Computing facilities for small physics analysis group: examples and consideration

Andrey Y Shevel
High Energy Physics Division
Petersburg Nuclear Physics Institute

Quick overview

- Small physics analysis group (10+ persons)
- Large scale computing for HEP
- Colocations and group owned computing cluster
- Consideration

Small physics group

- A few persons (10+ including students)
  - How many such the groups are in the World? I guess many.
- Limited resources (funding is very limited)
- Limited number of permanent staff (students come and go)
- Needs for group private data and analysis (i.e. private computing), other private activities.
- Which computing resources are available for such the group?

Group computing options

- Collaboration computing cluster
- Grid (just from desk/lap-top with UI packages)
- Clouds (several tests in a range of collaborations: ATLAS, CMS, STAR)
- Group owned ***colocated*** computing cluster
- Group owned computing cluster

Collaboration computing cluster

- Most of large collaborations do have owned computing facility (cluster) [STAR, PHENIX, and so on] rather complicated and large (thousands of cores)
  - There are rules:
    - who can be user;
    - which resources can be used;
    - how the user can start to use the resources;
    - changes in procedures can not be done fast.

The success of W-LCG

- Today:

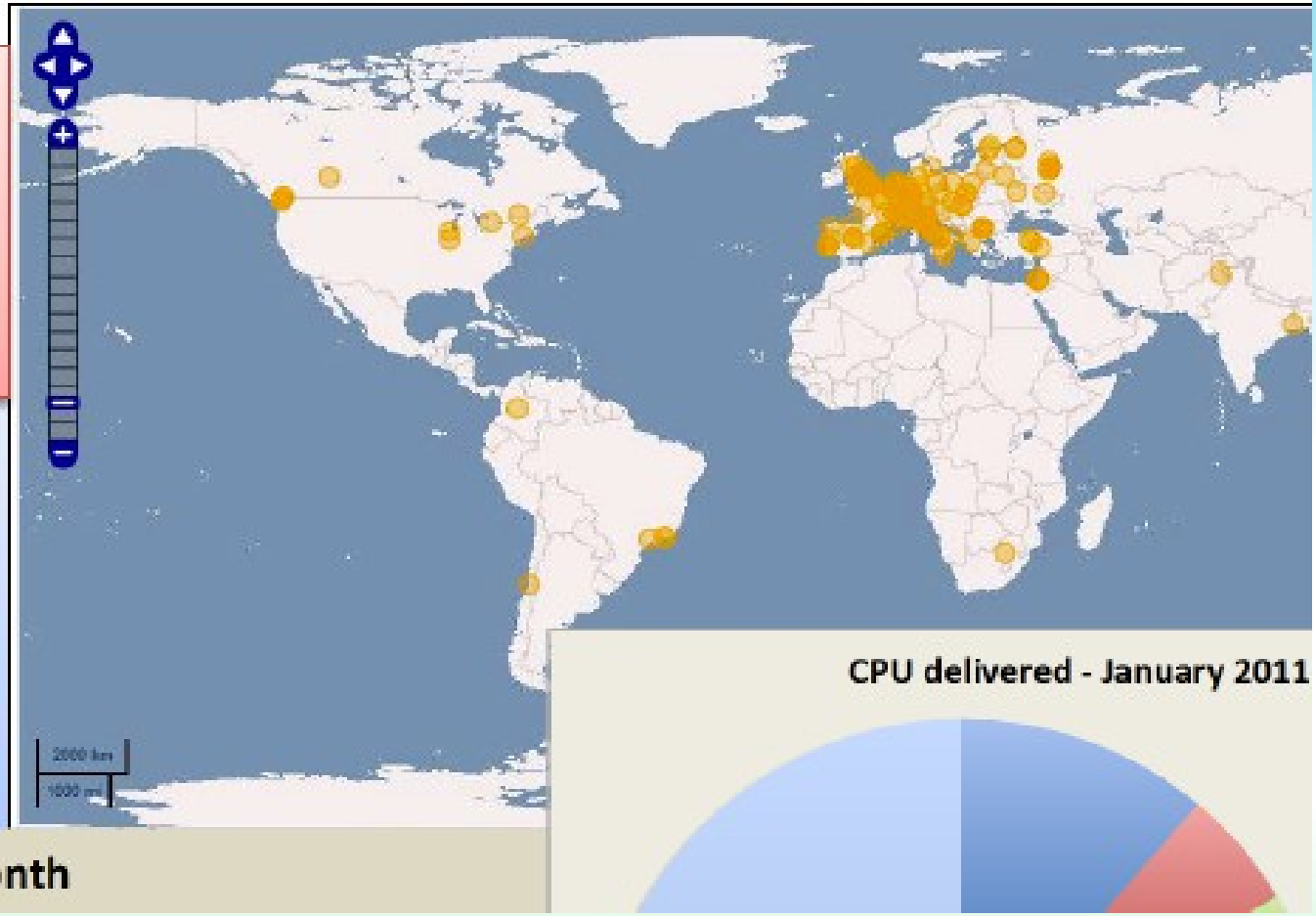
>140 sites

>250'000 CPU cores

>150 PB disk space

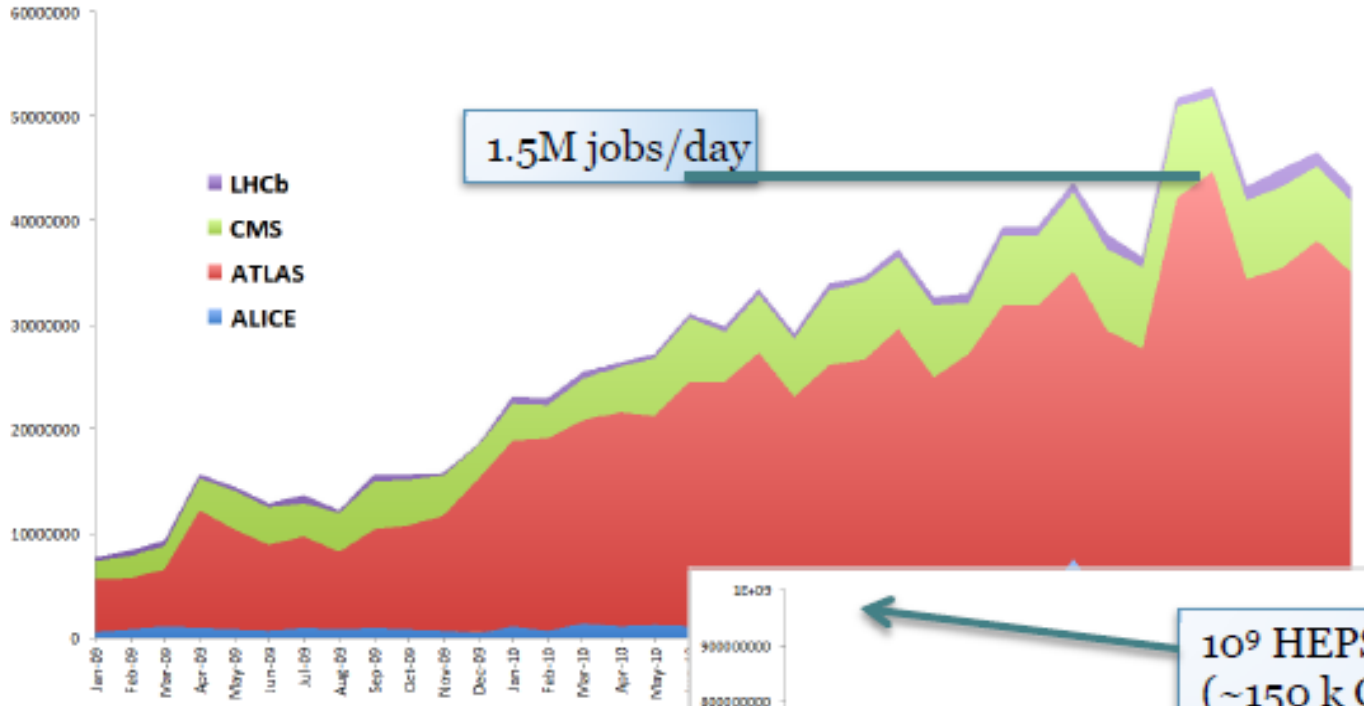
34 countries:

Australia, Austria, Belgium, Brazil, Canada, China, Czech Rep, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, India, Israel, Japan, Rep. Korea, Netherlands, Norway, Pakistan, Poland, Portugal, Romania, Russia, Slovenia, Spain, Sweden, Switzerland, Taipei, Turkey, UK, Ukraine, USA.

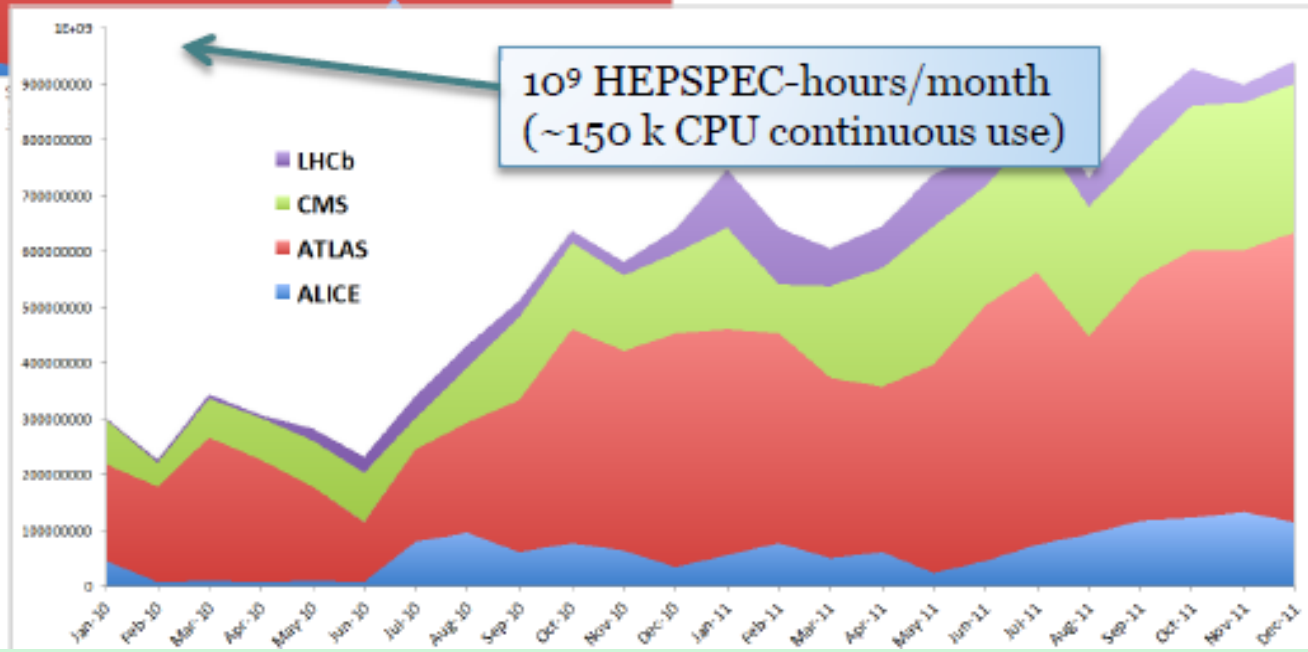


From Sverre Jarp on ACAT 5th Sep 2011

WLCG usage: continues to grow



- WLCG usage pattern:
  - Continuous
  - Ever increasing load
  - Some spikes
  - Mirrored by contributing national Grids



- Evolution:
  - Being discussed by *Technical Evolution Groups*
  - Operations is one of concerns

From Oxana Smirnova CHEP 22d May 2012

however ...

- «Reliable (Grid) operations are rather expensive»

Computing clouds

- Public computing clouds: many tests with use of Amazon EC2/EC3 and the like: *opinions are still divided*
- HEP private computing clouds: CERN has initiated such the project (Agile Infrastructure: Configuration and Operation Tools by Helge MEINHARD on 24 Apr 2012 - HEPIX)
  - See http://indico.cern.ch/contributionDisplay.py?sessionId=3&contribId=25&confId=160737.

Owned colocated computing cluster

- You buy the cluster and place it into rack in some commercial data center. The selection of the equipment and installed software is question of your choice (with little limitations like consumed power).
  - You need to pay for the space in rack
  - However you do not need to be careful for electricity, coolong, all technology.

Group owned computing cluster

- Buying, installation, support, and all operations is your own business.
- Very cheap at the start:
  - Buy and install just one machine in your office;
  - Deploy a range of VMs and debug any cluster architecture you like and apps inside VMs
  - After you can install several additional hardware machines and increase you cluster power almost automatically (for example with centralized data mounted with NFS) and vu a la you are done

Small physics group data analysis scenario

- To select some fraction of data on large cluster
  - For group owned cluster it is assumed that before start first analysis job on your cluster you need to replicate/copy the required data on the cluster.
  - The volume of the tranferred data does matter. You can discover that the data transfer speed around 100 MB/sec is not minimum!

Tested network speed

- Highest tested speed in HEP for the server now (summer 2012) is 40 Gbit
  - See http://indico.cern.ch/contributionDisplay.py?contribId=380&confId=149557
- In between cities: 100Gbit
- *«Really fast network are still expensive»*

SUNY NCG group cluster example

- Went into operation in 2000
  - 30+ machines 0.5 GB each, Pentium III (Katmai) 0.5GHz,
  - 6TB of disk space
  - Tape library 3TB
  - RH 6.x with PBS as batch processor
- Now (2012) https://sites.google.com/site/ramdata2009/
  - around dozen machines:
  - CPUs: mix from Pentium III (Katmai) to Intel(R) Xeon(TM) CPU 3.00GHz,
  - around 3 TB of disk space.
    - Slowest machine is used as NIS server.
  - SL 5.3, Torque/Maui, No VMs, all software were copied from PHEINX/BNL
  - Number of active users ~3 (registered 50)

SUNY NCG continuation

- No backup scheme for user home directories (users were recommended to keep critical files in clouds).
- No UPSs.
- No airconditioning.
- No permanent manager responsible for the cluster (the student is asked to switch on the machines if required)
- Web site and mailing list for users are in cloud (Google).
- No support for users.
- Initially Grid tools tested but later removed due to lack of maintenance and low user interest.
- Main user software components: root + PHENIX apps

PNPI HEPD cluster

- Went into operation in Feb 1998 with a few machines of Pentium III,
  - RH5x
  - Batch CODINE (predessor of SGE)
  - Apps were from CERNlib
- Now (2012) there are a few machines (aroud 2.5 GHz, 30 GB, 6-12 cores each) 13 TB of disk space
  - Fully virtualized (Xen)
  - SL5.7, Xen (virtual cluster consists of 16 VMs)
  - SGE 6.2u5p2, CERNlib
  - Number of active users ~15 (registered 150)

PNPI HEPD continuation

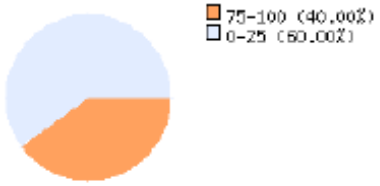
- Homemaid backup scheme for user directories.
- UPSs, UDP.
- Airconditioning.
- ~½ FTE to maintain/upgrade the cluster.
- Web site and mailing list for users is in the HEPD network domain.
- Support by mails.
- Grid UI tested but later removed due to lack of maintenance and low user interest.
- Main user software components: root, garfield, geant + apps from different collaborations

PNPI-HEPD Grid > pcfarm cluster > --Choose a Node

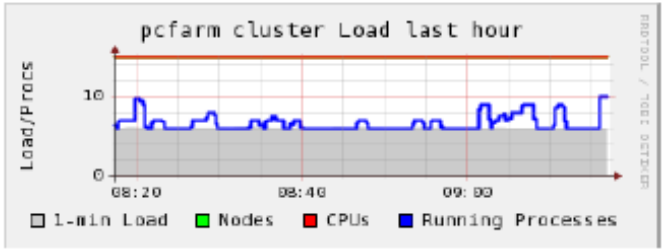
CPUs Total: 15
Hosts up: 15
Hosts down: 0

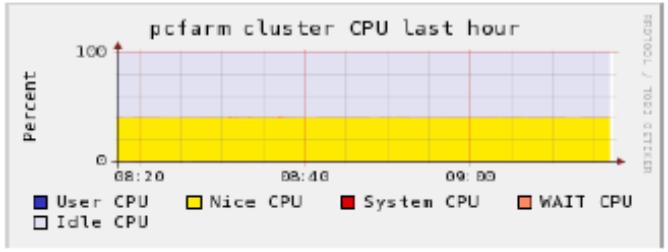
Avg Load (15, 5, 1m):
39%, 39%, 40%
Localtime:
2012-07-18 09:17

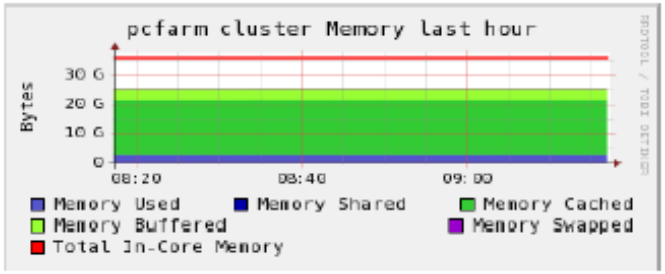
Cluster Load Percentages

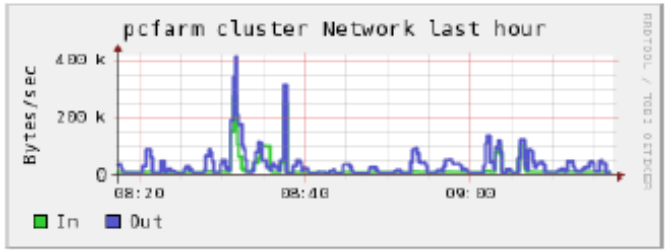


Overview of pcfarm cluster

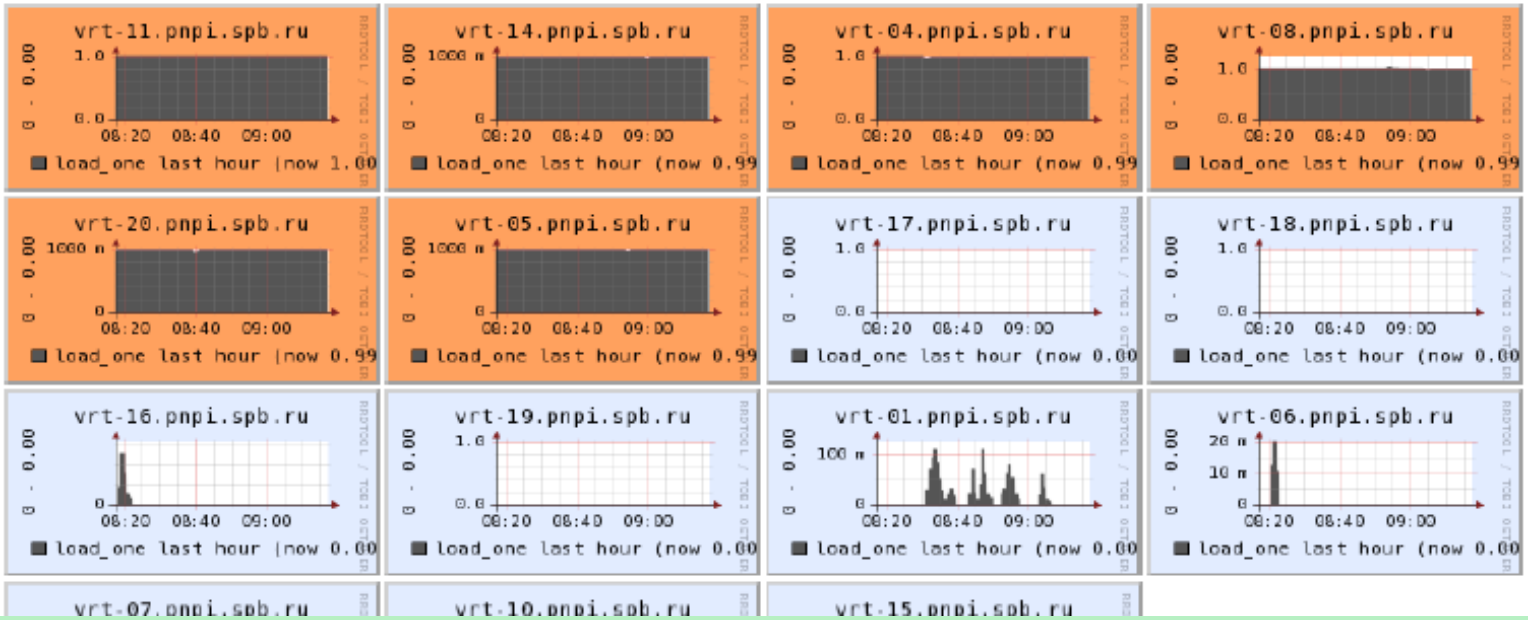








Show Hosts: yes no | pcfarm cluster load_one ( ) last hour sorted descending | Columns 4



Quick consideration

- Is it possible to build up and use the computing cluster of micro size for small group? **Yes.**
- Does it require serious support ? **Not really. You need to find minimum of support service which can be maintained by available minimum of manpower.**
  - Remark aside: complicated equipment like **Mars Rover** on remote planet is working many years almost autonomically, why not computing cluster?
- Is it possible to use it for private analysis or/and volunteer computing? **Yes**, with with Boinc/CoPilot toolset [volunteer LHC MC http://lhcathome.web.cern.ch/LHCathome/]

Analogy (large cluster vs micro cluster)

Conclusion

- Finally we can tell that the clusters of micro size is not alternative but complementation and gateway to all other computing options for small physics group.
- To my opinion the role of such the clusters is growing with the time.
  - Grouth of the CPU power per server (let us remember Intel Single Chip Cloud Computer)
  - Grouth of the disk drive capacity per spindle
  - Obvious possible hybrid co-processor architecture: GPU and FPGA
- Increased number of computing options for physics group is real challenge.

Thank you ! Questions?