Enabling Autonomic Computing on Federated Advanced CyberInfrastructures

Javier Diaz-Montes, Mengsong Zou, Ivan Rodero, Manish Parashar
Rutgers Discovery Informatics Institute
NSF Cloud and Autonomic Computing Center
Department of Electrical and Computer Engineering, Rutgers University
{javidiaz, mz228, irodero, parashar}@cac.rutgers.edu

ABSTRACT

We present a federation model to support the dynamic federation of resources and autonomic management mechanisms that coordinate multiple workflows to use resources based on objectives. We illustrate the effectiveness of the proposed framework and autonomic mechanisms through the discussion of representative use case application scenarios, and from these experiences, we discuss that such a federation model can support new types of application formulations.

Categories and Subject Descriptors

C.2.4 [**Computer - Communication Networks**]: Distributed Systems

Keywords

Cloud Computing, Federation, Autonomic Computing

1. INTRODUCTION

Cloud computing has emerged as a dominant paradigm that provides on-demand access to computing utilities, an abstraction of unlimited computing resources, and support for on-demand scale up, scale down and scale out. Clouds abstractions and infrastructure are rapidly becoming part of the overall research cyberinfrastructure, providing viable platforms for scientific exploration and discovery. Computational and Data-enabled Science and Engineering (CDS&E) applications enabled by advanced cyberinfrastructure are providing unprecedented opportunities for understanding and managing natural and engineered systems, and offering unique insights into complex problems and, in addition to support traditional enterprise data analytics services (e.g., those based on MapReduce) [1, 3, 5]. The simplicity of the cloud abstraction can alleviate some of the problems scientific applications face in current HPC environments. For example, many task computing applications can benefit from the easy access to elastic and readily accessible resources and the ability to easily scale up, down or out. Clearly, realizing these benefits requires the development of appropriate application platforms and software stacks [6].

In this paper, we present a model to federate distributed resources on demand and coordinate the execution of application workflows. These resources can be of different types of infrastructure including traditional HPC clusters, supercomputers, grids, and clouds. Additionally, the federation provides autonomic scheduling mechanisms that create an abstraction with cloud-like capabilities to elastically provision the resources based on user and application policies and requirements. We also discuss relevant usage scenarios of the proposed framework, including: (1) medical image research, which aims at achieving extended capacity, (2) molecular dynamics simulations using asynchronous replica exchange, which provides adaptivity and elasticity at the application-level, and (3) data analytics workflow based on clustering, which focuses on adaptation to achieve user objectives and requirements.

2. FEDERATING ADVANCED CYBERINFRASTRUCTURES

The federation model that we propose is aimed to orchestrate heterogeneous distributed resources using cloud-like capabilities and abstractions. Our federation can scale and extend its capacity by dynamically aggregating geographically distributed resources. Moreover, it is able to interoperate with different type of resources and take advantage of the different capacities that they can offer. Finally, it has to create an abstraction on top of the resources to provide users with on demand access to the resources and the ability to scale up, down or out as needed.

The federation supports master/worker, MapReduce and workflow programming models to ease the development of applications. Applications are typically described as workflows with multiple stages, where the output of one stage is the input of the next one. Each stage can run a different application or the same application with different length of tasks, computational requirements and data.

Applications are executed using the autonomic capabilities of the federation, which guides the process of resource provisioning to meet the objectives and constrains defined by the application. The autonomic capabilities are provided by the autonomic manager, which is responsible for managing workflows, estimating runtime and scheduling tasks at the beginning of every stage based on the resource view provided by the agents. At each stage, the adaptivity manager monitors tasks runtimes through results, handles the changes of application workloads and resource availability,

CAC'13, August 5–9, 2013, Miami, FL, USA.

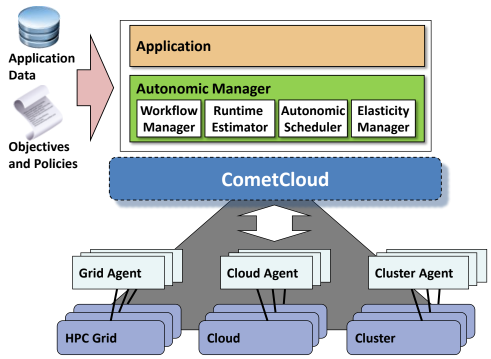

Figure 1: Architectural overview of the autonomic application management framework.

and adapts resource provisioning if required. Figure 1 shows the architecture of the autonomic management framework.

3. APPLICATION SCENARIOS

This section presents representative use case applications and experiences to illustrate the effectiveness of our proposed federation architecture, mechanisms and autonomic strategies. We believe the discussion of these experiences are useful to define next steps towards advanced cyberinfrastructure and clouds federation for different usage modes [4].

Medical image research: In this application use case we focus on Content-based image retrieval (CBIR) on digitized peripheral blood smear specimens using low-level morphological features. Specifically, we use CometCloud to execute CBIR in federated heterogenous advanced cyberinfrastructure and cloud resources with the goal of reducing the completion time. We observed that the application can dynamically scale across heterogeneous resources located in different HPC and cloud infrastructures. This presents many opportunities and challenges in the context of medical image research such as exploiting heterogeneous federated resources from the point of view of their capabilities. For example, the Matlab incarnation of CBIR can be run when licenses are available or an implementation for accelerators (e.g., GPU or Intel MIC) can be run when resources with accelerators are available.

Molecular dynamics simulations using asynchronous replica exchange: Typically molecular dynamics simulations are very static in terms of execution models. However, we evaluate an approach where we can monitor the progress of the replicas, using the secondary structure prediction methods and the radius of gyration, and perform actions based on the monitoring information. Here, we take advantage of an asynchronous formulation of the replica exchange [2] to implement a workflow on top of CometCloud to run simulations on dynamically federated large-scale distributed resources. Specifically, we can use clouds to quickly explore the application domain space saving the HPC allocations to compute only those replicas that are identified as relevant using the monitoring information.

Enterprise Business Data Analytics: Current enterprise business data analytics workflows combine different techniques in their stages such as MapReduce-like applications that aggregate large amounts of data from different sources for business intelligence with clustering techniques. These data points in the multi-dimensional information space can be clustered using Distributed Online Clustering (DOC) to search results and correlate them with known data sources, and allow visualizing and interpreting the results interactively through a GUI. The specific solution in this application use case is a federated hybrid cloud for handling "big data" through DOC. As part of this application scenario, we evaluated the autonomic manager by showing how to achieve user objectives such as time constraint and deadline using cloudbursts to a public cloud when the local resources are limited.

4. CONCLUSION

In this paper, we have presented a federation model that enables the orchestration of hybrid distributed infrastructures and the coordinated execution of application workflows using autonomic capabilities. We experimentally investigated, from an application's perspective, possible usage modes for integrating HPC and clouds as well as how autonomic computing can support these modes. In particular, we used three use case scenarios to highlight different aspects of the federation. First, we showed how medical image research applications can benefit from the federation of distributed resources and their aggregated computational power. Then, we exploited the principles of adaptivity and elasticity in the context of a molecular dynamics application. Finally, used a data analytic application to performed a deadline objective-driven workflow execution.

5. ACKNOWLEDGMENTS

This research is supported in part by NSF grants OCI 1310283, OCI 1339036, IIP 0758566 and DMS-0835436, and used resources from XSEDE (NSF grant OCI-1053575), FutureGrid (NSF grant 0910812), and NERSC (DoE contract DE-AC02-05CH11231).

References

[1] I. Gorton, Y. Liu, and J. Yin. Exploring architecture options for a federated, cloud-based system biology knowledgebase. In *IEEE Intl. Conf. on Cloud Computing Technology and Science*, pages 218–225, 2010.

[2] Z. Li and M. Parashar. Grid-based asynchronous replica exchange. In *IEEE/ACM GRID*, pages 201–208, 2007.

[3] S. Ostermann, R. Prodan, and T. Fahringer. Extending grids with cloud resource management for scientific computing. In *IEEE/ACM Grid*, pages 42–49, 2009.

[4] M. Parashar, M. AbdelBaky, I. Rodero, and A. Devarakonda. Cloud Paradigm and Practices for CDS&E. Technical report, Cloud and Autonomic Computing Center, Rutgers University, 2012.

[5] C. Vazquez, E. Huedo, R. Montero, and I. Llorente. Dynamic provision of computing resources from grid infrastructures and cloud providers. In *Grid and Pervasive Computing Conf.*, pages 113–120, 2009.

[6] D. Villegas, N. Bobroff, I. Rodero, J. Delgado, et al. Cloud federation in a layered service model. *J. Comput. Syst. Sci.*, 78(5):1330–1344, 2012.