The role of micro size computing clusters for small physics groups

A.Y. Shevel
High Energy Physics Division at Petersburg Nuclear Physics Institute, Gatchina, Orlova Roscha 1, 188300, Russian Federation
E-mail: Andrey.Shevel@pnpi.spb.ru

Abstract. A small physics group (3-15 persons) might use a number of computing facilities for the analysis/simulation, developing/testing, teaching. It is discussed different types of computing facilities: collaboration computing facilities, group local computing cluster (including colocation), cloud computing. The author discuss the growing variety of different computing options for small groups and does emphasize the role of the group owned computing cluster of micro size.

1. Introduction
A physics group 3–15 persons who really involved into program testing/developing often demands a number of computing facilities for analysis and related needs. The members of a physics group usually have their computer accounts on large computing facilities, which are supported by institutional collaborations. Such facilities have certain rules: who can access the computing installation, at what scale, and for which purpose. As a result, the registration procedure takes some time. On the other hand, short-term students and/or visitors might need a computer account just temporarily. Finally, a physics group needs in addition to the institutional computing infrastructure a more agile and flexible computing infrastructure completely under group’s control for several purposes:

- to keep common group data (papers, drafts, programs, fractions of experimental data, etc.);
- to test new/modified simulation or/and analysis software/algorithms;
- to provide an account for short-time visitors/students who need to do something in data analysis;
- for any other possible purpose, in particular, as a good gateway to remote large computing clusters.

We have to take into account the growth of the computing power of Computer Processing Units (CPU) every year. If we pay attention not only to CPU but to the whole computing power of a cluster, we can find, for example, such an estimation: “… DOE (U.S. Department of Energy) centers have historically delivered average improvements in computing capability of 40%–80% per year with relatively flat budget”. This means that dozens of modern computing nodes in 2013 are more powerful than hundreds of servers in 2002.

Obviously, such a small computing installation is a good complement to large computing facilities.

1 The Magellan Report on Cloud Computing for Science – U.S. DOE, Office of Advanced Scientific Computing Research, December, 2011, p. 125, http://www.science.energy.gov/~/media/ascr/pdf/programdocuments/docs/Magellan_Final_Report.pdf.
The computing needs can be considered in various ways (from the point of view of a small group):

- to use a centralized cluster (i.e. a collaboration cluster) of reasonable size;
- for cloud computing;
- as a group owned local cluster (might be in two instances):
  - location of the group computing cluster in the group office space with all responsibilities for air conditioning, electric power, hardware support, etc.;
  - colocation of the groups cluster hardware somewhere else.

Many pros and cons for each of the above listed options were discussed earlier. Here it is assumed that the physics group uses more than one cluster to get the computing task done. Within the scope of the paper, such group owned computing clusters are referred to as clusters Tier-3.

Usually, a small physics group has limited financial resources. This fact does impose a range of restrictions on the cluster architecture. The cluster has to be:

- cheap (a useful consideration on the true cluster ownership cost is in [3]);
- composed of reliable hardware;
- not demanding for intensive supervision/maintenance.

Other requirements are the implications of the desire to reduce the maintenance efforts: compatibility (architecture and base OS) with the collaboration cluster environment, in particular, with the same set of application software, as in the collaboration cluster.

From the above, we see that a group-owned computing cluster cannot be large or even mid-range, it is quite small, a micro cluster. A good configuration of a group-owned cluster might consist of 5–12 modern machines (multicore CPUs, 2–4 GB of the main memory per core, 10–20 TB or more of the disk space per machine). It is better to use a 10 Gbit network switch. Such a group cluster can help to get more flexibility when using several remote computing facilities: the collaboration cluster(s), public/private cloud computing, etc.

The situations in different physics groups might differ from each other. Here, the particular group cluster solutions for the Nuclear Chemistry Group (NCG) at SUNYSB/Chemistry and for the High Energy Physics Division (HEPD) at PNPI are discussed.

2. **Local computing cluster at SUNYSB/Chemistry**

The computing cluster in the NCG was organized in 2000, or a bit earlier. At that time, all the machines (30+) had 512 MB of the main memory and Dual 500 MHz CPUs. This cluster was used for program developments, test analysis, students work, etc. There were more than 70 registered users, and about 2–4 of them were quite active. Due to security reasons, the regular maintenance of the cluster is available only from specifically defined network domains. Because the cluster is located in a relatively large room with good ventilation, there is no need for air conditioning. After years of experience, we conclude that the University electric power supply is quite stable.

The basic OS (Scientific Linux with the same RPM set as on RACF) installation procedure and the basic configuration are semiautomatic: there are a couple of scripts which use the kickstart as the initial step, and another step consists of a script for a post kickstart configuration. No virtualization technique was used in the cluster. The pair of torque/maui was used as batch subsystem. More details about the cluster were discussed elsewhere.

3. **Local computing cluster at the High Energy Physics Division of PNPI**

The computing cluster at the High Energy Physics Division stemmed from a very small cluster consisting of three servers in February 1998. Details of the initial implementation are available in [4]. The cluster passed through multiple upgrades in hardware and software, though it remained quite

---

2 The cluster sizes: big, large = more than 1000 servers; middle size – up to 1000; small – up to 100; micro = O(10).

3 PHENIX technical note tn-452.0.

4 the same as above
small, or micro size. Now, the cluster consists of 5 hardware servers with 20 virtual machines (completely virtualized) and has around 26 TB of the disk space. The OS is Scientific Linux 5.7. There are about 150 registered users on the cluster; about 50 users logged many times per month, and about 15 persons use the cluster every day.

There is a home-made backup scheme for user home directories (not for the data). One experienced person spends part of his/her time to keep the cluster running.

In the two computing cluster examples for High Energy Physics (at SUNYSB and PNPI) we can see the main similar trend: to reduce the cluster Total Cost of Ownership (TCO). It includes everything: the cost of hardware and deployment, electric power, manpower, software and hardware maintenance, any operation cost, cost of upgrades, etc. In this context it is not bad to have a look at the “cloud” computing.

4. Cloud computing
The cloud computing is a main trend in world of computing infrastructure last 5-6 years (HEP is not exclusion). Many successful experiments with clouds were carried out [5–7]. Now it is more and more common paradigm, which has a lot of examples in government and private sectors. The quote below is a part of the most consistent computing cloud definition copied from [8].

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. A number of computing clouds are available with a some difference in type of service and in payment policy, e.g. [9, 10].

Some physicists are afraid to use public cloud computing service because the public cloud is out of their control (for instance, the service could be down forever due to business or/and political issues). In other words we can't consider public cloud service as 100% reliable all the time. However, we have to say the same about any other computing service of any kind. Small groups often have not so reliable local computing, which depends on unstable enthusiast activity. In many cases during even medium time frame (2–5 years), a local computing service is most probably less reliable than a public cloud computing service. If you are worrying about reliability of your data being safe, the obvious solution is to use all of available options.

Several successful testbeds with using the public cloud computing for production simulation in HEP were carried out, e.g., in ATLAS [9] and STAR [6]. The latter work has many deep and smart observations of the experience with computing in grid and cloud computing environments. The success does depend on a lot of details, in particular on the computing infrastructure components and their parameters which are “under hood” of the computing cloud. In [11] there were discussed the conversions of VM images may be due to the lack of open standards in the field. In other cases [6, 7], the authors found that the tested public cloud had not so good computer hardware parameters, as they expected. Also, the computing cloud initiatives and the government plans are to be taken into account [12].

5. Conclusion
Small computing/information installations are already on the way to use the clouds. For a small physics group, moving to the cloud does eliminate the cluster hardware maintenance task, but not the application software and data structure maintenance. Also, to achieve the maximal effect of using the cloud one should not ignore good understanding of the cloud hardware, architecture, and OS details.

To compare cluster of micro size and large cluster with many hundreds and more of servers someone might see a lot of similarities: everywhere you need security, proper OS and applications configurations, hardware, etc. However there are difference in between clusters of different size. For example, in micro cluster the common strategy is to buy and deploy computing nodes or/and components with parameters which do fit your task for this cluster (not always cheapest products). At the same time in large/huge clusters the usual strategy is to buy and use cheapest computing nodes and
other components.

We are emphasizing specifically the clusters of micro size, because if we take a look at a range of all size clusters, we might see more servers in the cluster, more spending and efforts to support it. With a more powerful cluster (bigger number of hardware servers), you need additional staff and additional activity to meet more complicated conditions including advanced cooling system, stronger regulations from the public authorities; fire safety, information security, insurance, etc; all these factors increase TCO significantly. In other words there are many reasons for TCO of mid-range computing clusters to grow reasonably faster with the number of hardware servers than total cost of servers in the cluster. This leads to the idea, that two main types of computing clusters will have long life: the huge clusters with many thousands of servers, often referenced as data centers which have a lot of users (actually such clusters are used as computing clouds), and micro clusters, which can be deployed in almost any office and used by a small group of users. Such the trend in the computing infrastructure is not unique only for HEP, especially if we take a look on a range of Micro Data Centers (Micro DC) [13-14].

In the light of the above experience, a group-owned cluster is to be used as sandbox for testings of any type and as an important gateway to public or private cloud computing. A number of public and private cloud computing instances is growing significantly every year, i.e. the importance of suitable gateways to different clouds for small physics group is growing, as well.

References

[1] Colocation centre – http://en.wikipedia.org/wiki/Colocation_centre.
[2] OSG Tier3 Twiki – http://twiki.grid.iu.edu/bin/view/Tier3/WebBook.
[3] The True Cost of HPC Cluster Ownership – http://www.clustermonkey.net//content/view/262/1/.
[4] Batch Computing Facility based on PCs – http://hepd.pnpi.spb.ru/CSD/CSDPublications/proc_043.pdf.
[5] Jan-Philip Gehrcke et al., ATALS@AWS, http://iopscience.iop.org/17426596/219/5/052020.
[6] Jerome Lauret et al., From grid to cloud: STAR experience. http://computing.ornl.gov/workshops/scidac2010/papers/data_j_lauret.pdf.
[7] Keith R. Jackson et al., Performance analysis of high performance computing applications on the Amazon Web Services Cloud – http://www.lbl.gov/cs/CSnews/cloudcomBP.pdf.
[8] NIST definition of Cloud Computing – http://www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf.
[9] http://www.dropbox.com/.
[10] http://www.amazon.com/ec2.
[11] Jerome Lauret et al., Contextualization in practice: the Clemson experience – http://pos.sissa.it/archive/conferences/093/027/ACAT2010_027.pdf.
[12] Vivek Kundra, U.S. Chief Information Officer, Federal Cloud Computing Strategy, http://www.cio.gov/documents/Federal-Cloud-Computing-Strategy.pdf.
[13] What Can a Micro Data Center Do For You?, The Journal From Rockwell Automation and Our Partners, October 2012, pp 46-49 http://www.rockwellautomation.com/thejournal
[14] MicroDC Solution http://enterprise.huawei.com/en/solutions/IT-solutions/cloud-data-centers/hw-193446.htm