BSC contributions in Energy-aware Resource Management for Large Scale Distributed Systems

Jordi Torres, Eduard Ayguadé, David Carrera, Jordi Guitart, Vicenç Beltran, Yolanda Becerra, Rosa M. Badia, Jesús Labarta and Mateo Valero

Barcelona Supercomputing Center (BSC)
Computer Architecture Department - Technical University of Catalonia (UPC)

{jordi.torres, eduard.ayguade, david.carrera, jordi.guitart, vbeltran, yolanda.becerra, rosa.m.badia, jesus.labarta, mateo.valero}@bsc.es

ABSTRACT

This paper introduces the work being carried out at Barcelona Supercomputing Center in the area of Green Computing. We have been working in resource management for a long time and recently we included the energy parameter in the decision process, considering that for a more sustainable science, the paradigm will shift from "time to solution" to "kWh to the solution". We will present our proposals organized in four points that follow the cloud computing stack. For each point we will enumerate the latest achievements that will be published during 2010 that are the basics for our future research. To conclude the paper we will review our ongoing and future research work and an overview of the projects where BSC is participating.

1. INTRODUCTION

Due to the escalating price of power, energy-related costs have become a major economic factor for ICT infrastructure and its host data centres. In addition to improving energy efficiency, companies are facing increasing pressure to reduce their carbon footprint due to EU regulations and campaigns demanding greener businesses. Our research community is therefore being challenged to rethink ICT strategies, adding energy efficiency to a list of critical operating parameters that already includes service performance or reliability.

In this paper we will present a brief overview of the current work that BSC is doing in the Green Computing field. Until now, our research was centered in performance management of distributed and parallel system [1]. Recently we included the energy parameter in the decision process, considering that for a more sustainable science, the paradigm will shift from "time to solution" to "kWh to the solution".

In order to be clearer, we decided to present the overview of our work, organizing the content of this paper following the well-known cloud computing stack based on three layers: Infrastructure-as-a-Service (IaaS), Platforms-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

We are considering that the current workloads that we should deal with are heterogeneous, including different types of jobs, not only CPU-intensive jobs, but also streaming, transactional, data-intensive, etc.

Regarding the resources, the current hardware that we should consider includes heterogeneous clusters of hybrid hardware (with different types of chips, accelerators, GPUs, …).

The research goals that will direct BSC research proposals, in addition to the goals that we have previously dealt with, like performance, includes different aspects: fulfilling the Service Level Agreements (SLA), considering the energy consumption or taking into account the new wave of popular programming models like MapReduce, amoung others.

However, these cloud goals have made the resource management a burning issue in today

systems. For BSC, Self-management is considered the solution to this complexity and a way to increase the adaptability of the execution environments to the dynamic behavior of Cloud Computing.

This is the BSC approach lead by one of our departments, the "Autonomic Systems and eBusiness Platforms" that is trying to build a "Smart Cloud" that can address the present challenges of the Cloud. The aim of this department is to research on autonomic resource allocation and heterogeneous workload management with performance and energy-efficiency goals for Internet-scale virtualized data centers comprising heterogeneous clusters of hybrid hardware.

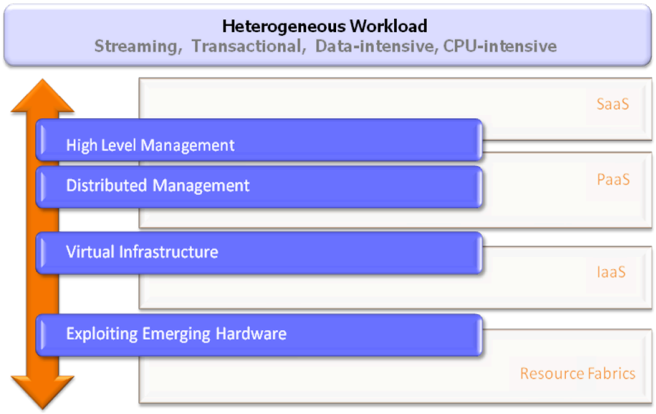

Fig 1. Cloud computing stack organization and summary of points.

As we mentioned previously, we will present our proposals under the cloud computing stack organization, where we can consider 4 main points. One giving solutions at the IaaS level, proposing a virtualized infrastructure. Another, offering a new proposal at the PaaS level. We also consider how the PaaS level functions can provide better support to the SaaS layer. Finally, we consider how emerging hardware can be exploited in an efficient way to reduce the whole energy consumption.

We are considering a whole control cycle with an holistic approach. By that, we mean that each level cooperates with the other levels through a vertical dialogue. Figure 1 shows a diagram that summarizes our proposals.

The rest of the paper revisits the different research points being conducted at the BSC, which emphasizes the main lines of work and the results obtained so far in each focus that will be published during 2010. We conclude the paper with the review of the ongoing work and we briefly list the projects in which the BSC is participating and are related to in this area.

2. VIRTUAL INFRASTRUCTURE

BSC is contributing the research community with the EMOTIVE framework, which allows to simplify the development of new middleware services for the Cloud. EMOTIVE framework is an open-source software infrastructure for implementing Cloud computing solutions that provides elastic and fully customized virtual environments in which to execute Cloud services. EMOTIVE abstracts a Cloud architecture using different layers and provides users with basic primitives for supporting the execution of services (features for resource allocation and monitoring, data management, live migration, and checkpointing, etc.). The core layer wraps each virtualized node and monitors its state, granting full control to the application of its execution environment without any risks to the underlying system or the other applications. One of the main distinguishing features of EMOTIVE framework is their functionalities that ease the development of new resource management proposals, thus contributing to the innovation in this research area. At the moment there are some scheduler implementations that take into account power-aware parameters. The EMOTIVE framework (www.EMOTIVEcloud.net) is build in collaboration with the Grid Computing and Clusters research group at BSC.

The latest achievements in this area are summarized in 2 contributions [2,3] that will be published during 2010.

3. DISTRIBUTED MANAGEMENT

At PaaS level we are working on application placement to decide where applications run and the allocated resources required. For this objective, applications must be modeled to make proper placement decisions in order to obtain a solution not only considering performance parameters, but also energy constrains. We are specially paying attention to MapReduce workloads (currently the most prominent emerging model for cloud scenario), working on the runtimes that allows control and dynamically adjusts the execution of these types of applications with energy awareness. Finally the energy awareness will be addressed at two levels: compute infrastructure (data placement and resource allocation) and network infrastructures (improving data locality and placement to reduce network utilization.

The latest research results in this point are summarized in 2 contributions [4,5] that will be published during 2010.

4. HIGH LEVEL MANAGEMENT

As we previously introduced, we are considering extending the Platform-as-a-Service layer functions to provide better support to Software-as-a-Service layer, according to high level parameters for resource allocation process. The main goal is to propose a new resource management aimed to fulfil the Business Level Objectives (BLO) of both the provider and its customers in a large-scale distributed system. We have preliminary results in the way that the decision-making is done in relation to several factors in a synergistic way depending on provider's interests, including business-level parameters such as risk, trust, and energy.

The most recent developments in this area are summarized in 2 contributions [6,7] that will be published during 2010.

5. EXPLOITING EMERGING HARDWARE

The fourth point deals with the study and development of both new hardware architectures that deliver best performance/energy ratios, and new approaches to exploit such architectures. Both lines of research are complementary and will aim to improve the efficiency of hardware platforms at a low level.

Preliminary results demonstrate that the energy modeling in real time (based on processor characterization) will be leveraged to make decisions.

The latest achievements in this point are summarized in 2 contributions [8,9] that will be published during 2010.

6. ON GOING WORK AND PROJECTS

Although we are considering a whole control cycle with a holistic approach, we will summarize our ongoing work by points in order to be more clear.

The major goal of the first focus was to be a testbed platform for our research, which is almost accomplished. We are currently working in extending the framework with plugins for third-party providers and the federation support for simultaneous access to several clouds that can take into consideration energy-aware parameters.

Our ongoing work in the second focus is devoted to Energy saving in MapReduce workloads. The key element to achieve energy efficiency at this level is the cooperation with the underlying platform and the dynamic modeling of application performance on real time, so that dynamic performance adaptation can be leveraged to control energy-consumption. The energy saving will be achieved by reducing network usage, placing DataNodes in low-consumption nodes or using hybrid data centers with suitable consolidation tasks strategy.

Our ongoing work in the third point, high level management, is continuing in the way of adding these new parameters (as Risk, Trust, Revenue,

Power efficiency, …) in the resource allocation process that a PaaS layer offers to SaaS layer.

Our work in the fourth focus is to leveraging hybrid systems to improve energy-saving. This work will address the problem of low-level programmability of hybrid systems that can result in poor resource utilization and, in turn, poor energy efficiency.

Finally, mentioning that BSC is participating in three projects in the area of Green Computing: the Spanish NUBA project (2009-2011), the EU OPTIMIS project (2010-2013), and the EU COST IC804 action (2009-2012).

7. ACKNOWLEDGMENTS

Many thanks to BSC researchers for their research effort and contributions to this field.

This work is partially supported by the Ministry of Science and Technology of Spain (contract TIN2007-60625), the NUBA project under contract MITyC TSI-020301-2009-30 , the Generalitat of Catalunya (2009-SGR-980), the European Commission in the context of the HiPEAC2 network of excellence (contract no. ICT-217068), the COST (European Cooperation in Science and Technology) framework, under Action IC0804, the MareIncognito and SoW Adaptive Systems project under the BSC-IBM collaboration agreement.

8. REFERENCES

[1] J. Guitart, J. Torres, and E. Ayguadé. *A Survey on Performance Management for Internet Applications*. Concurrency and Computation: Practice and Experience, Vol. 22 (1),pp.68-106, Jan 2010. ISSN: 1532-0634

[2] J. Ejarque, M. de Palol, I. Goiri, F. Julià, J. Guitart, R. Badia, and J. Torres. *Exploiting Semantics and Virtualization for SLA-driven Resource Allocation in Service Providers*. Concurrency and Computation: Practice and Experience, Vol. 22 (5), pp. 541-572, April 2010. ISSN: 1532-0634

[3] I. Goiri, F. Julià, J. Guitart, and J. Torres. *Checkpoint-based Fault-tolerant Infrastructure for Virtualized Service Providers*. 12th IEEE/IFIP Network Operations and Management Symposium (NOMS'10) Osaka, Japan, April 19-23, 2010, pp. 455-462 ISBN: 978-1-4244-5367-2, ISSN: 1542-1201

[4] J. Polo, D. Carrera, Y. Becerra, J. Torres, E. Ayguadé, M. Steinder, I. Whalley. *Performance-Driven Task Co-Scheduling for MapReduce Environments*. 12th IEEE/IFIP Network Operations and Management Symposium (NOMS'10). Japan, April, 2010.

[5] J. Polo, Y. Becerra, D. Carrera, V. Beltran, J. Torres and E. Ayguadé. *Towards Energy-Efficient Management of MapReduce Workloads*. Poster session. 1st Int'l Conf. on Energy-Efficient Computing and Networking (e-Energy 2010), 2010.

[6] J. Ll Berral, I. Goiri, R. Nou, F. Julia, J. Guitart, R. Gavaldà and J. Torres. *Towards energy-aware scheduling in data center using machine learning*. 1st Int'l Conf. on Energy-Efficient Computing and Networking (e-Energy), 2010.

[7] I. Goiri, J. Guitart, J. Torres. *Characterizing Cloud Federation for Enhancing Providers' Profit*. 2010 IEEE 3rd International Conference on Cloud Computing (CLOUD 2010).

[8] R. Bertran, M. Gonzalez, Y. Becerra, D. Carrera, J. Torres and E. Ayguade. *Towards Accurate Accounting of Energy Consumption in Shared Virtualized Environments*. Poster session. 1st Int'l Conf. on Energy-Efficient Computing and Networking (e-Energy'10). April, 2010.

[9] J. Polo, D. Carrera, Y. Becerra, V. Beltran, J. Torres, E. Ayguadé. *Performance Management of Accelerated MapReduce Workloads in Heterogeneous Clusters*. 39th International Conference on Parallel Processing (ICPP2010). San Diego, CA, USA. September 2010.