parallel Computing cluster) with 1100 jobs running: 120KWh

RO-07-NIPNE

THANK YOU!

On behalf of

Serban Constantinescu

Teodor Ivanoaica

Mihai Ciubancan