Thesis Summary

Federated Capacity Planning for Distributed Computing Infrastructures

Computing power has become the fifth utility for almost every application area in science and industry, and its ubiquitous availability is a key factor for advances in biotechnology, climate research, and product design. To cater this need, various concepts for the comprehensive, distributed provisioning of processing capacity have been developed over the last decade, and the most well-known representatives arguably are academic Grid computing and commercial Cloud computing. For that matter, both approaches share a certain tension between consumers and providers: While the former—researchers, engineers—express their desire for on-the-spot, reliable availability of computing power, the latter—academic computing centres, commercial data centres—care about economic and manageable provisioning.

In this work, we explore this tension with respect to capacity planning. To this end, we develop an appropriate model that reflects the needs of modern e-Infrastructures and embraces the results of three decades of distributed computing research. Using that model, we analyse whether—and if so, in how far—the building of federated infrastructures is reasonable from the stakeholders' (consumers and providers) perspective without disregarding their individual interests. In that context, we develop scheduling strategies for different architectures and approximate their limits. Finally, we evaluate the perquisites that need to be fulfilled in the technical architecture in order to transfer the afore discussed algorithms to real-world scenarios. To this end, we elicit the requirements from two production environments and develop a generalised interaction model that allows the application of the analysed strategies. There, we show that, using two newly designed protocols, not only the methods can be successfully transferred, but also—due to the extensibility of both the protocols and the architecture—envisage a manifold of other application scenarios.

Keywords: Scheduling; Resource Management; Distributed Computing; Grid Computing; Cloud Computing; Architecture; Standards