Escrito por Alan Vital Pereira

Sex, 26 de Março de 2021 14:56 - Última atualização Sex, 26 de Março de 2021 18:16

Titulo: **IoT Service Placement with Load Distribution and Service Migration in Edge Computing for 5G Networks**

Data: **31/03/2021**

Horário: **10:30**

Local**: Videoconferência**

Resumo:

Edge Computing (EC) is a promising concept to alleviate some of the cloud computing limitations in supporting Internet of Things (IoT) applications, especially time-sensitive applications, by bringing computing resources closer to end users, at the network edges. As promising as EC is, it also faces many challenges. These are mainly related to the resource management in the vast, distributed, dynamic, and heterogeneous setting brought by EC. A relevant issue for resource management is the service placement problem, which is the decision-making process of determining where to place different applications or services over the EC infrastructure according to some constraints, requirements, and performance goals. This decision- making process can thus be extended to include other related issues, such as load distribution and service migration. Load distribution relates to the "how" to distribute workloads of an application among its various instances placed over the EC infrastructure. Meanwhile, the service migration aims at tackling the "when" an application or service should be migrated to another location to keep satisfactory system performance. As such, these two issues are

Escrito por Alan Vital Pereira
Sex, 26 de Março de 2021 14:56 - Última atualização Sex, 26 de Março de 2021 18:16

integral parts of those that needs to be tackled by EC resource management and, more specifically, in the case of service placements. In this thesis, we investigate the IoT services placement with load distribution and service migration in the context of next generation networks with EC capabilities, such as the fifth-generation (5G) mobile system. First, we address service placement with load distribution as single and multi-objective problems and we the proposal to solve these using a well-chosen genetic algorithm. Analytical results show that through our proposed formulation and the associated proposed algorithms, we are able to outperform other benchmark algorithms in terms of multiple conflicting objectives, such as response deadline violation, operational cost, and service availability. Then, in order to handle load fluctuations, we propose a centralized limited look-ahead prediction control that periodically readjusts service placement and load distribution decisions by taking into account the performance-cost trade-off of service migrations. Evaluation results show that our predictive control has even better system performance regarding response deadline violations with a small additional migration cost compared to benchmark algorithms. Finally, we address the scalability issue faced by centralized decision- making process. This is achieved by designing a hierarchical distributed service placement solution. This solution reduces the dimensionality of the overall control decision problem by decomposing it into a set of local control problems solved in a hierarchical cooperative way. The evaluation of our distributed control indicates that the trade-off between the system performance and the scalability of decision-making depends on how the control decision problem is decomposed.

Banca examinadora:

- Miguel Franklin de Castro (Orientador)
- Yacine Ghamri-Doudane (Coorientador) (La Rochelle Université (ULR) - França)
- Dario Vieira Conceição (EFREI Paris - França)
- Miguel Elias Mitre Campista (Universidade Federal do Rio de Janeiro (UFRJ))
- Rami Langar (Université Gustave Eiffel - França)
- Leila Merghem-Boulahia (Université de Technologie de Troyes - França)
- Emanuel Bezerra Rodrigues (UFC)
- Roch Glitho (Concordia University - Canada)