Cloud Systems BOF

OGF 22, Cambridge, MA

Wednesday 27th February

Introduction

This BOF immediately followed the afternoon keynote session in which Irving Wladawsky-Berger presented "Cloud Computing, Grids and the upcoming Cambrian Explosion in IT".[1]

The meeting, attended by upwards of 60 people, was divided over two sessions with two talks in each, followed by discussion at the end of each talk. Geoffrey Fox, who defined the problem space from the viewpoint of the OGF, started the meeting:

What are Clouds? Geoffrey Fox[2]

Grid and clouds are clearly related both in goals and implementation but there are differences which are difficult to discuss as both terms do not have agreed definitions. Possible features of clouds are simple high level interfaces where the cloud (system) and not the system interface is exposed. Clouds are often associated with collections of virtual machines but this is not required. Further grids emphasize managing existing complex heterogeneous systems whereas clouds could be implemented from scratch with heterogeneity present as necessary to enhance system and not for historical reasons. Wladawsky-Berger emphasized that the simplicity of use and construction of clouds was aiming at the scalability to a factor of 100 larger than data centers. A critical difficulty of clouds is performance where affinity of computers to computers (as in parallel computing) or affinity of computers to data is needed.

There is currently no international group that is dealing with the necessary standardization of interfaces to cloud systems, the closest being the Computing Community Consortium who are organizing events to try to get parts of the cloud community together. This is therefore a relatively green field for the OGF, who have the expertise and mandate to show how interfaces defined within OGF can be used to access clouds. OGF also has the community to define use cases and develop core architectures/technologies.

There are several technical issues that must be addressed

- Performance overhead both internally to CPU (VM overhead) and between CPUs and storage.
- Cost gains including size efficiency which may include green considerations.
- Security has currently not been seriously explored. Note clouds do not cross administrative domains and this could simplify discussion compared to grids.
- What is the model of computing one can provide? What type of internal configuration is available? With suitable network connectivity between nodes for example can you do capability parallel computing (i.e. multiple modest size MPI jobs) in clouds?

- With major scientific problems there is normally a need for data-compute affinity due to the sheer data volume you need processing.
- Is there an underlying scheduler for the use case where demand exceeds supply and how would this affect externally available services?
- Is there going to be a situation where we would want to link clouds together either from different providers or between different functionalities (data cloud to compute cloud)?
- Are individual institutions or groups going to want to construct their own clouds, as they have campus grids??

Questions and Comments
It has also been shown that interoperability could be necessary for these systems, as for example the Amazon S3 system went down for a period recently. This could have catastrophic consequences for a business/project that has this as a core part of their business.
The point was made that clouds could give minority serving colleges and other education institutions throughout the US access to computing resource they normally would be able to muster.
Other examples that could be used to show where the arrival of computational resource has transformed research within the organization is Clemson University and their HTC system.

From the list of references we can find organization and people from outside the OGF that may be interested in what we are trying to do with these efforts.

We should ensure that the trusted Computing Community is involved since they will make the use of clouds be perceived from the enterprise community as safe.

Amazon provides services via clouds but also as web services. Are clouds just another service? If so, why can't we apply existing service standards to clouds? We don't know what the interface should be as we aren't sure what they will do and how they will look or be linked together. This is ideal for a best practice discussion and hence formation of an OGF group. This will be highlighted by the difference between the micro and macro interfaces that are exposed.

We must ensure that any standardization effort is on both compute and data clouds.

Cloud BoF, Steven Newhouse[3]
The next speaker was Steven Newhouse, who asked whether clouds were a new revolutionary concept or just the next logical evolution of distributed computing given the ubiquitous networking and computing that were now becoming available. Over the last 30 years networking had moved from being just about available between buildings on a single site, to providing multi megabit links to the home and multi gigabit links between machines, sites and countries. Due to the ubiquitous availability of networking we no longer have to carry our computing power with us, we can access it remotely from our laptop or portable device from almost whatever location we happen to be in.

The move to remote computing capability and networking has had significant implications on the way we work. The ‘knowledge worker’ has to collect and process information from many sources and redistribute it to their human network. Use of services to collect and process this gathered information and to distribute the resulting knowledge is key. These services may be available on my own local hardware, be run from my enterprise’s data center, or available as free or pay-to-use services from some remote service provider.

These remotely provided ‘cloud’ services are now becoming available to anyone as the bandwidth and reliability of networking improves. The significant issue, from a user perspective, then becomes as to how we access these cloud services:

- Through a web browser as a simple web portal?
- Through a web browser using Silverlight, AJAX, etc to give a richer client experience?
- Through a generic client application such as Excel, Word, etc?
- Through a bespoke client application that directly uses the cloud services?

In such an environment there are opportunities for standards at two levels:

- Infrastructure: low-level compute and data as well as management and monitoring
- Interface: AAA as well as functional apps and domains.

Developers, writing services that will sit in the interface layer, will need standards in order to be agnostic as to the implementation of the infrastructure services that they use from a particular service provider. OGF offers a set of standards to support the compute aspects of resource services (HPC-BP, BES, DRMAA) and an emerging set of standards to support data resource services. Access to many of these services can be encapsulated within APIs such as those produced from the SAGA-WG. Users, composing interface services together into their thin or thick client applications will also want standards to enable them to interact with ‘any’ cloud service. Standardization activities to this cloud service interface may be premature at this time as the usage model and the required services are not yet fully understood, but it is certainly an area that OGF could be involved in.

The overarching question from Steven is ‘What can OGF do?’ The obvious next steps are to better understand the connection mechanisms and usage of these cloud service instances from light weight client environments, and then to understand the connection from these cloud services to the ‘back-end’ infrastructure services exposed through OGF’s standardized interfaces.

Questions and Comments

Even though Clouds are not new, it would be an ideal opportunity to switch the direction in which we have tackled the problem from infrastructure/provider focused to user/problem focused.

There was the point made that through the provisioning of universal types of installers, the deployment within a cloud could become very much easier. The current interface other than as a whole VM doesn't lend itself to quick and easy deployments.

Cloud Computing – a view from the Network cloud, Martin Swany, U of Delaware/Internet2[4]

The third talk was intended to provoke discussion about the networking aspects Clouds. Martin started by questioning the impact that clouds could have on the existing Grid networking efforts. From the perspective of the network-related standards and research groups within OGF, the impact will be minimal. Their work will still be important as the issues for Clouds and Grids are basically the same.

Next he discussed the impact on existing networks and networking models. While the ideas in Grid and Cloud networks are essentially the same, the degree to which network models are coupled to models of computing and storage is becoming more important. The transparency of commercial networks will become a problem if we are trying to route as efficiently as possible between different compound services. From the ISP perspective, this will include the way that traffic patterns will dynamically change when different types of resources are available within the Cloud. This is the case both from the point of view of internal topology and performance measurement. The biggest step will have to be in the inclusion of the network within the complete view of the infrastructure. It must become an integrated part of the whole if we are to use clouds efficiently and reliably. It must be considered a dynamic entity, capable of reservation, alongside computing and storage

Thus, the issues of Cloud networks are essentially the same as for Grid networks, but with the increase in scale, the level of deployment and the degree to which the usage models need to be coupled become critical. This includes smarter, more visible network management infrastructure as well as dynamic circuit networking. This represents a paradigm shift in the way we consider the "network cloud."

First Steps in the Clouds, Kate Keahey, Argonne National Lab[5]

The final talk was then given by Geoffrey Fox on behalf of Kate Keahey who has been doing initial testing on use of cloud systems. The talk discussed two issues: (1) an example of the use of a cloud by an application community (the high-energy physics experiment STAR), and (2) challenges in providing existing scientific hardware resources as a compute cloud.

The HEP STAR experiment works with complex applications developed by a large community of scientists over many years. It has extremely stringent requirements on the environments in which it will run at all, and on top of that it also requires consistency across environments, so that it is now becoming very difficult to guarantee their availability at all participating institutions that provide platforms for STAR runs. This problem has been overcome through producing a virtual machine built to community requirements which can be deployed/overlayed over multiple resources. Using this as a virtual platform STAR has been able to run on multiple resources leveraging the

Amazon's cloud among others. The performance degradation (~10%) due to VM runtime overhead has been deemed acceptable by the community.

The second part of the presentation focused on challenges associated with making existing resources available as a cloud, specifically with obstacles that had to be overcome to create a science cloud within resources hosted in the University of Chicago. The typical challenge is that existing resources are dedicated to specific modes of usage (e.g. running a batch queue) required by the communities they serve that are not necessarily compatible with resource leasing for cloud computing. At the University of Chicago these challenges were overcome by developing an infrastructure called "workspace pilot" which adapts the existing batch scheduler to deploy VMs alongside jobs. It operates as a "Glide-in" mode that, once deployed, will take over the resource for the duration of its deployment and lease out the cycles cloud-computing style.
Her overall comments have been that at the moment moving a resource towards cloud computing or an application onto a cloud is not easy. Though once it is there moving between clouds (in her case EC2 -> Workspace Pilot) may be easier allowing the clients to leverage an expanded resource base. This has not been attempted for movements between different commercial suppliers. The talk also highlighted the Nimbus project will allows scientific applications to lease out cycles cloud computing style and is available to scientific users.

Summary
The overall feeling from the meeting was that a significant activity at the next OGF is essential with various suppliers of cloud resource invited as well technologists and users. There is also interest in interoperability between clouds, for example Avner Algom from Israel Grid Technologies (IGT) has a group (led by IBM) interested in clouds and standardization, who should also attend in Barcelona.

The formation of a community group though should wait until after the workshop, though we are starting a trawl for co-chairs. They must be geographically distributed.

We will setup a mailing list for this BOF to which the people that attended will hopefully subscribe. This will be used to advance the progress towards the workshop at OGF23. Group should make a list of important topics like security implications and separation of services

References

- http://en.wikipedia.org/wiki/Cloud_computing (Includes references to Amazon, Apple, Dell, Enomalism, Globus, Google, IBM, KnowledgeTreeLive, Nature, New York Times, Zimdesk -- Others like Microsoft Windows Live Skydrive important)
- http://en.wikipedia.org/wiki/Amazon_Elastic_Compute_Cloud
- http://uc.princeton.edu/main/index.php?option=com_content&task=view&id=2589&Itemid=1 Policy Issues
- http://www.cra.org/ccc/home.article.bigdata.html
- http://ianfoster.typepad.com/blog/2008/01/theres-grid-in.html OGF Thought Leadership blog

---

Presentations

[1] http://www.ogf.org/OGF22/materials/1137/Irving+Wladawsky-Berger+Keynote.pdf
[2] http://www.ogf.org/OGF22/materials/1055/OGFCloudBOFFeb27-08.ppt
[3] http://www.ogf.org/OGF22/materials/1055/CloudSystemsBoF.pdf
[4] http://www.ogf.org/OGF22/materials/1055/ogf22-cloudComp-net.ppt
[5] http://www.ogf.org/OGF22/materials/1174/ogf+clouds_Kate_Keahey_Feb26-08.ppt