Name: Marty Humphrey

Affiliation: Department of Computer Science, University of Virginia

Email addr: humphrey@cs.virginia.edu

2 best-fit topics from the list:

1. #1 – Cloud Architectures and Systems
2. #4 – Programming Models for the Cloud

Summary of Current Research Activities Related to Cloud Computing

Our existing NSF-supported research activity addresses performance of virtualized platforms, particularly as it applies to Cloud Computing. The use of virtual machines for Escience has been advocated both within the enterprise to replace aging machines and as the underlying technology of Cloud Computing. However, both scenarios can lead to inadequate performance. Within the enterprise, with incorrect planning or under unexpected heavy or even moderate load, there might not be enough physical capacity for every virtual machine to achieve reasonable performance. In cloud-computing-based scenarios, the "renters" are largely subject to the informal service promises of the cloud provider based on a granularity that can be too coarse or at the wrong level of abstraction. This existing project pursues a novel unified framework to ensure predictable Escience based on these two dominant emerging uses of virtualized resources. The foundation of the approach is to wrap an Escience application in a *performance container* framework and dynamically regulate the application's performance through the application of formal feedback control theory. The application's progress is monitored and ensured such that the job meets its performance goals (e.g., deadline) without requiring exclusive access to physical resources even in the presence of a wide class of unexpected disturbances. This project extends this foundation and early results in three important dimensions: creating support for non-specialists to use the framework; implementing these techniques in Eucalyptus, one of the major open-source cloud computing frameworks; and applying the techniques to "Software-as-a-Service" (SaaS), in which applications in the cloud are regulated to provide predictable performance.

We also have a strong collaboration with Microsoft Research on the use of Cloud Computing for eScience. The application in focus is the use of Windows Azure for processing satellite imagery. Through the design and implementation of this application – called *MODISAzure* – we have gained valuable insights into how best design cloud applications.

Future research problems that I think are important and I plan to pursue, if I am to be funded by the NSF, related to Cloud Computing.

The real-world application of Cloud technologies to (large-scale) science problems continues to be a challenging issue – that is, the early Cloud technologies have created a small collection of mechanisms and APIs, and it is not clear if these are really the right set for real-world scientists. To address this fundamental concern, my research group has had preliminary discussions with a number of domain scientists to begin to make this determination. A particularly strong potential collaboration I will highlight is with Prof. Jon Goodall of the University of South Carolina. In this partnership, we are attempting to advance hydrologic science and water resource management by leveraging cloud computing for modeling large watershed systems. Cloud computing offers the potential to transform how hydrologic systems are modeled by allowing modelers to include more detail in how they abstract large hydrologic systems within computational simulations through an on-demand access to computational resources. There are many important research questions that we would like to address in order to understand how a cloud computing paradigm can and should be used to model hydrologic systems. For example, it is unclear how to best divide the computational tasks within a complex hydrologic model so that the tasks can be performed efficiently within the cloud. It is also unclear to what extent cloud computing can be used to address other computationally-expensive tasks critical to the overall modeling process, including data preparation and model calibration. We do not know the practical limits and best practices when applying cloud computing for hydrologic systems, especially as these relate to quantifying costs associated with cloud resources. I believe that real partnerships such as this between computer scientists and domain scientists (in this case Hydrologists) are necessary to really understand and address the limits of the computational model provided by emerging Cloud technologies.

In general, I believe that a fundamental problem is that developers of cloud computing applications are faced with an increasingly difficult problem of comparing various cloud offerings and deciding which platform/vendor to choose. The core issue is to determine the functionality and performance of the services offered and how those match the requirements of the application. Simply, different cloud options offer different “built-in” mechanisms with different service-level agreements (SLAs). Comparing two cloud options based on cost per unit of compute, storage, and bandwidth can challenging – e.g., one cloud’s virtual machine (VM) capabilities can have subtle differences to another cloud’s so they cannot be directly compared, one cloud might charge bandwidth cost based on time of day of access, etc. It is even more difficult to compare based on the often ambiguously-worded SLAs presented as a PDF or an HTML page! I plan to conduct research to address this concern – simply, the cloud application developer needs to be able to decide the best cloud platform based on today’s financial costs and cloud capabilities, and this choice must not prevent the developer from exploiting the new capabilities offered in the future by *other* cloud platforms. This is a critical issue for the success of Cloud technologies, and touches upon performance issues, programmability, and evolvability.