A multipurpose computing center with distributed resources

IOP Conf. Series: Journal of Physics: Conf. Series **898** (2017) 082034 doi:10.1088/1742-6596/898/8/082034

A multipurpose computing center with distributed resources

J Chudoba[1], M Adam[1], D Adamová[2], T Kouba[1], A Mikula[1], V Říkal[1], J Švec[1], J Uhlířová[1], P Vokáč[1] and M Svatoš[1]

[1]Institute of Physics of the Czech Academy of Sciences, Na Slovance 2, 182 21 Prague 8, Czech Republic
[2]Nuclear Physics Institute of the Czech Academy of Sciences, Řež 130, 25068, Řež, Czech Republic

E-mail: Jiri.Chudoba@cern.ch

Abstract. The Computing Center of the Institute of Physics (CC IoP) of the Czech Academy of Sciences serves a broad spectrum of users with various computing needs. It runs WLCG Tier-2 center for the ALICE and the ATLAS experiments; the same group of services is used by astroparticle physics projects the Pierre Auger Observatory (PAO) and the Cherenkov Telescope Array (CTA). OSG stack is installed for the NOvA experiment. Other groups of users use directly local batch system. Storage capacity is distributed to several locations. DPM servers used by the ATLAS and the PAO are all in the same server room, but several xrootd servers for the ALICE experiment are operated in the Nuclear Physics Institute in Rez, about 10 km away. The storage capacity for the ATLAS and the PAO is extended by resources of the CESNET - the Czech National Grid Initiative representative. Those resources are in Plzen and Jihlava, more than 100 km away from the CC IoP. Both distant sites use a hierarchical storage solution based on disks and tapes. They installed one common dCache instance, which is published in the CC IoP BDII. ATLAS users can use these resources using the standard ATLAS tools in the same way as the local storage without noticing this geographical distribution. Computing clusters LUNA and EXMAG dedicated to users mostly from the Solid State Physics departments offer resources for parallel computing. They are part of the Czech NGI infrastructure MetaCentrum with distributed batch system based on torque with a custom scheduler. Clusters are installed remotely by the MetaCentrum team and a local contact helps only when needed. Users from IoP have exclusive access only to a part of these two clusters and take advantage of higher priorities on the rest (1500 cores in total), which can also be used by any user of the MetaCentrum. IoP researchers can also use distant resources located in several towns of the Czech Republic with a capacity of more than 12000 cores in total.

1. Introduction

The Computing Center of the Institute of Physics (CC IoP) of the Czech Academy of Sciences was founded in 2008 to formalize responsibilities for operations and user support activities. Groups from Division of Elementary Particle Physics, Division of Solid State Physics and Division of Condensed Matter Physics used several computing clusters or just individual powerful servers located in different rooms to satisfy their computing needs. Growing requirements of the ATLAS and ALICE experiments were the main driving force to build a dedicated server room in 2004. Computing clusters of different groups were moved to this facility and their maintenance was unified where it was possible. So the groups of solid state physics theoreticians could share resources with high energy physics

experimentalists. This unification was not forced in special cases of dedicated clusters for parallel computing, where user requirements were too diverse.

2. Hardware resources

2.1. Computing servers

The highest computing capacity is provided by the Golias cluster integrated in international Grids EGI/WLCG and OSG. This cluster consists of several nonhomogeneous subclusters as a consequence of gradual hardware upgrades. At the end of 2016, the total computing capacity of this cluster was 4600 jobslots for 1 core jobs provided by 224 servers. Each jobslot has at least 2 GB of memory. We use Hyper-Threading Technology on servers with enough memory, the total number of real cores was 2904. Available computing capacity in HEPSPEC06 units achieved 38600. The evolution of total capacity of resources connected to Grids is graphically visualized on Figure 1.

Three other clusters (Kalpa, LUNA and EXMAG) intended for parallel tasks are connected to the Czech national grid infrastructure MetaCentrum [1]. They are installed and monitored remotely with minimal requirements on a local administration. All nodes of these clusters are connected by Infiniband (each cluster in a separate domain) and allow running tasks requiring up to 736 cores (LUNA). Part of these clusters is shared with all users from MetaCentrum, part is exclusively reserved for local users belonging to projects which paid for these clusters. The fraction of reserved servers can be adjusted based on current or planned demands. Local users submit jobs to privileged queues with an absolute priority on all nodes, however already running jobs of other users are not suspended. The total capacity of these three clusters is 1440 cores. Local users can also use other MetaCentrum resources distributed in several places in the Czech Republic.

The computing center also hosts several group servers with relatively low total computing capacities.

2.2. Storage servers

LHC experiments ATLAS and ALICE determine technologies for most of the storage servers. They require high storage capacities with a moderate reliability. DPM was chosen as an SRM storage for ATLAS already in 2004 based on a personal involvement of one system administrator in the DPM development. It was advertised as a simple solution for relatively small capacities. Since then it has evolved in a solution capable to manage many petabytes and to support many transfer protocols. We deploy current version 1.9 (since December 2016) on all 10 disk pools and also on the head node. The MySQL database runs on a separate dedicated server. The total usable capacity of the DPM system 3.3 PB (raw capacity is 4.1 PB) is dedicated mostly to the ATLAS experiment. Smaller volumes are used by the Pierre Auger Collaboration and the Cherenkov Telescope Array project. Several hundreds of TB are kept as a reserve. We use this reserve when we move data from a certain server to do a hardware maintenance.

The ALICE experiment uses xrootd servers. Although DPM servers support xrootd protocol, which is now also used by the ATLAS collaboration, we deploy separate xrootd servers only for ALICE. Initial tests done several years ago did not show a good stability when the xrootd for ALICE was provided on a DPM server and although the software improved significantly since then we keep storages for these two major projects independent. The IoP hosts xrootd head node *xrdhead* and 5 disk servers. Three more disk servers are hosted in the Nuclear Physics Institute in Řež, a small town near Prague. A dedicated 10 Gbps optical link connects these two sites. The total usable capacity of xrootd servers exceeds 1.6 PB.

No backup of these servers is provided since usually other copies are available in other Tier centers. An exception is a LOCALGROUPDISK spacetoken reserved for users from Czech institutions. This spacetoken is used as a proxy storage of datasets available also on other sites but copied locally for faster processing on local servers. LOCALGROUPDISK is sometimes used to store outputs of users' jobs which are unique. They may be recovered by running the same procedure on the

input data, but this can be a difficult especially when input data are old or distributed on many sites. Our users can easily backup their LOCALGROUPDISK files by creating another copy on a LOCALGROUPTAPE spacetoken. They use standard rucio interface.

The LOCALGROUPTAPE spacetoken is provided by a dCache instance dcache.du.cesnet.cz. This server is operated by CESNET (the Czech NREN) Data Storage group as an extension to generic storage services for public research. Most of the offered capacity is based on tapes with disk storage for file staging. dCache solution was chosen because it offers a required SRM interface which can handle tape operations. A moderate 20 TB disk space is available but not used since the time when DPM capacity enabled increase of the ATLAS LOCALGROUPDISK to the current 350 TB. The dCache server is installed in Plzeň, about 100 km from Prague (measured by the length of the network connection). Storage facilities of the CESNET Data Storage group are located in other two geographically distant localities (Jihlava, Brno). We test an extension of a single dCache instance by capacities in Jihlava.

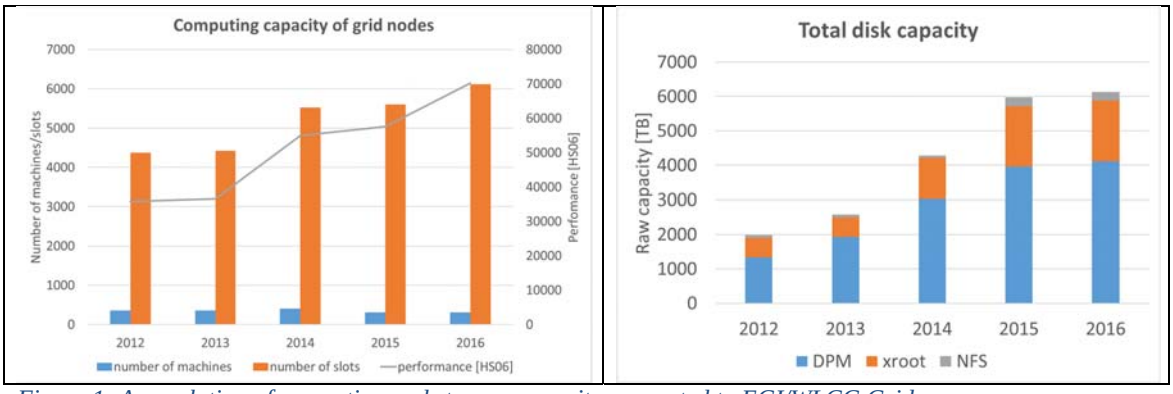

Figure 1: An evolution of computing and storage capacity connected to EGI/WLCG Grid.

2.3. Control servers

Many running services are needed to control resources, monitor them and report correctly to global grid services. We use separate virtual machines for such services to avoid unintended interactions between services. Our virtualization platform is based on KVM. Three identical physical servers, each with 2 CPU (16 cores) and 132 GB of memory are connected to an Optima 1500 disk array via two Fibre Channel channel switches. The design parameters of these servers were such that we were able to run all virtual machines on just 2 servers, so updates were done without downtime after a live migration. Since requirements on memory and disk space increases, we plan to deploy a new virtualization platform based on disk array Hitachi VSP G200 with 5 fast 1.6 TB SSD disks and 43 1.2 TB rotational disks.

Some servers like MySQL DB or syslog run on a bare hardware for performance and security reasons. The new disk array is fast enough to support also applications with high number of input/output operations per seconds (IOPS). It will enable us to setup two database servers in a failover mode. We will also move there monitoring infrastructure, which produces a lot of RTT graphs and suffers with performance problems.

Virtualization platform is extended by two older servers used originally for IPv6 testbed. Since we already deployed IPv6 addresses on all storage servers and worker nodes, we reused this platform for running non-critical services.

3. Networking

All server connections are realized by 1 or 10 Gbps ethernet, only clusters for parallel computing use also Infiniband.

3.1. Internal network

All grid computing servers (worker nodes) are connected by 1 Gbps ethernet to top of the rack switches. These are connected by 10 Gbps ethernet to a Cisco Nexus 3172PQ switch, which has 48 10 Gbps ports. All DPM and xrootd disk servers are also connected to this switch. All worker nodes and storage servers have IPv4 and IPv6 addresses. All local transfers via gridftp protocol utilize preferentially IPv6 protocol. Introduction of IPv6 addresses was done in cooperation with the HEPIX IPv6 working group and was described in several publications [1,2 and references there].

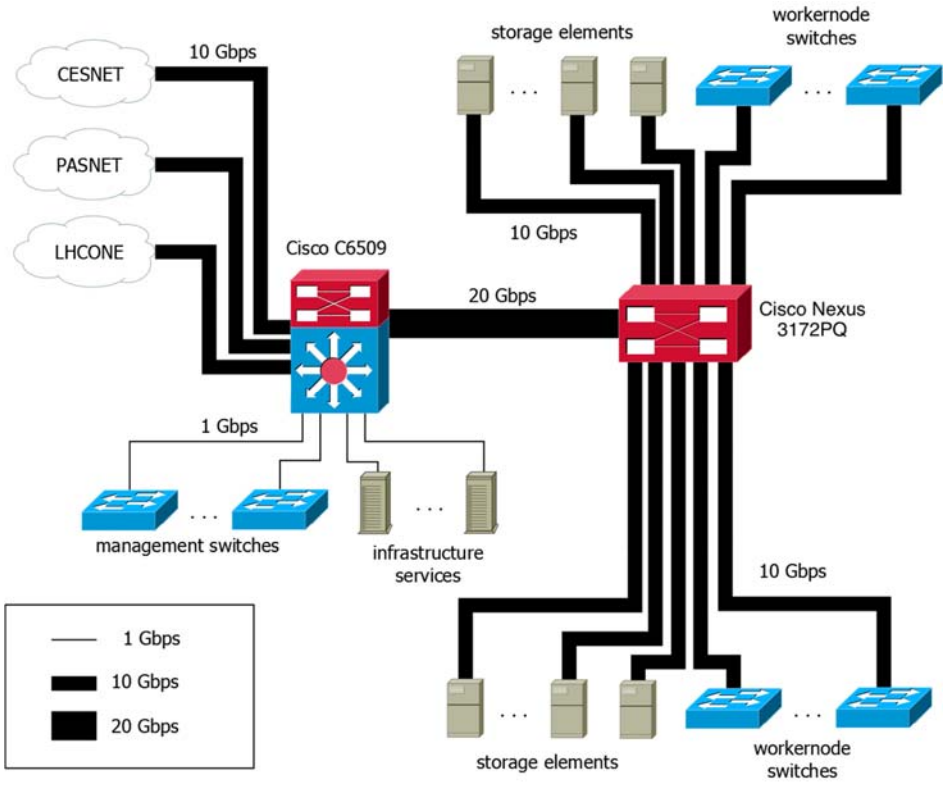

Figure 2: A schematic drawing of the main network components

3.2. Wide area network

Major load on external connection comes from LHC experiments ATLAS and ALICE. The 10 Gbps link provided by CESNET and announced to LHCONE was often saturated and was upgraded to 2x10 Gbps in 2016. Other 10 Gbps direct dedicated connections to several institutions involved in LHC program are prepared. Currently only a link to NPI Řež is used, because there are ALICE xrootd servers there. The special GLIF connection to FERMILAB was cancelled when a major user of this connection, experiment NOvA, joined the LHCONE in 2016. Another 10 Gbps link is available for transfers to sites not connected by LHCONE. Figure 2 displays a schema of internal end external network. A snapshot from the high level external network monitor is on Figure 3.

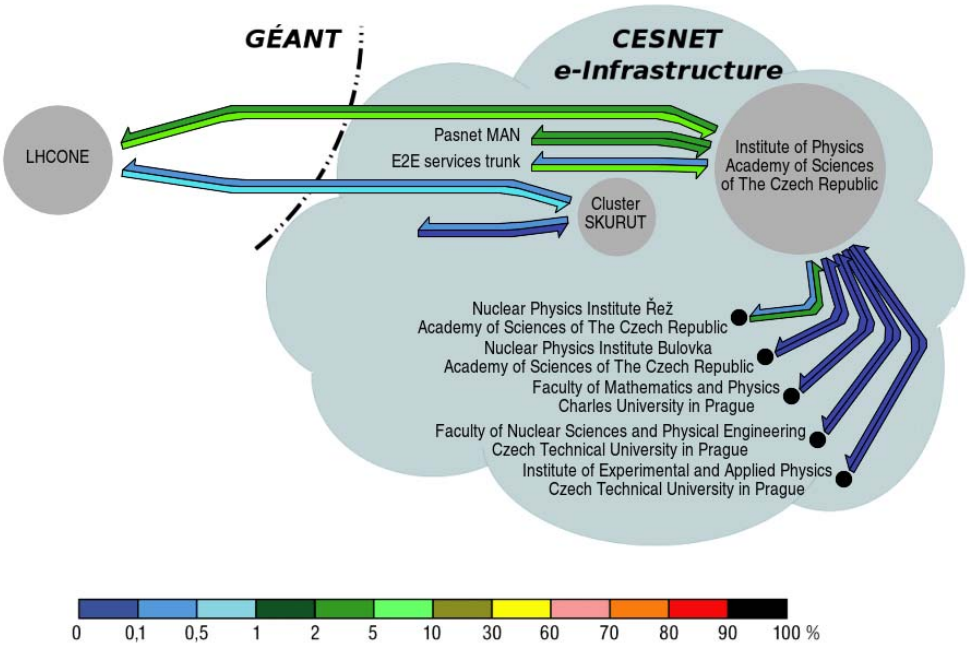

Figure 3: Monitoring of external connections is done by tools developed by CESNET, which also provides the connectivity. The colour of links marks their current utilization.

4. Monitoring and accounting

4.1. Local tools

We use many standard monitoring tools adjusted to our needs. Nagios determines status of all servers and check_mk interface is used to check results. A custom script creates a summary of errors and warnings and sends an email to administrators three times a day. Only critical errors of the most important services launch immediate emails and short text messages.

Another custom script queries the Torque server and displays utilization of worker nodes. The current occupancy of a worker node and its possible dedication to the special Atlas multicore queue is encoded by different colors and fonts; an example for one cluster is shown on Figure 4. Numbers of running and waiting jobs are also stored in RTT graphs. Munin and Ganglia graphs are mostly checked during debugging of various problems. Observium stores network traffic and produces summary graphs. Flow-tracker based on flow tools is also used for monitoring storage servers.

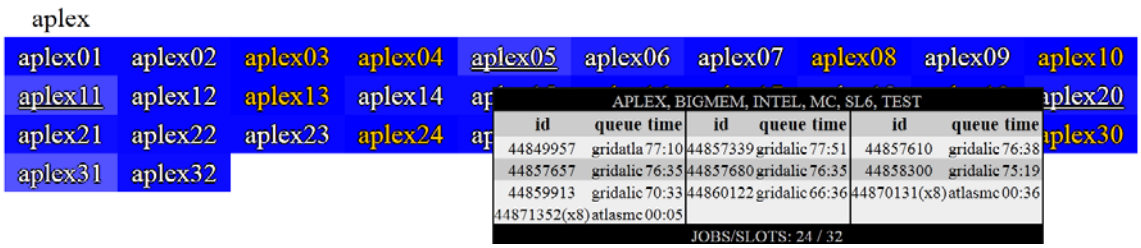

Figure 4: An example of a local monitoring of worker nodes' usage. The source information is parsed from the Torque server. Different colors denote different states of worker nodes. A mouse hover action displays details about running jobs on a selected node.

We use experimental installation of Elasticsearch, Logstash and Kibana stack for parsing of almost all log files. The current setup consists of 7 nodes, mostly reused several years old worker nodes or retired special servers. These nodes have 32 to 96 GB of memory, and 1 to 6 hard drives. A front-end server, which does not store data and serves as an ES master and Kibana server is realised as a virtual machine with 20 GB of memory, 8 cores and 10 GB virtual drive.

About 40.000.000 log entries are processed daily and increase the used space by 30 GB. The performance is satisfactory for standard daily processing but the interactive response of Kibana is not fast enough. Also any changes in log file processing result in lengthy operations. A re-indexing of Torque accounting data from year 2010 till 2016 took a day and half under normal cluster load.

4.2. External monitoring

We get an important feedback about resources' status directly from major users – the ALICE and ATLAS experiments. Both have their own extensive monitoring infrastructure based on frequent test jobs. The common EGI/WLCG monitoring, which uses *ops* Virtual Organization, gives less precise results, since it cannot cover all special cases. As an example we can mention a relatively frequent (several cases per year) failure due to a hardware problem on a disk controller on one of the disk servers. A generic storage test on a write and subsequent read cannot discover a single controller failure, because the system uses randomly chosen one and the number of controllers is high. Only those jobs which access files on a failed controller encounter an error and only parsing of results of jobs and feedback from their owners gives very detailed status overview. Status of controllers is also monitored by Nagios and most of the problems are discovered by local monitoring.

Local accounting data were checked with values published on the EGI accounting portal [4] and were found in a good agreement. Details about historical contributions to the ALICE and ATLAS experiments can be found in [5].

Clusters connected to MetaCentrum are monitored separately by a central MetaCentrum Nagios instance ran offsite. Any issues with these clusters are communicated by a MetaCentrum staff via emails. Registered users receive emails about downtimes and news and can check status of resources on a portal [6].

References

[1] Chudoba J et al.: Czech National Grid Infrastructure, *these proceedings*
[2] Kouba T et al 2012 J. Phys.: Conf. Ser. 396 042034
[3] Kelsey D et al 2013 J. Phys.: Conf. Ser. 513 062026
[4] EGI accounting portal, http://acounting.egi.eu
[5] Adamova D et al 2015 J. Phys.: Conf. Ser. 608 012035
[6] MetaCentrum Virtual Organization Portal, https://metavo.metacentrum.cz/en/index.html

Acknowledgments

The presentation of this work was supported by the Ministry of Education, Youth and Sports of the Czech Republic, project INGO LG15052. The work is also supported by projects LM2015058, LM2015038, LM2015046, LM2015042, and by the Institute of Physics of the Czech Academy of Sciences.