Appendix

1 Additional numerical results

1.1 KPI comparison of VERA and DQN

In continuation to the comparison of VERA with respect to DQN in terms of reward, we further present their KPI evolution for vRAN and livecast services in Fig. 1. As discussed, since DQN converges poorly due to the large number of actions, the observed KPI values in DQN for both the services fail to meet the target KPIs in Fig. 1. On the contrary, KPIs in VERA not only meet their respective target values, they also stick closer to the target to optimize the resource consumption.

Appendix A
Differential returns

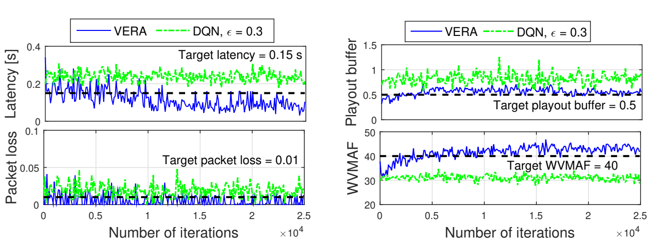
We adopt the definition of cumulative reward for service k, observed during decision window h, as the differential return $G_k^{(h)}$ defined in [?]:

$$G_k^{(h)} = \overline{r}(\mathbf{y}^{(h)}, a_k^{(h-1)} - \widehat{r}(k,\pi_k) + \overline{r}(\mathbf{y}^{(h+1)}, a_k^{(h)} \\ -\widehat{r}(k,\pi_k) + \overline{r}(\mathbf{y}^{(h+2)}, a^{(k,h+1)}) - \widehat{r}(k,\pi_k) + \cdots \quad (1)$$

where $\pi_k : \mathcal{X} \to \mathcal{A}$, denotes the greedy resource allocation policy of each service k that maps the context space into actions, and $\widehat{r}(k,\pi_k)$ is given by [?]:

$$\widehat{r}(k,\pi_k) = \lim_{h\to\infty} \frac{1}{h} \sum_{t=1}^{h} \mathbb{E}[\overline{r}(\mathbf{y}^{(t)}, a^{(k,t-1)}) | \overline{y}^{(1)}, a^{(k,0)} \sim \pi_k] \, .$$

In the above expression, $h = 0$ is the time at which the algorithm execution started, and $\overline{y}^{(1)}$ is the mean shared context computed averaging over the shared context observed in the N monitoring slots of the initial decision window. Thus, $\widehat{r}(k,\pi_k)$ is obtained as the average of the reward conditioned on $\overline{y}^{(1)}$ and subsequent actions taken according to policy π_k.



(a) (b)

Fig. 1. KPI evolution with respect to iterations: (a) vRAN and (b) livecast. The dashed dark line shows target KPI values.

Appendix B
Tile coding

Since the context space $\mathcal{X}$ is real, we use a practical method for action-value estimation using function approximation in an F-dimensional space, yielding the approximated function $\hat{q}_{\pi_k}(\overline{\mathbf{y}}^{(h)}, a_k^{(h)}, w) = \sum_{f=1}^{F} w_f s_f(\overline{\mathbf{y}}^{(h)}, a_k^{(h)})$, where $\mathbf{w}$ and $\mathbf{s}(\overline{\mathbf{y}}^{(h)}, a_k^{(h)})$ denote the F-size weight and feature vectors (resp.), with the latter being generated using tile coding [?], which converts the shared context representation into a binary feature vector such that vectors of shared contexts whose values are close by, have a high number of common elements. The continuous space of context variables is tucked up with tiles, and each tile has a corresponding index in the binary feature vector. Several offset grid of tiles, called tilings, are then stacked over the space to create regions of overlapping tiles. Thus, even though the context space by itself is not discretized, its shared representation is discretized as a binary vector for estimation of action values in the SARSA learning algorithm.

Appendix C
Convergence of Alg. 2 to a Pareto-efficient fair Solution

Here, we aim to prove that the iterative algorithm (Algorithm 2 of the manuscript) proposed for the joint allocation of primary and secondary resources is Pareto-efficient fair, and it converges at a sub-linear rate. The proof draws on existing results in [?], as the solution belongs to the class of evolutionary algorithms for solving a multi-criteria optimization problem. Since the multiple criteria may in general be of contradicting nature, the key idea in evolutionary algorithms is to establish dominance relations between different candidate solution sets (suiting different criterion), and select a trade-off solution that best meets all the criteria of the multi-criteria optimization problem. It has been shown that evolutionary algorithms guarantee convergence and diversity in their solutions, and are computationally efficient compared to other existing methods for solving multi-criteria optimization problems.[?].

We recall that in the VERA framework, the multiple criteria in the multi-criteria optimization problem are actually the rewards observed by different services. In view of this, we choose to go ahead with Pareto-dominance as in [?], since in this case a solution set is said to dominate the remaining candidate solution sets if at least one of the criteria (rewards) in the dominant solution is better and the rest may have the same values as in other candidate solution sets. Thus, in a Pareto-optimal solution some of the services gain and no service will lose, i.e., inherently it rules out the possibility of favouring a particular service, at the cost of degrading QoS of other services. It is possible that there are more that one Pareto-efficient solutions to a multi-criteria optimization problem, so to ensure fairness further, we choose that Pareto-dominant solution as final which maximizes the minimum value of criterion function amongst all Pareto-dominant solutions. However, unlike our proof of convergence (Proposition 1-4), [?] only talks about Pareto-efficiency (fairness not included), and does not consider primary and secondary capacity constrained resources and the interdependence between them. Neither does it have a mention about the convergence rate. Thus, our novel contribution here consists in showing through the contradiction argument that the proposed algorithm is Pareto-optimal as well as fair in jointly allocating both the primary and secondary capacity constrained resources.

Proposition 1. *Pareto-efficient allocation of the primary resource: Given a set of coefficients $u_k \geq 0, k \in C$, such that, $\sum_{k \in C} u_k = 1$, then the solution $S^\star = \{c_k^\star\}$, $k \in C$, that maximizes the multi-criteria optimization problem $\sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, is Pareto-efficient.*

Proof. Let $f(\mathbf{\Gamma}_1(S), \ldots, \mathbf{\Gamma}_K(S)) = \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, where $K = |C|$. We show that the solution which maximizes f is Pareto-efficient by contradiction. Let $S^\star$ be a solution to multi-criteria optimization problem $\max_{S \in \widehat{S}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, and is not Pareto-efficient. Then $\exists\, S'$ that dominates $S^\star$, i.e., relation $\mathbf{\Gamma}_i(S') > \mathbf{\Gamma}_i(S^\star), \mathbf{\Gamma}_j(S') \geq \mathbf{\Gamma}_j(S^\star), \forall i, j \in C, i \neq j$ is satisfied with strict inequality for at least one criterion. We first observe that the function f is strictly increasing on each component of the set $(\mathbf{\Gamma}_1(S), \ldots, \mathbf{\Gamma}_k(S)), \forall S \in \widehat{S}_s$, i.e., for each $k, j \in C$, $\mathbf{\Gamma}_k(S) > \mathbf{\Gamma}_k(S') \implies f(\mathbf{\Gamma}_j(S), \mathbf{\Gamma}_k(S)) > f(\mathbf{\Gamma}_j(S'), \mathbf{\Gamma}_k(S')), j \neq k$. Further, since the weights u_k are positive, any Pareto improved solution S' that dominates $S^\star$ would increase f, i.e., $\sum_{k \in C} u_k \mathbf{\Gamma}_k(S') > \sum_{k \in C} u_k \mathbf{\Gamma}_k(S^\star)$. This contradicts the definition of $S^\star$ which states that $S^\star$ is a solution to $\max_{S \in \widehat{S}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, hence $S^\star$ is a Pareto-efficient solution maximizing f. □

Proposition 2. *Pareto-efficient joint allocation of primary and secondary resource: Given a set of coefficients $u_k \geq 0, k \in C$, such that, $\sum_{k \in C} u_k = 1$, then the solution $S^\star = \{c_k^\star, \rho_m^\star(k)\ \forall m \in \mathcal{M}\}, \forall k \in C$, that maximizes the multi-criteria optimization problem $\sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, is Pareto-efficient.*

Proof. Let $\mathbf{\Gamma}_k(S)$ be the criteria function denoting the reward of the k-th service in a monitoring slot following the joint primary and secondary resource allocation strategy S. Since a service may in turn serve M units, it follows that, $\mathbf{\Gamma}_k(S) = \sum_{m \in \mathcal{M}} \Gamma_m(k)(S)$, where $\Gamma_m(k)(S)$ denotes the reward of m-th unit of the k-th service following S. Further, let $f(\mathbf{\Gamma}_1(S), \ldots, \mathbf{\Gamma}_K(S)) = \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, where $K = |C|$. From Proposition 1, it is already proved that the solution which maximizes f is Pareto-efficient with respect to the primary resource c_k. We show that the solution is also Pareto-efficient with respect to the secondary resource $\rho_m(k), \forall m \in \mathcal{M}, \forall k \in C$ by contradiction.

Let $S^\star$ be a solution set to multi-criteria optimization problem $\max_{S \in \widehat{S}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$ that is Pareto-efficient with respect to the primary resource, but not with respect to the secondary resource. Then $\exists\, S'$ that dominates $S^\star$, i.e., relation $\mathbf{\Gamma}_i(S') > \mathbf{\Gamma}_i(S^\star), \mathbf{\Gamma}_j(S') \geq \mathbf{\Gamma}_j(S^\star), \forall i, j \in C, i \neq j$ is satisfied with strict inequality for at least one criterion. We first observe that the function f is strictly increasing on each component of the set $(\mathbf{\Gamma}_1(S), \ldots, \mathbf{\Gamma}_k(S)), \forall S \in \widehat{S}_s$, i.e., for each $k, j \in C$, $\mathbf{\Gamma}_k(S) > \mathbf{\Gamma}_k(S') \implies f(\mathbf{\Gamma}_j(S), \mathbf{\Gamma}_k(S)) > f(\mathbf{\Gamma}_j(S'), \mathbf{\Gamma}_k(S')), j \neq k$. Further, since the weights u_k are positive, any Pareto improved solution S' that dominates $S^\star$ would increase f, i.e., $\sum_{k \in C} u_k \mathbf{\Gamma}_k(S') > \sum_{k \in C} u_k \mathbf{\Gamma}_k(S^\star)$. This contradicts the definition of $S^\star$ which states that $S^\star$ maximizes f, hence $S^\star$ is a Pareto-efficient solution with respect to primary as well as secondary resource allocation. □

Proposition 3. *Alg. 2 converges to a Pareto-efficient solution set at a sub-linear rate.*

Proof. The multi-criteria optimization problem defined as $\max_{S \in \widehat{S}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, where $\widehat{S}_s$ is the set of feasible solutions having cardinality $|\widehat{S}_s|$, and $\mathbf{\Gamma}_i$s are the criterion functions corresponding to the reward from each service. Let S be a solution from set of feasible solutions $\widehat{S}_s$, then $S = \{c_k, \rho_m(k)\ \forall m \in \mathcal{M}\}, k \in C$ s.t. $\sum_{k \in C} c_k \leq B_c$ and $\sum_{m \in \mathcal{M}} \rho_m(k) \leq B_\rho$, where each c_k is the primary resource allocated to service k, $\rho_m(k)$ is the secondary resource allocated to m-th unit of k-th service, B_c is the maximum computation capacity of the server, and B_ρ is total number of available RBs. The solution update proposed in Alg. 2 is similar to an iterative multi-objective search and update algorithm [?] where, at each iteration, the efficient solution to the multi-objective search problem replaces other solutions in the Pareto dominant solution set. Therefore, based on Sec. 4.3 and Thm. 2 in [?], convergence is ensured for the proposed Alg. 2.

Given the capacity-constrained CPU allocation is expressed in discrete units, in the worst case there can be many ways to obtain $S = \{c_k, \rho_m(k)\ \forall m \in \mathcal{M}\}, k \in C$, which may further increase as the number of services or UEs grow. However, subject to maximum capacity constraint, only those solutions are feasible that adhere to $\sum_{k \in C} c_k \leq B_c$ and $\sum_{m \in \mathcal{M}} \rho_m(k) \leq B_\rho$. This effectively reduces the possible number of solutions, thereby limiting the iterations executed in multi-criteria optimization problem over the set of feasible solutions $\widehat{S}_s$. Thus, the complexity of multi-criteria optimization problem is given by $O(|\widehat{S}_s|)$.

Let S_t be the solution to the multi-criteria optimization problem $\max_{S \in \widehat{S}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$ at iteration t, and the Pareto-efficient solution given by $S^\star$ is achieved at iteration $t^\star$. From Alg. 2, we know that at any iteration t, a solution will be added to $\mathcal{S}_d$ only if it dominates the existing solutions in $\mathcal{S}_d$. When a solution S dominates another possible solution S', then by definition of dominance it is mandatory that there is a Pareto improvement, such that some $\mathbf{\Gamma}_i$s will gain and none of them will loose. Consequently, (2) is satisfied over successive iterations, i.e., if $f(\mathbf{\Gamma}_1(S), \cdots, \mathbf{\Gamma}_k(S)) = \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$, where $K = |C|$, then,

$$f^\star(\mathbf{\Gamma}_1(S^\star), \cdots, \mathbf{\Gamma}_k(S^\star)) \geq f_{t+1}(\mathbf{\Gamma}_1(S_{t+1}), \cdots, \mathbf{\Gamma}_k(S_{t+1})) \geq f_t(\mathbf{\Gamma}_1(S_t), \cdots, \mathbf{\Gamma}_k(S_t))$$

It follows that

$$f^\star - f_{(t+1)} \leq f^\star - f_t \quad \text{and} \quad \frac{f^\star - f_{(t+1)}}{f^\star - f_t} \leq 1\,. \qquad (2)$$

Since $\frac{f^\star - f_{(t+1)}}{f^\star - f_t} \leq 1$ for $1 \leq t < t^\star \leq |\widehat{\mathcal{S}}_s|$, it implies that $f_{(t+1)}$ is closer to $f^\star$ compared to f_t, thus the algorithm is converging with increasing iterations at sub-linear rate. □

Proposition 4. *Fairness of Pareto-efficient primary resource allocation: The solution* $S^\star = \{c_k^\star, \rho_m^\star(k) \ \ \forall m \in \mathcal{M}\}, \forall k \in C$ *obtained using Algorithm 2 is fair with respect to primary resource allocation* $c_k^\star, \forall k \in C$.

Proof. Considering only the primary resource allocation $c_k, \forall k \in C$. Set $\mathcal{S}_d$ (in Alg. 2) includes the dominant Pareto-efficient solutions to the multi-criteria optimization problem $\max_{S \in \widehat{\mathcal{S}}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$. We prove that solution S^*, which maximizes $\min_{k \in C}(u_k \mathbf{\Gamma}_k(S^\star))$, is fair with respect to each criterion, by using contradiction. Let us assume that a Pareto-efficient solution $S \in \mathcal{S}_d$ is biased towards criterion $\mathbf{\Gamma}_i$, i.e., $\mathbf{\Gamma}_i(S) \gg \mathbf{\Gamma}_j(S), 1 \leq i, j \leq |C|, i \neq j$. Then, step 8 in Alg. 2 would maximize $u_j \mathbf{\Gamma}_j(S)$. This maximization ensures that for a Pareto-efficient solution $S^\star$, $\mathbf{\Gamma}_i(S^\star) \approx \mathbf{\Gamma}_j(S^\star)$. Thus, all solutions that are biased towards any single criteria are ruled out. Consequently, the Pareto-efficient solution $S^\star$ obtained in Alg. 2 is fair with respect to every criterion. □

Proposition 5. *Fairness of Pareto-efficient secondary resource allocation in the vRAN: For a given fair primary resource allocation* $c_k^\star$ *in the solution* $S^\star = \{c_k^\star, \rho_m^\star(k) \ \ \forall m \in \mathcal{M}\}$*, for* $k = 2$ *(denoting the vRAN service),* $S^\star$ *is fair with respect to secondary resource allocation* $\{\rho_k^\star(m) \ \ \forall m \in \mathcal{M}\}$.

Proof. Let the QoS satisfaction for m-th UE in the vRAN service be defined as $g(m) := l_t(m) - l_i(m)$, where $l_t(m), l_i(m)$ denote (resp.) the target traffic load and the instantaneous rate achieved by m-th UE. Set $\mathcal{S}_d$ (in Algorithm 2) includes the dominant Pareto-efficient solutions to the multi-criteria optimization problem $\max_{S \in \widehat{\mathcal{S}}_s} \sum_{k \in C} u_k \mathbf{\Gamma}_k(S)$. Here, we prove by the contradiction approach, that the solution $S^\star$ to the said multi-criteria optimization problem, wherein for a given fair primary resource allocation $c_k^\star$, the secondary resource allocation $\{\rho_m^\star(k) \ \ \forall m \in \mathcal{M}\}$ is obtained by minimizing $\max_{m \in \mathcal{M}} g(m)$ over all $S \in \mathcal{S}_d$, is fair with respect to the secondary resource allocation $\{\rho_m^\star(k) \ \ \forall m \in \mathcal{M}\}$.

Let us assume that a Pareto-efficient solution $S \in \mathcal{S}_d$ is fair with respect to the primary resource c_k, however it is biased in the secondary resource allocation $\{\rho_m(k) \ \ \forall m \in \mathcal{M}\}$. It follows that $\Gamma_k^{(i)}(S) \gg \Gamma_k^{(j)}(S), 1 \leq i, j \leq |\mathcal{M}|, i \neq j$. Then, step 8 in Algorithm 2 picks $S^\star$ such that for a given fair $c_k^\star$, the secondary resource allocation $\{\rho_k^\star(m) \ \ \forall m \in \mathcal{M}\}$ is obtained by minimizing $\max_{m \in \mathcal{M}} g(m)$ over all $S \in \mathcal{S}_d$. Here we recall that for vRAN service, the criteria $\Gamma_m(k)(S)$ denotes the reward of m-th UE following the resource allocation strategy S. Clearly, $\Gamma_m(k)(S)$ is a function of vRAN KPIs, i.e., latency and packet loss, and the QoS satisfaction $g(m)$ is indicative of observed KPI values for the m-th UE. To this end, the minimization ensures that for a Pareto-efficient solution $S^\star$, $\Gamma_k^{(i)}(S^\star) \approx \Gamma_k^{(j)}(S^\star)$. Thus, all solutions that are biased in the secondary resource allocation towards any single UE are ruled out, and the Pareto-efficient solution $S^\star$ obtained in Algorithm 2 is fair with respect to the secondary resource allocation, given a fair primary resource. □