COMPUTING FACILITIES FOR SMALL PHYSICS ANALYSIS GROUPS: EXAMPLES AND CONSIDERATION

A.Y. Shevel

1. Introduction

A small physics group (3–15 persons) might use a number of computing facilities for analysis/simulation, developing/testing, teaching. The most recent instances of small clusters for the Nuclear Chemistry Group at the State University of the New York campus, Stony Brook (http://www.sunysb.edu) and for the High Energy Physics Division at the Petersburg Nuclear Physics Institute (http://hepd.pnpi.spb.ru/) are briefly described here. Different types of computing facilities are discussed: collaboration computing facilities, local group computing clusters (including colocation), and cloud computing. A growing variety of the computing options and the growing role of the group-owned computing clusters of micro size are emphasized.

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, in what scale, and for what 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 analyses;
- 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) centres have historically delivered average improvements in computing capability of 40–80 % per year with relatively flat budget" [1]. This means that dozens of modern computing nodes in 2012 are more powerful than hundreds of servers in 2002.

Obviously, such a small computing installation is a good complement to large computing facilities. The computing needs can be considered in various ways (from the point of view of a small group):

- to use a big[1] centralized cluster (we mean a collaboration cluster);
- 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 [2].

Many pros and contras for each of the above listed options were discussed earlier [3]. Here it is assumed that the physics group uses more than one cluster to get the computing task done. Within the scope of papers, such group owned computing clusters are referred to as clusters Tier[2]-3 [4]. In the following in this paper, we will analyse our own local computing cluster and cloud computing facilities: now and in the nearest years.

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

- cheap (a useful consideration on the true cluster ownership cost is in Ref. [5]);
- composed of reliable hardware;
- not demanding for intensive supervision/maintenance.

---

[1] The cluster sizes: big, large – more than 1000 machines; middle size – up to 1000; small – up to 100; micro – around 10.

[2] In the grid, like the computing infrastructure around the LHC, several Tiers are defined: Tier-1, Tier-2, Tier-3, and so on. The difference is mainly determined by expected functionality (ability to accept and maintain the policy of Virtual Organization (VO), implement distinguished service for different VOs, availability of backup facilities, *etc*.).

Other requirements are the implications of the desire to reduce the maintenance efforts: compatibility (architecture and base OS) with the collaboration cluster environment (*e.g.*, as in the ATLAS or CMS and other CERN Collaborations), 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 5–15 modern machines (multicore CPUs, 2–3 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, we shall discuss the particular group cluster solutions for the Nuclear Chemistry Group (NCG) at SUNYSB/Chemistry and for the High Energy Physics Division (HEPD) at PNPI.

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 analyses, students work, *etc*. There were more than 70 registered users, and about 3–5 of them were quite active. More detailed information about the cluster is available in Ref. [3].

To reduce the downtime of the cluster, it is good to buy and install special equipment – a KVM switch over the Internet Protokol [6] to do many control actions (switching on and off any machine in the cluster, to get access to the console of any machine, *etc*.) remotely over Internet. So, the group might use a remote help from external experts. However, in the cases described here the idea has not been implemented yet.

For the batch system, we use a pair of torque/maui from http://www.supercluster.org. 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 Red Hat Package Manager set as on the RHIC&ATLAS Computing Facility at BNL) 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.

In our circumstances, the users mailing list does form a kind of a thinking engine for various methods how to use the cluster for concrete tasks. The mailing list is located in the Google.com (*i.e.* somewhere in a "cloud").

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

The computing cluster at the HEPD stemmed from a very small cluster consisting of three servers in February 1998. Details of the initial implementation are available in Ref. [7]. The cluster passed through multiple upgrades in hardware and software, though it remained quite small, or micro size. Now, the cluster consists of 5 hardware servers with 20 virtual machines (completely virtualized) and has around 16 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. Virtual tools permit to use specific configurations for specific user needs, *e.g.*, it is possible to use the CERNVM for a range of physics collaborations.

There is a custom-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. The cluster room is equipped with air conditioner, UPSs, and UDP.

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 support, 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 hot topic in the information technology for about 5 years. Many successful experiments with clouds were carried out (see Refs. [8–10]). It is not quite a common paradigm, though it has a lot of examples in government and private sectors. The quote below is a part of the most consistent cloud computing definition copied from Ref. [11].

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. This is just the beginning of the definition[3], but it gives the main idea. Scores of cloud services are available with a small difference in character and style of service and in policy of payment for the service, *e.g.* [12, 13].

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). That is true. On the other hand, we can consider the control capability as the reliability of the access to the cloud. Can we think that public cloud service is 100 % reliable all the time? The right answer is "no". Unfortunately, we have to say the same about any other case of computing service of any kind. At the same time, 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 the mentioned options.

Several successful testbeds with using the cloud computing for production simulation in HEP were carried out, *e.g.*, in ATLAS [12] and STAR [9]. 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. The authors of Ref. [14] were urged to do additional conversions of virtual machine images, may be for lack of open standards in the field. In other cases [9, 10], 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 [15].

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.

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 spendings and efforts to support it. With a more powerful cluster, you need additional staff and additional activity to meet more complicated conditions including stronger regulations from the public authorities: fire safety, information security, insurance, *etc.*; all these factors increase the TCO significantly. There are many reasons for the TCO of midrange computing clusters to grow faster with the number of hardware servers than the total cost of the servers in the cluster, that is the relative TCO (per server) is less for micro size clusters and for huge clusters (due to the large scale) than for midrange clusters (see Fig. 1). 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 centres 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.

[3] The whole definition is explained on two pages or so.

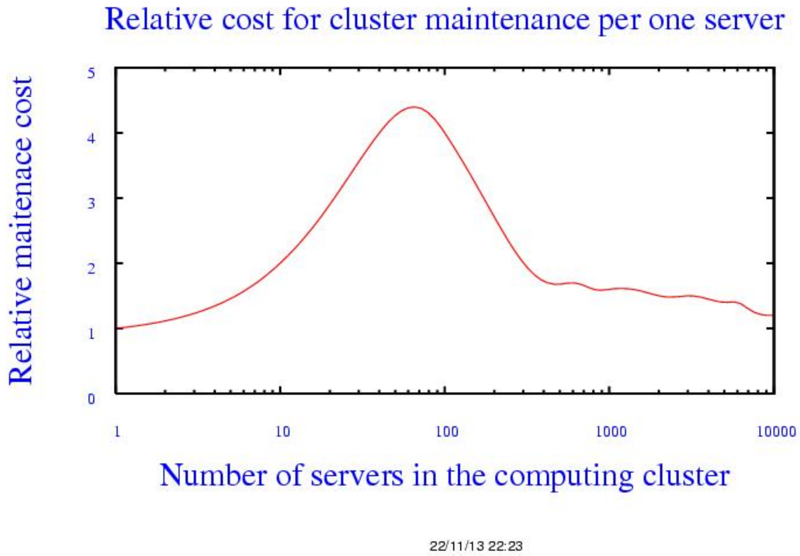

Fig. 1. Estimate of the relative maintenance cost per serever

In the light of the above experience, a group-owned cluster is to be used as an important gateway to public or private (*i.e*. collaborational) cloud computing. The number of public and private cloud computing instances is growing significantly every year. This means that the importance of suitable gateways to different clouds for small physics group is growing, as well.

References

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/program-documents/docs/Magellan_Final_Report.pdf
2. Colocation Centre, http://en.wikipedia.org/wiki/Colocation_centre
3. Damian Reynolds, Andrey Y. Shevel, PHENIX Technical Note tn-452.0, http://www.phenix.bnl.gov/phenix/WWW/publish/shevel/tech-reports/ClusterPaper-2011-04-11.pdf
4. OSG Tier3 Twiki, https://twiki.grid.iu.edu/bin/view/Tier3/WebBook
5. The True Cost of HPC Cluster Ownership, http://www.clustermonkey.net//content/view/262/1/
6. KVM Switch, http://en.wikipedia.org/wiki/KVM_switch
7. Batch Computing Facility Based on PCs, http://hepd.pnpi.spb.ru/CSD/CSDPublications/proc_043.pdf
8. Jan-Philip Gehrcke *et al*., ATALS@AWS, http://iopscience.iop.org/1742-6596/219/5/052020
9. Jerome Lauret *et al*., From Grid to Cloud: STAR Experience, http://computing.ornl.gov/workshops/scidac2010/papers/data_j_lauret.pdf
10. 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
11. NIST Definition of Cloud Computing, http://www.nist.gov/itl/cloud/upload/cloud-def-v15.pdf
12. Dropbox, http://www.dropbox.com/
13. ADrive, http://www.adrive.com/
14. Jerome Lauret *et al*., Contextualization in Practice: the Clemson Experience, http://pos.sissa.it/archive/conferences/093/027/ACAT2010_027.pdf
15. Vivek Kundra, U.S. Chief Information Officer, Federal Cloud Computing Strategy, http://www.cio.gov/documents/Federal-Cloud-Computing-Strategy.pdf