Notes from the Peak capacity focus meeting

Executive summary

Peak computing needs for research institutes could possibly be handled by a shared infrastructure. Actions have been identified to explore possibilities in this direction.

Why are there peaks in computing needs?

Computing needs in research companies as well as academic institutes vary considerably over time. The main use cases include genomics data analyses, e.g. genome wide association studies, and medical image analysis. The peaks are caused by the en masse data generation from big data projects. Pressure on the analysis for first publication or product development causes the need to run the analysis as fast as possible with the available infrastructure resources.

Firstly it is difficult to plan: at the time of peak needs, some jobs that are running could possibly wait, but job priorities do not always make that clear and it will require “social engineering” to properly prioritize. Reserving computing capacity beforehand, as was done when all scientific computing was done on mainframes in the 1960/70s, is no longer fitting current routines.

Secondly, unevenly distributed needs are another form of planning mistake: increased computing needs (e.g. when a new sequencer arrives in a lab) often come before the computing capacity is extended.

Research is intrinsically difficult to plan; *e.g.* based on analysis results analysis is repeated *ad hoc* with a different set of parameters.

Even with good planning peak capacity needs cannot be completely avoided: for instance if big and small jobs alternate on measuring equipment, the required successive analysis computing needs will overlap in time.

DIY solution: Buy for average needs or buy for peak needs?

Two in-house solutions to serve peak needs are possible:

- If the total computing capacity available is sufficient to deal with all needs over the course of the year, a relatively simple queuing system will take care of peak computing needs: if there are more tasks to do than there is capacity, jobs will be delayed until there is capacity available. This solution, however, is not always permissible: the peak needs may all need to be addressed quickly (e.g. because of what was promised to a customer, or in order not to lose their value).
- The obvious alternative is to expand the local infrastructure for peak needs. This, however will result in infrastructure that is idle for most of the time.

Or collaborate?

Another solution could be to offload the peak needs to facilities in other locations. This could be between any set of two or more partners (with some disadvantages) or between partners and a common infrastructure provider.

Jan Bot (SURFsara) indicated that both “large” (any kind of large) and distributed computing are complicated and the combination is very difficult in terms of research and data management. There is no silver bullet. It would be a dream that it would be possible to simply offer other centers access to your own spare capacity. Some problems that will come up if you try to realize such a solution are:

- The number of bilateral or multi-lateral agreements that would result is large
- There will be a large heterogeneity
- Universal access control would need to be instated
- Everyone would need to start performing proper accounting of use.

An alternative that was extensively discussed is to have a shared infrastructure, completely paid by contributions from different centers, that could be used by all participants for their peak needs. Different types of infrastructure could be imagined for this:

- For relatively stable work flows, the Grid was created by the high-energy physics community. It was designed not to have a single point of failure anywhere. For our kinds of jobs, the grid is not always the easiest to use:
    - Proper use of grid certificates is felt to be complicated.
    - Installing (constantly updated and diverse) software and access to the data on all the nodes requires forethought.
    - It seems the grid middleware has been buggy and sometimes jobs just disappeared.

  All this requires local help for users, and this is not easily accessible everywhere
- A more recently upcoming solution is “cloud” computing. In cloud computing a central infrastructure provides virtualized servers of the right size for a problem in minutes, under the complete control of the institute that needs the capacity. Servers can be deployed using only a web API (e.g. available through SurfConext) and can be accounted centrally. Cloud servers can be used to run specific kinds of work, but also to extend (on demand) a local general purpose compute cluster (if the network connection is fast enough). A cloud is *not* a way to do distributed computing, and cloud computing does *not* have the same robustness of grid computing. It is not (yet) possible to easily connect clouds in different locations.

The cloud idea was generally felt to be more attractive than Grid. Such a solution will require access control for the shared resources, software to make resources compatible between user-sites and the central cloud service, and solutions to deal with data stored at a distance.

A comparison was made with the situation in Finland, where CSC is offering cloud computing to different institutes (not as peak, but as base compute infrastructure). The Finnish model is that CSC gets the infrastructure money reserved for the institute and makes sure that enough modern infrastructure is available over the life time of the project. All access to the infrastructure is arranged via optical private networks, falls under the responsibility of the project's own IT people, and for all normal users the centrally run infrastructure is indistinguishable from infrastructure that is located with the project itself. This is a great advantage in the life sciences, where infrastructure users will be able to deploy a VM, but often are not interested in knowing about the most effective way or place to run a task (note that for challenging computing requirements, the difference made by expert knowledge of the infrastructure can make the difference between "impossible" and "doable").

Advantages of centralized peak capacity

- To be able to satisfy peak needs, a "fallover" capacity outside of the institute is much more efficient than running all the capacity that could be needed in-house and let it idle most of the time.
- There is economy of scale: buying as well as operating the infrastructure at a single site is cheaper than buying and operating at each site separately
- Adding capacity remotely will slightly reduce the risks in case of failures of local infrastructure.

Disadvantages

- A homogeneous cloud infrastructure provides services with a certain balance of CPU power, network bandwidth, disk speed and memory size. This may result in "overcapacity" for some kinds of jobs and may make it impossible to run some jobs that are disproportionately demanding in one aspect of computing (e.g. a job that needs only a single core but 512 GB of memory). The shared resources and the demands must be well thought out to avoid this.
- Network transport of data sets can add latency.
- How scalable will a solution be? Will there be scaling problems in I/O?

Points that need to be addressed

- We want infrastructure providers to make a proposal.
- The Service Level required on a peak capacity infrastructure has not been discussed in the meeting. It is obvious that 24/7 support will be very expensive, and on the other side that reliability of the infrastructure is very important (it becomes a single point of failure). This is an action item for the infrastructure provider that will host the infrastructure.
- Even though it is meant for peak capacity, most remote sites will probably choose to add permanent storage for their data. This is an action item for each participant to take into account.
- The cost of running compute infrastructure is not clear at the moment. Not at central academic sites, nor at local facilities. This must be addressed for both to make a fair comparison possible for dealing with the peak capacity. Furthermore, academics always can think of more calculations to do. The value should be counted against the cost. But for this to be possible, the costs must be known. This is an action item for each participant as well as for the infrastructure provider.

- It was proposed that we look for other countries that have had similar discussions and have come up with solutions. Should we potentially do this with other countries? Who can pick up this action item?

Points raised during the discussion that have no other place in the report

- Users need to know the details of the infrastructure they are using if they need to use the resources efficiently.
- Even: Small jobs do not fill a cluster in balance
- Our field has an infinite need for capacity. Calculate a cost to balance it with value
- We are not using “commodity” hardware
- In due time, biologists will acquire more IT skills. They will need it for all their work
- Cloud does not serve all computing needs

Side topic

Interesting side track of the discussion was: do no trust scientists and IT people that make claims about legal impossibilities unless they can show you where in the law it says that something is forbidden. Rumors about legal requirements for privacy and security are everywhere and hard to fight, in reality there is no legal obligation to keep all the data inside your in-house compute center. Complying with the requirements of a demanding customer may be a different matter, but a clear drawing of the network and system topology may help here.