The Czech National Grid Infrastructure

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082042 doi:10.1088/1742-6596/898/8/082042

The Czech National Grid Infrastructure

J Chudoba, I Křenková, M Mulač, M Ruda, J Sitera

CESNET, z. s. p. o., Zikova 1903/4, 160 00 Praha 6, Czech Republic

E-mail: Jiri.Chudoba@cern.ch

Abstract. The Czech National Grid Infrastructure is operated by MetaCentrum, a CESNET department responsible for coordinating and managing activities related to distributed computing. CESNET as the Czech National Research and Education Network (NREN) provides many e-infrastructure services, which are used by 94% of the scientific and research community in the Czech Republic. Computing and storage resources owned by different organizations are connected by fast enough network to provide transparent access to all resources. We describe in more detail the computing infrastructure, which is based on several different technologies and covers grid, cloud and map-reduce environment. While the largest part of CPUs is still accessible via distributed torque servers, providing environment for long batch jobs, part of infrastructure is available via standard EGI tools in EGI, subset of NGI resources is provided into EGI FedCloud environment with cloud interface and there is also Hadoop cluster provided by the same e-infrastructure.A broad spectrum of computing servers is offered; users can choose from standard 2 CPU servers to large SMP machines with up to 6 TB of RAM or servers with GPU cards. Different groups have different priorities on various resources, resource owners can even have an exclusive access. The software is distributed via AFS. Storage servers offering up to tens of terabytes of disk space to individual users are connected via NFS4 on top of GPFS and access to long term HSM storage with peta-byte capacity is also provided. Overview of available resources and recent statistics of usage will be given.

1. About the MetaCentrum project

MetaCentrum [1] is a CESNET project started already 20 years ago. It started at the Masaryk University in Brno as an activity to connect a very few supercomputing resources available for academic research at that time. Since then it evolved into the main Czech distributed computing infrastructure used by hundreds of researchers. It operates and manages computing and storage resources owned by e-Infrastructures CESNET and CERIT-SC as well as co-operative academic centers in the Czech Republic.

MetaCentrum is responsible for building the Czech National Grid Infrastructure (NGI) as a part of the European Grid Infrastructure (EGI). It also manages and develops the national ELIXIR infrastructure and integrates it to related international capacities.

In 2016 MetaCentrum operated resources owned by 10 academic institutions. The geographical distribution of resources is given in Figure 1. Users from more than 40 Czech universities and research institutions could use provided application software for wide range of disciplines: computing chemistry, bioinformatics, climate modeling, material sciences, astrophysics, etc. According their needs they could access Hadoop, Chipster or Galaxy platforms.

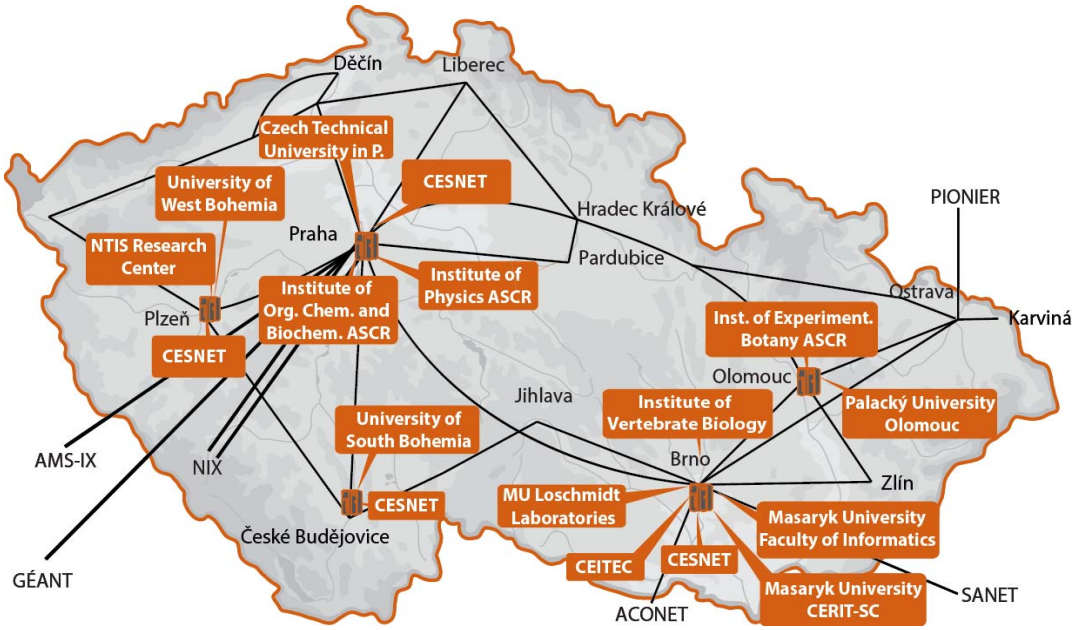

Figure 1: MetaCentrum resources are distributed in many towns within the whole Czech Republic.

2. Resources

A broad spectrum of computing servers is offered; users can choose from standard 2 CPU servers to large SMP machines with up to 6 TB of RAM or servers with GPU cards. Different groups have different priorities on various resources, resource owners can even have an exclusive access. The software is distributed via AFS. Storage servers offering up to tens of terabytes of disk space to individual users are connected via NFS4 on top of GPFS and access to long term HSM storage with peta-byte capacity is also provided.

Table 1: Usage statistics of MetaCentrum resources for 2015.

1. 1. - 31. 12. 2015	MetaCentrumVO (local resources)				EGI
	Total	Torque job scheduler	Cloud	Hadoop	Torque job scheduler
Number of jobs / VMs running	2943066	2837135	25674	80257	5059899
CPU time usage [CPU years]	7685	6529	1136	20	3228
Number of users	1338	1338	129	17	5607
Number of CPUs (end of 2015)	12408	11992 (shared resources between Torque and Cloud)		416	3312
Storage	2.4 PB	829 TB used	23 TB used	289 TB	3.8 PB
Acknowledgements to NGI in users' publications		362			n/a

3. Services

The most visible services for users are computing and storage resources. However many other resources are operated to administer these resources and provide a user support. Monitoring based on Nagios alerts administrators to various issues with hardware or non-optimal use of resources. Request tracker system RT [2] accepts user questions and requests and ensures that all are properly dealt with. The system is also interfaced to EGI GGUS request tracking system. Wiki server [3] can be used by any user to document procedures and best practices for their groups.

3.1. Batch system

Torque based batch system was used for many years. Many changes to the original Torque enhanced stability and added features required by users and geographically distributed resources. Some of the changes were propagated back to the main branch of the Torque project. The system was coupled to either a high-performance queue-based scheduler using backfilling or a planning-based scheduler that uses advanced schedule-optimizing metaheuristic [4]. The usage statistics for 2015 is given in Tab. 1; Figure 2 demonstrates that users come from many different institutions.

In 2016 we started to use free version of the PBSPro system for new clusters. Old clusters are being gradually migrated to the new system. The main reason for this change is a reduction of maintenance cost and long-term stability.

3.2. HPC cloud

Cloud site based on OpenNebula, oriented on high performance computing is available for all MetaCentrum users and via EGI Federated Cloud. The system is intended for use cases where the traditional grid approach with pre-installed nodes does not work.

3.3. Perun

Perun system [5] manages users, groups and access to resources in distributed environments. It has been integrated with a number of infrastructures and projects on national and international levels. It holds a database with all MetaCentrum users and manages a periodic renewal of membership. User accounts are kept active only for users who are still entitled for access to resources and summarize their work done in the previous period.

Accounts of users who do not respond are blocked with a possible later renewal. Perun is also used for a membership management in several EGI Virtual Organizations. Users apply for a membership or its prolongation via a Perun portal. Perun interacts with VOMS services, creates local accounts on shared User Interfaces and adds new users to mailing lists for VO users.

3.4. PBSmon

PBSmon is a web-based framework for monitoring and accounting in the Czech NGI. It visualizes the current state of computing resources, Torque (PBS) state and user personalized job information. The main purpose of PBSmon is to provide an intuitive interface to the complicated infrastructure consisting of clusters of virtualized machines assigned to various job queues.

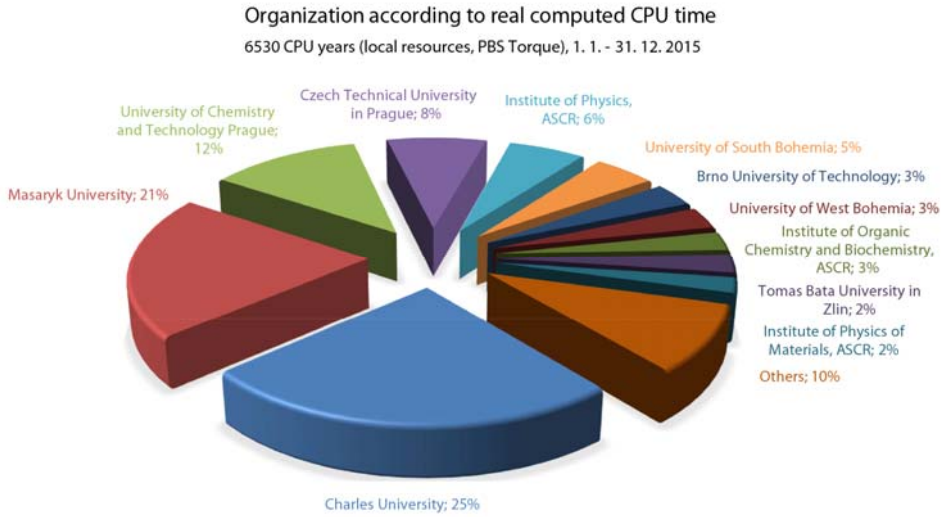

Figure 2: Distribution of users' home institutions according the used computing time.

4. Most important international projects

MetaCentrum researchers and administrators use their experience in high performance distributed computing in many international projects. Here we list selected major projects:

- **ELIXIR–Excelerate**: Development and integration activities in Compute Platform workpackage, with aim of AAI in clouds.
- **EGI–Engage**: Provisioning of all necessary services to allow seamless integration of national resources and local end users into pan-European grid infrastructure.
- **EGI Federated Cloud**: Coordination and implementation of technical solutions in cloud federation scenarios.
- **INDIGO DataCloud**: Coordination and development activities in project's build ant pilot services, virtual networks, and AAI.
- **AARC**: Expertise in the area of group management and federated access for non-web applications.

Leveraging the grid resources efficiently requires non-trivial effort by both the application area and IT experts. We look for such collaborations actively, involving also both under- and post-graduate students of computer science. The outcomes of this work are scientific results in the application areas, which would not be possible without the wide usage of computational resources, as well as more generally applicable results in the computer science itself. The scientific areas we work with include, but are not limited to, bioinformatics, chemistry, structural biology, neurology, earth observation, and astronomy.

5. Current challenges and future plans

The most often user complaint is a long waiting time for long jobs, which is given by a combination of a high user demand and a finite number of resources. We give higher priorities to productive users who publish papers where they acknowledge usage of Metacentrum resources. We also regularly increase the total computing and storage capacity. The gradually improved scheduling of jobs minimizes the number of idle computing cores.

We also plan to introduce IPv6 addresses for computing servers, because now we require a public IPv4 address for each computing node and the lack of IPv4 addresses starts to be an issue for several organizations. This effort is slowed down by a lack of IPv6 support in the batch scheduler and in AFS.

We expect higher future demand for running users' jobs in a specific environment under an OS given by users; for such requests we plan to extend cloud resources and enable usage of containers.

References
[1] Krenkova I et al: Yearbook 2011-12: National Grid Infrastructure Annual Report, 2013 ISBN 978-80-904689-7-9
[2] Request Tracker for CESNET, *https://rt4.cesnet.cz/rt/*
[3] MetaCentrum Wiki server *https://wiki.metacentrum.cz/w/index.php*
[4] Klusacek D: New Multi-ResourceFairness and Performance Metrics for Metacentrum, in [1] 2013
[5] Prochazka M, Licehammer S and Matyska L.: Perun – Modern Approach for User and Service Management, IST-Africa 2014 Conference proceedings, ISBN 978-1-905824-44-1

Acknowledgments
MetaCentrum is supported by the Ministry of Education, Youth and Sports of the Czech Republic under the project CESNET (Project No. LM2015042) provided within the program Projects of Large Research, Development and Innovations Infrastructures.