

Undergraduate Internship Program

Pedro Neves
Fall 2015

This report provides a short summary of my work during the UIP in Carnegie Mellon University (CMU) hosted by Professor David Garlan in the Institute for Software Research (ISR) from September 24th to December 20th, 2015.

My main goal here was to review the state-of-the-art of Big Data, Cloud and Architecture-based self-Adaptation techniques so that I can propose self-adaptable mechanisms for cloud data management. When I got to Carnegie Mellon I was immediately included in the self-adaption group that meet two times per week. Since I had few to none experience in self-adaption techniques, the group helped me a lot throughout my journey here. Besides Professor David Garlan and the remain members of the self-adaptation group, Bradley Schmerl (Principal Systems Scientist) Javier Camara (Systems Scientist) and Jorge Bernardino (Professor at ISEC and University of Coimbra), which were my advisors, had helped me a lot. With their help and orientation I started to review the state-of-the-art of Big Data systems and in particular Big Data systems in cloud environments. My goal was to understand what kind of dimensions these frameworks would typically allow to change. This research would reach its higher point with a presentation to the self-adaption research group, where we discussed several options and several dimensions that could be changed using Rainbow framework (a self-adaption platform developed by the group). Relying on the group suggestions, I focused on a particular set of Infrastructure, Platform and Application, which was Microsoft Event Hubs and Apache Storm as applications, and Microsoft Blob Storage as the platform, which is provided within the Microsoft HDInsight on the Azure cloud (Infrastructure). Soon I began searching for the dimensions that Microsoft would allow to change. The final result was a report which states the dimensions that can be changed at each level, provides the tactics, API and metrics so that further work can be performed to implement rainbow as a self-adaption mechanism to scale several dimensions of the Microsoft environment described. Furthermore, while enrolled in the UIP I was able to publish one paper, submit another one and participating in another one related with my work here. I am currently working on another paper over scalable and elastic techniques for Big Data systems in cloud environments.

The environment in the ISR and in particular among the self-adaption group was very stimulating. The university provided me the opportunity to assist to the 2015 SPLASH conference, hosted in Pittsburgh and access to many talks, among them some by: Jet propulsion Laboratory (JPL); NSA – with the presence of the NSA’s director, admiral Michael S. Rogers; and EA Games.

My stage at the Carnegie Mellon University was undoubtedly very enriching, I was able to be in contact with cutting-edge science, professionals and companies. In a nutshell, my participation in the UIP was very interesting and rewarding and I would definitely recommend it to anyone else.