Optimization of Optical Aggregation Network for 5G URLLC Service

Federico Tonini, Elisabetta Amato, Carla Raffaelli
DEI, University of Bologna - Bologna, Italy
f.tonini@unibo.it, e.amato@unibo.it,carla.raffaelli@unibo.it

***Abstract*—Optical aggregation network is a key subsystem to support high performance end-to-end services in 5G mobile networks. This paper presents a network design optimization methodology to implement ultra reliable low latency communications (URLLC) in a metro area. Novel optimization algorithms based on Integer Linear Programming (ILP) are defined to solve dedicated (DPP) and shared (SPP) path protection problems, in the presence of single failure with the aim to minimize the number of active nodes, by adopting the functional splitting options defined by 3GPP. The results prove the ability of the proposed algorithm to find optimal solutions which also minimize the number of high capacity node instances and the usage of bandwidth resources, especially when SPP is applied.**

***Index Terms*—5G, Optical networks, Slicing, Functional split, Optimization, ILP**

I. Introduction

5G technology has nowadays entered the experimental phase in many countries [1], [2], [3]. A few classes of services are envisioned to support use cases in different contexts which pose specific requirements on latency and bandwidth in the different segments of the end-to-end transport network [4], [5]. In particular, the Ultra Reliable Low Latency Communication (URLLC) class of service, as defined by the 3GPP initiative [4], is for sure one of the most critical one to be implemented, due to latency constraints down to 1 ms and the extremely high level of service reliability required. Use cases belonging to this service class can be identified in the automotive safety, Unmanned Aerial Vehicle control and smart health emergency interactive applications [5], just to mention a few of them.

The design of 5G network will extensively take advantage of Network Function Virtualization (NFV) coupled with the Software Defined Network (SDN) paradigm, to ensure unprecedent network flexibility and reconfigurability [2]. In order to meet use case specific requirements the slice concept is adopted [5], a slice being the set of physical or virtual resources able to offer a connectivity service meeting those requirements. A slice is embedded into the physical infrastructure whose bandwidth, delay and processing locations need to be identified and optimized.

Functional splits have been proposed for 5G to enhance the network efficiency and flexibility of the optical aggregation network segment [4] [6] giving rise to the fronthaul and midhaul sections in the optical aggregation network. This approach allows dynamic usage of resources based on statistical packet-based multiplexing of information, with consequent enhanced scalability. Differently from the conventional C-RAN approach, anyway included among possible functional splits as option 8, centralization is relaxed by assuming some processing capability distributed in the nodes of the aggregation network, thus achieving more relaxed bandwidth and latency constraints with respect to C-RAN [6]. The main problem is how to split the required functionalities into nodes in relation to the bandwidth available on the links and the processing capability in the nodes themselves, with the aim of achieving the full coverage of the packet-based metro area network. Each node is associated a virtual function hotel that perform the associated functionalities in relation to the chosen split option.

In the specific case of URLLC, extremely tight requirements in terms of latency and reliability need to be met, while optimally associate functions to nodes. To this end, ready-to-use additional backup resources need to be provided in case of link or hotel failure, able to meet the same quality of service of the primary ones. As a consequence, fast connectivity swap can be performed within the slice. This approach is known in literature as resource protection and, according to previous classifications, different protection schemes can be applied for fast slice resiliency [7]. Protection schemes can be either dedicated path (DPP) or shared path (SPP). In DPP, backup resources are dedicated and therefore cannot be shared with any other protection path. Conversely, SPP allows sharing backup resources among protection paths. SPP mechanisms are expected to require less additional resources compared to DPP, but are usually more complicated to operate. At the best of our knowledge the problem of reliable slice embedding with functional split still needs investigation. Option 8 referred to conventional C-RAN was studied in [8]. [9] address the flexible functional placement problem with no specific reference to the latency and reliability constraints relevant to URLLC. Anyway, a methodology to design the optical aggregation network while associating function splitting to nodes with URLLC is expected to greatly help the task of operators in the deployment of 5G network slicing. Some evaluations on functional splitting in optical aggregation networks have been presented based on heuristics in [10] for DPP.

With reference to the above scenario, this paper proposes an optimization procedure based on Integer Linear Programming (ILP) to design resilient latency constrained end-to-end slicing with 5G functional splits for the URLLC service class. The

TABLE I: Link capacity requirements for different 3GPP split options [12].

Layers	Split option	Bandwidth [Gbps]
L1	Opt.8	2.4
L2	Opt.6	0.152
L3	Opt.2	0.151
Core	Opt.1	0.150
Cloud	-	0.150

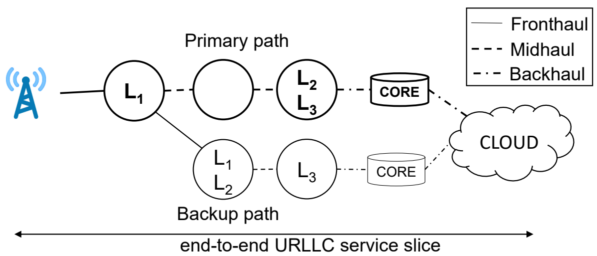


Fig. 1: Example of functional split for URLLC service slice.

target is to minimize the number of nodes where virtual functions are activated, given a potential set of node locations. Dedicated and shared path protection schemes, namely DPP and SPP, are proposed and related optimization procedures to minimize the physical network resources to be associated to URLLC network slice are described and applied to obtain evaluations of their effectiveness.

The paper is organized as follows. In section II the reference network architecture and problem formulation are introduced. Section III describes the optimization methodology for the design of URLLC slice using DPP and SPP with latency constraints to minimize the number of active hotels in the optical aggregation network. Section IV shows results and comparisons. Finally section V draws the conclusions and main achievements.

II. Reference Network Architecture and Problem Formulation

5G baseband functional splits have been investigated recently [6]. The functions composing the layers of the mobile network protocol stack can be split into multiple nodes and performed in sequence, forming a chain of functions. Function chaining has been investigated before, where service end points are known and only end to end bandwidth and latency constraints are applied [11]. However, when providing a service, the requirements of the baseband processing must be satisfied along with the one of the end to end service, usually performed in the cloud. The bandwidth requirement to carry different functions of the protocol stack are shown in Table I for a sample antenna configuration [12]. An example of a possible end to end URLLC service slice is presented in Fig. 1. The reliability required by this class of services implies the allocation of primary and backup path resources for each chain in the slice.

The formulation of the BBU hotel location problem with resiliency is as follows:

- **Given**: a set of nodes and related resources, which are candidates as hotels to host baseband, core and cloud functions, properly connected through a set of links.
- **To find**: a suitable functions placement, such that the number of active nodes (i.e., nodes hosting any function) is reduced to a minimum while reliability against single link or hotel failure is provided.
- **To ensure**: that each antenna is connected to URLLC service through a set of properly ordered functions forming a chain running in active nodes, so that the maximum allowed distance between the antenna and the cloud is not exceeded, the maximum bandwidth available on each link is not exceeded, and the available computational resources in each node are not exceeded.

III. Optimization Methodology for DPP and SPP in a URLLC Slice

This section first introduces the notation and rationale behind the models, then describes in detail the two strategies, namely DPP and SPP, in two separate sub-sections.

Let us consider a network characterized by a set of nodes N, each with capacity ρ_i, interconnected by links, captured by matrix $\gamma_{i,j}$, with bandwidth $\lambda_{i,j}$ and introducing a delay $\tau_{i,j}$, with $i,j \in N$. Each node is considered to be a source $s \in S$, with a set of antennas to be connected, through a chain of functions, to the service in the cloud. Let us consider an ordered set $T = \{t_1, .., t_q, .., t_k\}$ of k transport segments, with cardinality $|T| = k$ equal to the number of functions to be executed. Each transport segment corresponds to a couple of VNFs, one originating the transport flow and one terminating it. For instance, a generic transport segment t_q is originated by one VNF (v_{q-1}) and requires a VNF (v_q) performing the functions required to elaborate its traffic, and originates the traffic towards the next transport segment t_{q+1}. A binary variable $x^n_{i,t_q,s}$ is introduced to model the assignment of sources to nodes performing related functions. When $x^n_{i,t_q,s}$ is equal to 1, v_q (VNF function terminating transport segment t_q) is performed at node i for source s, and requires the activation of node i, modeled by the binary variable z_i, for primary and backup paths $n \in P = \{p, b\}$. Each transport segment t_q produces a primary (p) and backup (b) flow of data for each source s through the links, captured by the binary variable $w^n_{i,j,t_q,s}$, with a certain bitrate β_{t_q} and subject to latency requirements δ_{t_q}. Each VNF related to a transport segment t_q needs computational resources μ_{t_q}. The notation used in the two strategies is reported in Table II.

A. DPP model

Objective function:

$$Minimize \sum_{i \in N} z_i \tag{1}$$

Constraints:

$$\sum_{i \in N} x^n_{i,t_q,s} = 1, \quad \forall t_q \in T, s \in S, n \in P \tag{2}$$

$$\sum_{n\in P}\sum_{t_q\in T}\sum_{s\in S} x^n_{i,t_q,s} \leq M\cdot z_i, \quad \forall i\in N \tag{3}$$

$$\sum_{n\in P}\sum_{t_q\in T}\sum_{s\in S} x^n_{i,t_q,s}\cdot \mu_{t_q} \leq \rho_i, \quad \forall i\in N \tag{4}$$

$$\sum_{n\in P}\sum_{t_q\in T}\sum_{s\in S} w^n_{i,j,t_q,s}\cdot \beta_{t_q} \leq \lambda_{i,j}, \quad \forall i,j\in N \tag{5}$$

$$\sum_{t_c=1}^{t_q}\sum_{i\in N}\sum_{j\in N} w^n_{i,j,t_c,s}\cdot \tau_{i,j} \leq \delta_{t_q}, \quad \forall t_q,t_c\in T, s\in S, n\in P \tag{6}$$

$$w^p_{i,j,t_q,s} + w^b_{i,j,t_m,s} \leq \gamma_{i,j}, \quad \forall t_q,t_m\in T, s\in S, i,j\in N \tag{7}$$

$$\sum_{j\in N} w^n_{i,j,t_q,s} \leq 1, \quad \forall i\in N, t_q\in T, s\in S, n\in P \tag{8}$$

$$w^n_{i,i,t_q,s} \leq x^n_{i,t_q,s}, \quad \forall i\in N, t_q\in T, s\in S, n\in P \tag{9}$$

$$x^p_{i,t_m,s} + x^b_{i,t_q,s} \leq 1, \quad \forall i\in N, t_q,t_m\in T, s\in S \tag{10}$$

If $t_q = t_1$:

$$\sum_{j\in N} w^n_{s,j,t_q,s} = 1, \quad \forall s\in S, n\in P \tag{11}$$

$$\sum_{j\in N} w^n_{j,i,t_q,s} - \sum_{j\in N} w^n_{i,j,t_q,s} = x^n_{i,t_q,s}, \quad \forall i\in N, i\neq s, t_q\in T, s\in S, n\in P \tag{12}$$

If $t_q \neq t_1$:

$$\sum_{j\in N} w^n_{i,j,t_q,s} \geq x^n_{i,t_{q-1},s}, \quad \forall i\in N, s\in S, n\in P \tag{13}$$

$$\sum_{j\in N} w^n_{j,i,t_q,s} - \sum_{j\in N} w^n_{i,j,t_q,s} + x^n_{i,t_{q-1},s} = x^n_{i,t_q,s}, \quad \forall i\in N, t_q\in T, s\in S, n\in P \tag{14}$$

Constraint (2) ensures that only one node is active for each VNF and source along the whole path. Constraint (3) selects the active nodes (i.e., nodes that host at least one VNF). Constraint (4) ensures that computational resources required at node i to perform all VNFs from all sources and all the path are not exceeded. Constraint (5) guarantees that the bandwidth required over each link does not exceed the maximum link capacity for that link. Constraint (6) limits the delay of each path transport segment. Constraint (7) allows routing only over link of the physical topology. Also ensures that the links used for different path are different.

TABLE II: List of parameters of the ILP and corresponding definitions.

Parameter	Definition
T	set of transport segments.
S	set of source nodes.
N	set of network nodes, candidates to host virtual functions.
P	set of paths. P = {p,b} (p for primary, b for backup).
$\alpha_z, \alpha_c, \alpha_f$	tuning parameters for SPP objective function.
β_{t_q}	bandwidth requirement for transport segment $t_q \in T$.
δ_{t_q}	latency requirement for transport segment $t_q \in T$.
$\gamma_{i,j}$	1 if exists a link between nodes $i\in N$ and $j\in N$ in the physical network; 0 otherwise.
μ_{t_q}	capacity required to execute virtual function terminating transport segment $t_q\in T$.
ρ_i	computational resources available at node $i\in N$.
$\lambda_{i,j}$	available bandwidth over the link connecting nodes $i\in N$ and $j\in N$.
$\tau_{i,j}$	delay introduced by the link connecting nodes $i\in N$ and $j\in N$.
M	a large number.
$x^n_{i,t_q,s}$	1 if node $i\in N$ is performing VNF terminating transport segment $t_q\in T$ for source $s\in S$ for path $n\in P$; 0 otherwise.
$w^n_{i,j,t_q,s}$	1 if link connecting nodes $i\in N$ and $j\in N$ is carrying transport traffic $t_q\in T$ originated at source $s\in S$ for path $n\in P$; 0 otherwise.
z_i	1 if node $i\in N$ is selected to host at least one VNF; 0 otherwise.
c_i	computational capacity required at node $i\in N$ for backup purposes.
$f_{i,j}$	required bandwidth over the link $(i,j)\in N$ for backup purposes.
$y_{i,j,t_q,s}$	1 if the nodes $i\in N$ and $j\in N$ are performing VNF terminating transport segment $t_q\in T$ as primary and backup, respectively, for the source $s\in S$; 0 otherwise.
$l_{i,j,k,m,t_q,s}$	1 if the links $(i,j)\in N$ and $(k,m)\in N$ are used to transport data of segment t_q for primary and backup paths, respectively, for the source $s\in S$; 0 otherwise.
$d_{i,k,m,t_q,s}$	1 if the link $(k,m)\in N$ is used to transport data of segment t_q for backup and the primary path passes through node $i\in N$ for the source $s\in S$; 0 otherwise.

Function chaining is modeled as follows. Constraint (8) limits the sum of outgoing paths in each node, for each source and transport link. Constraint (9) forbids unnecessary loops.

Constraint (10) ensures that the nodes used for primary and backup are different, for all the transport segment.

For the first transport segment (t_1), constraint (11) ensures that there is one outgoing flow for each source s while constraint (12) represents the flow conservation towards the ending VNF for transport segment t_1. For the subsequent transport segments ($\{t_2, .., t_k\}$), constraint (13) ensures that there is a transport flow starting from the node (i) performing the previous transport function ($x_{i,t_{q-1},s}$) for each source. Constraint (14) represents the flow conservation of each transport segment t_q.

B. SPP model

In the SPP model, four additional variables have been introduced: c_i and $f_{i,j}$ which allow the reduction of computational resources and bandwidth reserved for backup, and $y_{i,j,t_q,s}$ and $l_{i,j,k,m,t_q,s}$ to find the pairs of nodes and links of the primary and backup. A new objective function is also introduced.

Objective function:

$$Minimize\ \alpha_z \cdot \sum_{i \in N} z_i + \alpha_c \cdot \sum_{i \in N} c_i + \alpha_f \cdot \sum_{i \in N} \sum_{j \in N} f_{i,j} \quad (15)$$

The constraints (4) and (5) have been replaced by: *Additional constraints:*

$$y_{i,j,t_q,s} \geq x^p_{i,t_q,s} + x^b_{j,t_q,s} - 1 \quad \forall i,j \in N, s \in S, t_q \in T \quad (16)$$

$$c_j \geq \sum_{t_q \in T} \sum_{s \in S} y_{i,j,t_q,s} \cdot \mu_{t_q} \quad \forall i,j \in N \quad (17)$$

$$\sum_{t_q \in T} \sum_{s \in S} x^p_{i,t_q,s} \cdot \mu_{t_q} + c_i \leq \rho_i, \quad \forall i \in N \quad (18)$$

$$l_{i,j,k,m,t_q,s} \geq w^p_{i,j,t_q,s} + w^b_{k,m,t_q,s} - 1 \quad \forall i,j,k,m \in N, s \in S, t_q \in T \quad (19)$$

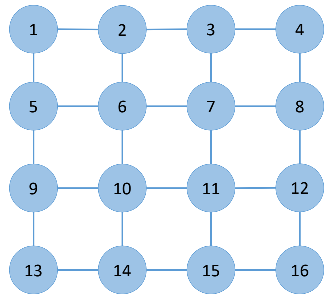
$$f_{k,m} \geq \sum_{t_q \in T} \sum_{s \in S} l_{i,j,k,m,t_q,s} \cdot \beta_{t_q} \quad \forall i,j,k,m \in N \quad (20)$$

$$d_{i,k,m,t_q,s} \geq w^b_{k,m,t_q,s} + \frac{\sum_{t_q \in T} x^p_{i,t_q,s}}{M} - 1, \quad \forall i,k,m \in N, s \in S, t_q \in T \quad (21)$$

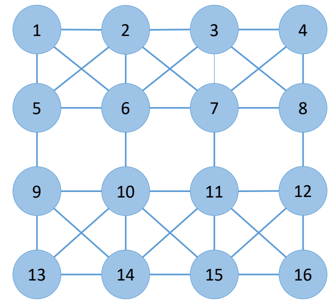
$$f_{k,m} \geq \sum_{t_q \in T} \sum_{s \in S} d_{i,k,m,t_q,s} \cdot \beta_{t_q}, \quad \forall i,k,m \in N \quad (22)$$

$$\sum_{t_q \in T} \sum_{s \in S} w^p_{i,j,t_q,s} \cdot \beta_{t_q} + f_{i,j} \leq \lambda_{i,j}, \quad \forall i,j \in N \quad (23)$$

Constraint (16) finds the primary (i) and backup (j) nodes performing the different VNFs for each source. Constraint (17) ensures that the capacity reserved for the backup node j is greater than or equal to the capacity required to perform primary functions at node i, to ensure reliability against single primary hotel failure. Constraint (18) ensures that computational resources required at node i to perform all VNFs from all sources and all the paths are not exceeded. Constraint (19) finds the primary link (i,j) and the backup link (k,m) carrying traffic of each transport for each source. Constraint (20) ensures that the bandwidth reserved for the backup path is greater than or equal to the bandwidth required in case of a single primary link failure. Constraint (21) finds the sources affected by a BBU hotel failure in i that are sharing the backup link (k,m) while constraint (22) counts the bandwidth required over link k,m in the case of hotel i failure. Constraint (23) guarantees that the bandwidth required over each link does not exceed the maximum link capacity for that link.



(a) Network A.



(b) Network B.

Fig. 2: Reference 16 node networks A and B with different connectivity.

IV. Numerical results and comparisons

In this section, the results obtained by running the two algorithms using the CPLEX commercial tool [13] are reported. The DPP strategy is firstly evaluated referring to the 16 node networks represented in Fig. 2. Each network node is connected to 10 antennas, collecting traffic from the radio section. The available bandwidth on each link is set to 40 Gbps (in each direction). In addition, each node is equipped with processing units (PUs) according to traffic generated at each layer of the functional splitting as shown in Table I. In particular, 0.5, 0.3, 0.2, 0.1, 0.1 PUs are assumed as requirements for L1, L2, L3, core and cloud virtual functions, respectively [9], [14], [15]. The length of each link or, equivalently, each hop is assumed to be 1 km, which results in a delay $\tau = 5 \mu s$. It should be noted that, given the limited size of the scenario, the latency constraints of each transport link are always satisfied. However, to satisfy the tight service requirements imposed by URLLC applications, all the nodes are allowed to host edge core and cloud functions. The SPP approach is evaluated then, assuming a 6 node network, not to incur in out of memory exception due to the higher complexity of the SPP algorithm.

A. DPP evaluation

The two networks A and B used to evaluate the DPP algorithm are presented in Fig. 2. The two networks differ for the connectivity represented by a different number of links.

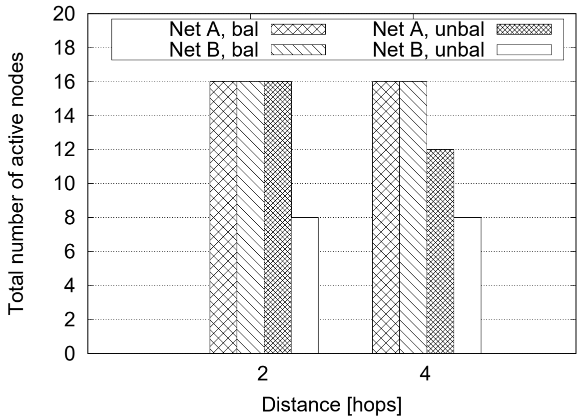


Fig. 3: Active nodes for the networks A and B in the balanced and unbalanced cases with 2 and 4 hops as distance constraint.

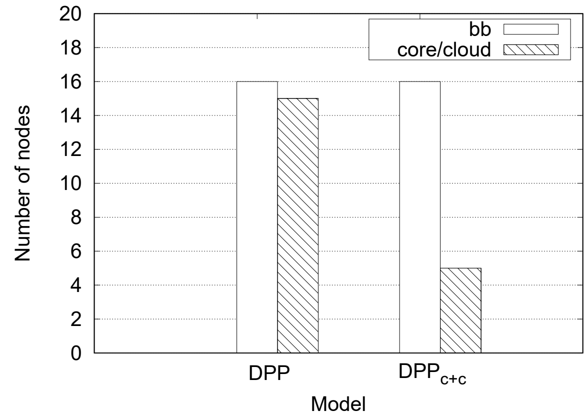


Fig. 4: Number of nodes performing at least one baseband (bb) processing (either L1, L2 or L3), core and cloud functions for the balanced case with 2 hops constraints for DPP and modified DPP$_{c+c}$ models in network A.

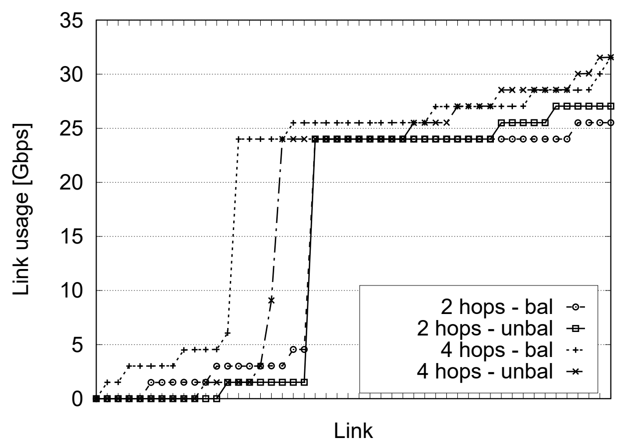


Fig. 5: Link usage of DPP in the network A, balanced and unbalanced, under different hop constraints. Links are sorted from the lowest to the highest usage one.

Two cases have been considered in the following. The case in which all the nodes candidate to host an hotel are equal to 25 PUs is referred to as balanced (bal). The unbalanced (unbal) case, instead, has nodes 6,7,10,11 with infinite capacity in terms of PUs, thus emulating centralized data centers, while all the other nodes are equipped with 10 PUs. The number of active nodes (i.e., nodes hosting baseband, core or cloud functions) obtained with the DPP model in the networks A and B in the balanced and unbalanced case under different hop constraints is reported in Fig. 3. The balanced case always requires the activation of all the nodes, as a consequence of the limitation of node resources, regardless the number of hops and network connectivity. Conversely, the unbalanced case shows a reduction in the number of active nodes. The effects of an increased network connectivity are evident, with only 8 active nodes required for both 2 and 4 hops. The network A allows a reduction of 4 nodes when moving from 2 to 4 hops, thanks to high capacity nodes that perform multiple functions in few nodes, while in network B no additional node reduction is possible due to the limited resources over links connecting high capacity nodes.

To facilitate the deployment of edge core and cloud functions, minimization of active nodes performing those functions can be added to the objective function of the DPP model. A new term is added to (1) with a lower priority, so that the primal objective remains the same (i.e., the minimization of the total active nodes). This case is referred to as DPP$_{c+c}$. Figure 4 shows the number of nodes performing at least one baseband processing (either L1, L2 or L3) and core/cloud functions for the balanced case with 2 hops constraints in the traditional DPP and modified DPP$_{c+c}$ formulation. While the number of nodes performing baseband processing is the same for the two cases, the nodes performing core and cloud functions in DPP$_{c+c}$ is considerably lower than the one of DPP, thus simplifying the deployment of these functions from a network operator and/or cloud provider point of view.

Figure 5 depicts the link usage of DPP in the network A, under 2 and 4 hop constraints, for both balanced and unbalanced cases. In the figure the links are sorted in increasing order of usage for each curve. Depending on the specific curve, links show different usage, which indicates potential statistical multiplexing gain when multiple slices are embedded on the same network. Some links exhibit a very low usage or even no usage, especially with 2 hop constraint. Many links needs 24 Gbps or slightly higher due to the capacity required for option 8 with 10 antennas (see Table I).

B. SPP evaluation

A 6 node network is considered to compare DPP and SPP (Figs. 6 and 7). All nodes are connected to 10 antennas. In the unbalanced case, nodes 2 and 5 have unlimited resources, while all the other node capabilities are limited to 10 PUs. The tuning parameters of the objective function are set as $\alpha_z >> \alpha_c = \alpha_f$ to prioritize the minimization of the overall active nodes.

Table III reports the number of active nodes, capacity and node savings for the 6 node networks in the balanced and

TABLE III: Active nodes, capacity and node savings for 6 node networks in the balanced and unbalanced cases under 2 and 3 hop constraints.

Network	Active nodes		Saved	
	DPP	SPP	Capacity	Nodes
2 hops - bal	6	4	66.6%	33.3%
3 hops - bal	6	4	66.6%	33.3%
2 hops - unbal	4	4	23.6%	0%
3 hops - unbal	4	4	27.8%	0%

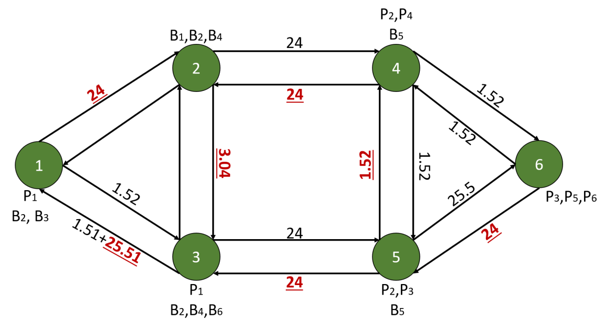


Fig. 6: Outcome of the DPP model in the balanced case for 6 node network and 2 hops as distance constraint. The bandwidth values are in Gbps, red and underlined for the backup path. Dark green color for active nodes.

unbalanced cases under 2 and 3 hop constraints. The SPP is capable of reducing the number of active nodes in the balanced case by 33.3%. In addition, the SPP approach allows to share backup node resources among antennas assigned to different primary paths, leading up to 66.6% and 27.8% node capacity savings in the balanced and unbalanced cases, respectively.

Figures 6 and 7 depict the outcome of DPP and SPP models, respectively, in the balanced case for 6 node network and 2 hops as distance constraint. The figures report the active nodes, that are the nodes performing primary and/or backup functions, eventually split in multiple nodes. The figures also report the link usage for both primary and backup (in underlined, red). In the case of SPP, the number of active nodes is lower than in the DPP case, thanks to the sharing of the backup paths. For instance, in the SPP case nodes 2, and 3 are able to reach node 5, for backup purposes, by sharing link 3-5. In the DPP case instead, they cannot reach node 5 due to the dedicated resource allocation and limited bandwidth over the links, thus requiring to activate additional nodes.

V. Conclusions

Optimization of functional split for URLLC service has been presented for DPP and SPP based on Integer Linear Programming. Two different sets of constraints have been defined with the objective to minimize the number of active nodes in the optical aggregation networks. The effectiveness of the algorithms has been shown also in terms of active nodes and bandwidth usage which is sensibly reduced with repsect to the conventional centralized approach and allows

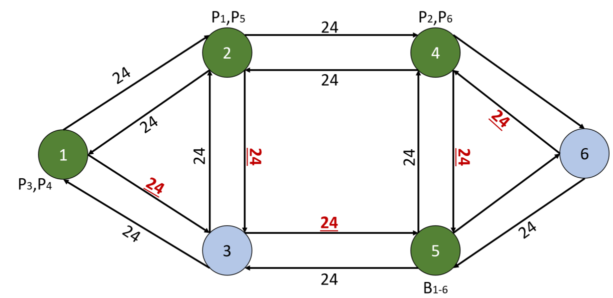


Fig. 7: Outcome of the SPP model in the balanced case for 6 node network and 2 hops as distance constraint. The bandwidth values are in Gbps, red and underlined for the backup path. Dark green color for active nodes.

statistical multiplexing gain for potential allocation of multiple slices. The further saving related to SPP has been shown in comparison with DPP. The SPP algorithm has some scalability limitations that, at this moment has reached optimization for a more limited size network with respect to DPP. In any case the evaluations result suitable for most metro contexts.

References

[1] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, 2017.

[2] A. Sgambelluri, M. Capitani, S. Fichera, K. Kondepu, A. Giorgetti, F. Giannone, B. Martini, F. Ubaldi, P. Iovanna, G. Landi, and L. Valcarenghi, "Experimental demonstration of a 5G network slice deployment exploiting edge or cloud data-centers," in *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, March 2019, pp. 1–3.

[3] A. Farhadi-Beldachi, E. Huques-Satas, A. Tzanakaki, Y. Yan, R. Nejabati, and D. Simeonidou, "Experimental demonstration of 5G fronthaul and backhaul convergence based on FPGA-based active optical transport," in *2018 European Conference on Optical Communication (ECOC)*, Sep. 2018, pp. 1–3.

[4] "Third Generation Partnership Project (3GPP) TR 38.801, v.14."

[5] "GSM Association - Network slicing use case requirements," 2018.

[6] L. M. P. Larsen, A. Checko, and H. L. Christiansen, "A survey of the functional splits proposed for 5G mobile crosshaul networks," *IEEE Communications Surveys Tutorials*, vol. 21, no. 1, pp. 146–172, 2019.

[7] D. M. M. Pioro, *Routing, Flow, and Capacity Design in Communication and Computer Networks*. Morgan Kaufmann, 2004.

[8] B. M. Khorsandi, F. Tonini, and C. Raffaelli, "Design methodologies and algorithms for survivable C-RAN," in *2018 International Conference on Optical Network Design and Modeling (ONDM)*, May 2018.

[9] D. Harutyunyan and R. Riggio, "Flex5G: Flexible functional split in 5G Networks," *IEEE Transactions on Network and Service Management*, vol. 15, no. 3, pp. 961–975, Sep. 2018.

[10] B. M. Khorsandi, F. Tonini, E. Amato, and C. Raffaelli, "Dedicated path protection for reliable network slice embedding based on functional splitting," in *2019 International Conference of Transparent Optical Network ICTON*, 2019.

[11] A. Hmaity, M. Savi, F. Musumeci, M. Tornatore, and A. Pattavina, "Protection strategies for virtual network functions placement and service chains provisioning," *Networks*, vol. 70, no. 4, pp. 373–387. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/net.21782

[12] "Small Cell Forum - Functional splits and use cases," 2016.

[13] *IBM ILOG CPLEX Optimization Studio V12.6.3*.

[14] B. M. Khorsandi, D. Colle, W. Tavarnier, and C. Raffaelli, "Adaptive function chaining for efficient design of 5G xhaul," in *2019 International Conference on Optical Network Design and Modeling ONDM*, 2019.

[15] H. Chang, B. Qiu, C. Chiu, J. Chen, F. J. Lin, D. de la Bastida, and B. P. Lin, "Performance evaluation of Open5GCore over KVM and Docker by using Open5GMTC," in *NOMS 2018 - 2018 IEEE/IFIP Network Operations and Management Symposium*, April 2018, pp. 1–6.