Load-aware Optimal Association for 5G Mobile Users in HetNets with S-BVT Backhaul

G. Otero, D. Larrabeiti, J. A. Hernández
Universidad Carlos III de Madrid
Madrid, Spain
{goterop, dlarra, jahgutie}@it.uc3m.es

J. P. Fernández-Palacios
Telefonica Global CTO
Madrid, Spain
jpfpg@tid.es

R. Martínez, M. Svaluto Moreolo
CTTC
Castelldefells, Spain
{ricardo.martinez, michela.svaluto}@cttc.es

Abstract—New paradigms are being proposed to cope with the increasing traffic demand in the access networks. One alternative is to leverage cost-efficient VCSEL-based Sliceable Bandwidth Variable Transceivers (S-BVTs) near the edge. In this work, we investigate the adoption of optical S-BVTs for the transport of cellular traffic. A dynamic and decentralized orchestration algorithm is proposed to achieve an optimal user association and resource allocation of the access network's resources while considering the load-balancing and availability of the underlying optical infrastructure. Two association and resource allocation techniques are explored in a two-tier heterogeneous network featuring different radio access technologies. The results show that the decentralized solution outperforms the coverage maps-based association and exhibits load balancing capabilities without the need of a centralized coordination, which can be leveraged by the S-BVT's adaptive capabilities.

Index Terms—S-BVT, Decentralized, Optimization, Access

I. INTRODUCTION

A recent focus of the research community are the technologies capable of supporting ultra-low latency services for 5G services. In order to address the flexibility and cost requirements of metro networks while providing high capacity, Sliceable Bandwidth Variable Transceivers (S-BVTs) have been proposed as a cost-effective transmission technology to interconnect MAN nodes with maximum elasticity and simplicity [1]. This is attained exploiting the intrinsic capability of generating multiple optical transmission flows. Thus, S-BVTs can be used to serve a high number of users and to leverage the available fiber bandwidth in an efficient way using the optical flexi-grid transport solutions. Sliceable variable allocation enables not only to move the traffic from a failing part of the network to another but also to quickly react to changes in the traffic demand, optimizing the resource utilization in real time. In this context, research initiatives such as PASSION [2] focus on the development low-cost photonic technologies based on direct-modulated large-bandwidth VCSELs, with an integrated modular design, supporting terabit capacities [3]. Their slicing capability allows hosting simultaneous WDM connections from other S-BVTs with a granularity of 50 Gb/s or even 25 Gb/s. Programmed by an control plane they can be dynamically configured to send/receive traffic from several nodes, simultaneously, with a single transceiver [4]. This is of particular interest in cellular networks, which exhibit a
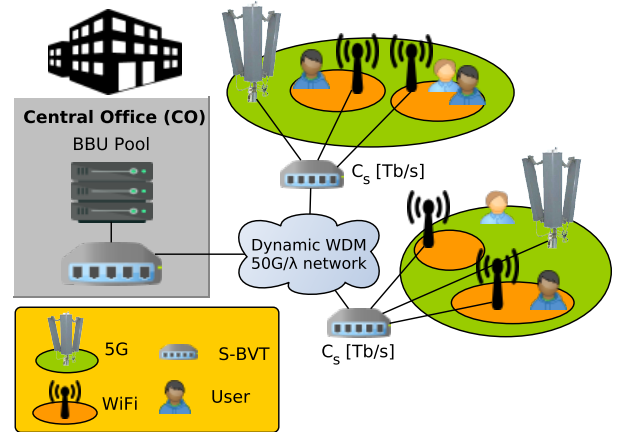
Fig. 1. Two-tier heterogeneous network deployment supported by S-BVTs.

very dynamic nature with high traffic variability. The provided capacity at any particular time can be released and further re-allocated to satisfy new demands in busier cells. This flexible capability can also explored for the transport of variable traffic as done in this work. The remainder of the paper explores the proposed orchestration scheme for a two-tier heterogeneous network featuring different radio access technologies and a S-BVT-based backhaul. Then we show the gain that can be obtained in terms of average user throughput by using the decentralized approach instead of the traditional one.

II. SCENARIO AND OPTIMIZATION

Figure 1 depicts the generic architecture of the proposed heterogeneous cloud radio access network (H-CRAN). A two-tier HetNet is considered with macrocells providing 5G connectivity and small cells featuring low-power and low-cost offloading with mmwave or Wifi technologies. We assume that the CRAN technology is used, meaning that the users' signals are not completely processed at the antenna site –remote radio heads (RRH)–, but they are transported via the fronthaul network to the processing units (BBUs). The optical backhaul supporting this transport makes use of the high-capacity provided by the S-BVTs, each aggregating the fronthaul traffic of a given set of macrocells and femtocells. Traditionally, mobile users are associated to the same base station in both uplink and downlink directions. Downlink/Uplink Decoupling (DUDe) is an alternative idea [5], [6] advocating independent decisions.

Assume a two-tier HetNet where the base stations (BS) in each tier are located following a two independent homogeneous Poisson Point Processes. Clustering the users according to the premise of maximizing the received power and minimizing the transmission distances, leads to the so-called Voronoi coverage maps. However, these blind decisions ignore the current state of the network, e.g., the base stations or S-BVTs loads. To solve this, we propose to model the association task as a Network Utility Maximization problem. To compute both the optimal association and allocation of resources from base stations to users and avoid the NP-hard combinatorial problem, we address this problem assuming a *Multi-Station Association* policy. Assume that the maximum achievable rate between user $u \in \mathcal{U}$ and base station b is given by the ergodic channel capacity:

$$r_{ub}^- = \text{BW}^{\text{DL}} \log_2(1 + \text{SINR}^-(u)) \quad (1)$$
$$r_{ub}^+ = \text{BW}^{\text{UL}} \log_2(1 + \text{SINR}^+(u)), \quad (2)$$

where BW^{DL} and BW^{UL} are the downlink and uplink channel bandwidths [7], SINR is the Signal-to-interference-plus-noise ratio, and the (+) and (-) signs represent uplink and downlink directions. Let $\mathcal{B}_{\text{DL}}$ and $\mathcal{B}_{\text{UL}}$ be a set of BSs in downlink and uplink, and y_{ub} the fraction of resources that BS b allocates to user u. This represents a portion of the total bandwidth. Normalizing to the total amount of resources of a BS we have $\sum_u y_{ub}^- \leq 1$ and $\sum_u y_{ub}^+ \leq 1$. Note that these variables indicate whether a user is associated to a given base station. The aggregated rate for user u in downlink and uplink are,

$$R_u^-(\mathbf{y}^-) = \sum_{b \in \mathcal{B}_{\text{DL}}} r_{ub}^- y_{ub}^- \quad (3)$$
$$R_u^+(\mathbf{y}^+) = \sum_{b \in \mathcal{B}_{\text{UL}}} r_{ub}^+ y_{ub}^+. \quad (4)$$

The ultimate goal is to maximize a utility function $U_\alpha(\cdot)$ of the family $\frac{x^{1-\alpha}}{1-\alpha}, \alpha \geq 0, \alpha \neq 1$, that is continuously differentiable, monotonically increasing and strictly concave. The special cases when $\alpha = 0$, $\alpha = 1$, and $\alpha \to \infty$ lead to linear, logarithmic and max-min utilities, respectively. We focus the study on the uplink, since it usually poses more stringent delay requirements than the downlink for the fornthaul. Let $\mathcal{S}_{\text{UL}}$ be the set of S-BVTs that aggregates the traffic of base stations in the uplink. Taking into account the restrictions on the base station and S-BVT capacities, we can formulate the optimization problem for the uplink as follows:

$$f_\alpha^{\text{Uplink}} \equiv \max_{\mathbf{y}^+ \in \mathcal{Y}_{\text{UL}}} \sum_u \left[U_\alpha(\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+) \right] \quad (5)$$

such that

$$\sum_{u \in \mathcal{U}(s)} \sum_{b \in \mathcal{B}(s)} r_{ub}^+ \cdot y_{ub}^+ \cdot c \leq C_s, \forall s \in \mathcal{S}_{\text{UL}}, \quad (6a)$$

$$\sum_{u \in \mathcal{U}} y_{ub}^+ \leq 1, \forall b \in \mathcal{B}_{\text{UL}}, \quad (6b)$$

where $y_{ub}^+ \geq 0$, C_s is the capacity of the S-BVT and c models the overhead of the fronthaul traffic added to each user's data

rate that is going to traverse a S-BVT. Note that the objective function is concave and the constraints are affine functions of $\mathbf{y}_u^+$. Using the Lagrange duality theory, we may relax the constraints to form the Lagrangian of the uplink problem as

$$L(\mathbf{y}^+, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = \sum_u \left[U_\alpha(\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+) \right] - \sum_{b_{\text{UL}}} \alpha_b (\sum_u y_{ub}^+ - 1) -$$
$$- \sum_s \lambda_s \left(\sum_{u \in \mathcal{U}(s)} \sum_{b \in \mathcal{B}(s)} r_{ub}^+ \cdot y_{ub}^+ \cdot c - C_s \right) =$$
$$= \sum_u U_\alpha(\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+) - \sum_{b_{\text{UL}}} \alpha_b \sum_u y_{ub}^+ + \sum_{b_{\text{UL}}} \alpha_b -$$
$$- \sum_s \lambda_s \sum_{u \in \mathcal{U}(s)} \sum_{b \in \mathcal{B}(s)} r_{ub}^+ \cdot y_{ub}^+ \cdot c + \sum_s \lambda_s C_s =$$
$$= \sum_u \left[U_\alpha(\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+) - \sum_{b_{\text{UL}}} y_{ub}^+ \alpha_b - \right.$$
$$\left. - \sum_{s \in \mathcal{S}(u)} \lambda_s (\sum_b r_{ub}^+ \cdot y_{ub}^+ \cdot c) \right] + \sum_{b_{\text{UL}}} \alpha_b + \sum_s \lambda_s C_s. \quad (7)$$

Now, the optimization separates into two different levels. On the one hand, each user must solve the u-th Lagrangian $L_u(y_{ub}^+, \boldsymbol{\alpha}, \boldsymbol{\lambda})$ for the given vector of multipliers $\boldsymbol{\alpha}$ and $\boldsymbol{\lambda}$.

$$\arg\max_{y_{ub}^+ \geq 0} \left[U_\alpha(\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+) - \sum_{b_{\text{UL}}} y_{ub}^+ \alpha_b - \right.$$
$$\left. - \sum_{s \in \mathcal{S}(u)} \lambda_s (\sum_b r_{ub}^+ \cdot y_{ub}^+ \cdot c) \right] \quad \forall u. \quad (8)$$

On the other hand, the master dual problem is

$$\min_{\boldsymbol{\alpha}, \boldsymbol{\lambda}} \quad g(\boldsymbol{\alpha}, \boldsymbol{\lambda}) = \sum_u g_u(\boldsymbol{\alpha}, \boldsymbol{\lambda}) + \boldsymbol{\alpha}^T \mathbf{1} + \boldsymbol{\lambda}^T C_s \quad (9)$$
$$\text{subject to } \boldsymbol{\alpha}, \boldsymbol{\lambda} \geq 0,$$

where $g_u(\boldsymbol{\alpha}, \boldsymbol{\lambda}) = L_u(y_{ub}^{+ *}, \boldsymbol{\alpha}, \boldsymbol{\lambda})$, that is, the u-th Lagrangian evaluated at the optimal point. Because the problem is differentiable in its domain, we may use the gradient projection method to update the multipliers of each BS and S-BVT as

$$\alpha_b(t+1) = \left[\alpha_b(t) - \theta \left(1 - \sum_{u \in \mathcal{U}(b)} y_{ub}^+ \right) \right], \forall b \in \mathcal{B} \quad (10a)$$

$$\lambda_s(t+1) = \left[\lambda_s(t) - \theta \left(C_s - \sum_{u \in \mathcal{U}(s)} r_{ub}^+ \cdot y_{ub}^+ \cdot c \right) \right], \forall s \in \mathcal{S}, \quad (10b)$$

where θ is the step size and t is the iteration index. To leverage the adaptability of S-BVTs, C_s can be dynamically changed, i.e., $C_s(t)$, without affecting the orchestration. The proposed algorithm will converge to the new optimal solution seamlessly. Finally, the Karush-Kuhn-Tucker conditions are

$$\begin{cases} \frac{\partial L_u}{\partial y_{ub}^+} = (\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+)^{-\alpha} \cdot r_{ub}^+ - \alpha_b^* - r_{ub}^+ \cdot c \cdot \lambda_s^* = 0 \to \\ \to \frac{1}{(\sum_{b_{\text{UL}}} r_{ub}^+ y_{ub}^+)^\alpha} \leq \frac{\alpha_b^* + r_{ub}^+ \cdot c \cdot \lambda_s^*}{r_{ub}^+}. \quad , \forall b \in \mathcal{B}_{\text{UL}}. \\ \alpha_b, \lambda_s \geq 0 \end{cases}$$
$$(11)$$

Note that the LHS of the summation in (11) is common for all BS. Thus, choosing b and s such that minimize $\frac{\alpha_b^* + r_{ub}^+ \cdot c \cdot \lambda_s^*}{r_{ub}^+}$, satisfies all the KKT conditions. It is worth mentioning that base stations can broadcast the value of their multipliers together with the value of their associated S-BVTs directly through the air. These values can be seen as a price of using a certain resource that will make the system converge to the optimal point by means of load balancing. Once that a combination of BS and S-BVT is chosen, the u-th Lagrangian can be simplified and solved locally by each user as:

$$y_{ub}^{+\,*} = \left(\frac{r_{ub}^{+\,1-\alpha}}{\alpha_b^* + r_{ub}^+ \cdot c \cdot \lambda_s^*} \right)^{1/\alpha}. \tag{12}$$

III. SIMULATION AND RESULTS

In the traditional association and resource allocation solution, the association of users to base stations follows the Voronoi coverage maps and the resource allocation is uniform among the users of a BS. For the optimized proposal, the decisions are taken based on the rules explained in the previous section.

The uplink coverage map of Figure 2a shows a deployment with 3 macrocells and 9 femtocells (P1 to P12). The simulator was validated in [6]. The transmission powers are set to 46 dBm and 20 dBm for macro and femtocells, the channel bandwidths are 1 Gbps for WiFi and 400 MHz for 5G cells, and the path-loss exponent is 4. The colors of the Voronoi cells explain how each one of the four S-BVTs aggregate the traffic of base stations (P1, P6, and P11), (P2, P5, and P12), and so on. Users follow random walk mobility model and the network performance in terms of the average user's uplink throughput is computed over 100 repetitions for each user load. Figure 2b shows the results for both the solutions as we increase the number of users. The decentralized solution outperforms the one based on coverage maps, achieving up to 3 times more user throughput in the high load scenario. Note that the higher the load of users, the bigger the gap between both solutions.

IV. CONCLUSIONS

In this study we considered the possible applications of S-BVTs in the access network. Their adaptability to fine match the data rate that cell's are producing in real time makes them good candidates to cope with the high variability of fronthaul traffic with load-dependent split options.

This paper focuses on improving the cell association and resource allocation mechanism aiming at improving the user's throughput and estimating the maximum traffic demand for an S-BVT at a given time and balancing the traffic over transceivers. We presented a decentralized algorithm that is able to maximize the network utility and exhibits load balancing properties. Most importantly, it outperforms traditional coverage-based solutions in terms of system and individual user throughput.
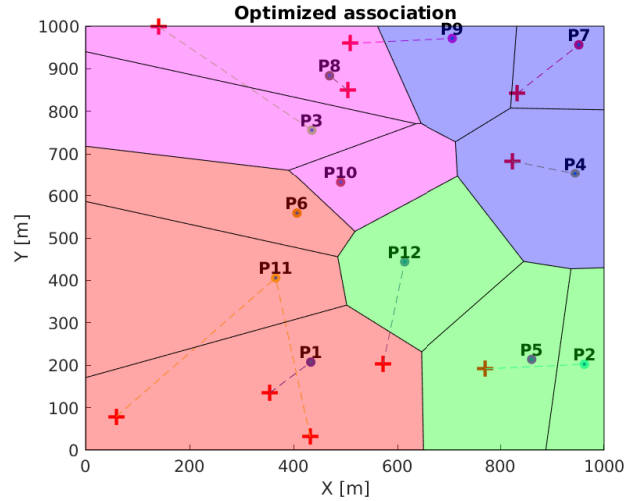
(a) Two-tier heterogeneous network deployment.
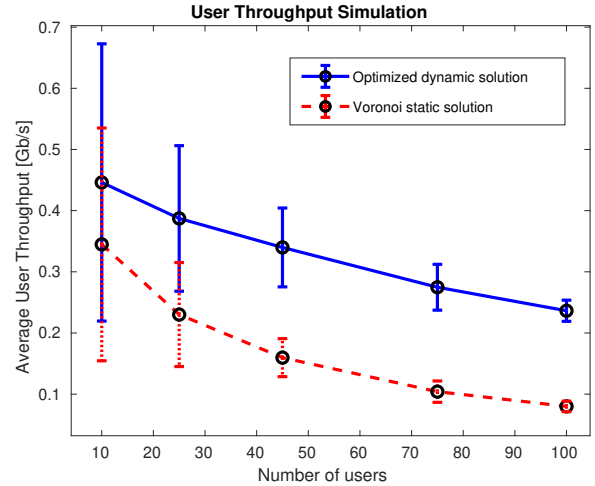
(b) Mean user's uplink throughput for both solutions (95% confidence interval).

Fig. 2. Simulated scenario and throughput results.

REFERENCES

[1] M. Svaluto Moreolo et al., "Multi-Tb/s Photonic Transceivers for Metro Optical Network Connectivity Evolution," Proc. SPIE, Metro and Data Center Optical Networks and Short-Reach Links IV, 117120A, March 2021.

[2] EU PASSION PROJECT. URL: http://www.passion-project.eu.

[3] M. Svaluto Moreolo, et al., "Programmable VCSEL-based photonic system architecture for future agile Tb/s metro networks," J. Optical Communications and Networking, OFC 2020 Special Issue, Vol. 13, No. 2, pp. A187-A199, Feb. 2021.

[4] R. Martínez et al., "Proof-of-Concept Validation of SDN-Controlled VCSEL-Based S-BVTs in Flexi-Grid Optical Metro Networks," OFC 2019, San Diego, CA, USA, pp. 1-3.

[5] H. Elshaer et al., "Downlink and uplink decoupling: A disruptive architectural design for 5G networks," in *Proc. IEEE GLOBECOM*, Austin, TX, 2014, pp. 1798-1803.

[6] G. Pérez et al., "Decentralized Coordination of Converged Tactile Internet and MEC Services in H-CRAN Fiber Wireless Networks," J. Lightwave Technol. 38, 4935-4947 (2020).

[7] A. S. Cacciapuoti, "Mobility-Aware User Association for 5G mmWave Networks," in *IEEE Access*, vol. 5, pp. 21497-21507, 2017.