ACM Ubiquity, Volume 9, Issue 31 *August 5 - 11, 2008*

Emergence of The Academic Computing Clouds

Kemal A. Delic – Martin Anthony Walker
Hewlett-Packard Co.
kemal.delic@hp.com, m.walker@hp.com

Abstract. *Computational grids are very large-scale aggregates of communication and computation resources enabling new types of applications and bringing several benefits of economy-of-scale. The first computational grids were established in academic environments during the previous decade, and today are making inroads into the realm of corporate and enterprise computing.*

Very recently, we observe the emergence of cloud computing as a new potential super structure for corporate, enterprise and academic computing. While cloud computing shares the same original vision of grid computing articulated in the 1990s by Foster, Kesselman and others, there are significant differences.

In this paper, we first briefly outline the architecture, technologies and standards of computational grids. We then point at some of notable examples of academic use of grids and sketch the future of research in grids. In the third section, we draw some architectural lines of cloud computing, hint at the design and technology choices and indicate some future challenges. In conclusion, we claim that academic computing clouds might appear soon, supporting the emergence of Science 2.0 activities, some of which we list shortly.

Keywords. Data Centers, Grids, Computing Clouds, Very Large-Scale Systems, Web Science

1. Emerging Computational Mega-Structures

We are witnessing today a massive shift towards economy of services. Technology plays here the major role, as it enables creation of new business platforms based on the new technology devices, and omnipresent, permanent connectivity which spawns some very relevant social phenomena and enables one to see cloud computing as a platform for creating new services [7].

Looking into the overall evolution of the Internet infrastructure and associated activities (Fig 1), we predict the emergence of the cloud as the generic infrastructural fabric enabling a huge number of services.

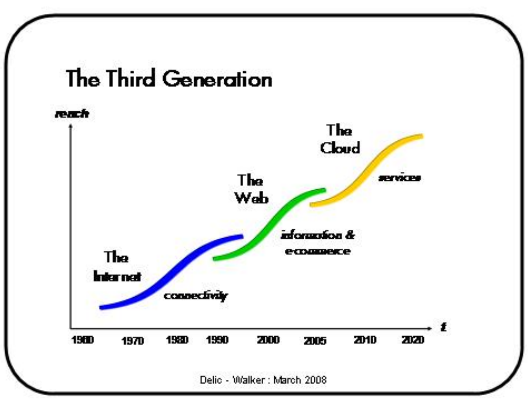

Figure 1. Internet Evolving into Cloud

Attracting hundreds of millions of users daily, these mega-structure service systems have a huge commercial value and are posed in the middle of a battle for the market share [9]. Beyond important commercial aspects, we assume that the academic world will see benefits and practical scientific uses. We will describe some of these service systems and infrastructure in more detail.

2. Data Centers Evolving into Grids

Grids are very large-scale virtualized, distributed computing systems. They cover multiple administrative domains and enable virtual organizations. The key characteristic of grids is their ability to upscale and downscale rapidly and gracefully. So, they provide utility type of computing which enables another type of business model and spawns start-up businesses.

Architecturally speaking, grids represent collections of data centers that are globally, regionally or locally placed and exhibit resilience and high availability [8]. Data centers are made of collections of racks containing various types of servers, typically arranged into clusters of various kinds. One should think that the data centers will typically host tens of thousands of servers and will be strategically placed on the intersection of the low energy supply and high-bandwidth network nodes.

Highly abstracted, the grid can be perceived as a large-scale application execution platform in which aggregation, virtualization and scheduling are the key abstractions enabling grid operations. Grids originated in the academic domain which was not so concerned about security and dependability [2], while we observe the emergence of commercial grids which are specifically designed for business purposes. They are named in analogy to other grids previously created in sectors of energy, communication and transportation.

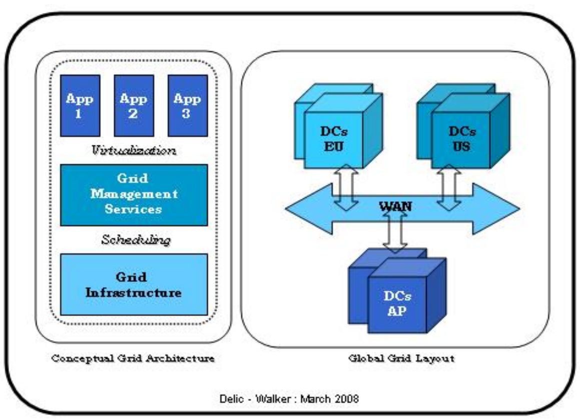

Fig 2. Conceptual Grid Architecture: Typical Layout

We sketch here (Fig 2) the global grid interconnected via a high-speed wide area network, so that applications can run flexibly over the entire grid – as needed. Somewhere else [3], we have argued about the type of grid applications which will lead to wider usage of grids with the important observation that the typical workloads for academic/scientific applications are very different from commercial workloads. Architects and designers of grids should be well aware of this important distinction and respect OGF and OASIS grid standards.

One important role of grids is their ability to transform large numbers of hardware and software assets into computational services for a multiplicity of clients and customers. This is a ‘software as service’ type of transformation which is becoming commercially very relevant as well as academically interesting [5].

3. Computing Clouds: The Power of Scale

At the high conceptual level we can represent the computing cloud (C2) as “a nexus of hardware, software, data and people which provides online services” [9]. Important to notice is that the participants are consumers and producers of C2 services. As such, they represent also a very complex system. Computing clouds are huge aggregates of various grids (academic, commercial), computing clusters and supercomputers (Fig 3). They are used by a huge number of people either as users (300 million users of Microsoft’s Live) or developers (330.000 application developers of Amazon EC2).

As hundreds of millions of users are connecting to Google daily (Google users have created 20 Petabytes of data) or using eBay services (1 Terabyte of logs is created by embay users in 2000 application servers per day) they are creating torrents of data and information. Thus, we can conclude that C2 denotes massive participation, collaboration and content creation by 100s of millions of users on an omnipresent, always available mega-structure. This poses big challenges, but offers some unprecedented opportunities.

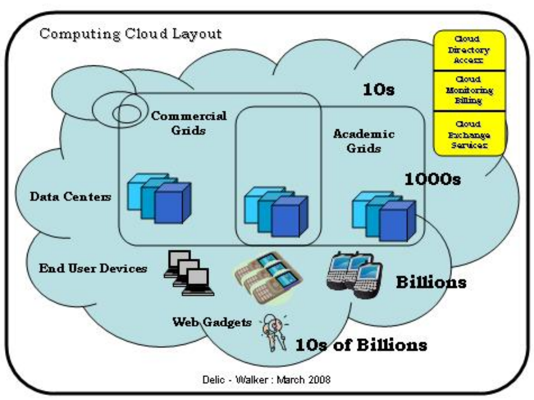

Figure 3. Computing Cloud

Computational grids originated in academic circles but it looks as if computing clouds may have a different growth path, as all important technology players have expressed their commercial interest to develop, contribute and play in this field.

4. The New Kind of Science(s)

The long history of science can be divided into three significant periods: empirical, theoretical and experimental/simulation – The first one was mainly experiential, without facilities to abstract, capture and share scientific knowledge. The second period of the great discoveries was followed by the formulation of the key scientific theories. They have been captured, shared and reused widely.

We are currently passing the period of the great experimental sciences creating torrents of data used for simulation and an unprecedented level of sharing and collaboration among scientists. Evolving mega-structures are hostingthe large scale social phenomena articulated as 'social computing'. There are suggestions to establish an entirely new type of scientific research based on the web fabrics called 'web science' [1].

We believe that the emergence of cloud fabrics will enable new insights into challenging engineering, medical and social problems. Previously, we had no easy means to tackle peta-scale type of problems, nor carry out mega-scale simulation - which cloud computing should enable. One interesting thread will be to revisit 50 year old problems in Artificial Intelligence or explore further the 'Science of Services' - both being under the auspices of Complex Systems research. In several ways the emergence of the computing cloud will invigorate academic research and will have strong potential to spawn innovative collaboration methods and new behaviors. To get an idea of this new type of project, one should look into SETI@home and FOLDING@home examples. Academic activities are well under way as the scientists have struggled for a long time with vast amounts of data coming not only from the web but also from a rising number of instruments and sensors [6]. Cloud computing has particular strength as it draws on many existing technologies and architecture and integrates centralized, distributed and 'software as service' computing paradigms into an orchestrated whole [10]. For now, C2 seems to be also a very promising computing technology paradigm [4]. We may yet observe emergence of Academic Cloud Computing (AC2) in the near future.

5. References

[1] Berners-Lee T. et al, Creating a Science of the Web, *Science*, Vol. 313, 11 August 2006. pp. 769-771
[2] Delic K., On Dependability of Corporate Grids, ACM Ubiquity, Volume 6, Issue 45 (December 7-13, 2005)
[3] Delic K.,Walker M., Architecting Enterprise Grids: Possible Inflection Points,
IADIS International Conference on Applied Computing, February 2007.
[4] Forrester Research Inc. Is Cloud Computing Ready for The Enterprise? March7, 2008
[5] Foster I., Service-Oriented Science, *Science*, Vol. 308, 6 May 2005. pp. 814-817
[6] Hand E., Head in the clouds, *Nature*, Vol. 449, 24 October 2007, p. 963
[7] Robison S., A Bright Future in the Cloud, Financial Times, March 4, 2008.
[8] Strong P., Enterprise Grid Computing, ACM Queue Magazine, July/August 2005.
[9] The Economist: When Clouds Collide, February 7, 2008.
[10] Weiss A., Computing in The Clouds, ACM netWorker, December 2007, pp. 16-25