HP-SEE

HPC cluster at IICT-BAS and HP-SEE infrastructure

www.hp-see.eu

Emanouil Atanassov
Grid Technologies and Applications
Institute of Information and Communication Technologies
Bulgarian Academy of Sciences
emanouil@parallel.bas.bg

HP-SEE

High-Performance Computing Infrastructure for South East Europe's Research Communities

Outline

HP-SEE

High-Performance Computing Infrastructure for South East Europe's Research Communities

- Objectives of HP-SEE operations
- Regional HPC Infrastructure – present and future
- High Performance Cluster at IICT-BAS
- Obtaining access
- Tools and services
- Conclusions

Objectives of HP-SEE operations

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- **WP5 – Regional HPC operations and interoperation** – will ensure that integrated services will be provided to end users, by deploying the regional HPC infrastructure on top of the existing networking infrastructure, complementary to the existing Grid infrastructure, and fused with end user services.
- Specifically, user community needs in terms of size and availability of HPC resources will be catered for. Current HPC installations in Bulgaria, Romania and Hungary will be integrated at the first stage, followed by the upcoming purchases in Greece and Serbia.
- These will form the backbone of the state-of-the-art regional infrastructure, which will be operated jointly by the project, using a set of operations and management tools that will be supported by WP5.
- The infrastructure will be open for use of the wider user community from the region, not only infrastructure provision countries

HPC Resource Centers

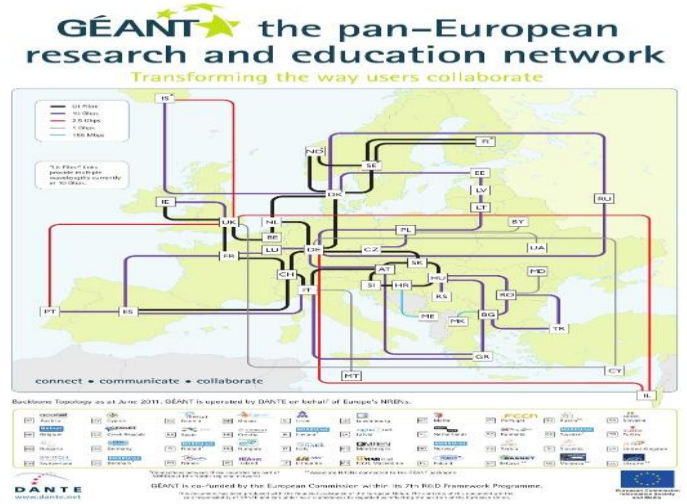
HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

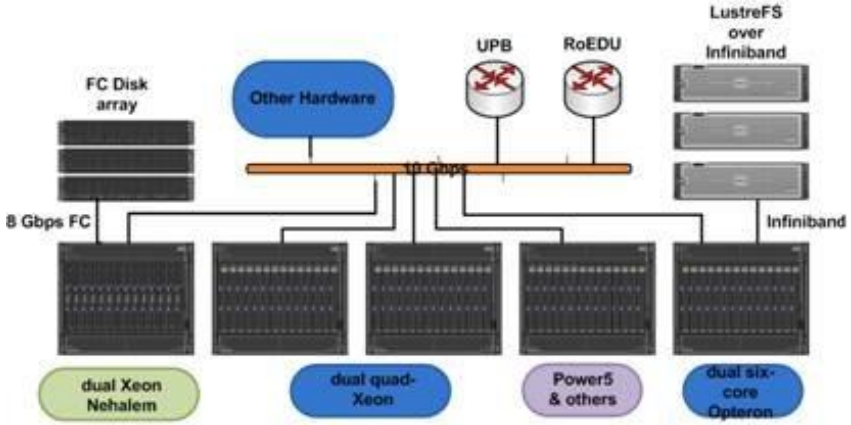
Country	Center	Computing Cores	Teraflops
Bulgaria			
	BG Blue Gene/P	8192	27.85
	HPCG	576	3.23
Macedonia			
	FINKI SC	2016	9
Hungary			
	NIIFI SC	144	0.5
	Pecs SC	1152	10
	Debrecen SC	3078	18
	Szeged	2112	14
Romania			
	InfraGRID	400	2.5
	IFIN_BIO	256	2.72
	IFIN_BC	368	3.9
	NCIT	562	3.4
	UVT Blue Gene/P	4096	13.9
Serbia			
	PARADOX	672	6.26
TOTAL		**23624**	**115.26**

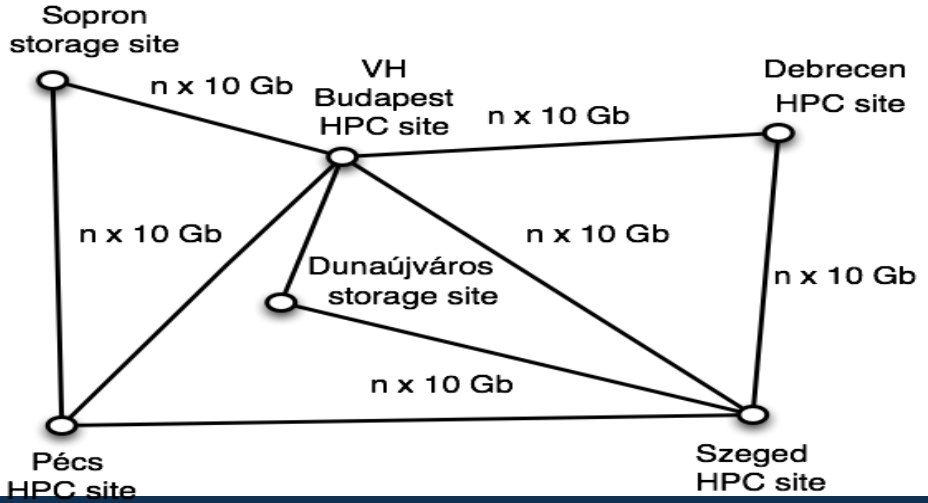
WP5.1 – Network Resource Provision

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- Most of the countries in the region are part of GÉANT. Interconnection of resource provider countries:
  - BG, RS – 1 Gbps
  - GR, RO, HU – 10 Gbps
  - MK – 300 Mbps
- HPC Centers interconnection example:







Blue Gene/P (BG)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Bulgaria**
Administrative Data	
System Name	BG
System Short Description	IBM Blue Gene/P
Computational Power	
Number of nodes	2048
CPU	PowerPC 450 processors (32 bits, 850 MHz)
RAM	4 GB per node
Max number of parallel processes	8192 cores
Interconnect type	IBM proprietary
Interconnect latency	2.5 μs
Interconnect bandwidth	10 Gbps
Peak performance (Tflops, double precision)	27.85
Operating system	CNL
Batch system	LoadLeveler

UVT Blue Gene/P (RO)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Romania**
Administrative Data	
System Name	UVT BlueGene/P
System Short Description	BG/P Supercomputer at UVT
Computational Power	
Number of nodes	1024
CPU	4x PowerPC 450 850Mhz L3 Cache: 8MB
RAM	4GB / node
Max number of parallel processes	4096
Interconnect type	Torus Network - peer-to-peer comm. Collective Network - all-to-all comm.
Interconnect latency	Torus Network: 100ns (32B packet), 800ns (256B packet) Collective Network: 3.0 µs
Interconnect bandwidth	Torus Network: 41Gbps Collective Network: 13.6Gbps
Peak performance (Tflops, double precision)	13.9
Operating system	Special Linux Operating System
Batch system	LoadLeveler

Debrecen SGI Altix (HU)

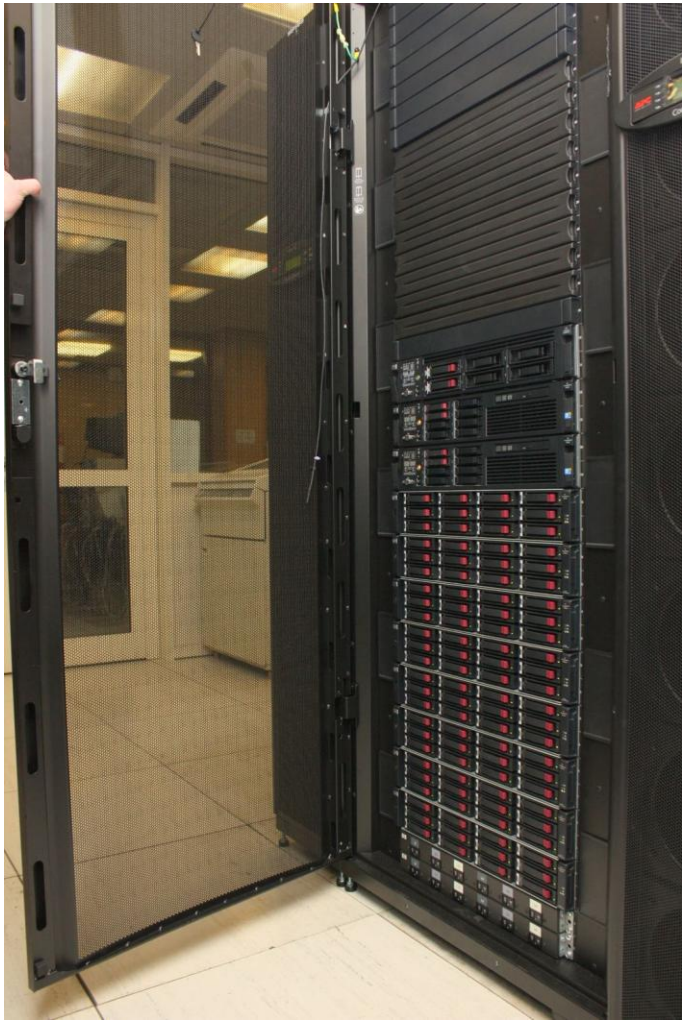
HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Hungary**
Administrative Data	
System Name	Debrecen Supercomputing Center
System Short Description	SGI Altix ICE8400EX supercomputer cluster at Debrecen's site
Computational Power	
Number of nodes	128
CPU	Intel Xeon X5680 (Westmere EP), @ 3.33GHz
RAM	48 GByte per node
Max number of parallel processes	3072
Interconnect type	QDR 4x Infiniband
Interconnect latency	1.1 μs
Interconnect bandwidth	40Gbps
Peak performance (Tflops, double precision)	18
Operating system	SUSE Linux Enterprise Server 11 SP1 (x86_64)
Batch system	Sun Grid Engine 6.2u5

Szeged HP CP4000BL (HU)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Hungary**
Administrative Data	
System Name	Szeged Supercomputing Center
System Short Description	HP CP4000BL blade supercomputer cluster at Szeged's site
Computational Power	
Number of nodes	44
CPU	AMD Opteron 6174
RAM	126 GByte per node
Max number of parallel processes	2112
Interconnect type	QDR 4x Infiniband
Interconnect latency	1.1 μs
Interconnect bandwidth	40Gbps
Peak performance (Tflops, double precision)	14
Operating system	Red Hat Enterprise Linux Server release 5.4 (Tikanga)
Batch system	Sun Grid Engine 6.2u5



Pécs SGI UltraViolet (HU)

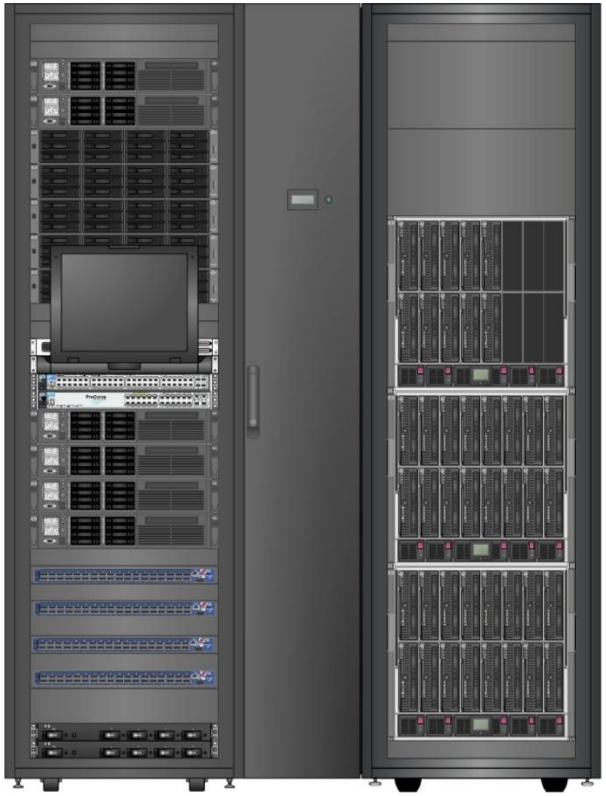
HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Hungary**
Administrative Data	
System Name	Pécs Supercomputing Center
System Short Description	SGI UltraViolet 1000 supercomputer at Pécs's site
Computational Power	
Number of nodes	1
CPU	Intel Xeon X7542 (Nehalem EX), @ 2.67GHz
RAM	6 TByte
Max number of parallel processes	1152 cores
Interconnect type	NUMAlink 5, paired node 2D torus
Interconnect latency	<1 μs
Interconnect bandwidth	15 GByte/sec
Peak performance (Tflops, double precision)	10
Operating system	SUSE Linux Enterprise Server 11 SP1 (x86_64)
Batch system	Sun Grid Engine 6.2u5

HP HPC Cluster(MK)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**FYR of Macedonia**
Administrative Data	
System Name	FINKI Supercomputing center
System Short Description	HP HPC Cluster at FINKI
Computational Power	
Number of nodes	84
CPU	Intel Xeon L5640 @ 2.26GHz
RAM	2016 MB
Max number of parallel processes	2016 (1008 cores with HT)
Interconnect type	QDR Infiniband, Fat tree
Interconnect latency	1 μs
Interconnect bandwidth	40 Gbps
Peak performance (Tflops, double precision)	9
Operating system	Scientific Linux
Batch system	



PARADOX (RS)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	Serbia
Administrative Data	
System Name	PARADOX
System Short Description	High performance cluster at the Institute of Physics Belgrade
Computational Power	
Number of nodes	84
CPU	Intel Xeon E5345@2.33GHz
RAM	8GB per node
Max number of parallel processes	672
Interconnect type	Gigabit Ethernet
Interconnect latency	50μs
Interconnect bandwidth	1 Gbps
Peak performance (Tflops, double precision)	6.26
Operating system	Scientific Linux 5.5 64 bit
Batch system	torque + maui

HPCG (BG)

HP-SEE
High-Performance Computing Infrastructure
Europe's Research Communities

Country	Bulgaria
Administrative Data	
System Name	HPCG
System Short Description	High performance cluster
Computational Power	
Number of nodes	36
CPU	Intel Xeon X5560 @2.8Ghz
RAM	24GB per node
Max number of parallel processes	576
Interconnect type	DDR Infiniband
Interconnect latency	2.5 μs
Interconnect bandwidth	20Gbps
Peak performance (Tflops, double precision)	3.23
Operating system	Scientific Linux 5.3 64 bit
Batch system	torque + maui

NCIT (RO)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	Romania
Administrative Data	
System Name	NCIT Computing Cluster
System Short Description	High Performance Computing Cluster at UPB
Computational Power	
Number of nodes	56
CPU	28*HS21 Dual Intel Quad-Core Xeon E5504 @2.00Ghz 20*LS22 Dual Opteron Six-Core AMD Processor 2435 @2.6Ghz 4*QS22 Dual IBM PowerXCell 8i @3.2GHz 4*HS22 Dual Intel Hex-Core Xeon E5630 @2.5GHz
RAM	16GB per node HS21,LS22 8GB per node QS22 32GB per node HS22
Max number of parallel processes	562
Interconnect type	LS22 - QDR 4x Infiniband Other – 10 GigabitEthernet
Interconnect latency	2.5 μs
Interconnect bandwidth	Infiniband – 40Gbps Cluster Backbone – 10Gbps Minimum inter-chassis bandwidth – 2x4Gbps
Peak performance (Tflops, double precision)	3.4
Operating system	Scientific Linux 5.5 64 bit
Batch system	SunGridEngine 6.2u5 torque + maui (on gLite)

IFIN_BC (RO)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Romania**
Administrative Data	
System Name	IFIN_BC
System Short Description	IBM BladeCenter
Computational Power	
Number of nodes	26 (16xQS22 + 10xLS22)
CPU	QS22: 2x IBM PowerXCell 8i @3.2GHz; LS22: 2x AMD Opteron Quad Core 2376 HE @2.3GHz
RAM	32 GB per QS22; 8GB per LS22
Max number of parallel processes	368 (32xPPE + 256xSPE + 80xAMD)
Interconnect type	Infiniband 4X DDR
Interconnect latency	2.5 μs
Interconnect bandwidth	20Gbps
Peak performance (Tflops, double precision)	3.28 for QS22; 0.62 for LS22
Operating system	Fedora 9 PPC for QS22, Fedora 9 x86_64 for LS22
Batch system	torque

IFIN_Bio (RO)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Romania**
Administrative Data	
System Name	IFIN_Bio
System Short Description	High Performance Biocomputing Cluster
Computational Power	
Number of nodes	32
CPU	2x Intel Xeon E5430 (Quad-Core) @2.67GHz
RAM	16GB per node
Max number of parallel processes	256 cores
Interconnect type	Myrinet 2000
Interconnect latency	3.2 μs
Interconnect bandwidth	2Gbps
Peak performance (Tflops, double precision)	2.72
Operating system	CentOS 5.4 64 bit
Batch system	torque

InfraGRID (RO)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Romania**
Administrative Data	
System Name	InfraGRID
System Short Description	High Performance Computing Cluster at UVT
Computational Power	
Number of nodes	50
CPU	50*HS22 Dual Intel Quad-Core Xeon E5504 @2.00Ghz
RAM	10GB per node HS22
Max number of parallel processes	400
Interconnect type	HS22 - QDR 4x Infiniband Other – 1 GigabitEthernet
Interconnect latency	1.1 μs
Interconnect bandwidth	Infiniband – 40Gbps Cluster Backbone – 8Gbps Inter-chassis bandwidth – 4x1Gbps
Peak performance (Tflops, double precision)	2.5
Operating system	CentOS 5.6 64 bit
Batch system	CondorHTC 7.4.4

NIIFI (HU)

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Country	**Hungary**
Administrative Data	
System Name	NIIFI Supercomputing Center
System Short Description	SUN E15K supercomputer cluster at NIIFI's site
Computational Power	
Number of nodes	2
CPU	US-III+ @1.2GHz, US-IV+ @1.8Ghz
RAM	158Gbyte, 286GByte
Max number of parallel processes	144 cores
Interconnect type	2 x Gigabit Ethernet
Interconnect latency	cca 300 µs
Interconnect bandwidth	1 Gbps
Peak performance (Tflops, double precision)	0,5
Operating system	SUN Solaris 9
Batch system	Sun Grid Engine

HPC Infrastructure - RS

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- PARADOX high performance cluster consists of 84 worker nodes (2 x quad core Xeon E5345 @ 2.33 GHz with 8GB of RAM). PARADOX is the largest HPC cluster in Serbia. Its computing nodes are interconnected by the star topology Gigabit Ethernet network through three stacked high-throughput Layer 3 switches, each node being connected to the switch by two Gigabit Ethernet cables in channel bonding. In terms of storage resources, PARADOX provides up to 50 TB of disk space to the HP-SEE community.
- tPARADOX training cluster for educational purposes, based on IBM's BladeCenter technology and it consists of IBM BladeCenter H chassis commonly used in high performance computing and different types of Blade servers that cover some of the major CPU architectures currently available: Intel's x86_64 and IBM's POWER and Cell/B.E.
- As part of the 10 Million Euro project for building supercomputing resources in Serbia, the PARADOX cluster will be expanded significantly in 2011.

HPC Infrastructure – GR Planned

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- GRNET has already submitted a proposal to GSRT for the creation of a national HPC center with a budget of around 10 Million Euro. Based on the conclusions of the HellasHPC project and the deliverable that describes the Greek strategy for HPC development the national HPC system is required to have the following technical characteristics
- Computational Power RMax (Linpack) 250-300 TFlops, RPeak: ~350 TFlops
- Number of processing elements 35.000 - 40.000 CPU cores
- Memory Size > 76 TB (at least 2GB per core)
- Storage Size 3 PetaByte
- Interconnect High Speed (>10 Gbit) with low lattency
- Initially a seed resource of about 40 Tflops will be integrated to HPSEE project.

HPC Cluster at IICT-BAS

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- HP Cluster Platform Express 7000 enclosures with 36 blades BL 280c with dual Intel Xeon X5560 @ 2.8Ghz (total **576** cores), 24 GB RAM per blade
- 8 controlling nodes HP DL 380 G6 with dual Intel X5560 @ 2.8 Ghz, 32 GB RAM
- Non-blocking DDR Interconnection via Voltaire Grid director 2004
- Two SAN switches for redundant access
- MSA2312fc with 48 TB storage, Lustre filesystem
- P2000 G3 with 48 TB storage added last week
- More than 92% efficiency on LINPACK (>3 TFlops, peak performance 3.2TFlops)

HPC Cluster at IICT-BAS

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- Extension cluster with 4 GPU cards NVIDIA GTX 295 (each card counts as 2 graphical devices), CPU Intel Core i7 @2.66 Ghz, 12 GB RAM.
- Total number of threads for GPU computing – 4x2x240=1920
- High performance Lustre filesystems:
  - /home – 22 TB
  - /scratch – 7 TB

Installed software at IICT

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

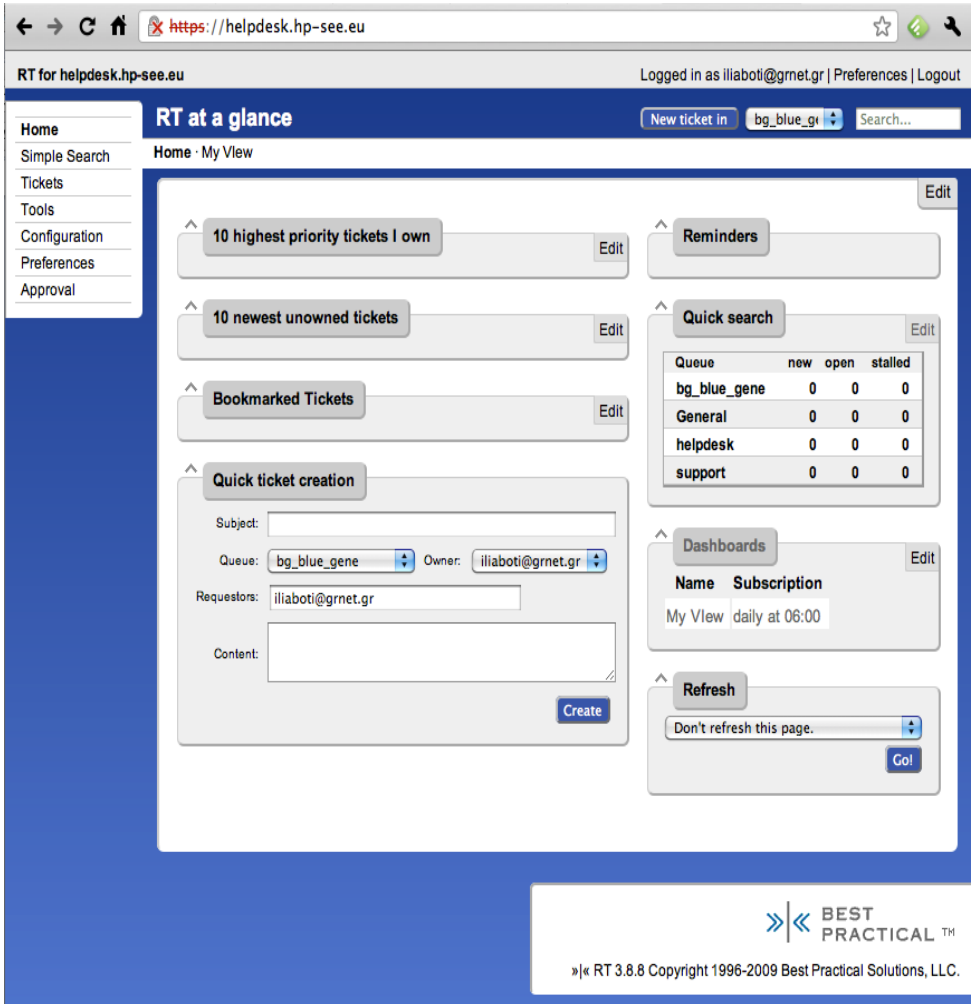
- Torque batch system
- Glite Grid middleware
- Unicore Grid middleware
- NVIDIA GPU computing SDK
- Compilers, MPI(mviapich1, mviapich2, openmpi), debuggers and profilers (MPE, scalasca, mpiP)
- ATLAS, LAPACK, HPL, ScaLAPACK, GotoBLAS, FFTW, SPRNG, MPI (MVIAPICH1/2, OpenMPI), BLACS, BLAS, CUDA, OpenCL, OpenFOAM, octave
- Charm++, GAMESS, GROMACS, NAMD, NWChem, ABINIT, WRF, CMAQ, SMOKE
- Compilers – Intel Cluster Studio, PGI
- Maple

Helpdesk

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

Helpdesk for HP-SEE users is operational at https://helpdesk.hp-see.eu

Other users of HPC cluster at IICT-BAS should use hpcg-support@bas.bg



Access to Grid

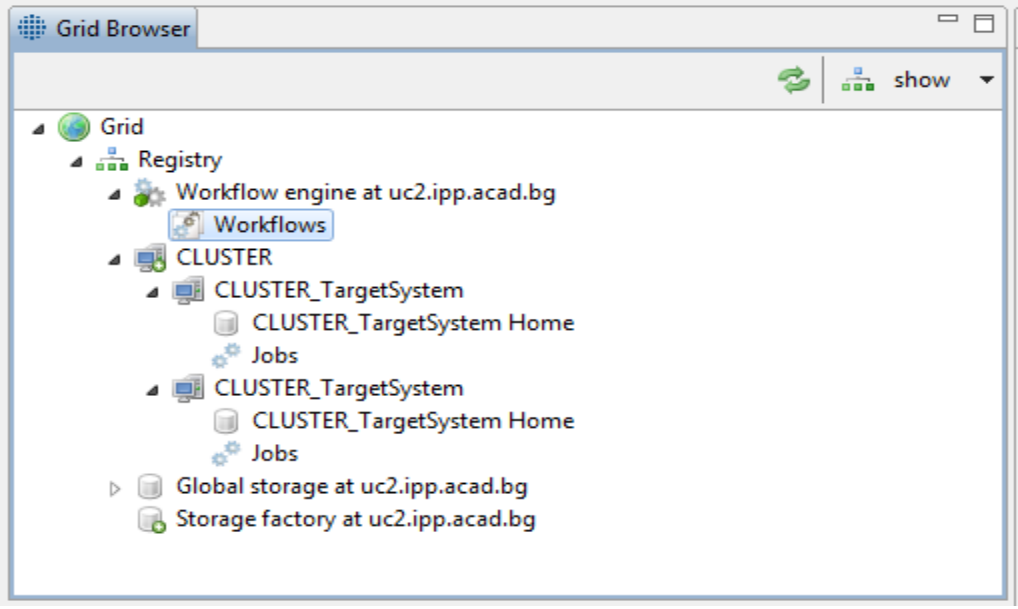
HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- Obtaining access to the infrastructure of the European Grid Intiative requires:
  - certificate (http://ca.acad.bg)
  - Membership in a virtual organization, that can be:
    - European
    - Regional
    - Bulgarian
- Most of the HPC clusters in the region are also available via Grid middleware
- The gateway node of the HPC cluster at IICT-BAS can serve as Grid User Interface

Unicore usage example

HP-SEE

Site name	Address	Requests served	Status	Message
WORKFLOW	https://uc3.ipp.acad.bg:7700	1637		OK
CLUSTER	https://uc3.ipp.acad.bg:7777	777		OK
REGISTRY	https://uc1.ipp.acad.bg:7778	3117		OK
SERVORCH	https://uc3.ipp.acad.bg:7701	2477		OK




```
[dgeorgiev@wn02 ucc-1.3.1]$ ./bin/ucc connect
You can access 1 target system(s).
[dgeorgiev@wn02 ucc-1.3.1]$ ./bin/ucc run samples/date.u
SUCCESSFUL exit code: 0
/home/dgeorgiev/ucc-1.3.1/./outs/3dc88cd4-0faf-48b6-ab0e-d7360a5031a3.stdout
/home/dgeorgiev/ucc-1.3.1/./outs/3dc88cd4-0faf-48b6-ab0e-d7360a5031a3.stderr
/home/dgeorgiev/ucc-1.3.1/./outs/3dc88cd4-0faf-48b6-ab0e-d7360a5031a3.properties
[dgeorgiev@wn02 ucc-1.3.1]$
```

Access to HP-SEE resources

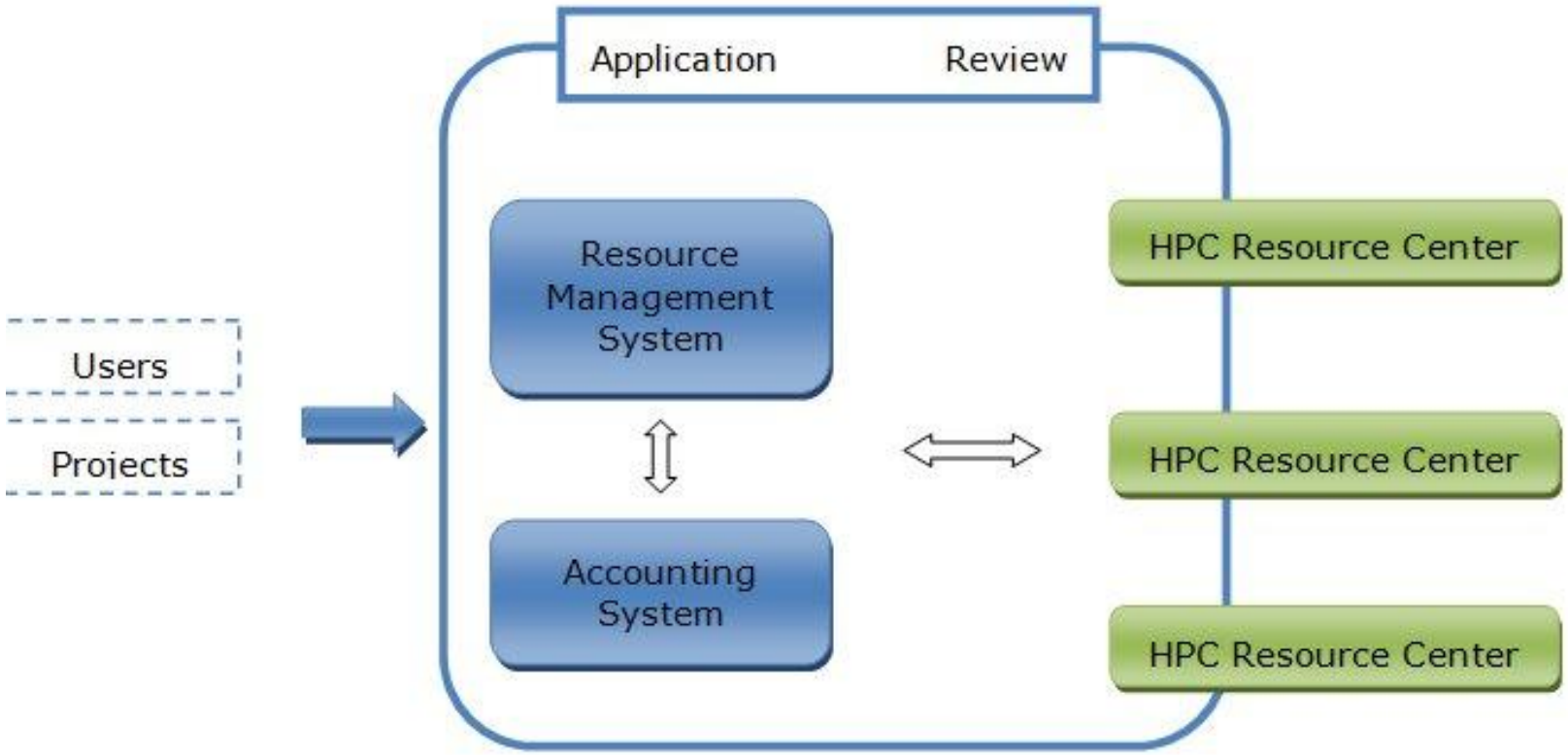
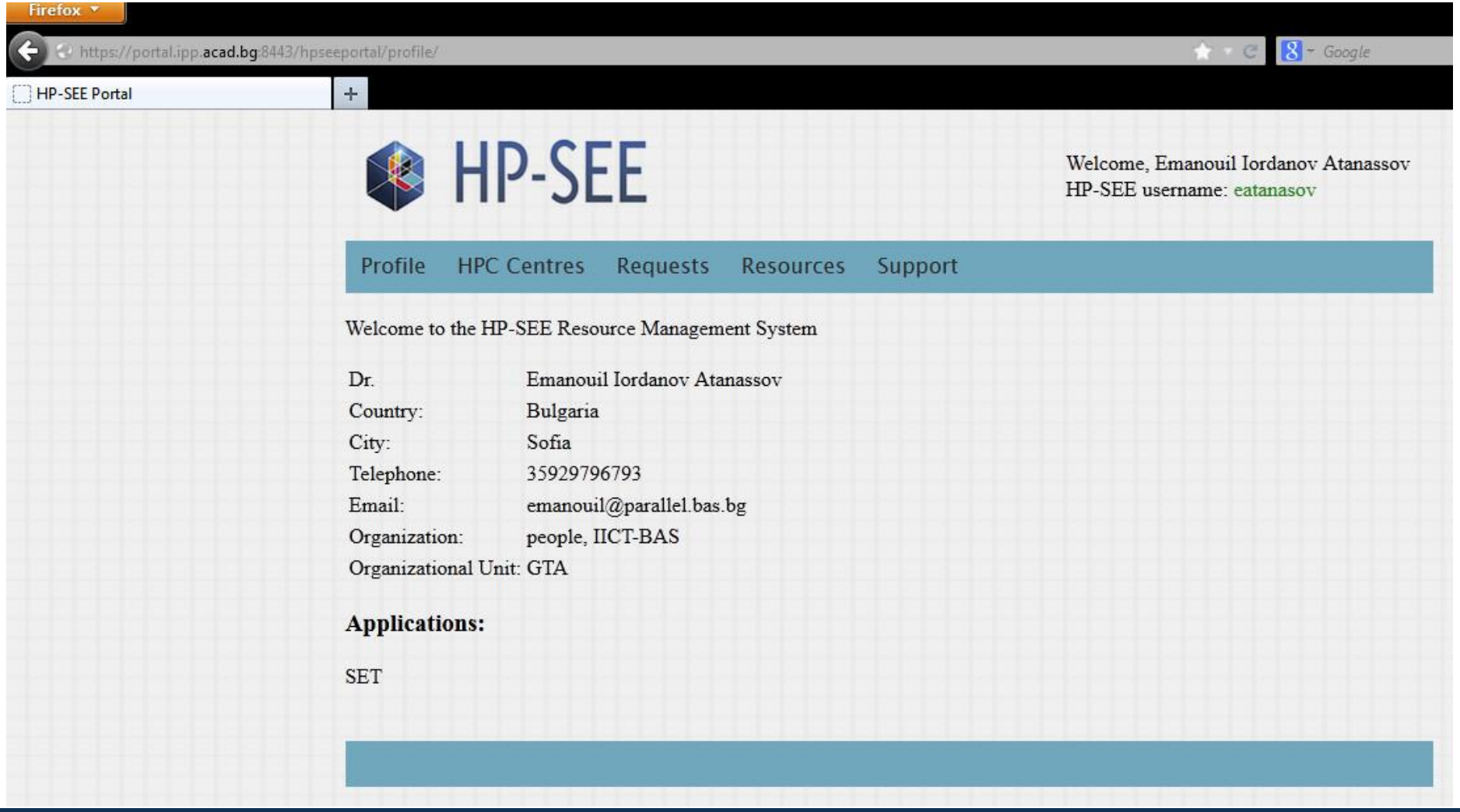


Figure 8. Resource Management

Access to HP-SEE resources

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities



Direct access to HPC cluster at IICT-BAS

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- User that have filled the access form http://www.grid.bas.bg/gta/projects/HP-SEE/form-accountHPCG_bg.doc
- and obtained direct access can log in to the user interface **gw.ipp.acad.bg** with username and password.
- All worker nodes are accessible with ssh without password
- The login node should be used only for compilation and light testing.
- Computational jobs should be submitted to the appropriate queue via **qsub**
- Additional software usually installed under /opt/exp_software

Direct access to HPC cluster at IICT-BAS

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- The /home file system is under raid6 and should be used for permanent storage.
- The /scratch file system should be used for temporary files and directories, especially when high amount of I/O is performed. Files can be removed by an administrator at any time. For each user there is /**scratch/*username***
- The GPU cluster is separate, but can be accessed with same username and password as the main one (see next presentation).
- An appropriate queue will be assigned

Rules for access

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- Do not waste resources – kill runaway jobs, etc.
- Report suspicious behavior immediately.
- Clean-up scratch space after usage
- Request priority reservations if necessary

Conclusions

HP-SEE
High-Performance Computing Infrastructure for South East Europe's Research Communities

- The current High Performance infrastructure in the region includes one supercomputer Blue Gene in Sofia and several clusters.
- Several neighboring countries plan significant expansion of their HPC capabilities.
- By sharing of the computational resources scientists from countries that do not posses such resources can develop, deploy and use HPC applications for computationally intensive research.