Tier-1 in Kurchatov Institute: first months of operations during Run-2

A. Berezhnaya, I. Lyalin, A. Rogovskiy, E. Ryabinkin, I. Tkachenko, V. Velikhov

NRC "Kurchatov Institute"

Tier-1 in Kurchatov Institute

Current status

Resources:

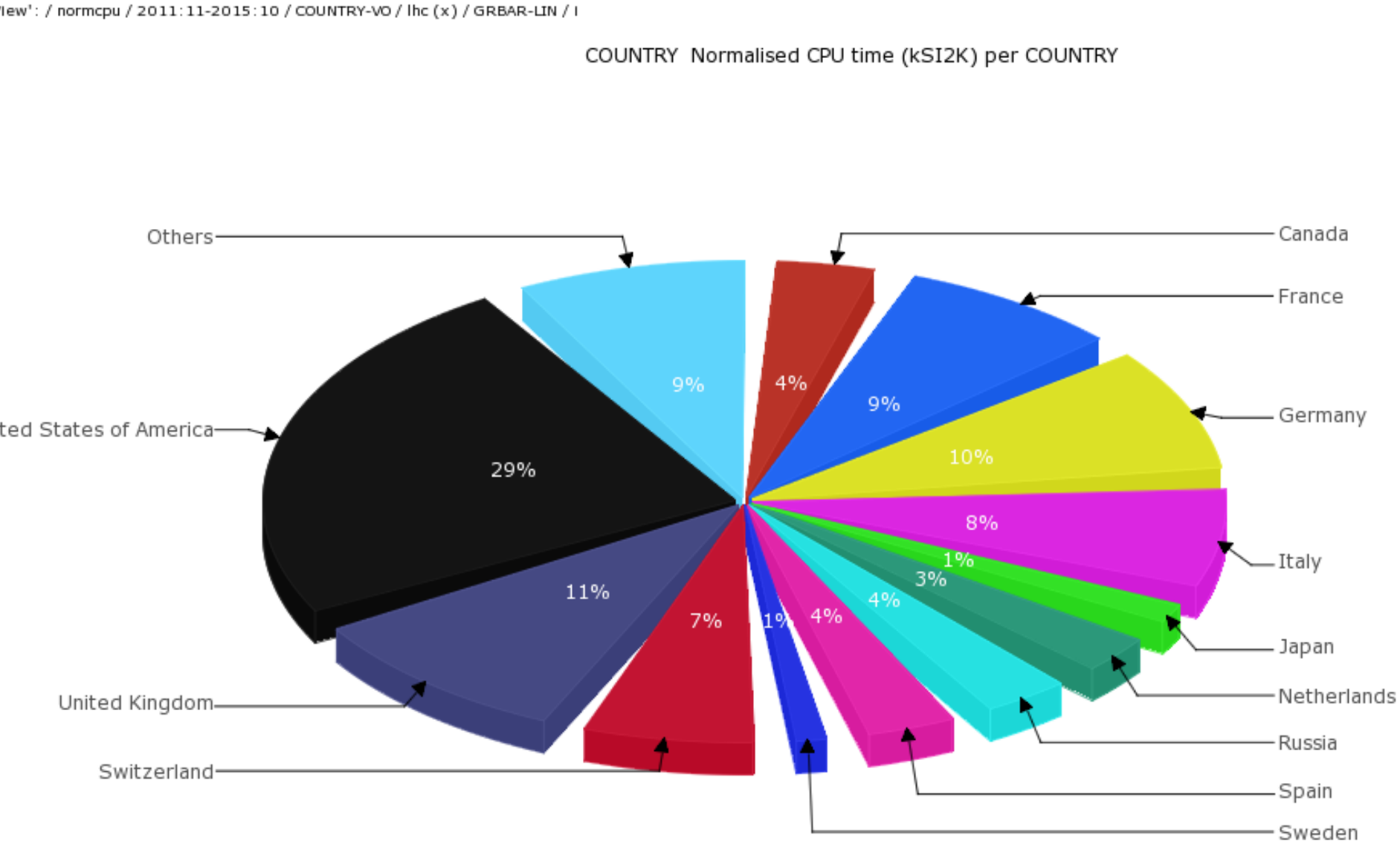
- 71 000+ HEP-SPEC06, 8000+ CPU slots
- 6.3 PB of disks;
- 7.4 PB of tapes;
- 10 Gbit/sec LHCOPN connection (two redundant triangles, shared with JINR);
- 5 Gbit/sec R&N-like connectivity.

Software:

- dCache for ATLAS, LHCb and tapes
- EOS for ALICE
- Enstore as the tape backend for dCache
- Still Torque/Maui as the scheduler
- Standard WLCG software stack

НИЦ «Курчатовский институт»

Look, ma: 4%


'View': / normcpu / 2011:11-2015:10 / COUNTRY-VO / lhc (x) / GRBAR-LIN / l
COUNTRY Normalised CPU time (kSI2K) per COUNTRY
Others
9%
Canada
4%
France
9%
Germany
10%
ted States of America
29%
8%
Italy
1%
3%
4%
Japan
Netherlands
Russia
United Kingdom
11%
7%
1%
4%
Switzerland
Spain
Sweden

Current developments: ATLAS and tape

ATLAS wants to save job logs to the tapes (they seem to be cheaper than disks). Job logs are small, so they should be aggregated either on the ATLAS side or before going to tapes. Why?

- 1MB file – 1-2 MB/sec write speed;
- 1GB file – 150-200 MB/sec write speed.

Recent poll made by ATLAS revealed that only 3 Tier-1 sites are able (and keen) to aggregate logs on the storage side. So, it was centrally decided that the VO aggregates the logs.

But ATLAS still wants Tier-1 sites to be able to do it on their side, so we agreed that we will do that. It will be another 80 M records in the dCache database, so we decided to split out the logs into a separate dCache/Enstore instance.

After consultations with dCache developers we decided that it will be the safe move, Enstore also seem to be able to withstand such a revolution.

We had already validated with ATLAS that logs are aggregated to our tapes without any visible problems and writing speed stays at 180-200 MB/sec.

НИЦ «Курчатовский институт»

Networking: trying to organize NREN in RU WLCG community

Current situation with networks for research and education in Russia is rather bad: actually nobody (apart from the sites) are interested in this.

And there is a dispute between Russia and GEANT that effectively cuts us from R&E networks in some european regions (we still enter them, but via commercial links).

So already for 1.5 years we in Kurchatov Institute are running LHCONE instance by ourselves. Last year brought us some progress: we organized direct traffic exchange with two major US NRENs, ESnet and Internet2.

Complemented with LHCONE connectivity to CERN, NDGF and NL T1 this gives as a possibility to use large part of LHCONE infrastructure. So we decided to continue this activity.

But to peer with some NRENs one should also become an NREN. So we're trying to become one, together with other members of RU WLCG community.

НИЦ «Курчатовский институт»

CA root transition

We also run national X.509 certification authority as the part of our Tier-1/Tier-2 activities. And this year the old CA root had expired: `Issuer: C=RU, O=RDIG, CN=Russian Data-Intensive Grid CA`

```
Validity
Not Before: Aug 9 14:44:37 2005 GMT
Not After : Aug 7 14:44:37 2015 GMT
Subject: C=RU, O=RDIG, CN=Russian Data-Intensive Grid CA
```

We took all appropriate measures, prolonged the certificate and put the new root to the EUGridPMA distribution 4 months before its expiration.

And seems like we made the right thing: it took just 4 GGUS tickets and a dozen e-mails to fix all places where the old CA root was still kept after 07.08.2015.

НИЦ «Курчатовский институт»

Log storage and analysis

We have many logs from our machines. Some are analyzed, some – aren't.

But they also need to be correlated with each other, because it is easier to find the inter-relations between various parts of the system (e.g. EOS pool, DNS failures, overloaded routers, etc).

We're now experimenting with ELK (Elastic Search, Logstash and Kibana) to analyze dCache logs.

We also have a project for central storage of the log files. It will be based on NoSQL, since it allows good DB scalability, logs are append-only objects and we can run data mining on top of the NoSQL storage.

НИЦ «Курчатовский институт»

24x7 operations team

Now we have team of 7 people doing 24x7 monitoring and alerting for our Tier-1 resources (and currently just for them). Since Grid is the open system, failure diagnostics is harder than, e.g. for HPC. From the other hand, brilliant system administrators won't go into 24x7 duty (some will, but they are the rare birds in our forest).

We have some automation for various tasks and that's one area where progress has been made.

But in general, we have a problem here: we are not satisfied with our current operations. And the blame is on us, not the monitoring people. So, that's the problem we're currently trying to solve.

Thanks for your attention!

НИЦ «Курчатовский институт»