Application to NSF PI meeting on "The Science of Cloud Computing"

Name, affiliation and email address

José A.B. Fortes, University of Florida, E-mail: fortes@ufl.edu

Topics of research interests in Cloud Computing

Network Support for the Cloud, Fault Masking in the Cloud, Cloud Self-Monitoring and Autonomic Control, Cloud Inter-operability and Standardization

Summary of current research activities related to Cloud Computing

I am the Director of the NSF-funded four-university Center for Autonomic Computing (CAC) and a co-PI of the NSF-funded 10-university FutureGrid (FG) project. In this context my activities span the following areas:

1. Virtual networking for intercloud computing: this work focus on the design and development of ViNe, a virtual networking technology that enables the deployment of overlay networks on WANs across firewalls and NAT devices. The solution is platform independent and application agnostic and can be deployed at either user-level or system-level. It has been demonstrated in cross-cloud bioinformatics applications involving up to five commercial, public and private clouds.

2. Datacenter management for elastic computing: this work focus on the problem of on-demand provisioning of virtual machines and dynamic resource allocation to virtual machines in order to provide Quality of Service to dynamic workloads. The approaches adopted in our work rely on two-level autonomic managers (at the application VM level and at the physical resource layer) implemented using fuzzy logic and genetic algorithms. Multi-objective criteria including performance, power and thermal dissipation metrics are used to guide the optimization process. Cross-layer approaches are also used to integrate these criteria.

3. Management of the health of IT cloud systems: this work considers the design of IT systems whose components are provided by IaaS clouds and the challenge of ensuring fault-free behavior by such systems. The approach relies on modeling of system behavior so that faults can be predicted and remaining useful life can be managed so to prevent faults and the associated recovery actions if they were allowed to occur.

4. Bioinformatics applications on multicloud systems: this work uses virtual networking (ViNe), VM migration technologies (from collaborative work with INRIA, Rennes), contextualization techniques (Nimbus-based), resource consumption prediction and Hadoop to scale bioinformatics applications across multiple clouds.

5. Intercloud testbeds: this work focuses on the deployment of hardware and middleware to implement NSF-funded and industry-funded testbeds across the CAC and FG universities to study standards, middleware and applications for intercloud systems. These testbeds include all major middleware stacks for cloud computing and interfaces to several public and commercial cloud providers of infrastructure as a service. Their initial implementations are already being used.

Abstract of future research problems related to Cloud Computing

The above-listed current research activities in Cloud Computing provide the basis for future research activities which will include:

1. Virtual clouds: clouds built by using, combining or connecting clouds are starting to emerge and will become common and necessary in the future. Autonomic virtual networks will be essential for the deployment of services that transparently use multiple clouds. This technology will enable self-managed virtual networks as a service as well as user-created autonomic networks.

2. Intercloud standards: we have engaged with IEEE and several industrial partners to develop and deploy an infrastructure for research that creates, evaluates and integrates proposals for cloud computing standards targeting interoperability and other requirements of intercloud computing. This testbed builds on existing capabilities already in place but additional work will be done to provide all the required functionality for the standards needed at all levels of the cloud middleware stack.

3. Cloud component modeling: resources and applications provided as a service are the elemental components of future IT systems. To engineer such systems it is essential to have their components described by models that capture important characteristics, e.g., structure, behavior, reliability and security. These models should be usable for purposes of analysis, validation and design of cloud-based IT systems. Service-level agreements are a starting point for this endeavor but much additional work is needed to extend current SLAs to the envisioned models.

4. Self-caring IT systems using cloud components: proactive self-management of the health of cloud-based IT systems will require systematic methods to design, model, and manage the reliable operation of these systems. The resulting designs need to include autonomic controllers that can anticipate faults and redirect system operation and/or reconfigure system structure so that such faults never occur. Event-based modeling (e.g. Petri nets) are being investigated to provide both the modeling capability and the controller architecture for such purposes.

5. Applications of cloud computing: applications based on workflows across multiple clouds and applications with real-time constraints are being considered in the areas of bioinformatics, collaboration work and human interface modeling.