The role of micro size clusters for small physics laboratory

Andrey Y Shevel (Andrey.Shevel@pnpi.spb.ru)
High Energy Physics Division
Petersburg Nuclear Physics Institute

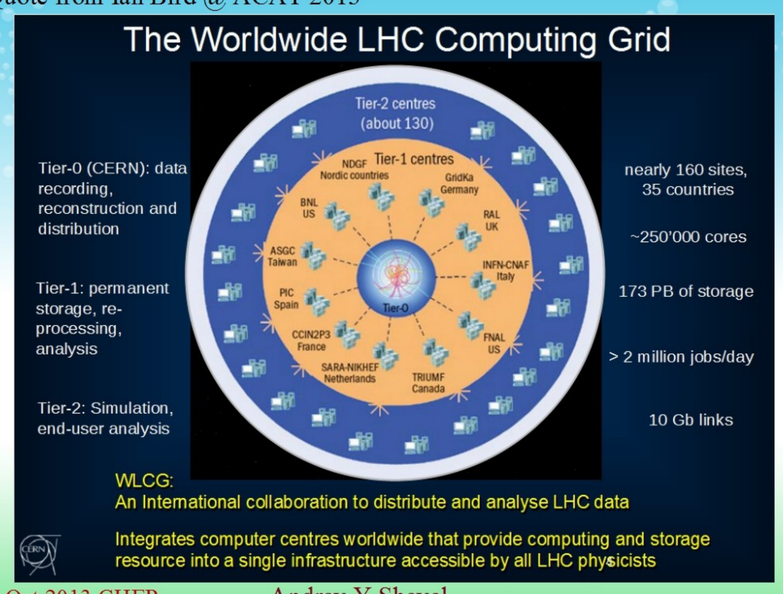
Small physics group

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

Group computing options

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

Quote from Ian Bird @ ACAT 2013



Computing options

- Computing clouds:
  - Public computing clouds: many tests with use of Amazon EC2/EC3 and the like: *opinions are still divided*
  - HEP private computing clouds: several sites have such the projects (e.g. 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 equipment 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, cooling, other technology details.

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

SUNY NCG group cluster example

- Went into operation in 2000, passed through many modernizations)
- Now (2013) https://sites.google.com/site/ramdata2009/
  - A few machines: CPUs: mix from Pentium III (Katmai) to Intel(R) Xeon(TM) CPU 3.00GHz, around 3 TB of disk space (*recently went down*).
  - SL 5.3, Torque/Maui, No VMs, all software were copied from PHEINX/BNL
  - Number of active users ~3 (registered ~50)
  - No backup scheme for user home directories (users were recommended to keep critical files in clouds).
  - No UPS, 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), just crowd support through mailing list.
  - Main user software components: root + PHENIX apps

SUNY NCG group cluster example

- Went into operation in 2000, passed through many modernizations)
- Now (2013) https://sites.google.com/site/ramdata2009/
  - A few machines: CPUs: mix from Pentium III (Katmai) to Intel(R) Xeon(TM) CPU 3.00GHz, around 3 TB of disk space (*recently went down*).
  - SL 5.3, Torque/Maui, No VMs, all software were copied from PHEINX/BNL
  - Number of active users ~3 (registered ~50)
  - No backup scheme for user home directories (users were recommended to keep critical files in clouds).
  - No UPS, 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), just crowd support through mailing list.
  - Main user software components: root + PHENIX apps

PNPI HEPD cluster

- Went into operation in Feb 1998. Passed through many modernizations.
- Now (2013) there are a few machines (aroud 2.5 GHz, 30 GB, 6-12 cores each), 26 TB of disk space
  - Fully virtualized (Xen), SL5.7, Xen (virtual cluster consists of 16 Vms), SGE 6.2u5p2, CERNlib, # of active users ~15 (registered ~150)
  - 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 L1/L2 and crowdsourced.
  - 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.

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 rovers** on remote planet is working many years almost autonomically, why not computing cluster?
- *Micro* size cluster is relatively cheaper than other clusters.

Advantages of micro size clusters for small physics groups and others

- Easy to start;
- Not difficult to support;
- Good instrument for:
  - Software and architecture development, testing, etc
  - Teaching
  - Private analysis
  - Using as the gateway to large computing facilities.
- Large clusters have a chance to see the pool of students/scientists who get initial experience on micro size clusters

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.
  - Growth of the CPU power per server.
  - Growth 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.