Ubiquitous Cloud Computing

Ghislain Fouodji Tasse¹ and Karen Bradshaw² Department of Computer Science and Information Systems Rhodes University, P. O. Box 94, Grahamstown 6140 Tel: +27 46 6038291, Fax: +27 46 6361915 email: g09f5474@campus.ru.ac.za¹; k.bradshaw@ru.ac.za²

Abstract—With the tremendous development in computing technology, it has become possible for most users to enjoy the benefits of high performance or multipurpose computing. This is achieved through cloud computing, which is a term that generally refers to delivering hosted services (e.g. Platform as a Service) over the Internet. Because, we are interested in designing a general purpose computing platform, we will focus on the trend Infrastructure as a Service (IaaS). Such infrastructure will need to be self adaptive, fault-tolerant and easy to use. A solution is to use a dynamic grid, where diverse computer resources are put together to constitute a uniform resource in terms of processing power and memory (a cloud). We demonstrate how this can be particularly useful in an institution, where an internal cloud is built by harnessing workstations' processors idle times.

Index Terms—Cloud computing, high performance computing (HPC), computer resources.

I. INTRODUCTION

PARALLEL and distibuted computing are two concepts that were introduced in the past few decades to take advantage of the increase in computer hardware capabilities. These concepts have revolutionised the approach to programming and their implementation can be fairly complex. But, with the amount of work done in the area, their usage can be facilitated by abstracting certain complexities. This is why trends such as grid or cloud computing have arisen, which provide in their own way, a certain level of abstraction to one's computation. These trends can also be distinguished by the granularity (fine-grained or coarse-grained) of the computational problems they were natively developed to solve. This paper studies the use of a dynamic Grid as a candidate for provisioning computational resources in a ubiquitous computing environment.

In this paper, we present a background study on related research fields in Section II. Based on this study, we identify and describe in Section III our key research questions. Then, in Section IV, we describe our intended approach for resolving the issues revealed by the research questions. Section V presents our motivation behind the research, and shows how it can contribute to the world of computing. Finally, we conclude this paper and highlight some plausible extensions to this research in Section VI.

II. RELATED WORK

A cloud is an emerging IT delivery model, taking its place in the industrial scene. It seeks to enable users to gain access to their applications from anywhere, via any connected device [1]. Technically, clouds can be said to be virtual servers over the internet. Among others, users are provided with the ability to start and stop servers or use compute cycles only when needed, often paying only upon usage. Additionally, live migration of a virtual machine (VM) is possible, thus avoiding any interruption to ongoing activities, critical for real-time computing. It also permits multiple and distinct operating systems to safely coexist on the same platform, thereby enhancing productivity [2]. It is a key element for cloud computing because it provides important advantages in sharing, manageability and thus has an important business value [3]. In fact, a recent study (September 2009) performed by the Aberdeen Group found that some companies achieved on average an 18% reduction in their IT costs and a 16% reduction in data center power consumption by adopting cloud computing [4].

On the other hand, ubiquitous computing is considered as the third wave of computing after mainframe and desktop computing, where many computers serve one person. Effectively, to go beyond desktop computing the heterogeneity of the actual underlying infrastructure (heterogeneous network and computers) should be transparent to the user [5].

III. RESEARCH QUESTIONS

Based on the literature survey, we focus on the following questions throughout our research.

A. How can we increase computing performance without investing in a new infrastructure?

In other words, we need to figure out ways of using available computer resources more efficiently. In fact, hardware is currently under-utilised and it is believed that adequate software frameworks can be created to provide a set of new services to users. As defined in the previous section, the use of Cloud computing in a way, is an answer to the question.

B. Can a reliable grid computation be performed on a nondeterministic and heterogeneous network?

Though we know our answer to the previous question will be related to the trend cloud computing, we still need to cogitate on the design and the desired structure of the underlying computing infrastructure. Because our research implementation will attempt to generate a cloud on top of a set of disparate computer resources, via grid technologies, reliability is an expected concern and we will have to resolve it. The ownership and the spatial distribution of the computers are most likely to be different, but a safe collaboration needs to be established between them.

The authors would like to acknowledge the financial support of Telkom, Comverse, Stortech, Tellabs, Amatole Telecom Services, Bright Ideas 39, and THRIP through the Telkom Centre of Excellence in the Department of Computer Science at Rhodes University.

C. Is this research a step towards "personal computing"?

The term "personal computing" refers natively to the usage of personal computers (PCs). However, in this new age of technology, we believe it is no longer limited to PCs, and has invaded the area of grids. A generic term that names this type of computing was introduced in Section II, namely the term ubiquitous computing. Effectively, this research question is related to how super-computing can be made accessible to an average PC user. A model which is compliant with the initial state of the user computing environment is obviously required.

IV. RESEARCH DESIGN

Our initial aim is to design a robust cloud environment with the underlying hardware being a set of peer nodes with different abilities (network of workstations). In other words, we aim to build a Multiple System Image (MSI) where a node in the system sees the rest of the system as a single image. We are planing to use the open source cloud computing toolkit OpenNebula and other APIs like Libvirt to implement such a system [6], [7].

Fig. 1 shows an overview of the architecture for each node that enables them to cluster their resources and/or to use the other's nodes resources transparently (i.e., as a cloud service, typically IaaS). The next subsections describe the key building blocks of the intended design.

- 1) **Clustering mechanism**: Nodes are stand-alone computers and initially, they have no explicit connections between them. The idea is to cluster these nodes via grid technologies, and gain control on their computing resources via their virtualization drivers.
- 2) Scheduling mechanism: Nodes are not expected to be of the same nature, and will definitely not possess the same real-time capabilities. Due to the heterogeneity of such an environment, an adequate resources management policy has to be designed, and means for enforcing it have to be implemented. This will be achieved by writing a scheduling module, which will be used via the OpenNebula toolkit.
- 3) Information policy: Because of the decentralised nature of the components in our infrastructure, and its unpredictable behaviour, suitable means of communication between nodes should be provided. Information will be passed through the cloud, to keep it aware of ongoing events (e.g. removal of a node).



Figure 1. Overview architecture of a participating node.

The general idea is to set up a private network between the nodes (peers) that are willing to join the cluster and ensure the load is distributed fairly among them. For the user view (See Fig. 1), an administrative GUI is used to provide facilities such as monitoring the status of user's tasks running on the cloud.

V. MOTIVATION AND POSSIBLE APPLICATIONS

HPC is far from reaching regular computing users, but the ease of a cloud interface allows us to make this possible. Moreover, the extensible nature of the constructed infrastructure allows individuals or groups to donate their local infrastructure either on a permanent or on-demand basis. Our motivation is the ability to harness a distributed set of computer resources, and build an intelligent cluster with the following characteristics:

- Self adaptive, scalable, self optimizing.
- Transparency (ease of use): A user can easily run a computation on a supercomputer at a national centre while dynamically visualising the results on a local machine.
- Safety or reliability of the infrastructure (fault tolerance, multiple points of failure).

A. High availability

By clustering many computers, we can design an environment that provides reliability and high availability, which is key to applications like web servers or databases.

B. High performance

A high performance infrastructure is needed in many institutions (e.g. research institutions, hospitals). Performance can be achieved via the potential of parallelism offered by the platform.

VI. CONCLUSIONS AND FUTURE WORK

We observe throughout this paper that it is possible to build an internal cloud by harnessing groups of PCs within an institution, thereby providing a uniform and robust resource (IaaS). It is also possible to extend it to other areas. For instance, mobile devices have not yet been integrated into Grid computing platforms mainly due to their inherent limitations in processing and storage capacity, power and bandwidth shortages. But, we look forward to a future where such devices can participate actively in the ubiquitous cloud environment. Finally, like some cloud's service providers, a pricing policy can be integrated in the system.

REFERENCES

- [1] I. Corporation, "Seeding the clouds: Key infrastructure elements for cloud computing," February 2009.
- [2] M. F. Mergen, V. Uhlig, O. Krieger, and J. Xenidis, "Virtualization for high-performance computing," *SIGOPS Oper. Syst. Rev.*, vol. 40, no. 2, pp. 8–11, 2006.
- [3] S. Technologies, "Server virtualization and cloud computing: Four hidden impacts on uptime and availability," June 2009.
- [4] (2009, September) Business adoption of cloud computing: Reduce cost, complexity and energy consumption. Aberdeen Group.
 [Online]. Available: http://www.aberdeen.com/Aberdeen-Library/6220/ RA-cloud-computing-sustainability.aspx
- [5] M. Weiser. (1996) Ubiquitous computing. [Online]. Available: http://www.ubiq.com/hypertext/weiser/UbiHome.html
- [6] The open source toolkit for cloud computing. [Online]. Available: http://www.opennebula.org/
- [7] Libvirt: The virtualization api. [Online]. Available: www.libvirt.org

Mr Ghislain Fouodji Tasse completed his Honours in Computer Science in 2009 under the guidance of Dr Karen Bradshaw in the Department of Computer Science at Rhodes University. Ghislain is now in his first year of MSc in Computer Science and working for the Centre of Excellence (CoE) at Rhodes University.