ENCQOR 5G Technology Development Challenge Statement

# Algorithms for virtualized network embedding

| | |
|---|---|
| **Challenge Launch Date** | • December 13, 2018 |
| **Deadline for Open Call Applications** | • January 24, 2019 |
| **Challenge Statement** | 5G network slicing encompasses the placement and interconnection of virtualized network functions (VNFs) on a physical network to create end-to-end network services. The physical network provides compute, connection, and storage resources to the VNFs. A successful application of this approach would produce a cost-effective allocation of resources to users with very different quality-of-service (QoS) requirements. However, the science of path computation and resource assignment still needs to be fully discovered to ensure that this approach is effective.<br><br>This project will devise algorithms to perform the allocation of physical resources to virtual network functions. Ideally, the algorithms should ensure that network resources are used “optimally” in the sense that the network is serving the maximum number of users it can serve, or to minimize network capital cost without compromising QoS. In practice, the algorithms must be computationally feasible in that they are fast and scale well. The final algorithms must therefore provide quantifiably good approximations of any optimal solution. |
| **Project Partner** | • Ciena |
| **Timeline** | • 2 years |
| **Available funding** | • Up to $130 000 CDN |
| **Applicant Type** | • Ontario based university |
| **Location** | • Ontario |
| **Project Details** | 5G network slicing encompasses the placement of virtualized network functions on islands of compute in the network and connecting them together with the physical network to create end-to-end network services. This goes well-beyond current practices of dividing up network bandwidth into virtual private network connections or the limited capabilities of current orchestration and virtualization solutions.<br><br>The concept of this type of network slicing was introduced by the 3rd Generation Partnership Project (3GPP) to enable cost-effective allocation of resources to users with very different quality-of-service (QoS) requirements. In 5G, network slices are completely connected virtualized networks running on general and specialized compute in-network infrastructure. While some parts of a network slice may need to use hardware acceleration in physical network functions (PNFs): wireless forward error correction, network functions such as forwarding, load-balancing and packet filtering, all other functionality is implemented in software with virtual network functions (VNFs) and running on general purpose compute in service providers’ data centers. |

Network slices achieve cost savings by pooling users with the same QoS requirements into network slices providing similar QoS. Cost savings are achieved by (1) designing the network with PNFs only where they are needed and (2) by not mixing network loads with different QoS's together. For example, all ultra-reliable and low-latency (uRLL) users may be in the uRLL slice, while massive broad band (MBB) users may be in an MBB slice. Generally speaking, uRLL users will have their network functions (virtual or physical) allocated in data centers near the network edge, thus resulting in little network traffic to the core. On the other hand, the less stringent latency requirements of MBB users means that they can have their virtual network functions served by data centers further away from the edge, but they may induce in a significant amount of network traffic towards the core.

From the point of view of network slice orchestration - allocating compute and PNF resources and instantiating virtualized network functions - it is important to know, which PNFs to assign to which network slice and which computing resources to assign to which VNF, so that network resources are optimized. Network resources are used "optimally" when the network is serving the maximum number of users it can serve, or when the network capital cost is minimized while serving the same number of users.

We identify three problems that need to be solved in this setting:

- Resource assignment when instantiating network slices that meets the QoS requirements of the network slice and optimizes network resources. A 5G network slice is a whole virtual network that needs to be embedded into the physical network and the available compute resources. From a mathematical graph perspective, network slice is a tree that needs to be embedded in a graph, with special restrictions on meeting QoS requirements and availability of resources. A sub-problem of this problem is to dynamically extend the network slice (tree) by adding services in not as yet covered area, which is equivalent to extending the existing tree with a path to the area requiring coverage.
- Defragmentation of network resources, which re-optimizes the network slice allocation. The defragmentation is required when new physical resources are added to the network, or when a network slice is reduced or removed from the network. A network slice may be reduced at different times of the day to save on energy use in the data centers. The defragmentation process re-assigns network resources to return the network to its optimal state and provides an order list of VNF and network connection moves that accomplishes it without disrupting services.
- Network design and resource planning, which uses analytics derived from network slice dynamics to design and plan the expansion of the bandwidth and computing resources.

Our expected outcome of the research collaboration is a suite of algorithms, which find the answers to the three optimization problems. Practically, the algorithms should be fast, robust, and easy to debug in the field. Ideally, the algorithms should take into account the practicalities of networks, such as the multi-layer aspects (IP over DWDM), protection schemes, and statistical multiplexing of packet and compute resources.

Due to the computational complexity of the stated problems, we expect that the final algorithms will not actually be able to find the optimal solutions to our stated problems in a timely fashion. We expect the academic partner to develop heuristics approximating the optimal solution. The project will be structured in a way, which

| | |
|---|---|
| | allow us to evaluate how well the delivered algorithms work with respect to the optimal solutions. |
| **Project Goals/ Outcomes** | We expect this project to consume the time of at least two PhD candidates over a period of 2 years. As such, we expect this research to result in at least four conference and three journal publications.<br><br>Another goal of the project is to deliver a working prototype of the research to run on the ENCQOR testbed. This goal will be subject to successfully solving the underlying research problems and the availability of appropriate automation tools on the ENCQOR testbed. |
| **Applicant Capabilities** | Applicants should be able to assign any intellectual property discovered during this project to Ciena.<br><br>PhD Candidates<br>• Minimum 2 candidates available to perform the research<br>• Students should have knowledge of networking concepts<br>• Students should be comfortable with network simulations<br>• Students should have strong statistical knowledge and mathematical abilities<br>• Candidates should be able to independently use complex open-source and commercial packages (GNU Linear Programming Kit, Gurobi, Google or-tools) to create network optimization simulations<br><br>**Principal investigator(s)**<br>Optimization expertise<br>• Comprehensive understanding of optimization theory<br>• Experience applying optimization to problems in wireless networks<br>• Experience developing effective heuristics<br>Networking expertise<br>• Research experience in networks<br>• Research experience in modelling and predicting QoS<br>• Research experience with discrete-time network simulations<br>AI expertise<br>• Research experience in machine learning |