Topic 6: Grid, Cluster and Cloud Computing

Erik Elmroth, Paraskevi Fragopoulou, Artur Andrzejak, Ivona Brandic, Karim Djemame, and Paolo Romano

Topic Committee

Grid and cloud computing have changed the IT landscape in the way we access and manage IT infrastructures. Both technologies provide easy-to-use and on-demand access to large-scale infrastructures. Grid and cloud computing are major research areas with strong involvement from both academia and industry. Although significant progress has been made in the design, deployment, operation and use of such infrastructures, many key research challenges remain to achieve the goal of user-friendly, efficient, and reliable grid and cloud infrastructures. Research issues cover many areas of computer science to address the fundamental capabilities and services that are required in a heterogeneous environment, such as adaptability, scalability, reliability and security, and to support applications as diverse as ubiquitous local services, enterprise-scale virtual organizations, and internet-scale distributed supercomputing. While there are several differences, grid and cloud computing are closely related in their research issues. Both areas will greatly benefit from interactions with the many related areas of computer science, making Euro-Par an excellent venue to present results and discuss issues.

The issues to be covered include but are certainly not limited to the following: middleware; applications and platforms; interoperability and portability; aggregation and federation of grids and clouds; efficient energy usage of resources; resource/service/information discovery; resource management and scheduling; programming models, tools, and algorithms; dependability, adaptability, and scalability; security for grids and clouds; workflow management; accounting, billing and business models; automated or autonomic management of resources and applications; quality-of-service and Service-Level-Agreement.

- The paper entitled *Scalable Reed-Solomon-based Reliable Local Storage for HPC Applications on IaaS Clouds* by Leonardo Bautista Gomez, Bogdan Nicolae, Naoya Maruyama, Satoshi Matsuoka and Franck Cappello, introduces a novel persistency technique that leverages Reed-Solomon (RS) encoding to save data in a reliable fashion on IaaS Cloud computing platforms to be used for HPC applications. Compared to traditional approaches that rely on block replication, this technique demonstrates about 50% higher throughput while reducing network bandwidth and storage utilization by a factor of 2 for the same targeted reliability level. This is achieved both by modeling and real life experimentation on hundreds of nodes.
- Subsequently, the paper *Caching VM Instances for Fast VM Provisioning: A Comparative Evaluation*, by Pradipta De, Manish Gupta, Manoj Soni

C. Kaklamanis et al. (Eds.): Euro-Par 2012, LNCS 7484, pp. 311–312, 2012.

and Aditya Thatte from IBM Research India, presents a method to overcome the delays in transfer and booting time for the preparation of VMs in cloud environments. Alternatively, a VM is prepared a priori, and saved in standby state in a "cache" space collocated with the compute nodes. On receiving a matching request, the VM from cache is instantly served to the user, thereby reducing service time. Based on usage data collected from an enterprise cloud, and through simulation, it is shown that a reduction of 60% in service time is achievable.

- The paper *Improving Scheduling Performance using a Q-Learning-based Leasing Policy for Clouds* by Alexander Fölling and Matthias Hofmann from TU Dortmund University, presents a reinforcement learning-based policy which controls the maximum leasing size of cloud computing resources with regard to the current resource/workload state and the balance between scheduling benefits and costs in an online adaptive fashion. Furthermore, it provides an appropriate model to evaluate such policies and presents heuristics to determine upper and lower reference values for the performance evaluation under the given model. Using event driven simulation and real workload traces, the authors were able to investigate the dynamics of the learning policy and to demonstrate the adaptivity on workload changes.
- Finally, the paper *Impact of Variable Priced Cloud Resources on Scientific Workflow Scheduling* by Simon Ostermann, Radu Prodan from the University of Innsbruck, analyzes the problem of provisioning Cloud instances to large scientific workflows that do not benefit from sufficient Grid resources as required by their computational requirements. An extension is proposed to the dynamic critical path scheduling algorithm to deal with the general resource leasing model encountered in today's commercial Clouds. The availability of the cheaper and unreliable Spot instances is analyzed and their potential to complement the unavailability of Grid resources for large workflow executions are studied. Experimental results demonstrate that Spot instances represent a 60% cheaper but equally reliable alternative to Standard instances provided that a correct user bet is made.

We would like to take the opportunity of thanking the authors who submitted a contribution, as well as the Euro-Par Organizing Committee, and the external referees with their useful comments, whose efforts have made this conference and this topic possible.